1-2016

Use of Qualitative Behaviour Assessment as an Indicator of Welfare in Donkeys

Michela Minero
University of Milan

Emanuela Dalla Costa
University of Milan

Francesca Dai
University of Milan

Leigh Anne Margaret Murray
University of Milan

Elisabetta Canali
University of Milan

See next page for additional authors

Recommended Citation

Authors

Michela Minero, Emanuela Dalla Costa, Francesca Dai, Leigh Anne Margaret Murray, Elisabetta Canali, and Françoise Wemelsfelder

Use of Qualitative Behaviour Assessment as an Indicator of Welfare in Donkeys

Michela Minero[1], Emanuela Dalla Costa[1], Francesca Dai[1], Leigh Anne Margaret Murray[1], Elisabetta Canali[1], Francoise Wemelsfelder[2]

[1] *University of Milan*
[2] *Scotland's Rural College*

ABSTRACT

One of the objectives of the Animal Welfare Indicators project was to develop animal-based indicators to assess donkey welfare, including their emotional state. This study aimed to develop a fixed rating scale of Qualitative Behaviour Assessment (QBA) for donkeys, to evaluate the inter-observer reliability when applied on-farm, and to assess whether the QBA outcomes correlate to other welfare measures.

A fixed list of 16 descriptors was designed on the basis of a consultation in a focus group. The fixed list was then used by four trained observers to score nine 2 min videos of groups of donkeys owned by six farms and on-farm to score 11 donkey facilities representative of the most common type of donkey facilities in Western Europe. On each farm one experienced assessor collected different welfare measures on all the adult donkeys. The QBA scores and welfare measures were analysed using Principal Component Analysis (PCA, correlation matrix, no rotation). Kendall's W and ANOVA were used to assess inter-observer reliability. PCA revealed three main components explaining 79% of total variation between them. PC1 ranged from at ease/relaxed to aggressive/uncomfortable, suggesting that this Component is important in the description of the valence of donkeys' affective states. PC2 was more related to the level of arousal of donkeys, ranging from apathetic to distressed/responsive. The four assessors showed a good level of agreement on the first two dimensions of the PCA (Kendall's W varying from 0.61 to 0.90), and there was no significant effect of observer on donkey QBA scores (ANOVA p > 0.05), both for the videos and on-farm. PCA of all measures together showed positive QBA descriptors on PC1 (relaxed, at ease, happy, friendly) to be associated with positive human–donkey interaction indicators (absence of tail tuck, no avoidance, and positive reaction to an assessor walking down the side of the donkey).

Our findings suggest that QBA is a suitable tool to identify the emotional state of donkeys on-farm. A fixed list of descriptors can be used consistently by different trained assessors as a valid addition to a number of animal welfare assessment indicators.

1. Introduction

In Europe, even if the use of donkeys in agriculture has consider-ably decreased after World War II, from the year 2000 the number of these animals was reported to be growing (Faostat, 2011), thanks to their

adaptability to very different types of activity, ranging, for instance, from onotherapy, to garbage collection. Given its versatility, the donkey sector is fragmented and gathering objective information regarding donkey welfare may be challenging. Over the last decade a lot of effort was placed in developing valid and objective methods to assess animal welfare on-farm (EFSA Panel on Animal Health and Welfare, 2012; Knierim and Winckler, 2009; Visser et al., 2014). One of the aims of Animal Welfare Indicators(AWIN), an international animal welfare research project funded by EU FP7, was to develop welfare indicators that are supported by scientific evidence for donkeys, among other species ("Animal Welfare Indicators project," 2012). These welfare indicators are largely animal based and reflect the animal's perception of its situation (EFSA Panel on Animal Health and Welfare, 2012). Positive welfare indicators consider the presence of positive emotions and present the advantage of enabling a better communication of the commitment to reach "higher welfare" standards in a more proactive manner. However, investigating affective states of animals might be a difficult task, especially when the evaluation has to be performed on-farm. Differently from humans, where verbal language helps to assess emotional experiences, in animals only behavioural and physiological measurements help to evaluate the emotions that are assumed to correspond to "opportunity situations" where the pleasure conferred by being able to perform a behaviour or enjoy a resource motivates the animal (Berns et al., 2012; Fraser and Duncan, 1998).

Qualitative Behaviour Assessment (QBA) is a relatively new scientific method to evaluate the expressive quality of animal behaviour and emotions. It integrates and summarises the different aspects of an animal's dynamic style of interaction with the environment and can be used in addition to other welfare indicators or classical ethological measures (Wemelsfelder et al., 2000).The use of QBA enables the identification of the main dimensions of mood states (Mendl et al., 2010) and facilitates bridging the gap that traditionally exists between subjective judgments and scientific measurement approaches (Minero et al., 2009; Wemelsfelder,2007). This method relies on the ability of humans to integrate observed details of behaviour and to address the animal's experience through the expressive nature of its dynamic demeanour. QBA scoring uses a selected list of terms to describe the different elements of an animal's expressive repertoire (Wemelsfelder,2007). These terms have an expressive, emotional connotation and can be individually generated by observers, as in the case of the Free-Choice-Profiling methodology (FCP), or they are chosen by researchers first from literature and then discussed in focus groups of experts and tested on-farm (Andreasen et al.,2013). FCP is unsuitable for on-farm welfare assessment, as it requires a minimum of 10 observers and extensive data analysis; hence, the second approach using a fixed list of terms is generally adopted for on-farm assessment. In research, when using both quantitative and qualitative methods for assessing behaviour, it is essential to avoid anthropomorphism and possible observer bias; for this reason it is fundamental to know the animal species, in this case donkeys, well. A growing body of research indicates that QBA can be rigorously applied to answer different research questions in horses (Fleming et al., 2013; Minero et al., 2009;Napolitano et al., 2008) and other farm animals (Bassler et al.,2013; Napolitano et al., 2012; Rousing and Wemelsfelder, 2006;Rutherford et al., 2012; Wemelsfelder, 2012; Wemelsfelder et al.,2000). Many studies showed that, when correctly applied, QBA shows good correlations to behavioural, physical and physiological measures, thus confirming the validity of the observers' assessments (Brscic et al., 2009; Rutherford et al., 2012; Stockman et al.,2011; Wickham et al., 2012). Importantly, QBA is reported to be a method that can either be applied retrospectively, e.g. to assess animals on video footage, or has the potential for immediate use, for example in on-farm welfare assessments (Fleming et al., 2013).To date, no authors have yet published works where QBA has been applied to donkeys. QBA has the potential to indicate the positive aspects of the welfare, however most researchers agree that, with welfare being a complex multidimensional concept, no single indicator can be considered as an exhaustive system to evaluate the welfare of animals, and it is always preferable to integrate and cross-validate QBA with other measures of

welfare. In fact, QBA cannot be used as a stand-alone welfare indicator, as it does not cover all the aspects of the welfare of the animal (Andreasen et al., 2013).

The aim of the present study was to develop a fixed QBA rating scale for donkeys and to evaluate the inter-observer reliability of trained assessors using the fixed QBA rating scale from videos and on-farm. Furthermore, we aimed to assess if the QBA outcomes correlate to other measures of donkey health and welfare, taken at the same time on the same farms.

2. Material and methods

2.1. Development of the rating scale

A first selection of QBA descriptors was made from a list of terms derived from papers where qualitative terms were used to describe donkey behaviour. The list contained 27 terms, given in English, that were then discussed during a focus group.

2.2. The focus group

On February 7th and 13th 2013 a focus group on Qualitative Behaviour Assessment (QBA) in donkeys took place on-line, thanks to technical support offered by the University media technology department. A group of seven people internationally acknowledged for their expertise with donkeys (veterinarians, breeders, donkey welfare experts) were selected as partners and collaborators in the AWIN project and participated in the focus group. At the start of the first session the QBA method was introduced to them, and it was explained how to assess animal behaviour as expressive body language.

The participants then discussed the list of 27 descriptors chosen from the literature on donkeys, and agreed on a brief general characterisation of each selected term. Participants were asked to give examples and describe situations in which different terms could be used, to be able to create term characterisations that were widely applicable. In a second round, they refined some of the term descriptions and removed 12 terms which they felt may be difficult to interpret, or may not be very relevant to the on-farm assessment of donkey welfare. Participants discussed the possible differences in the interpretation of descriptors between different languages; in order to overcome linguistic barriers, the use of English as well as bilingual dictionaries proved useful for reaching consensus among participants about the brief characterisations of the terms.

As a practical exercise, the participants of the focus group then watched seven videos of donkeys, filmed individually or in groups for 1 min, and used QBA to score them using the list of descriptors. After this, one term was added to the list and was given a characterisation. It was agreed that this list of 16 terms (Table 1) would be used to score donkeys at 11 farms/facilities in Italy, with the understanding that following observations carried out on farms, it may be revised further. The QBA rating scale to be tested was construed by putting each of the descriptors next to a continuous visual analogue scale of 125 mm length where the terms 'minimum' (this expressive quality is absent) and 'maximum' (this quality could not be present more strongly) represented the ends of the scale.

2.3. Training of assessors

The four assessors were all female, aged between 25 and 36 years, consisting of two veterinarians who were researchers in the field of applied ethology, and two zoologists. Before the first assessment, the four assessors, all experienced with donkeys, and skilled in assessing animal behaviour, were made familiar with the concept of QBA by reading relevant scientific literature and participating as auditors in the focus group. The training of assessors then consisted of two subsequent phases: a dialogue on the meaning of

each descriptor, starting from the definition produced by experts, and a discussion among assessors about how they would quantitatively score the responsiveness styles of donkeys, based on the intensity of their behaviour observed both from videos and live. Finally, to test inter-observer reliability of the QBA term list, assessment was carried out from nine video clips selected because representative of the donkey expressive repertoire, each of 2 min duration and showing groups of donkeys owned by six farms. The number of donkeys in each video varied from two to 20. Assessors scored the videos in a classroom, independently and without talking to each other during the entire procedure.

Table 1. List of QBA descriptors and definitions agreed by the expert focus group.

Aggressive	Behaving in an angry or rude way, fighting or attacking another donkey
Agitated	Restless, an animal can stand still and be agitated, fidgety, worried or upset, excited, disturbed, troubled
Anxious	Worried/tense, troubled, apprehensive, distressed
Apathetic	Having or showing little or no emotion; indifferent
At ease	In a relaxed attitude or frame of mind
Curious	Eager to learn, inquisitive, wishing to investigate
Distressed	Much troubled, upset, afflicted, panicking
Fearful	Having fear, afraid, even not linked with something going on in the environment, flight response, look anxious, back up/away, not move further.
Friendly	On the same side; not hostile, showing positive feelings toward another animal or person/the donkey approaches another animal/person and expressing grooming behaviour
Happy	Feeling, showing or expressing joy, pleased
Playful	Very active, happy, and wanting to have fun, mischievous
Pushy	Offensively assertive or forceful, bossy, dominant
Relaxed	To make less tense or rigid
Responsive	Receptive, aware of the environment
Uncomfortable	Not comfortable, not relaxed
Withdrawn	Secluded or remote, shy, not searching for contact with others

2.4. Farm visits

2.4.1. Qualitative Behaviour Assessment

QBA assessments were carried out on 11 donkey farms representative of the most common type of donkey facilities in Western Europe: four dairy donkey farms, three facilities where donkeys were used for Animal Assisted Activities, one donkey sanctuary and three farms where donkeys were kept as companion animals. The average number of animals per farm was 20 (min 10 max 150).The assessments were performed in the morning, in case of dairy donkey farms at least 2 h after milking. Straw or hay was always available to the animals. Assessors were expressly unaware of the different backgrounds of the farms, they had never entered them before and did not have any expectations about the outcome of the assessment. QBA took place immediately after entering the farms and letting the animals adapt to the observers' presence. The four assessors were always dressed in the same type and colour of clothes at all the farms. The assessment took place outside of the paddocks where animals were kept, without disturbing them. On each farm, observers assessed all adult donkeys at the same time without talking to each other, observation sessions lasted from 10 to 15 min. Depending on how the farm

was structured, observers needed to move in order to be able to observe all animals, so one or two points of observation per farm were used. After observing the donkeys, the assessors moved to a place where they were not visible to the animals and independently scored the animals on the 16 qualitative descriptors. For this they used a dedicated electronic Android application, specifically developed at SRUC for QBA automated data recording and analysis. The assessors ticked the visual analogue scale next to each descriptor at the appropriate point. The score was automatically recorded as the measure of the distance in millimeters between the left 'minimum' point of the scale and the point where the observer's thick crossed the line. Thus, for each observer and each farm, a data spreadsheet was automatically created containing the scores of observed donkeys on each of the 16 qualitative descriptors.

2.4.2. Welfare assessment

A further welfare assessment was carried out after completing QBA scoring. On each farm, one trained assessor scored all adult donkeys individually. Data was gathered on relevant animal based indicators selected or developed by AWIN researchers (Table 2),related to the four principles used in the Welfare Quality® frame-work (Dalla Costa et al., 2014).

Table 2. List and description of animal-based measures, their score and prevalence at farm level.

Measure taken on individual donkeys	Description	Score	Prevalence of measure at farm level
Ear position	Ear position while assessed by the observer	Relaxed, flat back-aggressive	Relaxed ears: proportion of donkeys with relaxed ears
BCS	Body Condition Score	1–5 according to Quaresma, Payan-Carreira, & Silva (2013)	BCS = 3: Proportion of donkeys with good BCS (score = 3)
Skin lesions	Presence of skin lesions (alopecia, superficial or deep wounds)	Yes, no	No lesions: proportion of donkeys with no skin lesions
Joint swellings	Presence of joint swellings	Yes, no	No joint swellings: proportion of donkeys with no joint swellings
Hoof condition	Presence of signs of neglecting e.g. hoof overgrowth	No signs of neglecting, clear signs of overgrowth	Good hooves condition: proportion of donkeys with no signs of neglecting of the hooves
AD	Presence of any avoidance distance behaviour while approached	Distance (cm) of the first avoidance behaviour	No AD: proportion of donkeys with no avoidance signs (0 cm)
WDS	Walking down the side of the donkey towards its tail and assess the behavioural reaction	Negative or neutral/positive reaction	Positive WDS: proportion of donkeys with neutral/positive reaction
Tail tuck	Presence of tail tuck during WDS	Yes, no	No tail tuck: proportion of donkeys with neutral/positive reaction (no tail tuck)

2.5. Statistical analysis

When the video scoring and the farm visits had been completed, the QBA scores provided by the four assessors were automatically downloaded from the QBA App to an Excel file. The other animal based welfare measures collected on single donkeys were aggregated at farm level as described in Table 2 and entered in an Excel file. IBM SPSS Statistics 21 software (IBM Corp., 2012) was used for statistical

analysis. Data was tested for normality using a Kolmogorov-Smirnov test. As variables were not normally distributed, the scores were transformed using $x'_{ij} = \log(1 + x_{ij})$ transformation. To analyze QBA scores, a Principal Component Analysis (PCA, correlation matrix, no rotation) of the four observers' data sets was conducted separately for every phase of the research (videos and on-farm assessment). The PC scores attributed to the animals on the observed farms on the first three main Principal Components were then tested for inter-observer reliability, using Kendall Correlation Coefficient W. Kendall *W* values can vary from 0 (no agreement at all) to 1 (complete agreement), with values higher than 0.6 showing substantial agreement. In order to test whether there were any significant effects of observer on the PCA farm scores, a one-way analysis of variance was conducted on the PC1 and PC2 scores of the four observers (separately for videos and on-farm assessments), with observer as fixed effect and farm as random factor. Subsequently the inter-observer reliability for each descriptor separately was calculated using Kendall's *W*. To assess how Qualitative Behaviour Assessment related to the other animal-based welfare indicators, welfare measures aggregated at farm level and QBA scores were merged in a new file and analysed using a Principal Component Analysis (PCA, correlation matrix, no rotation).

Table 3. PCA outcomes and inter-observer reliability for the QBA rating scales.

		PCA Factor 1	PCA Factor 2
Videos	% of variation explained	40%	16%
	Kendall's W (N = 4, df = 8)	0.90	0.66
On-farm assessment	% of variation explained	45%	14%
	Kendall's W (N = 4, df = 8)	0.61	0.69

3. Results

3.1. Reliability testing

Regarding the results on inter-observer reliability of QBA, Table 3 reports the variance explained by the first two Principal Components of PCA analysis (separate for videos and on-farm assessment), and the Kendall's W values for the four observer scores on these components. The assessors overall showed a good level of agreement for the first two PCA components, with W varying between 0.61 and 0.90. There was no significant effect of the observer on mean QBA scores on either dimension (ANOVA PC1 $F = 2.22$; $p = 0.11$; ANOVA PC2 $F = 1.32$; $p = 0.28$), indicating that observers not only ranked the different donkey farms in similar ways, but also gave them similar scores on the rating scales. A significant observer effect was only found on the third PC (ANOVAPC3 $F = 12.5$; $p = 0.000$) suggesting that this component explains the residual variance ascribable to human individual variability in assessment processes. ANOVA indicated that there was a significant effect of farm on both PC1 and PC2 (ANOVA PC1 $F = 10.68$; $p = 0.00$; ANOVA PC2 $F = 5.39$; $p = 0.00$, ANOVA PC3 $F = 2.48$; $p = 0.02$) indicating that donkeys housed in a given farm were perceived as in a different emotional state from donkeys in other farms.

Table 4 shows the Kendall's *W* values for each of the QBA don-key descriptors separately. For video assessments, the assessors showed good overall agreement with 13 out of 16 descriptors showing Kendall *W* values higher than 0.60. For on-farm assessment, the observers' agreement in using single descriptors varied depending on whether all 11 farms visited were analyzed together(for seven out of 16 terms *W* > 0.60), or only the last six farm (for 12 out of 16 terms *W* > 0.60), indicating the importance of growing experience in reaching agreement on the use of single terms.

Due to this good overall inter-observer reliability in applying QBA, and considering also the increasing agreement in the observers' scores for separate terms after some on-farm experience, in the subsequent analysis of QBA assessment of the farms we considered only the data of one observer scoring the 11 donkey farms visited.

Table 4. Kendall's *W* correlation coefficients for all descriptors used by four observers from videos and on-farm. Values larger than 0.60 are bold typed.

Descriptor	Kendall's *W*		
	Videos	**11 donkey facilities**	**last 6 donkey facilities entered**
Aggressive	0.36	**0.70**	**0.86**
Agitated	**0.60**	**0.63**	**0.74**
Anxious	**0.56**	0.31	0.22
Apathetic	**0.67**	0.42	**0.58**
At ease	**0.84**	0.51	**0.76**
Curious	**0.85**	**0.65**	**0.84**
Distressed	**0.65**	**0.69**	**0.82**
Fearful	0.55	0.39	0.33
Friendly	**0.75**	0.51	**0.61**
Happy	**0.91**	0.49	**0.66**
Playful	**0.60**	0.51	0.51
Pushy	**0.71**	**0.60**	**0.74**
Relaxed	**0.79**	0.51	**0.71**
Responsive	0.50	0.29	0.48
Uncomfortable	**0.84**	**0.58**	**0.67**
Withdrawn	**0.63**	**0.70**	**0.82**

3.2. Outcomes for QBA assessment of the farms

The PCA of the QBA assessments for the 11 farms visited identified five main factors with Eigen value greater than 1; the first three Components together explain a good proportion of variation (79.00%) between donkey farms. Fig. 1 shows the distribution of the descriptors along the first two PCA factors. Many of the terms load strongly on the first Principal Component accounting for 43.70%of the total variance and range from at ease/relaxed to aggressive/uncomfortable, suggesting that this Component is important in the description of the valence of donkeys' affective states. Animals with high positive scores on this Component can be described as in a more positive emotional state than donkeys with high negative scores. Component 2 counts for 22.49% of variance and seems to be more related to the level of arousal of donkeys ranging from apathetic to distressed/responsive. The third Component, counting for 12.80% of the total variance is characterized by anxious/withdrawn and playful with opposite signs. As play is certainly linked with a good relationship with other group mates and a positive emotional state, donkeys with high negative scores on the third Component can be described as much more in harmony with their mates and the environment they live in.

3.3. Relationship between QBA and other welfare indicators

QBA scores were analyzed together with the other welfare indicators measures through PCA, revealing three main components explaining 71.79% of total variation between donkey farms (Table 5). QBA descriptors appear to be correlated to some welfare measures: PC1 shows that QBA descriptors linked

with positive emotional state (relaxed, at ease, happy, friendly) are associated with 'no tail tuck', 'positive WDS response', and 'no avoidance distance from the assessor'. On PC2, QBA descriptors indicating higher arousal (distressed, agitated, responsive, playful) seem to be aligned with 'relaxed ears', an apparently contradictory outcome. On PC3, good health indicators such as 'no lesions', 'good hooves condition' and 'no joint swellings' appear associated with curious and playful demeanour, linking good health with more positively assertive expressions.

4. Discussion

The first objective of this research was to develop a fixed QBA rating scale for on-farm assessment of welfare of donkeys. In the focus group there was great discussion on the differences in interpretation between different languages, with many analogies being used to convey particular descriptors. Participants reported that the discussion was very useful and suggested that assessors take time to discuss the terms on a list in order to develop a common understanding of these terms. The outcome of this discussion was a fixed list composed of seven positive and nine negative descriptors and brief characterisations of each term. To date, this was the first time that comprehensive characterisations for each descriptor of a fixed list of QBA terms for a particular species was generated as an aid for new assessors. The participants suggested that the fixed list could be adapted should assessors be interested in evaluating donkeys under situations very dissimilar to the ones described in the present study.

A central characteristic for any measurement tool is consistency in measurements when applied by different assessors (Martinand Bateson, 2007). According to the results, assessors consistently used the QBA descriptors to distinguish between expressions of positive and negative donkey emotions in an on-farm environment, confirming what was already known for other species (Minero et al., 2009; Napolitano et al., 2012; Phythian et al., 2013; Rutherford et al., 2012; Sant'Anna and Paranhos da Costa, 2013). A holistic approach underpins the QBA method, which essentially refers to dynamic patterns of demeanour by the whole animal, and so descriptors have to be considered together, as an integrated whole in describing patterns of demeanour. Following from this, the agreement on the use of single terms should not be regarded as paramount. Assessors in this study mostly reached satisfactory agreement using individual QBA descriptors when scoring videos, but they found it more difficult to score some terms (i.e. friendly, happy, playful) in a similar way when on-farm. One possible explanation is that scoring live is more complex, and poses different challenges from scoring from videos, and that more training and experience are needed on-farm in order to reach a better level of agreement. This supposition is confirmed by the results of the analysis performed on the last six farms visited, where the level of agreement for individual descriptors improved, with only three of them (anxious, fearful and responsive) showing a lower level of agreement and one (playful) showing a moderate level of agreement. Thus, on-farm training of new assessors proved to be efficient and useful in improving the reliability of individual terms. These results highlight the importance of the use of a clear definition of descriptors and training in the use of a fixed list of QBA terms for on-farm welfare assessment (Bokkers et al., 2012; Meagher, 2009). The choice of assessors is a key element as well: in this study all of them had good previous experience in observing the behaviour of donkeys. As QBA works by relative comparison of samples and depends on contrasting expressions to anchor quantification of intermediate values, it can be suggested that the more different samples are, the better the method works. This is fundamental especially during on-farm training of assessors. In the present study, the voluntary participation of donkey facilities may have interfered with the variability and representativeness of the farm sample. In fact, it might be argued that only facilities achieving acceptable welfare of donkeys would intentionally take part in a study on welfare assessment. Previous QBA studies conducted on-farm report that they also may have been limited by this factor (Andreasen et al., 2013). In future on-farm studies it would be preferable to enlarge

the number of visited farms and to make sure that the selected sample of farms shows a sufficiently large spread in levels of welfare.

Fig. 1. Loadings of QBA descriptors along the first two PCA factors.
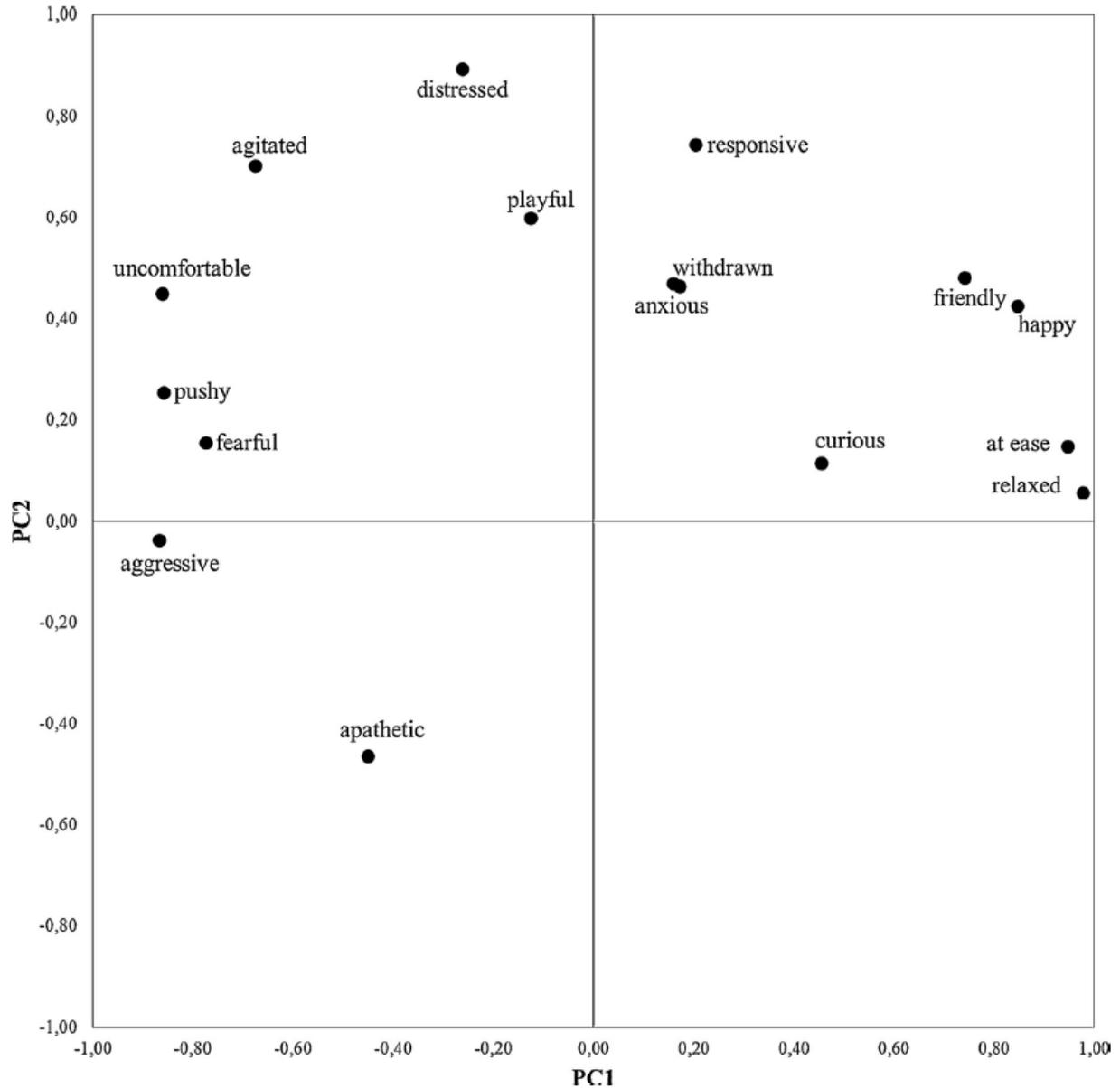
The relation found between positive QBA PC1 'mood' scores and the absence of any avoidance distance from an approaching human, an absence of tail tuck, or any negative reactions to the Walking Down Side test, reinforces what has previously been observed in other species (Brscic et al., 2009; Ellingsen et al., 2014). Ellingsen et al. (2014) found that cows described as tense, fearful, scared and nervous were primarily handled by stockpersons that were scored as aggressive/dominating and insecure/nervous. On the other hand, confident, calm and friendly cows were handled by calm/patient stockpersons and received more positive inter-actions (e.g. talking quietly, petting and touching). It is worth noticing that in

the present study, QBA was performed as the first assessment on-farm, on undisturbed animals, before taking any other welfare measure or interviewing the farmer. This means, for instance, that observers rated donkeys as fearful or friendly before having seen their reaction to humans. Furthermore, assessors were expressly unaware of the farms welfare characteristics before entering them. The association of high PC2 'arousal' scores with relaxed ears does not seem easy to explain, however the 'relaxed ear' measure is primarily in contrast to a 'flat-back aggressive' measure, and the QBA term 'aggressive' does not load at all on PC2. So this association perhaps rather arbitrarily reflects the absence of aggression in the donkeys' agitated and/or playful demeanour. The association between curiosity, playfulness and good health on PC3 seems to make more sense, and may indicate how good health can have wider positive implications for animal well-being.

Table 5. Total variance explained and Loadings for QBA descriptors and other donkey welfare measures on PC1 and PC2 of a combined PCA. The highest loadings for each factor are bold typed.

	PC1	PC2	PC3
Eigen value	10.31	4.25	2.65
% of variance explained	43.00	17.75	11.04
% cumulative variance explained	43.00	60.75	71.79
Items	**PC1**	**PC2**	**PC3**
Agitated	-0.563	**0.728**	-0.001
Aggressive	**-0.877**	0.051	0.196
Anxious	0.129	0.489	-0.096
Apathetic	-0.512	-0.368	-0.306
At ease	**0.947**	0.068	0.239
Curious	0.314	0.122	**0.654**
Distressed	-0.131	**0.888**	0.101
Fearful	**-0.792**	0.229	-0.207
Friendly	**0.736**	0.410	0.103
Happy	**0.887**	0.328	0.247
Playful	-0.055	**0.598**	**0.492**
Pushy	**-0.815**	0.341	0.137
Relaxed	**0.992**	-0.038	-0.021
Responsive	0.325	**0.699**	**-0.539**
Uncomfortable	**-0.798**	0.501	-0.104
Withdrawn	0.161	0.522	-0.414
Relaxed ears	-0.452	**0.554**	0.033
No AD	**0.846**	0.085	0.085
Positive WDS	**0.883**	0.337	0.068
No Tail tuck	**0.931**	0.090	0.058
BCS = 3	-0.637	0.104	0.433
No joint swellings	-0.435	0.306	**0.510**
Good hooves condition	0.247	-0.056	**0.591**
No lesions	-0.589	-0.430	**0.595**

Welfare is a complex concept, that encompasses different aspects of physical and mental health of animals (Broom, 2011).All these aspects are important and at least partially independent, and for this reason welfare assessment cannot be summarized by just assessing one indicator, as was also suggested for dairy cows by Andreasen et al. (2013). Notably, while the physical health of donkeys (and other farm animals) can be monitored with the use of indicators that are already in place, at the moment there are no other objective and feasible measures to assess their emotional state. An interesting aspect of QBA is that it mostly relies on long-standing engagement and experience with a particular species, rather than on particular professional qualifications or expertise, which gives it a relatively wide range of application. After suit-able training on its use, QBA can allow stockpersons to reach a better level of awareness about the emotional state of their animals; welfare assessors can detect subtle shifts in demeanour that may be overlooked by isolating and quantifying individual physical behaviours, but that nonetheless are important for welfare assessment.

5. Conclusions

In conclusion, our findings suggest that Qualitative Behaviour Assessment is a suitable tool to identify the emotional state of donkeys on-farm. A fixed list of descriptors was created and used consistently by different trained assessors as a valid addition to a number of other donkey welfare indicators. However, our results also indicate that it is important to invest time in training assessors, to ensure that both their interpretation of terms and their use of the visual analogue scales are properly aligned. The future development of on-line training material, including reference videos scored by experts, would further increase the applicability of the method in general practice. QBA appears particularly useful in evaluating positive aspects of the life of donkeys, adding complementary and relevant information to their on-farm welfare assessment.

Conflict of interest

All authors of the manuscript "Use of Qualitative Behaviour Assessment as an indicator of welfare in donkeys" declare no actual or potential conflict of interest including financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could inappropriately influence, or be perceived to influence, their work.

Acknowledgements

References

Andreasen, S.N., Wemelsfelder, F., Sandøe, P., Forkman, B., 2013. The correlation of Qualitative Behavior Assessments with Welfare Quality® protocol outcomes inon-farm welfare assessment of dairy cattle. Appl. Anim. Behav. Sci. 143, 9–17.

Animal Welfare Indicators project [WWW Document], 2012. URL www.animal-welfare-indicators.net.

Bassler, A., Arnould, C., Butterworth, A., 2013. Potential risk factors associated with contact dermatitis, lameness, negative emotional state, and fear of humans in broiler chicken flocks. Poultry Sci., 2811–2826.

Berns, G.S., Brooks, A.M., Spivak, M., 2012. Functional MRI in awake unrestrained dogs. PloS One 7, e38027.

Bokkers, E., de Vries, M., Antonissen, I., de Boer, I., 2012. Inter- and intra-observer reliability of experienced and inexperienced observers for the Qualitative Behaviour Assessment in dairy cattle. Animal Welfare 21, 307–318.

Broom, D.M., 2011. A history of animal welfare science. Acta Biotheoret 59,121–137.

Brscic, M., Wemelsfelder, F., Tessitore, E., Gottardo, F., Cozzi, G., Van Reenen, C.G.,2009. Welfare assessment: correlations and integration between a Qualitative Behavioural Assessment and a clinical/health protocol applied in veal calves farms. Italian J Anim Sci 8, 601–603.

Dalla Costa, E., Murray, L., Dai, F., Canali, E., Minero, M., 2014. Equine on-farm welfare assessment: a review of animal-based indicators. Animal Welfare 23,323–341.

EFSA Panel on Animal Health and Welfare, (AHAW), 2012. Statement on the use of animal-based measures to assess the welfare of animals. EFSA J. 10, 1–29.

Ellingsen, K., Coleman, G.J., Lund, V., Mejdell, C.M., 2014. Using qualitative behaviour assessment to explore the link between stockperson behaviour and dairy calf behaviour. Appl. Anim. Behav. Sci. 153, 10–17.

Faostat, 2011. www.faostat.fao.org [WWW Document]. URL Faostat.fao.org.

Fleming, P., Paisley, C., 2013. Application of Qualitative Behavioural Assessment to horses during an endurance ride. Anim. Behav. 144, 80–88.

Fraser, D., Duncan, I.J.H., 1998. Pleasures, pains and animal welfare: toward a natural history of affect. Animal Welfare, 383–396.

IBM Corp., 2012. IBM SPSS Statistics for Windows.

Knierim, U., Winckler, C., 2009. On-farm welfare assessment in cattle: validity, reliability and feasibility issues and future perspectives with special regard to the Welfare Quality approach. Animal Welfare 18, 451–458.

Martin, P., Bateson, P., 2007. Measuring Behaviour: An Introductory Guide, 3rd ed.Cambridge University Press, Cambridge, Massachusetts, USA.

Meagher, R.K., 2009. Observer ratings: validity and value as a tool for animal welfare research. Appl. Anim. Behav. Sci. 119, 1–14.

Mendl, M., Burman, O.H.P., Paul, E.S., 2010. An integrative and functional framework for the study of animal emotion and mood. Proc. Biol. Sci./R. Soc.277, 2895–2904.

Minero, M., Tosi, M.V., Canali, E., Wemelsfelder, F., 2009. Quantitative and qualitative assessment of the response of foals to the presence of an unfamiliar human. Appl. Anim. Behav. Sci. 116, 74–81.

Napolitano, F., De Rosa, G., Braghieri, A., Grasso, F., Bordi, A., Wemelsfelder, F.,2008. The qualitative assessment of responsiveness to environmental challenge in horses and ponies. Appl. Anim. Behav. Sci. 109,342–354.

Napolitano, F., De Rosa, G., Grasso, F., Wemelsfelder, F., 2012. Qualitative behavior assessment of dairy buffaloes (Bubalus bubalis). Appl. Anim. Behav. Sci. 141,91–100.

Phythian, C., Michalopoulou, E., Duncan, J., Wemelsfelder, F., 2013. Inter-observer reliability of Qualitative Behavioural Assessments of sheep. Appl. Anim. Behav.Sci. 144, 73–79.

Rousing, T., Wemelsfelder, F., 2006. Qualitative assessment of social behaviour of dairy cows housed in loose housing systems. Appl. Anim. Behav. Sci. 101,40–53.

Rutherford, K.M.D., Donald, R.D., Lawrence, A.B., Wemelsfelder, F., 2012. Qualitative Behavioural Assessment of emotionality in pigs. Appl. Anim. Behav.Sci. 139, 218–224.

Sant'Anna, A.C., Paranhos da Costa, M.J.R., 2013. Validity and feasibility ofqualitative behavior assessment for the evaluation of Nellore cattle temperament. Livestock Sci. 157, 254–262.

Stockman, C.A., Collins, T., Barnes, A.L., Miller, D., Wickham, S.L., Beatty, D.T., Blache, D., Wemelsfelder, F., Fleming, P.A., 2011. Qualitative behavioural assessment and quantitative physiological measurement of cattle naïve and habituated to road transport. Anim. Prod. Sci.

Visser, E.E.K., Neijenhuis, F., de Graaf-Roelfsema, E., Wesselink, H.G.M.H., de Boer,J., van Wijhe-Kiezebrink, M.C.M., Engel, B., van Reenen, C.G., 2014. Risk factors associated with health disorders in sport and leisure horses in the Netherlands.J. Anim. Sci. 92, 844–855.

Wemelsfelder, F., 2007. How animals communicate quality of life: the qualitative assessment of behaviour. Anim. Welfare 16, 25–31.

Wemelsfelder, F., 2012. Assessing pig body language: agreement and consistency between pig farmers, veterinarians, and animal activists. J. Anim. Sci. 90, 3652–3665.

Wemelsfelder, F., Hunter, E., Mendl, M., Lawrence, A., 2000. The spontaneous qualitative assessment of behavioural expressions in pigs: first explorations of a novel methodology for integrative animal welfare measurement. Appl. Anim.Behav. Sci. 67, 193–215.

Wickham, S.L., Collins, T., Barnes, A.L., Miller, D.W., Beatty, D.T., Stockman, C.,Blache, D., Wemelsfelder, F., Fleming, P.A., 2012. Qualitative behavioral assessment of transport-naïve and transport-habituated sheep. J. Anim. Sci. 90, 4523–4535.