JZAR JOURNAL OF ZOO AND AQUARIUM RESEARCH
www.jzar.org

Research article

Recognising emotional expressions in captive bottlenose dolphins: Can lay observers agree using qualitative behavioural assessment?

Emma Warner[1], Sabrina Brando[2] and Françoise Wemelsfelder[3]

[1]Royal (Dick) School of Veterinary Studies, Easter Bush, Midlothian EH25 9RG, United Kingdom.
[2]AnimalConcepts, P.O. Box 378, 03725 Teulada, Alicante, Spain.
[3]Animal and Veterinary Sciences Group, SRUC, Roslin Institute Building, Bush Estate, EH25 9RG, United Kingdom.

Correspondence: Emma Warner, email; emma.warner@ed.ac.uk

Keywords: animal welfare assessment, aquarium, captive dolphins, dolphin behaviour, free choice profiling, qualitative behaviour assessment, *Tursiops truncatus*, zoo

Article history:
Received: 24 Jan 2021
Accepted: 8 Jul 2022
Published online: 31 Jul 2022

Abstract

This study applies qualitative behavioural assessment (QBA) to bottlenose dolphins *Tursiops truncatus*. Ten observers recruited through convenience sampling, who were unfamiliar with bottlenose dolphins and managed care of cetaceans, were instructed to use a Free Choice Profiling methodology to assess 20 video clips showing captive dolphins in a range of environmental and contextual settings. QBA scores were analysed using Generalised Procrustes Analysis, which shows a high level of agreement between observers (74.21% P<0.001) and generated three main consensus dimensions together explaining 61.9% of the variation between observer scoring patterns. Dimension 1 was characterised as ranging from 'energetic/active/excited' to 'calm/bored/sad', Dimension 2 as ranging from 'happy/playful/calm' to 'frustrated/aggressive/annoyed', and Dimension 3 from 'focused/engaged/curious' to 'unwilling/shy/nervous'. Dolphin behaviours observed in the QBA clips were scored by the experimenter using an ethogram of 37 behavioural categories, and then correlated with the three consensus dimensions using Spearman's rank correlation. Dimension 1 correlated with 'porpoising' (r_s=0.484), 'wait horizontally' (r_s=0.481) and 'face object'(r_s=-0.469), all at P<0.05; Dimension 2 with 'spy hop' (r_s=0.480), 'head following' (r_s=0.463), 'bubble single' (r_s=0.463), 'jaw clap' (r_s=-0.521) and 'bubble stream' (r_s=-0.518), all at P<0.05; and Dimension 3 with 'ball toss' (r_s=0.621), 'dynamic swim' (r_s=-0.632) and 'avoid trainer' (r_s=-0.624) at P<0.01, and 'carry object with mouth' (r_s=0.523) and 'touch trainer' (r_s=-0.558) at P<0.05. These findings indicate that QBA appears to be a suitable tool for assessing emotional expressivity in captive dolphins. Potentially meaningful associations between QBA dimensions of dolphin expressivity and ethogram-based behaviours are discussed but need further substantiation in future research.

Introduction

There is growing public awareness of and concern for animals in human care, with a particular focus on marine mammals, including cetaceans and polar bears *Ursus maritimus*. Zoos and aquariums (henceforth zoos) strive to achieve optimal welfare standards for the animals in their care. Professional zoos recognise and are committed to their responsibility to promote optimal welfare, and the benefits of doing so to research and conservation, public education programmes and commercial and charitable interests (Brando et al. 2018).

Zoo animal welfare assessment has historically been limited to resource-based measures. Various zoo associations including the American Zoological Association (AZA) created animal care manuals detailing resources that must be provided as part of the duty to animal care, including requirements on nutrition, enclosure design, veterinary care, animal training, husbandry, reproduction and transport (Whitham and Wielebnowski 2009). The provision of these resources is often assumed to maximise the potential for an animal to experience good welfare, but does not address whether an individual animal's needs (Barber 2009) and preferences (Brando and Buchanan-

Smith 2018; Brando and Herrelko 2021) are actually met. Therefore there is a growing need for animal-based assessments to gain insight into whether available resources meet animals' welfare requirements (Whitham and Wielebnowski 2013). These assessments may consist of physical parameters including body condition, clinical assessment, activity level, food and water intake, and psychological parameters. Psychological parameters include undesired or aversive responses to animal training, husbandry, catching events, group social disruption and use of provisions and enrichment (Justice et al. 2017), and positive responses such as dolphins' willingness to participate (Clegg et al. 2019; Delfour et al. 2020) and anticipatory behaviour prior to training sessions (Clegg et al. 2017, 2018) and presentations (Jensen et al. 2013; Miller et al. 2011). Similar parameters have been adapted for the welfare assessment framework for captive bottlenose dolphins known as "C-Well" (Clegg et al. 2015). Delfour and Charles (2021) provide a recent review of current welfare indicators in marine mammals in human care.

In recent years, interest in more holistic approaches to animal care and welfare has grown, including the use of different types of assessment that combine assessment of input resources—what is provided, the care that is given—and animal-based indicators that address an animal's welfare status. One promising indicator for assessing an animal's experience of its environment is Qualitative Behaviour Assessment (QBA). QBA is an integrative, 'whole-animal' approach, describing and quantifying the dynamic demeanour shown by animals in how they move around their environment—for example in a way that is relaxed, lively and playful, or tense, agitated and evasive. Such different 'styles' of behaviour can be seen as emotionally expressive 'body language', giving more direct access to animals' experience than would be possible through measurement of distinct physical elements of behaviour (Wemelsfelder et al. 2001; Wemelsfelder 2007). A growing number of studies support the validity of QBA by reporting significant and meaningful associations between QBA and a range of recognised quantitative behavioural and physiological measures, indicating that the different types of measure can provide complementary information, and, when used in conjunction, can mutually reinforce understanding of animal welfare (Wemelsfelder and Mullan 2014 ; Fleming et al. 2016). This approach has been successfully applied to a number of livestock and companion animal species (Arena et al. 2019; Fleming et al. 2016). QBA has also recently been applied to elephants in zoos (Yon et al. 2019) and to the relationship between zoo keepers and giraffes (Patel et al. 2019). This work led Rose and Riley (2019) to suggest that QBA could play a useful role in evaluating the effect of environmental design and enrichment on zoo animal welfare.

The aim of the present study was to investigate the applicability of QBA to characterising emotional expressivity in bottlenose dolphins *Tursiops truncatus* in human care. It was an exploratory study with a methodological focus, and as such did not address any specific dolphin welfare issues nor aimed to interpret any associations between qualitative and quantitative behaviour assessments in terms of underlying mechanisms governing the association between emotional states and behaviour. When applying QBA to a novel species, it is important to gain a sense of which descriptors may be optimal for characterising that species' emotional expressivity, and so a 'Free Choice Profiling' method was used to enable observers to generate and quantify their own terms for describing dolphin expressions presented to them on video (Wemelsfelder et al. 2001).

The objectives of this study were to investigate (i) whether a group of observers unfamiliar with bottlenose dolphins showed consensus in their qualitative assessments of dolphin expressivity from video; (ii) the extent to which dimensions of dolphin expressivity correlated with physical behaviours observed in the same video footage; and (iii) whether any such associations indicated a potential for meaningful alignment between qualitative and quantitative assessment approaches.

Materials and methods

Study animals

In this study 36 captive bottlenose dolphins were filmed in two European dolphinaria: Boudewijn Seapark, Belgium (n=8) and Dolfinarium Harderwijk, Netherlands (n=28). The animals comprised 7 smaller groups housed in various locations within these dolphinaria, including both outdoor and indoor pools. Groups were a combination of mixed sex and single sex animals, and included ages from 1 to 52 years.

Video recording

Video footage of dolphins in both dolphinaria was collected using a Canon Legria HFM52 camcorder, mounted on a tripod. Filming occurred opportunistically over a period of six full working days, for three days in each dolphinarium between 0830 and 1700. The primary aim of filming was to capture examples of as wide a range as possible of emotional expressivity displayed by the dolphins. Suitable locations and times to film the different expressions were identified with the assistance of local dolphin trainers. Filming across all areas of the dolphinarium was permitted, including areas not accessible to the public.

The dolphins were filmed, with sound, either above the surface of the water, or below the surface through glass. Three distinct contextual categories were videoed: 'training session', when members of the animal training staff are training specific behaviours with the animals; 'enrichment session', when enrichment objects have been provided to the animals for a period of time by the trainers; and 'out of session', when the animals are in their free time, the trainers are not asking for behaviours and there are no objects in the water. These categories provided enough variation to allow a diverse range of behaviours and expressions to be exhibited.

The video footage was edited to create a sample size of n=20 video clips (plus one additional clip for observer practice which was not included in the analysis), that ranged from 1 min to 1 min 54 sec in duration. Video duration was not standardised, but limited to a maximum of 2 minutes to ensure the dolphins were in view for long enough for observers to score properly. This particularly applied to clips filmed above the surface where dolphins were not physically in the frame for the full duration of the video, such as when they were spy hopping. Spy hopping is an investigatory behaviour where a dolphin raises half of its body out of the water in a vertical position (Lauderdale 2017).

Qualitative Behaviour Assessment (QBA)

Observers

Ten observers (seven females, three males) were recruited through convenience sampling via email invitation. Eight observers came from the MSc in Applied Animal Behaviour and Animal Welfare from the University of Edinburgh (one from the 2016/17 cohort, seven from 2017/18), and two PhD students from Scotland's Rural College, SRUC. All observers had experience and knowledge of animal behaviour, but none had experience working with dolphins nor knowledge of dolphin behaviour. All observers were familiar with the concepts of qualitative behaviour assessment (QBA) and Free Choice Profiling (FCP), but only three had previous experience using both QBA and FCP in an experimental context.

All experimental procedures were approved by the Human Ethical Review Committee (HERC) and the Veterinary Ethical Review Committee (VERC), from the Royal (Dick) School of Veterinary Studies at the University of Edinburgh.

Experimental procedures

FCP methodology as described in Wemelsfelder et al. (2001) was used to generate data. Observers attended two sessions each, spaced at least three days apart. Sessions were run in seminar rooms at the Royal (Dick) School of Veterinary Studies, and videos were played using a projector and smart board with the sound on.

FCP Session 1 (generation of terms)

The aim of this session was to generate descriptors for dolphin expressivity. At the start of the session observers were provided full instructions, background and context to the study. Observers were shown one practice video and afterwards were asked to say out loud which adverbs in their view described the expressive qualities of the dolphins, to ensure they were describing expressive qualities rather than physical behaviours. They were then shown 20 video clips in an order designed to create contrast between expressions.

Before each clip, the instructor read out a brief explanation clarifying the overall context in which the dolphins were filmed, for example whether a video clip was obtained during a training or enrichment session, or off-session. Providing such context for each video clip helped remove the ambiguity of the poorly visible (underwater) conditions in certain clips, so observers could focus on the animals' expressions rather than trying to decipher why the animals were there. Contextual information is acknowledged to form a crucial element of qualitative assessments, and, when conducted in person under 'live' assessment circumstances, is normally abundantly available. However, in this study the observers, none of whom were experienced in practical dolphin management, were only presented with very brief video clips. If they were to try and guess the context, this would distract from and hinder focusing on the dolphins' expressions. Therefore, it was decided to prevent such ambiguity by providing basic contextual information, enhancing the observers' ability to create appropriate terms for assessing dolphin expressivity in a range of environments. When more than one animal was present in a clip, observers were instructed to focus on a specific animal, or animals within a group. For example, "Clip 3. These dolphins were provided with enrichment objects. Please pay attention to the dolphin that first appears in the clip".

After each video, observers were given approximately 2 min to write down on a paper form adverbs that in their view described the various aspects of dolphin emotional expression observed in that clip. After 10 clips, a refreshment break was provided.

FCP Session 2 (quantification of terms)

The aim of Session 2 was for observers to watch the same clips as in Session 1 and quantify observed dolphin expressions by scoring their own personal descriptors on visual analogue scales (VAS). Observers were provided an individual pack containing 20 scoring forms, with each form showing a list of the descriptors they had generated individually in Session 1. Terms which labelled physical behaviour or environmental surroundings were excluded, and in some cases the negative form of a word was removed (e.g. 'unhappy') if the positive form ('happy') was also present. A positive term has a wider scoring range as it ranges from 'not at all happy' to 'could not be happier' whereas the negative term 'unhappy' would only range from 'not at all unhappy' to 'could not be unhappier'. In addition, the use of double negatives may confuse observers and cause errors in scoring. Words for which a negative term is commonly used such as 'unwilling' or 'unsure' were retained.

Each term was placed next to a visual analogue scale of 125 mm length, which is the standard length used in food science where Free Choice Profiling was first developed (Arnold and Williams 1986). The order in which terms were listed on the form was designed to enhance contrasts in meaning, in order to avoid scoring contagion caused by the proximity of similar terms and to encourage observers to focus on each term separately. Observers were shown the same 20 videos in the same order as in Session 1, and the same context was given before each clip. After each clip, observers were given time to score the observed dolphin on each of their terms. The VAS was described to observers as a continuous incremental scale where the minimum point (0) indicated that a dolphin "was not at all e.g. excited". The maximum point (125) on the VAS indicated that a dolphin "could not be more e.g. excited". A short break was provided after every five clips to avoid fatigue and to keep observers engaged.

Quantitative behaviour assessment

After the QBA sessions were completed, quantitative data were collected by the experimenter for the same 20 video clips that were shown to observers. Thirty-seven behavioural events (Table 1) were recorded using continuous sampling and analysed with Observer XT software (Noldus Information Technology). This produced frequencies and durations for the 37 behavioural events in each clip, which were subsequently divided by the number of animals observed in each clip to produce an average value per clip. Because clip length was not standardised in QBA sessions, these values were then expressed as rates per minute (frequencies) or proportions of time (durations). In clips showing dolphins filmed in groups, individual animals were sometimes clearly identifiable through their colouring or size and so could be reliably assessed throughout the clip. However, where this was not feasible, observers were instructed to assess the group as one unit, where events would be recorded as and when they occurred, regardless of the individual identities or number of animals performing a behaviour.

Statistical analysis

Measurement of VAS scores

The observers' scores on each VAS were measured with a ruler and entered manually into a data matrix for each individual observer in Microsoft Excel, before being imported into Genstat (2008, VSN International, UK). Each matrix was laid out with the 20 dolphin clips in the first column, and an individual observer's list of terms along the top row. Each score was entered for the corresponding clip and term in the ensuing matrix.

Generalised Procrustes Analysis

The degree of alignment among the 10 observer data matrices was analysed using Generalised Procrustes Analysis (GPA), a multivariate technique that has been described in detail elsewhere (Wemelsfelder et al. 2000, 2001). GPA focuses on recognising complex scoring patterns produced by the distances between the dolphin scores generated by each individual observer on their terms. Through a complex iterative process of matching individual observer scoring patterns, a consensus profile was generated. How well each individual observer's scoring profile fitted this consensus profile was quantified by the Procrustes statistic and illustrated by an 'observer plot'. The statistical significance of this consensus profile was evaluated by a randomisation test. GPA was re-run 100 times randomising all observer profiles, to generate a randomised mean consensus profile. The difference between original and randomised mean consensus profiles was then tested for significance, using a one-tailed Student t-test (n=99).

Interpreting GPA dimensions

The next step was to reduce the number of dimensions produced using principal component analysis (PCA), as initially the consensus profile has as many dimensions as the maximum number of terms generated by any of the observers (n=49 in this

Table 1. Ethogram listing all behavioural events; those scored as frequencies are indicated by $_F$ and those scored as durations are indicated by $_D$.

Behavioural events	Description of behavioural patterns
Social behaviours	
Jaw Clap $_F$	Dolphin suddenly and forcefully closes its jaws, producing a detectable sound towards a recipient (Miller et al. 2011)
Water Toss $_F$	Dolphin tosses water at human using rostrum or open mouth
Bubble Ring $_F$	Dolphin produces a bubble ring with the blowhole, directed at a recipient
Bubble Single $_F$	Dolphin produces a single bubble with the blowhole, directed at a recipient
Bubble Stream $_F$	Dolphin produces a bubble stream with the blowhole, directed at a recipient
Flipper Flap $_F$	Dolphin is stationary, and moves its flippers up and down, or side to side in the water, in the direction of a recipient
Ramming $_F$	Dolphin forcefully hits or attempts to hit another individual with the rostrum or melon (Miller et al. 2011)
Tail Slap $_F$	Dolphin makes, or attempts to make contact with another dolphin using its fluke, usually smacking them with force (adapted from Miller et al. 2011)
Chase $_D$	A dolphin swims quickly and actively after one or more dolphins causing the receiver to flee (adapted from Miller et al. 2011)
Body Slam $_F$	Dolphin forcefully contacts another dolphin using the side of its body (adapted from Samuels and Spradlin 1995)
Circle $_F$	Dolphin makes an abrupt circular turn vertically or horizontally in the water to face a recipient
Swim behaviours	
Directional Swim $_D$	Dolphin swims in one direction dorsal side up, for more than 3 s (Miller et al. 2011)
Side Swim $_D$	Dolphin swims on its side, for more than 3 s (Miller et al. 2011)
Ventral Swim $_D$	Dolphin swims upside down with ventral side pointing towards the surface for more than 3 s (Miller et al. 2011)
Dynamic Swim $_D$	Dolphin sustains an increased speed, swimming in one or multiple directions, including corkscrews, for more than 3 s, forcefully using its tail for propulsion (adapted from Miller et al. 2011)
Rest $_D$	Dolphin remains stationary at the surface or the bottom, or drifts very slowly (adapted from Lauderdale 2017)
Training behaviours	
Correct Behaviour $_F$	Dolphin is requested by trainer to perform specific behaviour, and does so correctly (adapted from Samuels and Spradlin 1994)
Avoid Trainer $_F$	Dolphin is requested to touch trainer's hand target or perform specific behaviour, but turns the head or full body away, hides underwater, or swims away from the trainer
Touch Trainer $_F$	Dolphin touches the trainer when presented with a hand, knee or foot target
Feed $_F$	Dolphin is fed fish or other food by a trainer
Object interations	
Ball Throw $_F$	Dolphin throws a ball in the air, across the water, abruptly moving its head and body to propel the ball
Ball Toss $_F$	Dolphin tosses a ball in the air, across the water or to a human using its mouth, rostrum or tail (Kuczaj et al. 2006)
Bite Object $_F$	Dolphin bites an object with its mouth
Carry Object with Flipper $_D$	Dolphin carries an object using its flipper for more than 3 s
Carry Object with Mouth $_D$	Dolphin carries an object using its mouth for more than 3 s
Carry Object with Rostrum $_D$	Dolphin carries an object using its rostrum for more than 3 s
Pick Up Object $_F$	Dolphin picks up an object using its mouth, rostrum, flipper or tail flukes, and drops it again within 3 seconds
Observing behaviours	
Face Object $_D$	Dolphin is stationary, but may orient itself with its head facing object of interest
Floating One Eye $_D$	At the surface, dolphin floats on its side, with one eye above the water to observe (Jensen et al. 2013)
Head Following $_D$	Dolphin is stationary, facing person, and its head follows the person's movements
Head Scan $_D$	Dolphin moves its head side to side or up and down whilst performing a directional swim. Dolphin may also be carrying an object
Wait Horizontal $_D$	Dolphin floats horizontally underwater, or at surface, where it may place its rostrum on the side of the pool, in anticipation of a response from another dolphin or human
Wait Vertically $_D$	Dolphin floats vertically underwater, or at surface, where it may place its rostrum on the side of the pool, in anticipation of a response from another dolphin or human
Other behaviours	
Spy Hop $_F$	Dolphin raises and lowers half of its body out of the water in a vertical position (Lauderdale 2017)
Jump $_F$	Dolphin's entire body comes out of the water and re-enters smoothly (Samuels and Spradlin 1995; Miller et al. 2011; Lauderdale 2017)
Porpoising $_F$	Dolphin jumps partially out of the water (flukes remain in water) and re-enters head first (Lauderdale 2017)
Beach $_D$	Dolphin slides more than one-third of its body onto the flat side of the pool (Lauderdale 2017)

study). The reduction of dimensions to two or three 'principal components' explained the majority of the variation between observed dolphins.

The coordinates produced by the consensus profile for each dolphin clip were correlated with each observer's data matrix to produce 10 individual 'word-charts'. The three principal dimensions produced two two-dimensional word charts for each observer, Dimensions 1 against 2, and Dimensions 1 against 3. Each term generated by observers was plotted against these dimensions, and terms with the highest correlations were considered the strongest representative descriptors of the dimensions. The three highest loading terms at both ends of the three dimensions for all observers were collated and placed together in a table. From the pool of 30 terms now presented for each end of the three main dimensions, the experimenter subsequently chose three representative terms to label the meaning of these dimensions.

Correlation between QBA dimensions and physical dolphin behaviours

In order to provide behavioural context to the expressive qualities of the dolphins, the animals' scores on the three main QBA consensus dimensions were correlated with the ethogram-based behavioural data, using Spearman's rank correlations. Given the large numbers of correlations calculated (3 QBA dimensions × 37 behaviours = 111 correlations) there is a concern that some could be significant by chance—with an alpha rate of 5% potentially giving five or six correlations. However, in this study a correlation's significance is not used to support any kind of causal association or inference. Instead, the focus of investigation is the extent to which qualitative and quantitative types of assessment appear to align. Therefore, for exploratory reasons a 5% rather than 1% alpha rate is reported, with full acknowledgment that outcomes must be considered provisional and requires further research.

Results

Qualitative Behaviour Assessment

Observer consensus

The level of agreement amongst observer profiles as reflected by the Procrustes statistic (74.21%), was significantly higher than that for 100 mean randomised profiles (58.26±0.27%; t_{99}=30.67; P<0.001), indicating that the consensus was a meaningful feature of the data set rather than an artefact of the statistical GPA procedures. The majority of observers fell within the 95% confidence region, but three were outliers. There was no discernible reason why these three observers were outliers. The data were reanalysed excluding their scores, however this did not affect outcomes in any substantial way. In the remainder of the paper the data from all observers are included.

Dimensions of dolphin expression

The majority of variation between observed dolphins was explained by three main dimensions of dolphin expression, explaining 32.2%, 21.5% and 8.2% of the variation for Dimensions 1, 2 and 3 respectively, adding up to 61.9% of the total variance. Figures 1 and 2 show examples of word charts for these three dimensions, displaying all terms generated by Observer 2 and their level of correlation with the three consensus dimensions. Observer 2's word charts are considered a representative example because this observer's terms loaded highly onto both ends of the three consensus dimensions. For Observer 2, the highest loading terms for the top and bottom ends of Dimension 1 were 'excited/energetic/interested' and 'bored/listless/calm' respectively, while for Dimension 2 they were 'happy/entertained/playful' and 'aggressive/angry/frustrated', and for Dimension 3 were 'focused/interested/curious' and 'rebellious/shy/timid'.

Overall, Dimension 1 was characterised as ranging from 'energetic/active/excited' to 'calm/bored/sad', Dimension 2 from 'happy/playful/calm' to 'frustrated/aggressive/annoyed' and Dimension 3 from 'focused/engaged/curious' to 'unwilling/nervous/shy' (Table 2). The experimenter judged that these terms adequately reflected the diversity of expression observed in the dolphins.

Correlations of QBA dimensions with ethogram-based categories of dolphin behaviour

Table 3 shows the significant correlations of QBA Dimensions 1, 2 and 3 with the ethogram-based categories of dolphin behaviour in Table 1. Significantly correlating behaviours along Dimension 1 were porpoising (r_s=0.484, P<0.05), wait horizontal (r_s=0.481, P<0.05) and face object (r_s=-0.469, P<0.05), indicating that when porpoising and waiting horizontally, dolphins were perceived as doing so in an 'energetic/excited/active' manner, and when facing objects, in a 'calm/bored' manner.

Significantly correlating behaviours along Dimension 2 were spy hop (r_s=0.480, P<0.05), bubble single and head following (r_s=0.463, P<0.05). Bubble single and head following were always performed in the same clips and were perceived by observers as being performed in a 'happy/playful/calm' manner. Conversely jaw clap (r_s=-0.521, P<0.05) and bubble stream (r_s=-0.518, P<0.05), also always performed in the same clips, were perceived by observers as performed in an 'aggressive/frustrated/annoyed' manner.

Dimension 3 was significantly correlated with ball toss (r_s=0.621, P<0.01), and carry object with mouth (r_s=0.523, P<0.05) indicating that interactions with an object were perceived by observers to be executed in a curious and engaged manner. Conversely correlations with dynamic swimming (r_s=-0.632, P<0.01), avoid trainer (r_s=-0.624, P<0.01) and touch trainer (r_s=-0.558, P<0.05), indicated that these behaviours were perceived by observers to be performed in an 'unwilling/nervous/shy' manner.

Discussion

The main aim of this study was to investigate whether naïve observers could show agreement in describing the emotional expressivity of bottlenose dolphins in human care, and to subsequently examine whether and how such expressivity correlated with known dolphin behaviours. The results indicate that 10 observers show a high degree of inter-observer reliability in their qualitative assessments of dolphin body language, with the consensus profile explaining 74.21% of the variation between observer scoring profiles. The analysis identified three main consensus dimensions, together explaining 61.9% of the variation between dolphins. Dimension 1 was characterised as ranging from 'energetic/active/excited' to 'calm/bored/sad', Dimension 2 from 'happy/playful/calm' to 'frustrated/aggressive/annoyed', and Dimension 3 from 'focused/engaged/curious' to 'unwilling/nervous/shy'.

These three QBA dimensions showed a number of significant correlations with ethogram-based assessments of dolphin behaviour made for the same QBA video clips. As acknowledged in the Methods, given the large number of calculated correlations, some may have been found by chance. However, in this study a correlation's significance is not considered to indicate any kind of causal association. Instead a methodological focus has been used in exploring whether and how qualitative and quantitative assessments might align. If meaningful associations occur, then the different methods might contribute complementary types of information to studies of animal welfare, an outcome repeatedly reported for QBA studies across a range of animal species (Fleming et al. 2016; Rutherford et al. 2012). This hypothesis is discussed in further detail below.

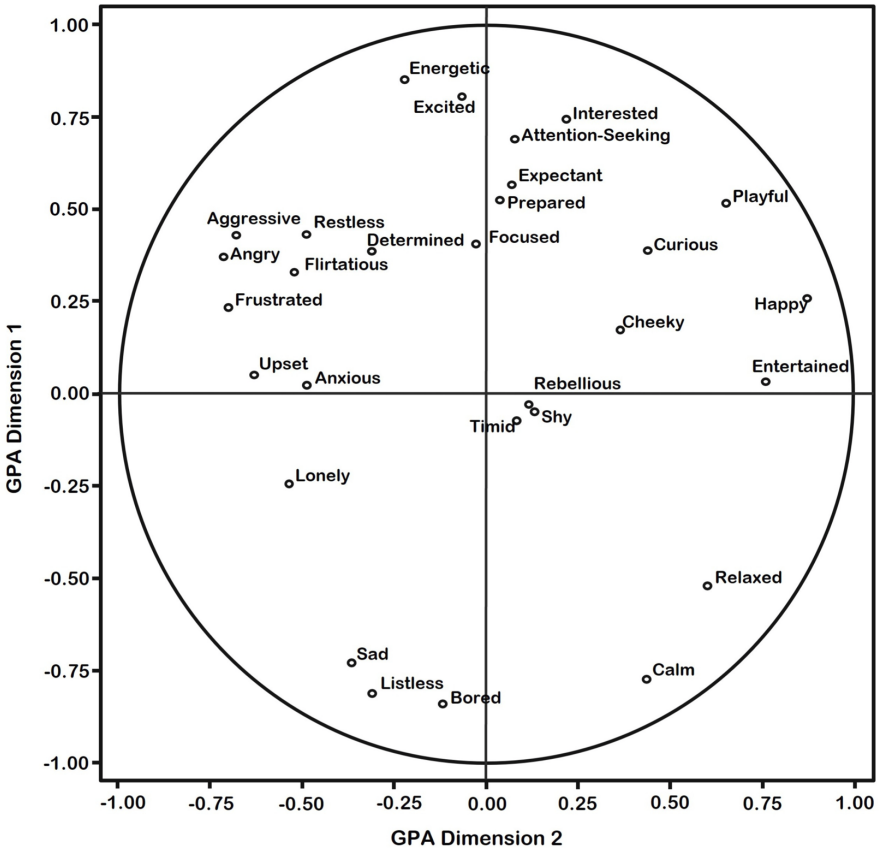

Figure 1. Observer 2-word chart for Dimension 1 versus Dimension 2.

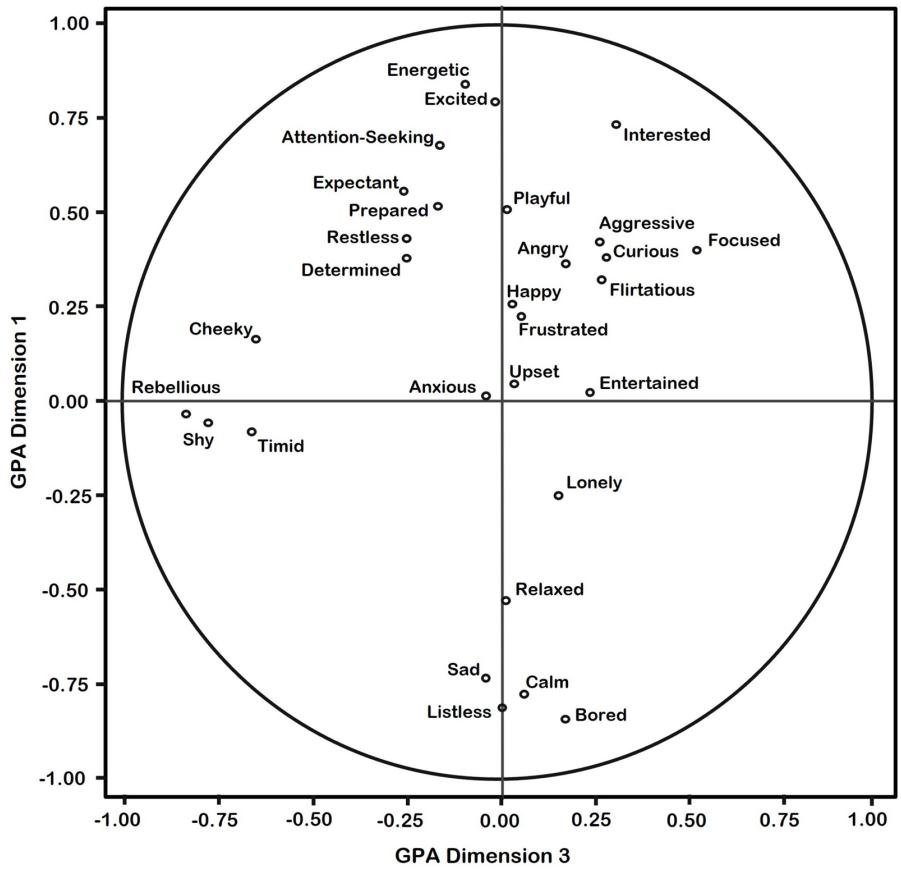

Figure 2. Observer 2-word chart for Dimension 1 versus Dimension 3.

Table 2. Overview of the highest loading terms provided by all 10 observers for the three main consensus dimensions; the three high-loading terms most frequently used by observers were used as labels for the dimensional axes.

Consensus dimension	Top end	Bottom end
Dimension 1	Energetic (6), Active (4), Excited (3), Agitated, Alert, Aroused, Assertive, Attention-Seeking, Bold, Busy-Minded, Confident, Curious, Enthusiastic, Hurried, Interested, Inviting, Keen, Lively, Social, Tense	Calm (5), Bored (4), Sad (3), Lethargic (2), Passive (2), Relaxed (2), Subdued (2), Hopeless, Lifeless, Listless, Lonely, Patient, Peaceful, Reserved, Suffocated, Tedious, Tired
Dimension 2	Happy (5), Playful (5), Calm (3), Entertained (2), Friendly (2), Fulfilled (2), Relaxed (2), Achieved, At-Ease, Comfortable, Content, Engaged, Having-Fun, Reserved, Satisfied, Stimulated	Frustrated (6), Aggressive (3), Annoyed (3), Restless (2), Agitated, Aimless, Angry, Anxious, Bold, Depressed, Desperate, Dominant, Fed-Up, Irritated, Longing-For, Nervous, Robust, Sad, Stressed, Temperamental
Dimension 3	Focused (4), Engaged (3), Curious (2), Interested (2), Angry, Assertive, At-Ease, Bored, Calm, Confident, Dedicated, Eager, Habituated, Inquisitive, Interactive, Motivated, Obsessive, Patient, Perseverant, Possessive, Repetitive, Tense, Tired	Unwilling (3), Nervous (2), Shy (2), Timid (2), Unsure (2), Agitated, Annoying, Attention-Seeking, Defiant, Distracted, Embarrassed, Fearful, Friendly, Frustrated, Hesitant, Insecure, Mischievous, Outgoing, Passive, Rebellious, Restless, Silly, Tired, Worked-Up

Qualitative Behaviour Assessment

Agreement amongst observers was high in this study. The presence of three clear dimensions is an indicator that dolphins, to humans, are expressive animals that observers found numerous and complex ways to describe, despite having no marine mammal experience. As future QBA may be primarily conducted by dolphin trainers, ensuring proper training in the QBA process will be essential to maintain observer reliability and consistency (Cooper and Wemelsfelder 2020; Fleming et al. 2015).

Considering the high-loading descriptors for either end of Dimension 1 as presented in Table 2, it appears that the primary expressive variation amongst observed dolphins in this study related to the animals' level of arousal, ranging from high-energy descriptors such as 'energetic', 'excited' and 'agitated', to low-energy terms such as 'calm', 'bored' and 'relaxed'. Clustering of terms at either end of this dimension is expected to contain both positively and negatively valenced terms, because Dimension 2 intersecting Dimension 1 was perceived to relate to the animals' mood, contributing both positive and negative mood terms to the low and high arousal ends of Dimension 1 (Figure 1). However, from Table 2 it becomes apparent that although there are some negative mood terms on the high arousal end of the dimension (e.g. agitated) and some positive mood terms on the low arousal end (e.g. relaxed), the overall mood tone of the high arousal end seems predominantly positively valenced, and the overall tone of the low arousal end negatively valenced. Thus it appears that Dimension 1 contains both an arousal and a valence aspect, describing the contrast between 'energetic enthusiastic' and 'calm bored' dolphins. The presence of a few apparently ill-fitting terms at either end of Dimension 1 should not be interpreted as anomalous, in that an animal's expressive qualities can be multi-faceted and contain different subtle shades rather than the either/or distinctions made by ethograms. It seems perfectly feasible to suggest that both bored/sad (Dimension 1), and happy/playful (Dimension 2) animals can be calm, and equally, that agitation can be a part of both energetic/active/excited (Dimension 1) and frustrated/aggressive (Dimension 2) expressions. What matters is to select descriptive labels that best characterise the overall expressive pattern depicted in the observers' word charts.

A combined mood/arousal pattern also seems to characterise Dimension 2, but this time with inverse pairings compared to Dimension 1. It appears that the positive mood end of Dimension 2 has a somewhat greater emphasis on low arousal terms such as 'calm', 'relaxed', 'at ease' and 'comfortable', while the negative mood end has a somewhat greater emphasis on high arousal terms such as 'restless', 'irritated', 'nervous' and 'stressed'. Thus Dimension 2 seems to characterise the contrast between 'happy relaxed' and 'stressed irritable' dolphins, with some apparently ill-fitting but not disruptive terms such as playful and sad at the top and bottom ends respectively.

High-loading terms for Dimension 3 appear to indicate a contrast in attentiveness and motivation, ranging from focused, engaged and curious to unwilling, nervous and shy. As this dimension explained only 8.2% of variation between animals, the pattern of high-loading descriptors is not as clear as for the primary two dimensions, however overall, the contrast between terms at either end of this dimension seems clear enough to include it as meaningful.

Thus in accordance with many QBA studies of farm and companion animal species (Fleming et al. 2013; Grosso et al. 2016; Minero et al. 2016; Phythian et al. 2016; Arena et al. 2017), the present study found dimensions that reflect and/or combine various aspects of valence and arousal.

Correlation between qualitative and quantitative assessments

The three QBA dimensions reported in this study were found to correlate significantly with a number of ethogram-based behaviours across the categories 'social behaviours', 'swim behaviours', 'anticipatory behaviours', 'training behaviours' and 'object interaction'. Most of these correlations were weak, with *r*-values between 0.4 and 0.6, and three just under 0.65, indicating only tentative alignments between the two types of assessment in the context of this study. However, it is often difficult to find associations between different measures in uncontrolled, highly variable environmental conditions so this is worth discussing. Acknowledging the caveats previously stated, the identified associations are considered in more detail below.

Social behaviours

In the category 'social behaviours' two different types of 'bubble behaviour' were associated with opposite ends of Dimension 2: 'bubble single' with 'happy/playful/calm' and 'bubble stream' with 'frustrated/aggressive/annoyed'. These associations are supported by recent dolphin research which suggests that single bubble bursts indicate surprise, excitement or curiosity, and that bubble streams, often produced in agonistic social interactions,

Table 3. Spearman rank *r*-values between ethogram behaviours and the top and bottom ends of QBA dimensions 1, 2 and 3. * indicates P<0.05 and ** indicates P<0.01.

		Dimension					
Behaviour category	Behaviour	1		2		3	
		Energetic/ Active/Excited	Calm/ Bored/ Sad	Happy/Playful/ Calm	Frustrated/ Aggressive/ Annoyed	Focused/ Engaged/Curious	Unwilling/ Nervous/Shy
Anticipatory	Wait horizontal	0.481 *					
	Spy hop			0.480 *			
	Porpoising	0.484 *					
Swim	Dynamic swim						-0.632 **
Social	Head following			0.463 *			
	Bubble single			0.463 *			
	Jaw clap				-0.521 *		
	Bubble stream				-0.518 *		
Training	Avoid trainer						-0.624 **
	Touch trainer						-0.558 *
Object	Face object		-0.469 *				
	Ball toss					0.621 **	
	Carry object with mouth					0.523 *	

can be considered markers of distress (Moreno et al. 2019). In the current study 'bubble single' and 'head following' occurred mostly together during playful interactions between animal and experimenter through the glass. 'Bubble stream' frequently occurred together with 'jaw clap' which is also considered a form of aggressive social behaviour in dolphins (Clegg et al. 2015 McCowan et al. 2000; Kyngdon et al. 2003; Miller et al. 2011; Clegg et al. 2015).

Swim behaviours

In this category the behaviour 'dynamic swim' was associated with the 'unwilling/nervous/shy' end of Dimension 3. This behaviour is considered a form of high-energy swimming (Lauderdale 2017; Miller et al. 2011) and in the current study was filmed in a group of dolphins during a 6 min period before the start of a training session. The association with Dimension 3 indicates that the animals' energetic swimming was perceived as an expression of nervousness, which in the context of impending training could be a form of nervous, rather than happy, anticipation. That dynamic swimming might be an indication of nervousness is worth investigating further.

Anticipatory behaviours

In this category 'wait horizontal' and 'porpoising' were associated with the 'energetic/excited/active' end of Dimension 1 and 'spy hop' with the 'happy/playful/calm' end of Dimension 2. All three behaviours are generally understood to be anticipatory in nature, with potentially either a positive or negative meaning to the animal (Clegg et al. 2018; Jensen et al. 2013). In the current study the three behaviours, like dynamic swimming, were all filmed just before training events, however in contrast to dynamic swimming observers perceived them as part of positive expressivity. Thus QBA might be capable of picking up expressive aspects of anticipatory behaviour that can assist in interpreting the meaning of these behaviours for dolphins in particular contexts. As such it could potentially help trainers to evaluate how dolphins emotionally perceive different forms of training and demonstration events. Clegg et al. (2017) suggest that frequencies and durations of anticipatory behaviours may indicate whether animals experience positive or negative affective states, to which QBA could add a more direct indication of the animal's anticipatory experience.

Object interaction behaviours

In this category 'face object' was associated with the 'calm/bored/sad' end of Dimension 1, while 'ball toss' and 'carry object with mouth' were associated with the 'focused/engaged/curious' end of Dimension 3. Why 'face object' was not associated with a more positive expressivity, such as 'happy/playful/calm', is difficult to justify but it signifies that facing objects was not a behaviour associated with playful excitement. On the other hand, the association of tossing and carrying balls in the mouth with a focused, engaged and curious expression is easier to interpret, and supports the theory that providing objects to dolphins may stimulate their cognitive engagement and improve their welfare (Brando et al. 2016, 2018; Clark 2013; Clark et al. 2013; Lauderdale 2017). With further substantiation of these relationships, QBA

could help trainers to evaluate the efficacy of different types of objects in environmental enrichment programmes in ways that might be more effective than simply recording whether animals interact with an object (Clark et al. 2013; Delfour and Beyer 2012; Duncan 1997; Jensen et al. 2013).

Training behaviours

In this category the behaviours 'touch trainer' and 'avoid trainer' were associated with the 'unwilling/nervous/shy' end of Dimension 3. For 'avoid trainer' this association appears to make sense, but for 'touch trainer' it needs some explanation. The behaviour 'touch trainer' occurred in a context where touching the trainer's hand target signified a neutral starting position indicating the dolphin's willingness to initiate learning procedures in training sessions. Its association with 'unwilling/nervous/shy' expressivity suggests that being repeatedly asked to touch the trainer's hand target in order to initiate learning became increasingly aversive to the dolphins, potentially indicating their reluctance to participate at various points in this specific training session. This is supported by the association of 'avoid trainer' with this dimension, a behaviour shown frequently by the dolphins in the same session. Thus as noted for anticipatory behaviour above, this is an example where QBA could add affective information to dolphins' performance of behaviours in training situations, specifying the animals' experience of those situations at different moments in time.

The discussion above suggests a generally meaningful alignment between qualitative assessments of dolphin emotional expressivity and what is known in the scientific literature about a range of physical behaviours in dolphins, for the specific contexts included in this study. This would in turn suggest that qualitative assessments of dolphin expressivity do not stand apart from quantitative behaviour assessments but have the potential for meaningful integration with them—a hypothesis that requires further substantiation in future research.

Future directions

Overall the feedback given by the dolphinaria involved in the current study was positive in that trainers were interested in the QBA methodology and the concept of scientifically assessing the emotional expressions of the dolphins in their care. If QBA were to be included in recognised dolphin welfare assessment protocols, the next step would be to develop a standardised list of QBA descriptors for dolphins and to validate this list for practical use, similarly to what has been done for other species such as donkeys and shelter dogs (Arena et al. 2019; Minero et al. 2016). Integration of QBA with other indicators in larger welfare assessment protocols would provide additional information on dolphins' emotional experience and enrich understanding of the animals' overall welfare state (Wemelsfelder and Mullan 2014). Zoos have recently become more interested in QBA as a tool to monitor animal welfare, and have suggested it may help to guide the design of animal habitats and enrichment programmes across a range of taxa (Rose and Riley 2019). It is preferable to involve dolphin trainers and caretakers in development of QBA tools from the start, to ensure the practical relevance of these tools.

Conclusions

The outcomes of this study demonstrate a high level of consensus between observers in their qualitative assessments of dolphin behavioural expressions, as filmed in two different European dolphinaria. Observers identified three main dimensions of expression that showed a number of meaningful associations with ethogram-based behaviours, suggesting QBA has the potential to assist with understanding the experience of dolphins in captive situations. However, given the exploratory nature of this study and the uncertainty imposed by calculating large numbers of correlations, further research is needed to substantiate the findings reported here.

Acknowledgements

The authors would like to thank Dennis Vink, Jurrien Wouters, Sander Van der Heul and the members of staff at Boudewijn Seapark and Dolfinarium Harderwijk for granting access to the animals to obtain the video materials used for this study. Thank you also to the student observers from the University of Edinburgh who volunteered their time to participate in this study. The work reported in this paper was a part of the first author's MSc in Applied Animal Behaviour and Animal Welfare at the University of Edinburgh, and was partly sponsored by the Douglas Houghton Memorial Fund.

References

Arena L., Wemelsfelder F., Messori S., Ferri N., Barnard S. (2017) Application of free choice profiling to assess the emotional state of dogs housed in shelter environments. *Applied Animal Behaviour Science* 195: 72–79. doi.org/10.1016/j.applanim.2017.06.005

Arena L., Wemelsfelder F., Messori S., Ferri N., Barnard S. (2019) Development of a fixed list of terms for the qualitative behavioural assessment of shelter dogs. *PLoS ONE* 14(10): e0212652. doi.org/10.1371/journal.pone.0212652

Arnold G.M., Williams A.A. (1986) The use of generalized Procrustes analysis techniques in sensory analysis. In: Piggot J.R. (ed.). *Statistical Procedures in Food Research.* London, UK: Elsevier Applied Science, 233-253.

Barber J.C.E. (2009) Programmatic approaches to assessing and improving animal welfare in zoos and aquariums. *Zoo Biology* 28(6): 519–530.

Brando S., Bowles A.E., Böye M., Dudzinski K.M., van Elk N., Lucke K., McBain J., Singer R., Wahlberg M. (2016) Proceedings of marine mammal welfare workshops hosted in the Netherlands and the USA in 2012. *Aquatic Mammals* 42(3): 392–416.

Brando S., Broom D.M., Acasuso-Rivero C., Clark F. (2018) Optimal marine mammal welfare under human care: Current efforts and future directions. *Behavioural Processes* 156: 16–36.

Brando S., Buchanan-Smith H.M. (2018) The 24/7 approach to promoting optimal welfare for captive wild animals. *Behavioural Processes* 156: 83–95. doi.org/10.1016/j.beproc.2017.09.010

Brando S., Herrelko E.S. (2021) Wild animals in the city: Considering and connecting with animals in zoos and aquariums. In: Bovenkerk B., Keulartz J. (eds.). *Animals in Our Midst: The Challenges of Co-existing with Animals in the Anthropocene*. The International Library of Environmental, Agricultural and Food Ethics, vol. 33. Springer Cham, 341–360. doi.org/10.1007/978-3-030-63523-7_19

Clark F.E. (2013) Marine mammal cognition and captive care: A proposal for cognitive enrichment in zoos and aquariums. *Journal of Zoo and Aquarium Research* 1(1): 1–6. doi.org/10.19227/jzar.v1i1.19

Clark F.E., Davies S.L., Madigan A.W., Warner A.J., Kuczaj II S.A. (2013). Cognitive enrichment for bottlenose dolphins (*Tursiops truncatus*): Evaluation of a novel underwater maze device. *Zoo Biology* 32(6): 608–619. doi.org/10.1002/zoo.21096

Clegg I.L.K., Borger-Turner J.L., Eskelinen H.C. (2015) C-Well: The development of a welfare assessment index for captive bottlenose dolphins (*Tursiops truncatus*). *Animal Welfare* 24(3): 267–282.

Clegg I.L.K., Rödel H.G., Cellier M., Vink D., Michaud I., Mercera B., Böye M., Hausberger M., Lemasson A., Delfour F. (2017) Schedule of human-controlled periods structures bottlenose dolphin (*Tursiops truncatus*) behavior in their free-time. *Journal of Comparative Psychology* 131(3): 214–224. doi.org/10.1037/com0000059

Clegg I.L.K., Rödel H.G., Boivin X., Delfour F. (2018) Looking forward to interacting with their caretakers: Dolphins' anticipatory behaviour indicates motivation to participate in specific events. *Applied Animal Behaviour Science* 202: 85–93.

Clegg I.L.K., Rödel H.G., Mercera B., van der Heul S., Schrijvers T., de Laender P., Gojceta R., Zimmitti M., Verhoeven E., Burger J., Bunskoek P.E., Delfour F. (2019) Dolphins' willingness to participate (WtP) in positive reinforcement training as a potential welfare indicator, where WtP predicts early changes in health status. *Frontiers in Psychology* 10: 2112. doi.org/10.3389/fpsyg.2019.02112

Cooper R., Wemelsfelder F. (2020) Qualitative behaviour assessment as an indicator of animal emotional welfare in farm assurance. *Livestock* 25(4): 180–183. doi.org/10.12968/live.2020.25.4.180

Delfour F., Beyer H. (2012) Assessing the effectiveness of environmental enrichment in bottlenose dolphins (*Tursiops truncatus*). *Zoo Biology* 31(2): 137–150. doi.org/10.1002/zoo.20383

Delfour F., Monreal-Pawlowsky T., Vaicekauskaite R., Pilenga C., Garcia-Parraga D., Rödel H.G., García Caro N., Perlado Campos E., Mercera B. (2020) Dolphin welfare assessment under professional care: 'Willingness to participate', an indicator significantly associated with six potential 'alerting factors'. *Journal of Zoological and Botanical Gardens* 1(1): 42–60. doi.org/10.3390/jzbg1010004

Delfour F., Charles A. (2021) Understanding and assessing emotions in marine mammals under professional care. *International Journal of Comparative Psychology* 34. doi.org/10.46867/ijcp.2021.34.00.01

Duncan A.E. (1997) A veterinary assessment of the risks and benefits of environmental enrichment. *AZA Annual Conference Proceedings* 39: 208–215.

Fleming P.A., Paisley C.L., Barnes A.L., Wemelsfelder F. (2013) Application of qualitative behavioural assessment to horses during an endurance ride. *Applied Animal Behaviour Science* 144(1–2): 80–88. doi.org/10.1016/j.applanim.2012.12.001

Fleming P.A., Wickham S.L., Stockman C.A., Verbeek E., Matthews L., Wemelsfelder F. (2015) The sensitivity of QBA assessments of sheep behavioural expression to variations in visual or verbal information provided to observers. *Animal* 9(5): 878–887. doi.org/10.1017/S1751731114003164

Fleming P.A., Clarke T., Wickham S.L., Stockman C.A., Barnes A.L., Collins T., Miller D.W. (2016) The contribution of qualitative behavioural assessment to appraisal of livestock welfare. *Animal Production Science* 56(10): 1569–1578. doi.org/10.1071/AN15101

Grosso L., Battini M., Wemelsfelder F., Barbieri S., Minero M., Dalla Costa E., Mattiello S. (2016) On-farm qualitative behaviour assessment of dairy goats in different housing conditions. *Applied Animal Behaviour Science* 180: 51–57. doi.org/10.1016/j.applanim.2016.04.013

Jensen A.L.M., Delfour F., Carter T. (2013) Anticipatory behavior in captive bottlenose dolphins (*Tursiops truncatus*): A preliminary study. *Zoo Biology* 32(4): 436–444. doi.org/10.1002/zoo.21077

Justice W.S.M., O'Brien M.F., Szyszka O., Shotton J., Gilmour J.E.M., Riordan P., Wolfensohn S. (2017) Adaptation of the animal welfare assessment grid (AWAG) for monitoring animal welfare in zoological collections. *Veterinary Record* 181(6): 143–143. doi.org/10.1136/vr.104309

Kuczaj S.A., Makecha R., Trone M., Paulos R.D., Ramos J.A. (2006) Role of Peers in Cultural Innovation and Cultural Transmission: Evidence from the Play of Dolphin Calves. *International Journal of Comparative Psychology* 19(2): 223–240.

Kyngdon D.J., Minot E.O., Stafford K.J. (2003) Behavioural responses of captive common dolphins (*Delphinus delphis*) to a 'Swim-with-Dolphin' programme. *Applied Animal Behaviour Science* 81(2): 163–170. doi.org/10.1016/S0168-1591(02)00255-1

Lauderdale L.K. (2017) *Efficacy of Cognitive Enrichment for Bottlenose Dolphins (*Tursiops truncatus*): Evaluation of Planning Abilities Through the Use of a Novel Problem Solving Task.* PhD thesis, University of Southern Mississippi.

McCowan B., Marino L., Vance E., Walke L., Reiss D. (2000) Bubble ring play of bottlenose dolphins (*Tursiops truncatus*): Implications for cognition. *Journal of Comparative Psychology* 114(1): 98–106. doi.org/10.1037/0735-7036.114.1.98

Miller L.J., Mellen J., Greer T., Kuczaj S.A. (2011) The effects of education programmes on Atlantic bottlenose dolphin (*Tursiops truncatus*) behaviour. *Animal Welfare* 20(2): 159–172.

Minero M., Dalla Costa E., Dai F., Murray L.A.M., Canali E., Wemelsfelder F. (2016) Use of qualitative behaviour assessment as an indicator of welfare in donkeys. *Applied Animal Behaviour Science* 174: 147–153. doi.org/10.1016/j.applanim.2015.10.010

Moreno K.R., Macgregor R.P. (2019) Bubble trails, bursts, rings, and more: A review of multiple bubble types produced by cetaceans. *Animal Behavior and Cognition* 6(2): 105–126. doi.org/10.26451/abc.06.02.03.2019

Patel F., Wemelsfelder F., Ward S.J. (2019) Using qualitative behaviour assessment to investigate human-animal relationships in zoo-housed giraffes (*Giraffa camelopardalis*). *Animals* 9(6): 381. doi.org/10.3390/ani9060381

Phythian C.J., Michalopoulou E., Cripps P.J., Duncan J.S., Wemelsfelder F. (2016) On-farm qualitative behaviour assessment in sheep: Repeated measurements across time, and association with physical indicators of flock health and welfare. *Applied Animal Behaviour Science* 175: 23–31. doi.org/10.1016/j.applanim.2015.11.013

Rose P., Riley L. (2019) The use of qualitative behavioural assessment to zoo welfare measurement and animal husbandry change. *Journal of Zoo and Aquarium Research* 7(4): 150–161. doi.org/10.19227/jzar.v7i4.423

Rutherford K.M.D., Donald R.D., Lawrence A.B., Wemelsfelder F. (2012) Qualitative behavioural assessment of emotionality in pigs. *Applied Animal Behaviour Science* 139(3–4): 218–224. doi.org/10.1016/J.APPLANIM.2012.04.004

Samuels A., Spradlin T.R. (1995) Quantitative behavioral study of bottlenose dolphins in swim-with-dolphin programs in the United States. *Marine Mammal Science* 11(4): 520–544. doi:10.1111/J.1748-7692.1995.TB00675.X

Wemelsfelder F., Hunter E.A., Mendl M.T., Lawrence A.B. (2000) The spontaneous qualitative assessment of behavioural expressions in pigs: First explorations of a novel methodology for integrative animal welfare measurement. *Applied Animal Behaviour Science* 67(3): 193–215. doi.org/10.1016/S0168-1591(99)00093-3

Wemelsfelder F., Hunter T.E.A., Mendl M.T., Lawrence A.B. (2001) Assessing the 'whole animal': A free choice profiling approach. *Animal Behaviour* 62(2): 209–220. doi.org/10.1006/anbe.2001.1741

Wemelsfelder F. (2007) How animals communicate quality of life: The qualitative assessment of animal behaviour. *Animal Welfare* 16(1):25–31.

Wemelsfelder F., Mullan S. (2014) Applying ethological and health indicators to practical animal welfare assessment. *OIE Scientific and Technical Review* 33(1): 111–120. doi.org/10.20506/rst.33.1.2259

Whitham J.C., Wielebnowski N. (2009) Animal-based welfare monitoring: Using keeper ratings as an assessment tool. *Zoo Biology* 28(6): 545–560. doi.org/10.1002/zoo.20281

Whitham J.C., Wielebnowski N. (2013) New directions for zoo animal welfare science. *Applied Animal Behaviour Science* 147(3–4): 247–260. doi.org/10.1016/j.applanim.2013.02.004

Yon L., Williams E., Harvey N.D., Asher L. (2019) Development of a behavioural welfare assessment tool for routine use with captive elephants. *PLoS One* 14(2): e0210783. doi.org/10.1371/journal.pone.0210783