Do you see the same cat that I see? Inter- and intra-observer reliability for Qualitative Behaviour Assessment as temperament indicator in domestic cats

IC Travnik†‡, DS Machado†‡ and AC Sant’Anna*‡§

† Programa de Pós-Graduação em Comportamento e Biologia Animal, Instituto de Ciências Biológicas, Universidade Federal de Juiz de Fora, 36.036-330, Juiz de Fora, Minas Gerais, Brazil
‡ Núcleo de Estudos em Etologia e Bem-estar Animal, Departamento de Zoologia, Instituto de Ciências Biológicas, Universidade Federal de Juiz de Fora, 36.036-330, Juiz de Fora, Minas Gerais, Brazil
§ Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil
* Contact for correspondence: aline.santanna@ufjf.edu.br

Abstract

Qualitative Behaviour Assessment (QBA) is used to assess animals’ emotional expressions and its potential for serving as an indicator of temperament has been explored. This method is open to assessors’ interpretation and it is therefore necessary to evaluate the observers’ reliability for different species and contexts. We aimed to assess the intra- and inter-observer reliability of QBA as an indicator of cat (Felis catus) temperament. The QBA was applied by 19 observers with divergent profiles of contact with cats (cat owners vs non-owners) and experience in behavioural assessment (experienced vs inexperienced). Forty-two, 12-min videos were assessed, composed of footage of four behavioural tests: unfamiliar person, novel object, conspecific reaction, and food offering tests. By using Principal Component Analysis, we found three principal components (PC) that were considered the main dimensions of cat temperament. According to Kendall’s coefficient of concordance, intra-observer reliability was high to very high in PC1 (0.80–0.90) and moderate to high in PC2 and PC3 (0.50–0.82). Inter-observer reliability for the 19 observers was high in PC1 (0.71) and low in PC2 and PC3 (0.21–0.29). The individual concordances with the gold observer (defined based on greater experience with the QBA) ranged from moderate to high. We concluded that QBA could be a reliable tool to assess cat temperament, given the high values of intra- and inter-observer reliabilities in PC1, which is the dimension that most explains the behavioural variations in the cats’ temperament. The same did not occur for PC2 and PC3, showing that reliability varied among the different dimensions and observers.

Keywords: animal welfare, behaviour, companion animals, personality, rating method, shelter cats

Introduction

The use of rating methods for assessing animal behaviours has increased in recent years (Finka et al 2019; Fukimoto et al 2019, 2020; Salonen et al 2019). A widely used type of rating method is the Qualitative Behaviour Assessment (QBA) (Wemelsfelder et al 2001). The QBA was developed by Wemelsfelder et al (2000) and allows the assessment of animals’ emotional expressions, behaviour and patterns of interactions with their environment, holistically and integratively instead of via analysis of discrete and isolated categories of behaviour (Wemelsfelder & Lawrence 2001). Thus, it is possible to identify subtle variations in the animals’ body language that are hard to identify using other coding methods (Wemelsfelder et al 2001). Additionally, behaviours presented by one or a few animals, which would be disregarded in usual coding methods, can be gathered with QBA. The QBA is based on the use of descriptors to quantify positive or negative mental states on visual analogue scales.

The QBA can be used in one of two ways, from the free choice profile, where an observer uses descriptors chosen by him/her at the moment of the assessment, or the fixed-list, in which the observer uses a list of predefined descriptors (Bokkers et al 2012; Phythian et al 2013; Diaz-Lundahl et al 2019). The latter method has been regarded as valid, mainly for assessments with more practical purposes (Arena et al 2019; Travnik & Sant’Anna 2021). The QBA has been applied through video footage, usually making use of short-time videos (1–2 min) (Subejsen et al 2020), or longer observations of animals in the field (around 20 min) incorporated into welfare assessment protocols (Welfare Quality® 2009; Dwyer et al 2015). Recently, the QBA has also been shown to be valid for accessing temperament dimensions using 12-min videos showing inter-individual behavioural differences in animals’ reactions to three different stimuli (Travnik & Sant’Anna 2021).

The assessment of consistency among observers for different species and contexts is an important step so that