Applied Animal Behaviour Science xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Applied Animal Behaviour Science

journal homepage: www.elsevier.com/locate/applanim

Development of an observational quantitative temperament test in three common parrot species

Thomas Coutant[a,*], Sophie Bagur[b], Caroline Gilbert[c]

[a] ENVA, Ecole Nationale Vétérinaire d'Alfort, 7 avenue du Général de Gaulle, 94700, Maisons-Alfort, France
[b] Team Memory, Oscillations and Brain States (MOBs), Brain Plasticity Unit, CNRS, ESPCI Paris, PSL Research University, 10 rue Vauquelin, Paris, France
[c] Université Paris-Est, ENVA, Ecole Nationale Vétérinaire d'Alfort, UMR 7179, CNRS/MNHN, 1 avenue du Petit Château, 91800, Brunoy, France

ARTICLE INFO

Keywords:
Temperament
Reliability/validity
African grey parrot (*Psittacus erithacus*)
Amazon parrot (*Amazona* spp.)
Anxiety
Curiosity

ABSTRACT

Although temperament has been studied in a wide range of animal species such as primates, dog or a rodents, it has not yet been well documented in birds and in particular in psittacids. Since parrots possess developed communicative and cognitive skills, the study of personality traits is of particular interest. The aim of our study was to develop a reliable and valid temperament test by measuring quantitative behavioural parameters in two genera of medium-sized parrots: the African Grey Parrot (*Psittacus erithacus, n = 15*) and the Amazon parrot (*Amazona spp., n = 16*). We selected a set of 26 behavioural parameters based on a high intra-observer reliability. A principal component analysis was used to establish two reliable and valid temperament traits: anxiety/vigilance and curiosity/neophilia. Our test meets 5 out of 6 reliability and validity criteria which could be assessed. The two identified traits might be related to those found in other animal species, i.e. neuroticism and extraversion. These traits allowed us to demonstrate differences in the temperament of two species from two different genera: Blue-Fronted Amazon Parrots were significantly more anxious/vigilant and more curious/neophilic than African Grey Parrots. We found that parrots were more curious/neophilic when the test was repeated after six weeks, suggesting that a brief exposure to the experimental conditions resulted in a process of habituation. However, a further test eighteen months after the initial test revealed a high consistency in the two temperament traits. To our knowledge, this is the first report of an objective observational temperament test applied on two parrot species in order to compare their temperament. We believe that comparison of different avian species' temperament using similar testing procedures opens an interesting avenue of research which could be used to link temperament, phylogenetic and ecological data.

1. Introduction

Although behavioural variations between individuals may have been considered as an undesirable source of variation in biological studies, nowadays this variability has given rise to an entire field of research in human and animal psychology and ethology, which has increasingly developed over the past decades: the study of temperament and personality (Jones and Gosling, 2005). In human research, temperament is defined as heritable behavioural tendencies, established during early development that persist throughout life and provide the basis for the construction of personality (McCrae et al., 2000). Several studies have shown that such inter-individual behavioural differences also exist in animals. These differences are contextually and temporally consistent and are essentially dependent on genetics, pre-natal development and early post-natal experiences (Groothuis and Carere, 2005; Svartberg et al., 2005; Van Oortmerssen et al., 1984).

Despite numerous recent studies, a consensual definition of temperament has not yet arisen. There is also a lack of consensus in the terms used to define behavioural inter-individual differences in animals: "personality" (Ledger and Baxter, 1997), "temperament" (Jones and Gosling, 2005), "character" (Ruefenacht et al., 2002), "emotional predispositions" (Sheppard and Mills, 2002), "coping style" (Koolhaas et al., 2010; van Zeeland et al., 2013), "behavioural syndrome" or "behavioural profile" (Groothuis and Carere, 2005). In this study, we chose to use the term of temperament, which is defined as "individual behavioural differences which are consistent over time and across situations" according to Réale et al. (2007) or according to Diederich and Giffroy (2006) as "differences in behaviour between individuals that are relatively consistently displayed when tested under similar situations". In order to determine the temperament of an animal, one possibility consists in using questionnaires (Diederich and Giffroy, 2006), filled out by the owner of the animal or the person in charge of its care. However,

* Corresponding author at: CHV Frégis 43 Avenue Aristide Briand, 94110, Arcueil, France.
E-mail addresses: thomas_c_93@hotmail.fr (T. Coutant), caroline.gilbert@vet-alfort.fr (C. Gilbert).

https://doi.org/10.1016/j.applanim.2018.01.007
Received 15 August 2017; Received in revised form 11 January 2018; Accepted 14 January 2018

questionnaires may be subjective, since they rely on the perception of the person answering the questionnaire (Diederich and Giffroy, 2006; Ley et al., 2009). Another possibility is to perform several behavioural tests, to collect data by direct observations of the studied animal (Diederich and Giffroy, 2006). Animals are confronted to a series of standardized and novel situations to reduce the influence of their previous experiences, and their reactions are observed. It is usually recommended to create a diversified set of situations to allow the animal to express its full range of behaviours (Réale et al., 2007). The construction of such behavioural tests varies across studies, and can rely on two types of methodology. The first one is a qualitative and subjective approach. Experts of the studied animal species establish a list of personality items that covers the ethogram of the species, then the observer uses this list to rate the animals' behaviours for each item and analyses the ratings with dimensionality reduction techniques (Ruefenacht et al., 2002). This approach avoids subjective biases due to personal involvement with the animal but not those associated with the observer's pre-conceptions about the temperament of a species, a breed or a population of animals (Groothuis and Carere, 2005). The second approach, which can be described as quantitative and objective, is the one currently recommended to avoid biases due to observer subjectivity (Groothuis and Carere, 2005). This method does not rely on rating animals on pre-established scales during tests, but instead uses a wide-range of quantitative behavioural parameters such as the frequency or duration of different patterns of behaviour (Diederich and Giffroy, 2006) that are collected by one or several observers through a pre-established ethogram.

Since temperament tests are more and more widely used, several authors insist on the necessity of evaluating them according to reliability and validity criteria (Jones and Gosling, 2005; Diederich and Giffroy, 2006; Taylor and Mills, 2006). Reliability (comprising three criteria) concerns the degree to which the test scores are free from errors of measurement (Taylor and Mills, 2006). Tests must be based on trustworthy data collection (criteria 1: verifying intra- and inter-observer reliability), they must be temporally consistent (criteria 2: test-retest verification) and traits must assess the same behaviour (criteria 3: internal consistency). Validity (comprising three criteria) concerns the appropriateness, meaningfulness, and usefulness of the specific inferences made from the test results (Taylor and Mills, 2006). Content validity (criteria 4) evaluates whether the test measures ethologically relevant aspects of behaviour. Construct validity (criteria 5) assesses the relationship between temperament traits (i.e. whether traits should or not be correlated). Criterion validity (criteria 6) is the extent to which an association between the scores for each factor and an external criterion (behavioural or physiological measurement) can be demonstrated (Taylor and Mills, 2006). Despite the fact that these 6 criteria are essential to correctly conclude on the temperament of an animal, very few studies tested and assessed all of these criteria (Diederich and Giffroy, 2006; Jones and Gosling, 2005; Taylor and Mills, 2006).

Temperament tests have already been applied to many animal species (Brydges et al., 2008; Carlstead et al., 1999; Gosling and John, 1999; Lowe and Bradshaw, 2001; Pruitt et al., 2013; Valenchon et al., 2013) with a great variety of purposes. Studies on temperament aim for example at 1) predicting performances of utility dogs (Goddard and Beilharz, 1986; Slabbert and Odendaal, 1999; Weiss and Greenberg, 1997), 2) choosing partners optimally for reproduction in endangered species (Carlstead et al., 1999), 3) understanding the evolution of temperament across animal species, including human personality (Gosling, 2001; Groothuis and Carere, 2005), or 4) understanding how temperament influences the evolution of animal species and ecosystems (Pennisi, 2016). However, although temperament studies are now more and more numerous, only few temperament tests have been performed in avian species and even fewer in psittacine species, despite the fact that they are becoming popular pet species and possess complex cognitive capacities (Pepperberg et al., 1998; van Zeeland et al., 2013; Wilson, 1999). Until now, to our knowledge, only two studies developed temperament tests and identified traits in psittacine species. A study on parakeets (cockatiels, *Nymphicus hollandicus*) compared subjective and objective methods (Fox and Millam, 2010) and identified three temperament traits (aggression, sociability and sensitivity playfulness), and a study on parrots (orange-winged Amazons, *Amazona amazonica*) used a subjective temperament test (Cussen and Mench, 2014) to identify two traits (neuroticism and extraversion). However, these studies did not verify all criteria of reliability and validity.

A first aim of the present study was therefore to develop a quantitative and objective temperament test meeting all criteria of reliability and validity in three medium-sized parrots species: African grey parrots (*Psittacus erithacus*), blue-fronted Amazons (*Amazone aestiva*) and red-fronted Amazons (*Amazona automnalis*). We hypothesized that temperament traits revealed by the test would be similar with some of the previous results in parrots (Cussen and Mench, 2014; Fox and Millam, 2010) or other species (Réale et al., 2007). A second aim was to compare the temperament of two of these three different species to reveal potential species' related temperament differences. Such a comparison has been discussed several times (Réale et al., 2007; Uher, 2008) but has, to our knowledge, never been undertaken. Indeed, inter-species temperament comparisons have only been realised retrospectively by reviewing multiple studies (Gosling, 2001; Gosling and John, 1999) or by directly comparing behavioural parameters but not temperament traits (Capitanio, 2004; Mettke-Hofmann et al., 2002). We hypothesize, considering the broad socio-ecological differences between the studied species, that differences will arise for temperament traits.

2. Materials and methods

2.1. Subjects

All parrots tested in this study belong to the « Association de Sauvegarde et d'Accueil des Perroquets (ASAP) », a shelter for parrots that is accredited by French administrative authorities and rescues parrots which were abandoned by their owners, confiscated because of illegal detention or mistreatment, or brought by border patrol when illegally wild-caught and brought into the French territory. We included in the present study all 31 healthy parrots of three parrot species held in the centre: 15 African Grey parrots (*Psittacus erithacus*), 12 blue-fronted Amazons (*Amazone aestiva*) and 4 red-fronted Amazons (*Amazona automnalis*). No systematic information regarding the sex, age, or holding conditions prior to arrival at the shelter were available. The origin (captive or wild-caught) was the only available information (2 African Grey parrots and 2 red-fronted Amazons included in the study were wild-caught).

2.2. Procedure

2.2.1. Temperament tests

The testing procedure was performed three times at the parrot shelter, the second test 6 weeks after the first test and the third one 18 months after. The parrots were tested in a random order for the first test and the same order was kept for the two subsequent tests. The testing procedure took place in a room of approximately 8 m^2 where none of the parrots had been before and which was visually isolated from the other parrots of the shelter. At the beginning of the test, a parrot was placed on a table (area = 1.6 m^2) at the centre of the room and the test began as soon as the experimenter left the room. The room was closed and blind except for a window in the middle of the ceiling. The whole test was recorded using two cameras (HD Webcam C270, Logitech, Switzerland), one in a corner of the room that gave a side view of the table, and one on a corner of the ceiling giving a view from above, both connected to a computer during experiments.

The temperament test consisted of a series of five different subtests, lasting in total 27 min. The experimenters were the same for all tests, a familiar experimenter, and an unknown one, naive with ethological