Animal Behaviour 77 (2009) 771–783

Contents lists available at ScienceDirect

Animal Behaviour

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

Review

The repeatability of behaviour: a meta-analysis

Alison M. Bell*, Shala J. Hankison, Kate L. Laskowski

School of Integrative Biology, University of Illinois, Urbana-Champaign

ARTICLE INFO

Article history:
Received 8 August 2008
Initial acceptance 28 October 2008
Final acceptance 12 December 2008
Published online 12 February 2009
MS. number: ARV-08-00541R

Keywords:
behavioural syndrome
coping style
courtship
individual difference
mate preference
personality
temperament

There is increasing interest in individual differences in animal behaviour. Recent research now suggests that an individual's behaviour, once considered to be plastic, may be more predictable than previously thought. Here, we take advantage of the large number of studies that have estimated the repeatability of various behaviours to evaluate whether there is good evidence for consistent individual differences in behaviour and to answer some outstanding questions about possible factors that can influence repeatability. Specifically, we use meta-analysis to ask whether different types of behaviours were more repeatable than others, and if repeatability estimates depended on taxa, sex, age, field versus laboratory, the number of measures and the interval between measures. Some of the overall patterns that were revealed by this analysis were that repeatability estimates were higher in the field compared to the laboratory and repeatability was higher when the interval between observations was short. Mate preference behaviour was one of the best studied but least repeatable behaviours. Our findings prompt new insights into the relative flexibility of different types of behaviour and offer suggestions for the design and analysis of future research.

Within the field of animal behaviour, there is growing interest in consistent individual differences in behaviour (Dall et al. 2004; Sih et al. 2004a, b; Dingemanse & Reale 2005; Bell 2007; Reale et al. 2007). Accumulating evidence from a wide variety of species suggests that some individuals are consistently more aggressive, more exploratory, or more bold than other individuals and that these consistent individual differences in behaviour are often heritable (Boake 1994; Stirling et al. 2002; Kolliker 2005; van Oers et al. 2005) and related to fitness (Dingemanse & Reale 2005; Smith & Blumstein 2008). However, to date, the published data have not been summarized in a way that allows us to assess the evidence for consistent individual differences in behaviour and to explain why the magnitude of individual differences is greater in some studies compared to others.

Many studies over the past several decades have already quantified consistent individual differences in behaviour by measuring the behaviour of individuals on more than one occasion. A variety of statistics have been used to estimate behavioural consistency such as the product moment correlation or the Spearman rank correlation, but the most widely used statistic is the intraclass correlation coefficient, which estimates repeatability (Hayes & Jenkins 1997). Repeatability is the fraction of behavioural variation that is due to differences between individuals. Formally, repeatability is $r = s_A^2/s^2 + s_A^2$ where s_A^2 is the variance among individuals and s^2 is the variance within individuals over time. Behaviours that show relatively low within-individual variance compared to high among-individual variance are more repeatable. In other words, when individuals behave consistently through time and when individuals behave differently from each other, then the behaviour is repeatable. In the past, most studies measured repeatability as a first step towards studying the genetic basis for a behaviour in order to set an upper bound to heritability (Boake 1989; Dohm 2002). A different rationale for estimating repeatability is to assess interobserver reliability and the internal consistency of an instrument, (Hoffmann 2000).

From a different perspective, the large collection of repeatability estimates provides an opportunity to evaluate whether there is good evidence for consistent individual differences in behaviour and to determine whether there are systematic factors that can explain variation in behavioural consistency. Therefore summarizing this literature (previously reviewed in part in: Boake 1989; Hayes & Jenkins 1997; Forstmeier & Birkhead 2004) will provide a strong foundation for moving the study of animal personality forward.

Here, we perform a meta-analysis of the large number of published estimates of repeatability that are based on observations of a single behaviour measured on the same individuals on more than one occasion. Although closely allied with concepts of behavioural

* Correspondence: A. M. Bell, School of Integrative Biology, 505 S. Goodwin Ave., Urbana, IL 61801, U.S.A.
E-mail address: alisonmb@life.uiuc.edu (A.M. Bell).

0003-3472/$38.00
doi:10.1016/j.anbehav.2008.12.022

syndromes (Sih et al. 2004a), temperament (Reale et al. 2007), personality (Gosling 2001) and coping styles (Koolhaas et al. 1999), all of which generally refer to behavioural consistency through time and across situations, repeatability is more restrictive than these concepts because it ideally refers to consistency of a particular behaviour through time, not necessarily behavioural consistency across situations or contexts. However, in many cases, the specific environmental situations in which a behaviour is being measured are not known. As a result, repeatability estimates reflect both consistency through time and consistency across unmeasured situations (Martin & Reale 2008). Obviously, using a similar framework to assess the evidence for behavioural correlations across contexts is a promising subject for future meta-analyses.

In addition to assessing the claim that individual differences are common, we wish to know whether there are generalizations that can be made about the factors influencing repeatability. We perform an exploratory analysis to address the following questions.

Are Certain Types of Behaviour More Repeatable Than Others?

Studies have estimated the repeatability of behaviours ranging from mate preference to exploratory behaviour to parental behaviour. Therefore, we have an opportunity to ask whether certain types of behaviour are more repeatable than others. One prediction is that behaviours that are more sensitive to the environment (more plastic) are less repeatable. For example, we might assume that behaviours under morphological or physiological constraint should be relatively stable compared to behaviours influenced by energetic needs or the immediate social environment (Castellano et al. 2002; Smith & Hunter 2005). However, if all individuals respond to the environment in a similar way, the behaviour can still be repeatable despite this plasticity. Instead, repeatability estimates are especially affected by individual*environment interactions, or when individuals respond differently to the environment (Nussey et al. 2007; Martin & Reale 2008). Therefore comparing the repeatability of different types of behaviour has the potential to reveal new insights about the flexibility or canalization of different types of behaviour.

Are Certain Taxa More Repeatable Than Others?

Reviews of heritability estimates have found strong taxonomic differences (Mousseau & Roff 1987). Among vertebrates, for example, the heritability of morphological traits is significantly lower for ectotherms than it is for endotherms (Mousseau & Roff 1987), perhaps because ectotherms are more influenced by their environment. Here, we follow Mousseau & Roff's lead and test whether the same pattern applies to repeatability. We compared patterns of repeatability variation within four major phylogenetic groupings: invertebrates versus vertebrates and endothermic vertebrates versus ectothermic vertebrates. The invertebrate–vertebrate comparison allows us to evaluate the suggestion that the behaviour of taxa with less flexible nervous systems is less plastic, leading to higher repeatability estimates for invertebrates.

Does Repeatability Decrease with the Interval Between Observations?

From a genetic perspective, repeatability might decrease with the interval between measurements because the 'same' phenotypic trait may be influenced by different sets of genes at different ages. Therefore increasing the interval between measurements should decrease repeatability of the phenotypic traits because the two measures do not represent exactly the same trait at the genetic level.

Environmental effects might also cause repeatability to decrease with the interval between observations. For example, when the interval between observations of behaviour is short, it is likely that the animals are of similar state (hunger, size, age, condition, dominance, etc.) during both observations and are experiencing similar environments. For example, we might expect reproductive effort in birds to be more repeatable within broods rather than across seasons (Potti et al. 1999; Moreno et al. 2002). In contrast, when the interval between observations is long, there is more opportunity for developmental change; individuals are more likely to undergo dramatic change such as sexual maturity or a niche shift over a longer period of time. Indeed, consistency decreases with the interval between observations in humans (Roberts & DelVecchio 2000) and great tits, *Parus major* (Dingemanse et al. 2002). Published estimates of repeatability have used a wide variety of intervals between measurements; therefore, they provide an opportunity to test this intuitive suggestion.

Does Repeatability Increase with the Number of Observations Per Individual?

Several studies have measured the same individuals more than two times to calculate repeatability. On one hand, increasing the number of measurements per individual can decrease the measurement error associated with each observation, and therefore might increase repeatability (Hoffmann 2000). On the other hand, when individuals are measured repeatedly, they might habituate to the behavioural assay and become less responsive, or alternatively, become sensitized (Martin & Reale 2008). It is of practical importance to evaluate the relationship between the number of measures and repeatability for the design and analysis of future experiments (i.e. if there is much to be gained by measuring individuals more than twice; Adolph & Hardin 2007).

Does Repeatability Vary Among Age Groups?

In humans, behavioural consistency increases with maturity (Roberts & DelVecchio 2000); older people behave more consistently than younger ones, perhaps because the cumulative experience of the environment leads to increasing consistency with age. Other mechanisms that could cause repeatability to increase with age are the process of consolidated identity or reputation (Roberts & DelVecchio 2000; Dall et al. 2004). Because some studies have estimated the repeatability of behaviours in juveniles while others have measured adults, here, we have an opportunity to ask whether the same trend applies to nonhuman animals.

Do Repeatability Estimates Differ Between the Field and the Laboratory?

Presumably, environmental variance is greater in the field compared to the stable conditions in the laboratory. To the extent that a changing environment is associated with behavioural plasticity, we might expect repeatability to be lower in the field, as has been found for estimates of heritability in *Drosophila* (Hoffmann 2000).

Do Males and Females Differ in Repeatability?

Two lines of thought in the literature suggest that males might be more repeatable than females. First, the older literature on the persistence of aggression (e.g. Andrew 1972; Wingfield 1994) suggests that testosterone can cause males to be more predictable than females. Second, honest indicator models of sexual selection predict that the behaviours indicated by a sexually selected trait are predictable because females use the trait as a reliable cue for how