



Volunteer experience influences the conclusions of behavioural experiments

Luke Mangaliso Duncan, Neville Pillay*

School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, South Africa

ARTICLE INFO

Article history:

Accepted 3 June 2012

Available online 29 June 2012

Keywords:

Zoo volunteers

Data collection

Behavioural observation

Hypothesis testing

Chimpanzees

ABSTRACT

Volunteers offer an inexpensive and rapid means of collecting behavioural data, but their reliability is often overlooked. Past research has suggested that observers that were inexperienced are equally adept at recording behaviour as experienced observers, and inexperience was regarded as being merely unfamiliar with a sampling technique but not unknowledgeable about behaviour. The aims of our study were (i) to investigate the reliability of relatively naïve volunteers (i.e. those with no prior behavioural scoring experience) as behavioural data collectors; and (ii) to test the influence of the strength of inter-observer concordance on the outcome of testing a specific ethological hypothesis. Two cohorts of volunteers (high school and university students) conducted observations on a group of captive chimpanzees, simultaneously with an experienced observer (LD), recording behaviour and the location of the chimpanzees in their enclosure. Kendall's Tau agreement scores and odds ratios indicated poor agreement between inexperienced volunteers and the experienced observer, regardless of the educational experience of the volunteers and difficulty of the behaviour scored. We compared the data between the volunteers and experienced observer by independently testing each dataset with regard to the hypothesis that the chimpanzees were stressed by being in close proximity to the public. The school cohort data supported the hypothesis, while the time-matched experienced observer data suggested no relationship between public proximity and stress in chimpanzees. A separate analysis of the university cohort and time-matched experienced observer data both indicated that chimpanzees were more stressed at locations away from the public. These findings suggest that inter-observer agreement scores offer insights into the precision of data but not accuracy. Furthermore, the use of volunteers as data collectors should be assessed in relation to the aims of the study in question, since volunteers may be appropriate for studies of general patterns of behaviour but not for detailed ethological examinations.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Ethology assumes that behaviour recorded by observers reflects behaviour performed by animals (Coelho and Bramblett, 1981). Consistency in data collection is particularly important when datasets are compiled over a long

timeframe (Coelho and Bramblett, 1981), when data are used for complex and detailed analyses (Jones et al., 2001) or when ratings or recordings of behaviour encompass comparatively subjective measures, such as animal welfare (Meagher, 2009) or personality (Scott et al., 2009). Biased or unreliable recording of behaviour decreases both accuracy and precision of the data and thus compromises the scientific integrity of a study.

One way to ensure accuracy and precision of the data gathered by observers is to compare human observers against an absolute measure of behaviour. For

* Corresponding author at: 1 Jan Smuts Avenue, Braamfontein, Johannesburg 2050, South Africa. Tel.: +27 117176459; fax: +27 117176494.

E-mail address: neville.pillay@wits.ac.za (N. Pillay).

example, O'Driscoll et al. (2008) compared the recordings of human observers of the lying behaviour of cattle against an automated data logging system and found that agreement between the two systems was high. However, data loggers are not feasible for all behaviours. For example, domestic cats *Felis catus* adopt a sedentary posture for grooming, engaging in bites, licks or scratches (Trulson, 1976), behaviours which a data logger might record as lying or resting. Furthermore, behaviour is often defined both qualitatively and quantitatively. A mouse moving through an enclosure in a circuit may not be abnormal unless it is performed repeatedly which may be considered stereotypic (Jones et al., 2008). Therefore, behaviour recording requires that observers use a certain amount of subjective judgement in order to accurately record behaviour.

Given the complex and subtle nature of behaviour, many studies compare records between multiple observers as a means of assessing inter-observer reliability (Jones et al., 2001). Most studies have shown good inter-observer agreement with a variety of methodologies ranging from ethogram-based continuous sampling (Coelho and Bramblett, 1981) to subjective assessment of acute pain experience in lambs (Molony et al., 2002). In addition, observer agreement ratings do not appear to decay over time (Coelho and Bramblett, 1981).

There is some debate as to the role of experience in observer reliability. Some suggest naïve individuals may make better observers as they are less likely to be biased (Meagher, 2009) and may provide novel insights that trained ethologists might overlook (Shyan-Norwalt, 2005; Tami and Gallagher, 2009). Others argue that experienced observers make more accurate observations due to their familiarity with the methodology and/or study subjects (Margulis and Westhus, 2008). Also, individuals that regularly engage with animals are more likely to achieve good agreement than individuals with no experience of the subjects in question (Lloyd et al., 2007).

Despite the importance placed on observer experience, studies have found that inexperienced individuals are as good as experienced individuals in scoring behaviour (Coelho and Bramblett, 1981; Wemelsfelder et al., 2000; Jones et al., 2001). However, these outcomes should be interpreted with caution as Coelho and Bramblett (1981) had the observers undergo formal study of primate ethology and extensive training prior to the beginning of observations. Furthermore, both the studies of Wemelsfelder et al. (2000) and Jones et al. (2001) utilised observers who, while lacking background in the specific scenario being tested, had extensive zoological and ethological experience, using psychology and zoology graduates and trained ethologists respectively. Therefore the term 'inexperienced' is misleading and thus we hereafter consider inexperienced individuals as having no experience with behavioural data collection or training in zoology, psychology or ethology.

Many zoo-based studies use volunteer data collection (Newman et al., 2003; Shyan-Norwalt, 2005; Margulis and Westhus, 2008) because rigorous behavioural observation is often impractical for staff (Margulis and Westhus, 2008). However, in general, volunteers typically have a variety of skill sets and experience levels (Arbon et al., 2006) and

thus may misinterpret instructions (Jones et al., 2001), find behaviours difficult to identify (Tami and Gallagher, 2009) or require a degree of training to maintain accuracy (Molony et al., 2002).

To date, no empirical investigation into the reliability of zoo volunteers has been conducted. The first aim of our study was to assess the reliability of inexperienced volunteers as data collectors in a zoo setting with regard to education and experience effects on volunteer efficacy. As experience level appears to have little influence on inter-observer agreement (Coelho and Bramblett, 1981; Wemelsfelder et al., 2000; Jones et al., 2001; Tami and Gallagher, 2009), we predicted that data collected by inexperienced volunteers and those collected by an experienced individual (LD) would generate good inter-observer agreement scores. Age and aptitude do not influence the ability of volunteers to conduct wildlife surveys (Newman et al., 2003) and thus, we predicted that the level of education of inexperienced volunteers would have no influence on inter-observer agreement scores. Finally, as some behaviours may be more difficult to recognise than others (Tami and Gallagher, 2009), we predicted that agreement would be better for easily identifiable behaviours than for more difficult behaviours.

There are various standardised statistical measures of inter-observer agreement, including Spearman's correlation coefficient (O'Driscoll et al., 2008), Cohen's Kappa coefficient (Rousing et al., 2005), Kendall's coefficient of concordance and effective percentage agreement (Jones et al., 2001). Given the variation in the calculation of inter-observer agreement scores, there are no standardised units of agreement, nor does a distinct cut-off rule exist for the interpretation of these agreement estimates, but scores of 0.7 or more are generally considered to represent acceptable agreement between observers (Meagher, 2009) but some studies have considered agreement scores as low as 0.5, effectively representing an approximate 50% agreement (Bolig et al., 1992; Wielebnowski, 1999). In contrast, hypothesis testing typically requires a significance level of 95% in order to reject the null hypothesis. Data from several observers at only 50% agreement may introduce noise into the dataset, creating a bias that may compromise the outcomes of hypothesis testing. This presents a troubling question: does inter-observer agreement, or a lack thereof, influence the outcome of hypothesis testing?

Thus, the second aim of our study was to determine whether agreement scores influence the outcome of hypothesis testing. For this study, all observations by both the experienced observer and all volunteers were conducted with a group of captive chimpanzee *Pan troglodytes* study subjects, recording behaviour and location within the enclosure. Some suggest that primates are stressed by public interactions (Chamove et al., 1988), and thus we predicted that the chimpanzees would engage in more anxiety-related behaviours (self-directed grooming, abnormal behaviour and vigilance) in locations within their enclosures that place them into close proximity with or unsheltered from the public compared to locations further from or less exposed to the public. Conversely, when away from the public, the chimpanzees would engage in "relaxation" behaviours (social grooming, social play, resting).

Download English Version:

<https://daneshyari.com/en/article/4522902>

Download Persian Version:

<https://daneshyari.com/article/4522902>

[Daneshyari.com](https://daneshyari.com)