Animal Behaviour 85 (2013) 281–285

Contents lists available at SciVerse ScienceDirect

Animal Behaviour

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

Commentary

Is the magnitude of the group-size effect on vigilance underestimated?

Guy Beauchamp*

Faculty of Veterinary Medicine, University of Montréal, St-Hyacinthe, QC, Canada

ARTICLE INFO

Article history:
Received 30 July 2012
Initial acceptance 19 September 2012
Final acceptance 10 October 2012
Available online 10 November 2012
MS. number: AS-12-00590

Keywords:
funnel-shaped scatter
group-size effect
heteroscedasticity
linear regression
quantile regression analysis
statistical modelling
type II error
unmeasured variable
vigilance

In birds and mammals, individuals often interrupt their foraging activities to scan their surroundings for signs of predator activity (Caro 2005). Such vigilance may take up a large amount of time for solitary foragers. When foraging in groups, however, individuals can rely on one another to detect predators and predation risk is also diluted among more companions (Pulliam 1973; Bertram 1978). As a consequence, individuals in groups can decrease their own investment in vigilance at no increased risk to themselves. The group-size effect on antipredator vigilance represents this downward adjustment in vigilance as group size increases. Since it was predicted in the early 1970s, hundreds of studies have sought changes in vigilance with group size in birds and mammals. In a recent meta-analysis in birds, for instance, a decrease in vigilance with group size was documented in a large proportion of the species (Beauchamp 2008) and was found to be medium in size, in line with the magnitude reported for other phenomena in evolutionary biology (Møller & Jennions 2002).

While support for the group-size effect on vigilance represents a success story in animal behaviour research, many issues remain problematic. For instance, why the magnitude of the effect varies from species to species is not clear and, in addition, several of the assumptions underlying traditional vigilance models have been challenged recently (Lima 1995; Sirot & Touzalin 2009). Here, I want to draw attention to a methodological issue that arises when fitting statistical models to vigilance data. Typically, researchers regress the percentage of time spent vigilant by focal individuals in a group as a function of group size. The relationship between vigilance and group size is not expected to be linear (Dehn 1990), and while this issue is sometimes ignored, transformation of the data often produces a linear relationship. The most common statistical tool used to model the group-size effect is the simple linear regression. The word ‘simple’ is used in the sense that group size is the only independent variable considered in the model. Multiple regression models have also been employed and include cofactors such as temperature and sex, which may influence vigilance levels on their own (Liley & Creel 2008; Sansom et al. 2008). An underlying assumption in regression models, be they simple or multiple, is homoscedasticity, which means that for each possible group size, the dispersion of the data about the mean should be the same (Sokal & Rohlf 1995). Heteroscedasticity, its opposite, can be clearly visualized in a scatterplot: the data will be more spread out at one end producing a funnel-shaped scatter. Estimates of confidence intervals about the slope of the regression can be biased

* Correspondence: G. Beauchamp, Faculty of Veterinary Medicine, PO BOX 5000, St-Hyacinthe, QC J2S 7C6, Canada.
E-mail address: guy.beauchamp@umontreal.ca.

http://dx.doi.org/10.1016/j.anbehav.2012.10.023

when the assumption of homogeneity of variances is not met. However, this is not an issue in vigilance research since the regression is rarely used for that purpose.

More relevant to students of animal behaviour, a funnel-shaped scatter in a linear regression may suggest that at least one unmeasured variable influences the response variable (Rosenbaum 1995). In many observational studies, such as those that document the group-size effect on vigilance, it is probably the rule, rather than the exception, that more than one independent variable may be involved. This raises the question: what would happen if a third variable comes into play and interacts with group size? The concept of interaction between two independent variables is well known in statistics: an interaction occurs when the effect produced by one variable varies depending on the value taken by the other variable. In a typical two-way ANOVA setting, an interaction indicates that the difference between the means associated with one factor will not be the same at all levels of the other factor. For instance, the effect of a drug on body temperature may be higher in older than in younger patients, resulting in an interaction between drug and age. In an ANCOVA setting, the slope between two variables will vary depending on the level of the cofactor. In an ANCOVA with an interaction, the scatter of data will typically be wider at one end of the plot, producing a funnel shape.

The influence of interactions in linear regression has been examined extensively in the ecological literature. Ecologists have long been interested in such issues because interaction effects are pervasive in the form of limiting factors that prevent organisms from fully responding to changes in the independent variable of interest (Cade et al. 1999). Consider the relationship between the number of fish in a stream and the ratio between the width and the depth of the stream (Dunham et al. 2002). Ideally, the scatter in the data should be minimal about the trend line if the ratio alone explained variation in fish number. However, this relationship may vary as a function of another unmeasured factor, say temperature, such that the expected relationship is apparent only in a certain range of temperatures. Because of such limiting factors, scatterplots will often look funnel-shaped.

In addition to the suggestion that unmeasured factors may be at play, heteroscedasticity may also imply that the magnitude of the relationship between two variables, which is measured by the slope of the linear regression, may be poorly estimated. This is because the standard linear regression only looks at the effect of an independent variable on changes in the mean of the dependent variable, and such changes may not be representative of the real effect at work (Cade & Noon 2003). To illustrate this point, I simulated the group-size effect on vigilance and then introduced a negative interaction with a third variable that would be unknown to the field researcher but whose effect is known in the simulation. A negative interaction makes sense in the context of vigilance since the effect of any third factor is more likely in small groups than in large groups. In large groups, vigilance levels are expected to be quite low, and the scope for adjustments in vigilance is therefore more limited.

In the first step of the simulation, I created 200 groups of varying sizes and evaluated vigilance in each simulated group using the following equation: intercept $-$ β1 $\times$ group size $+$ error. In this equation, β1 represents the magnitude of the group-size effect. To the vigilance level in each group, a random number, the error term, was added to create random scatter in the plot (see Fig. 1 for details). The scatterplot for this simulation shows in a simple way the expected decrease in vigilance with group size (Fig. 1a). In the second step, a negative interaction term with an unmeasured variable was added to the above model. The scatterplot for the second step of the simulation (Fig. 1b) clearly shows that the addition of a negative interaction term produces a funnel-shaped

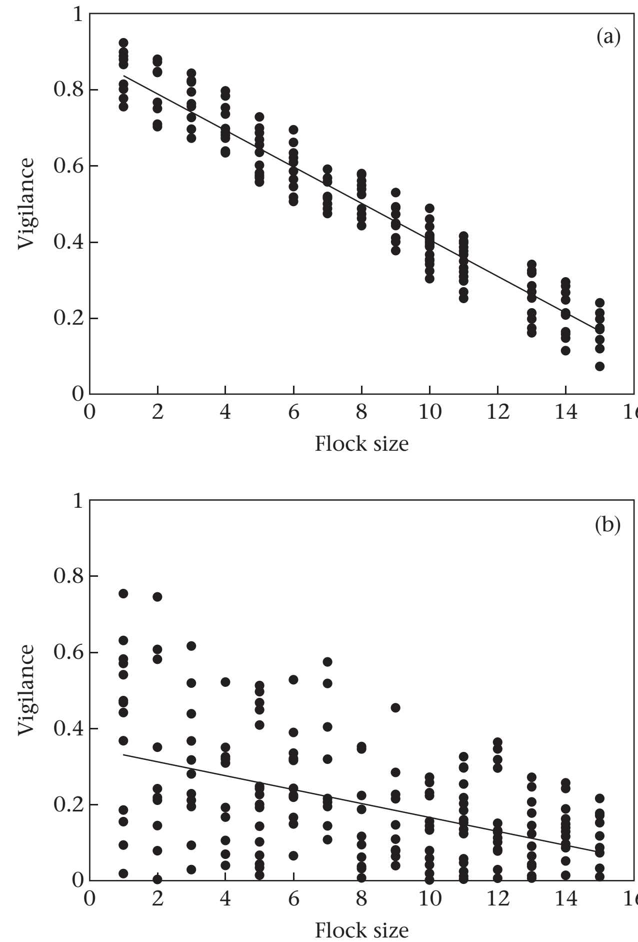

Figure 1. Changes in vigilance levels with group size in 200 simulated groups. (a) Vigilance for each group was simulated using the following equation: $0.9 - 0.05 \times$ group size $+$ error. The size of each group was randomly selected from a uniform distribution ranging between 1 and 15. A random number from a uniform distribution, ranging from -0.1 to 0.1, was added to the vigilance level of each group for error. The linear regression line is shown and the slope was equal to -0.05. (b) The same number of groups was simulated and the same equation was used to determine vigilance levels, but this time a negative interaction term with a third variable, *y*, was added with parameter 0.1. Values for the third variable *y* were selected from a uniform distribution ranging from 0 to 1. The linear regression line for the relationship between vigilance and group size is shown and the slope was -0.02, a value that underestimates the predicted value of -0.05.

scatter, and that the value of β1 is underestimated by a simple linear regression that ignores the interaction effect, as would be the case if this factor was not measured in the field. Extensive simulation studies in the ecological literature confirm that interactions produce heteroscedasticity and that the value of the slope (the magnitude of the effect) is underestimated (Cade et al. 2005).

Is the problem of heteroscedasticity an issue in the literature on vigilance? And, if it is the case, is the magnitude of the group-size effect also underestimated? To answer these questions, I examined scatterplots of the relationship between vigilance and group size in the published literature over the past 40 years. In plots where individual data points were quite apart, I recreated the data set using coordinates from the plots and ran a standard linear regression between vigilance and the logarithm base 10 of group size. Taking the logarithm produced a more linear relationship between vigilance and group size. To produce comparable slope estimates from plot to plot, I used *z* scores for vigilance values.

To assess heteroscedasticity, I used quantile regression analysis. The purpose of a quantile regression is the same as a standard linear