PROCEEDINGS B

Invited reply

CrossMark
click for updates

Author for correspondence:
Charles Perreault
e-mail: cperreault@asu.edu

The accompanying comment can be viewed at http://dx.doi.org/10.1098/rspb.2015.2184.

THE ROYAL SOCIETY
PUBLISHING

Cultural history, not ecological environment, is the main determinant of human behaviour

Sarah Mathew[1,2] and Charles Perreault[1,2]

[1]School of Human Evolution and Social Change, Arizona State University, PO Box 872402, Tempe, AZ 85287, USA
[2]Institute of Human Origins, Arizona State University, Tempe, AZ 85287-4101, USA

Towner *et al.* [1] question the methods and the theoretical framework of our study of behavioural variation among Native American tribes of Western North America [2]. Here we show that their concerns are unfounded and that our results are robust. We also clarify the theoretical issues that motivated our paper, and explain why it is critical to disentangle the role of ecology and cultural inheritance in a cultural species like humans.

Towner *et al.* contend that the higher summed absolute value of the cultural–historical betas (i.e. C) relative to that of ecology (i.e. E) is due to the fact that cultural history has 10 potential predictors and ecology has seven. Their argument is based on the fact that the expectation of the absolute beta of a predictor representing stochastic noise is not zero, but $(2\sigma^2/\pi)^{1/2}$ (assuming the coefficients are drawn from a Gaussian distribution with mean zero and standard deviation σ). This is because the probability density on the negative side of the distribution of beta is shifted to the positive side in the distribution of the absolute beta. Thus, any predictor in a model will contribute to C or E, even when they represent stochastic noise. Towner *et al.* suggest that the correct measure of the effect size of cultural history and ecology is $C - M_C(2\sigma^2/\pi)^{1/2}$ and $E - M_E(2\sigma^2/\pi)^{1/2}$, where M_C and M_E are the number of cultural historical and ecological predictors, respectively.

We show here that this correction is not necessary, and neither does it change th results. The model selection approach shields our results from the effect of predictor class size. Predictors that represent stochastic noise will have betas with large standard error relative to their effect size. Such predictors are unlikely to be included in the best model, as the Akaike information criterion (AIC) penalizes models for the number of predictors they contain. As a result, the absolute value of a β coefficient in the best model is a good measure of the true absolute magnitude of the effect of that predictor, even if it does not incorporate explicitly the standard error of the β estimate.

However, explicitly incorporating the standard error of the β estimates, as Towner *et al.* suggest, is a valid alternative to the approach we used. But Towner *et al.*'s correction is incomplete, because it is only applicable to the worst-case scenario where the predictors represent stochastic noise. Just as the expectation of the absolute beta of a predictor representing stochastic noise is shifted from zero to $(2\sigma^2/\pi)^{1/2}$, the expectation of the absolute effect of a predictor with mean β and standard deviation σ is

$$E(|\beta|) = \beta - 2\int_{-\infty}^{0} x \, \frac{1}{\sigma\sqrt{2\pi}} e^{((x-\beta)^2/2\sigma^2)}.$$

Thus, the absolute effect of a predictor, β', is the difference between $E(|\beta|)$ and the null expectation

$$\beta' = \beta - 2\int_{-\infty}^{0} x \, \frac{1}{\sigma\sqrt{2\pi}} e^{-((x-\beta)^2/2\sigma^2)} - \left(\frac{2\sigma^2}{\pi}\right)^{1/2}. \tag{1.1}$$

The extent to which β' differs from $|\beta|$ depends on the extent to which the density of the distribution of the β coefficient encompasses both negative and

2

positive values. When the magnitude of an effect size is large relative to the standard error, $\beta' - |\beta|$ will be small, and vice versa. Because the best model is unlikely to contain predictors with small effect size relative to standard error, it is not surprising that when we reanalyse the data using Towner *et al.*'s approach, we get the same results as before (electronic supplementary material, figures S1–S3).

Towner *et al.* claim that our results are driven by the fact that 29 out of 457 traits have extremely large betas. This is not the case. Our results are unchanged by the removal of these 29 traits (electronic supplementary material, figures S4 and S5).

Towner *et al.* are confused about our use of the arc tan transformation. The arc tan transformation is a natural way to calculate the ratio between effect sizes when effect sizes can be zero. Consider a trait with two summed absolute values, *C* and *E*. The arc tan transformation represents the relative magnitude of *C* and *E* as the slope, ranging from $0°$ to $90°$, of the line that goes from the origin to the coordinates of a trait (electronic supplementary material, figure S6).

Towner *et al.* argue that analysing the linguistic matrices with the *prcomp* function is problematic, because the *prcomp* function cannot operate on pairwise datasets, but only on datasets consisting of sampling units in rows and measured attributes in columns. It is obvious from the dimensions of the linguistic matrices (172×116, 172×85, etc.) that the matrices are not pairwise. More importantly, the rows of the linguistic matrices are our sampling units (i.e. tribes) and the entries in the columns are attributes (i.e. the language group to which a tribe belongs). Similarly, the rows of the spatial distance matrix represent the sampling units (i.e. tribes), and the entries in the columns are attributes (i.e. the distance from potential sources of diffusion).

Towner *et al.* suggest that we should assess the importance of ecology and cultural history by finding which of a set of *a priori* models performs the best according to some information criterion. For example, cultural history will be considered an important determinant of a trait if the AIC of the model that includes both cultural historical predictors and ecological predictors is lower than the AIC score of a model including only ecological predictors. Given our research question, this approach is inadequate, as it provides no information about the relative effect sizes of ecology and cultural history. Our analysis not only shows that the best models include both ecology and culture predictors for almost all of the traits, but it also specifies the relative magnitude of their effects.

Towner *et al.* also misunderstand the theoretical issues that motivated our paper. They argue that our comparison of the effect of ecological environment and cultural history is flawed (i) because it fails to acknowledge that social learning is itself a mode of adaptation to the environment, (ii) because nobody has ever seriously proposed that behavioural strategies emerge de novo with each generation, and (iii) because the interactions between the mechanisms underlying behaviour are too complex to be studied.

First, we did acknowledge that social learning leads to behaviours that are adapted to local environments. This is what we meant when we wrote 'social learning can also lead to behaviours that are adapted to local environments'. Therefore, the effect of ecology may not only be due to single-generation adaptive mechanisms, such as trial-and-error learning and reaction norms, but also due to the effect of cultural adaptations to local ecology. In contrast, the effect of cultural history can only be attributed to social learning, because only social learning (genetic evolution aside) operates over multiple generations. Thus, our estimate of the effect size of cultural history is a conservative measure of the effect of social learning as it excludes rapid cultural adaptations to environments.

Second, behavioural strategies do emerge de novo with each generation among all animal species, including humans. No one doubts that animals respond to their environments through non-social adaptive responses, such as evolved heuristics, trial-and-error learning, reasoning and developmental plasticity. These non-social adaptive mechanisms will lead to de novo emergence of behaviour each generation as individuals independently converge on the same behaviour. Given the importance of these mechanisms in the world of non-human animals, it is important to ask what role they play in humans. Thus, it is not surprising that a number of researchers have stressed the importance of these mechanisms at the expense of cultural mechanisms [3,4].

Third, Towner *et al.* would like us to give up on comparing the effect of cultural and non-cultural mechanisms, because their interactions are too complex. Their logic is puzzling. Anthropologists have long been disentangling the effect of cultural history from that of the environment. As early as the late nineteenth century, social scientists recognized that societies can be similar not because they have converged on the same behavioural strategies independently, but because they share a cultural ancestor. Recently, anthropologists, including some of the co-authors of the comment, have advocated for the use of cultural phylogenetic methods in order to control for shared ancestry [5]. The premise of cultural phylogenetic methods is that the effect of shared ancestry can be separated from that of the other mechanisms that shape human behaviour. If Towner *et al.*'s argument is valid, then these efforts at controlling for shared ancestry are misguided.

The difference between our approach and cultural phylogenetic methods is that our approach puts shared cultural ancestry on equal footing with the other predictors. This allows us to quantify its effect size on the same scale as the other classes of predictors. In contrast, cultural phylogenetic methods treat shared ancestry as a factor that needs to be muted in order to reveal what is scientifically interesting. Whereas these methods were developed to test adaptive hypotheses about human behaviour, they may be silencing the main mechanism that gives rise to human adaptation, cultural evolution over multiple generations.

Towner *et al.*'s view is also at odds with the study of trait heritability in biology. Like culture, genes form an inheritance system. Researchers routinely partition the effect of genes and environment on all sorts of traits, including behavioural traits. In doing so, they compare the effect of shared genetic ancestry with that of other sources of phenotypic variation, such as developmental plasticity, reaction norms and learning. It would make no sense for biologists to refrain from efforts to disentangle the effect of genes and environments on the basis that they both interact in complex ways to give rise to phenotypes.

References

1. Towner MC, Grote MN, Borgerhoff Mulder M. 2016 Problems modelling behavioural variation across Western North American Indian societies. *Proc. R. Soc. B* **283**, 20152184. (doi:10.1098/rspb. 2015.2184)

2. Mathew S, Perreault C. 2015 Behavioural variation in 172 small-scale societies indicates that social learning is the main mode of human adaptation.

Proc. R. Soc. B **282**, 20150061. (doi:10.1098/rspb. 2015.0061)

3. Pinker S. 2010 The cognitive niche: coevolution of intelligence, sociality and language. *Proc. Natl Acad. Sci. USA* **107**, 8993 – 8999. (doi:10.1073/pnas. 0914630107)

4. Tooby J, Cosmides L. 1992 The psychological foundations of culture. In *The adapted mind:*

evolutionary psychology and the generation of culture (eds JH Barkow, J Tooby, L Cosmides), pp. 19 – 136. New York, NY: Oxford University Press.

5. Borgerhoff MM. 2001 Using phylogenetically based comparative methods in anthropology: more questions than answers. *Evol. Anthropol. Issues, News Rev.* **10**, 99 – 111. (doi:10.1002/evan.1020)