Animal Behaviour 110 (2015) e9–e12

Contents lists available at ScienceDirect

Animal Behaviour

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

Forum

Why does costly signalling evolve? Challenges with testing the handicap hypothesis

Szabolcs Számadó [a], Dustin J. Penn [b,*]

[a] *MTA-ELTE Theoretical Biology and Evolutionary Ecology Research Group and, Department of Plant Systematics, Ecology and Theoretical Biology, L. Eötvös University, Budapest, Hungary*
[b] *Konrad Lorenz Institute of Ethology, Department of Integrative Biology and Evolution, University of Veterinary Medicine, Vienna, Austria*

ARTICLE INFO

Article history:
Received 26 March 2015
Initial acceptance 29 April 2015
Final acceptance 26 May 2015
Available online 20 October 2015
MS. number: 15-00259R

Keywords:
action-response games
animal communication
handicap principle
receiver psychology
strategic costs

Zahavi's handicap hypothesis (Grafen, 1990; Zahavi, 1975; Zahavi & Zahavi, 1997) is a popular explanation for the evolution of honest and costly signalling. The general idea is that individuals honestly signal their quality because signalling is costly and therefore low-quality individuals cannot afford to produce dishonest signals. However, this hypothesis is controversial for several reasons. (1) Zahavi suggested that selection favours the evolution of honest signalling because (and not despite) of their costs, and he made the radical suggestion that when it comes to the evolution of signalling, natural selection favours waste rather than efficiency. (2) Zahavi argued that this idea is a general principle, not merely a hypothesis, which explains honest signalling in most or all contexts. (3) There are several versions of the handicap hypothesis, but attempts to provide theoretical support have largely failed. The main exception is a model proposed by Grafen (1990), which has become widely accepted among behavioural ecologists; however, his conclusions have been challenged (Bergstrom, Számadó, & Lachmann, 2002; Getty, 1998, 2006; Hurd, 1995; Lachmann, Számado, & Bergstrom, 2001; Számadó, 1999, 2000, 2011). (4) There have been many attempts to empirically test the handicap hypothesis, but there is no consensus regarding how it might be tested (Kotiaho, 2001).

Despite these difficulties, Polnaszek and Stephens (2014) recently conducted a study with trained blue jays, *Cyanocitta cristata*, to experimentally test the handicap hypothesis. They concluded that their findings provide the first experimental evidence that signal costs enforce honesty, and they interpreted their results to support the handicap principle. This experiment is unusually clever and insightful, and the findings provide important implications for honest signalling and receiver psychology (Guilford & Dawkins, 1991). However, we raise several caveats about the theoretical background, interpretations and conclusions of the study, and we explain why this study and other attempts to test the handicap hypothesis will be problematic as long as there is not a clear theoretical model to test.

THE JAY TRAINING EXPERIMENT

In this experiment, pairs of blue jays occupying adjacent cages were trained to play a communication game in which one bird, the sender, could choose to hop onto one of two perches, which could be used as a signal about the state of the environment, and the receiver responded by selecting a perch on the same or opposite side of the sender, depending upon the signal it perceived (Polnaszek & Stephens, 2014). The sender could choose to send an honest or dishonest signal about the environment, depending on whether one of the two red lights in the signaller's cage (visible only for the signaller) were turned on or off indicating the state for the given trial as either true or false. The birds were experimentally rewarded depending on their choices and they were tested under two conditions. In the incentives-aligned treatment, there was mutual interest between signaller and receiver, as both birds were rewarded for choosing a response that corresponded to the state of the environment. In the incentives-opposed treatment, there was a conflict of interest, as the signaller was interested in selecting the signal state regardless of the state of the environment, whereas the receiver was only rewarded if the response corresponded to the state of the environment. The authors also experimentally manipulated the cost of signalling by forcing the sender to take loops of

* Correspondence: D. J. Penn, Konrad Lorenz Institute of Ethology, Department of Integrative Biology and Evolution, University of Veterinary Medicine, Vienna, Savoyenstraße 1a, A-1160 Vienna, Austria.
E-mail address: Dustin.Penn@vetmeduni.ac.at (D. J. Penn).

http://dx.doi.org/10.1016/j.anbehav.2015.06.005

shuttle flights between a third perch and its current position before it could use the signalling or the nonsignalling perch. The authors showed that when there was no conflict (incentives-aligned treatment), the jays produced honest signals, and increasing cost of the signals had no effect on honesty. However, when they increased the conflict (incentives-opposed treatments), increasing the signalling costs affected their honesty: when the costs of signalling were low, they were often dishonest (not corresponding to the state of the environment), whereas when the costs of signalling increased, the jays produced more honest signals. The study also showed that the receivers followed or trusted the signals more often when they were reliable. The authors concluded that their study provides the first experimental evidence demonstrating that signal costs stabilize honesty, and they imply that this finding confirms the handicap principle.

ZAHAVI'S HANDICAP PRINCIPLE

Rather than supporting Zahavi's handicap principle (Zahavi & Zahavi, 1997), the findings in this study contradict this proposal. The costs of signalling stabilized honesty, but only when there was a conflict of interest between signaller and receivers. To our knowledge, this study provides the first experimental evidence that signals need not be costly to be honest under shared interests, and that signal cost has no effect on honesty under such conditions. This result is theoretically expected, but it contradicts suggestions that the handicap hypothesis is a general principle that explains honest signalling (with and without conflicts of interest; Zahavi & Zahavi, 1997). Also, Zahavi assumed that honest signals must be perceptibly costly or wasteful, since this is the only way to demonstrate honesty, and yet the birds' shuttle flights (the costs that maintained honesty) could not be seen by the receivers. There are other restricted versions of the handicap hypothesis, but as we explain next, these models were not supported either.

HANDICAPS AS STRATEGIC COSTS

The jay study was also interpreted to support a version of the handicap hypothesis proposed by Maynard Smith and Harper (1995), which views handicaps as strategic costs of signalling, and Polnaszek and Stephens (2014, p. 2) defined handicaps accordingly, i.e. 'any signal whose reliability is ensured by costs that exceed the minimal cost necessary to make the signal'. All signals have production or efficacy costs, which are necessary for a trait to transmit information or influence the behaviour of conspecifics, and the Maynard Smith and Harper (1995) version crucially predicts that they have additional strategic costs (the cost component that maintains honesty under conflict of interests). A cricket's song is costly to produce to reach females from afar (production costs), but the question is whether the males' songs are more costly than they need to be to reach female receivers. Do gazelles jump higher than they need to jump to signal their health to predators when stotting? No one has proposed how to measure such strategic costs, and the jay experiment did not attempt to distinguish strategic versus efficacy costs of signalling, which is the basis for this definition of handicaps.

THE STRATEGIC HANDICAP HYPOTHESIS

Polnaszek and Stephens (2014, p. 6) also cited Grafen's (1990) strategic handicap hypothesis as the 'authoritative mathematical statement of the handicap principle'; however, criticisms of his model (Getty, 1998, 2006) and conclusions (Hurd, 1995; Lachmann et al., 2001; Számadó, 1999, 2011) were too lightly brushed off. Grafen's (1990) main results were that (1) signals are honest, (2) signals are costly and (3) signals are costlier for worse signallers, and yet these conditions have all been challenged by later models and empirical results (see Számadó, 2011 for a review). Signals need not be honest, not even on average, to evolve (Számadó, 2000). Honest signals need not be costly even under conflicts of interest (Bergstrom et al., 2002; Hurd, 1995; Lachmann et al., 2001; Számadó, 1999) and honest costly signals need not be costlier for poor-quality signallers (Getty, 1998, 2006).

It is also unclear how the jay experiment provides evidence or a test of Grafen's strategic handicap model. The versions of the model proposed by Grafen (1990) and Zahavi and Zahavi (1997) assume that the costs of signalling that enforce honesty are a strategic choice (where individuals can choose their level of investment) rather than an unavoidable constraint imposed on the signallers, for example high-quality signallers could use low-intensity signals but they 'choose' not to and vice versa. However, in the jay experiment costs of shuttle flights were artificially forced on the signallers: the birds could not use the signalling perch before paying the full cost of the signal. In addition, an experimental test requires showing that the marginal cost of producing the same signal is greater for low- than high-quality individuals, but this hypothesis was not tested for two reasons. First, the quality or condition of the birds was not known or examined, and quality was only mimicked by imposing two different conditions ('true' versus 'false') on the jays, which were signalled by red lights. This implementation is irrelevant to the jays' ability to bear the cost of signalling. Second, the model in the jay study is a differential benefit model (like the Sir Phylip Sydney game, Maynard Smith, 1991), rather than a differential cost model (Grafen, 1990). The costs imposed on the signallers were the same in the two different conditions, and thus, by definition, there cannot be any difference in the marginal costs.

ACTION-RESPONSE GAME VERSUS HANDICAP MODEL

The authors constructed a simple model to derive the conditions of honesty for the jay experiment, and they cited Grafen's model (1990) as the 'authoritative cost condition' (Polnaszek & Stephens, 2014, p. 3) of honesty. However, the authors' model is an example of an action-response game (Hurd, 1995; Számadó, 1999) rather than a handicap model, and the conditions of honesty that can be derived from these games are different (see Appendix). The results of action-response games show that honest signals need not be costly not even under conflict of interest for high-quality signallers (Bergstrom et al., 2002; Hurd, 1995; Lachmann et al., 2001; Számadó, 1999), contrary to previous authors' claims (Grafen, 1990; Maynard Smith & Harper, 1995; Zahavi & Zahavi, 1997), assuming that signal costs vary as a function of quality. The explanation is that it is not the cost paid by 'high-quality', i.e. true condition, signallers at the equilibrium that maintains honesty, but the potential cost of cheating for 'low-quality signallers', i.e. false condition (Hurd, 1995; Számadó, 1999). This potential cost of cheating will be paid at the equilibrium for high-quality signallers only if there is a constraint linking the signal cost paid by low-quality signallers to the cost paid by high-quality signallers.

In terms of the jay experiment, if the experimenters impose a cost only on the 'false' condition, the system still remains honest and individuals under the 'true' condition (i.e., 'high-quality' individuals) do not have to pay a cost at the equilibrium. Consequently, if individuals pay a cost under the 'true' condition, then it is only because the constraint imposed by the experimenters was chosen that way (i.e. they implemented a differential benefit model). Therefore, results of the experiment cannot be used as evidence in favour of the necessity of such cost (as assumed by the handicap models), as it only reflects the choice made by the experimenters.