EDITORIAL

The Use of Theory in Behavioural Research

Michael Taborsky
University of Bern, Bern, Switzerland

Scientific research ought to be based on a theoretical framework. This truism is reflected also in the ‘aims and scope’ of Ethology, stating that this journal of behavioural biology ‘contains scientific articles of general interest … that are based on a theoretical framework’. But what exactly does this mean? How do behavioural researchers base their studies on theory? And is this basement well founded?

The use of theory may come in different flavours. In ethology we aim to understand the causal mechanisms underlying behaviour, at both the ultimate (evolutionary) and proximate (mechanistic) levels (Tinbergen 1963; Bateson 2003; Taborsky 2006). Often empirical research is simply based on plausibility arguments. This may be adequate especially at the early stages of an investigation, e.g. when investigating the largely unknown recognition abilities of a species (Jouventin et al. 2007), or when a research question is based on very general expectations. For example, when the relationship between predators and prey is at issue, one may wish to confirm that conspicuousness or other handicaps increase predation risk (Bateman & Fleming 2006; Husak et al. 2006) or that vigilance and other behaviours are influenced by the risk of predation (Baxter et al. 2006; Kutsukake 2006). In a different context, it seems logical that rearing conditions and social learning influence social preferences and discrimination of kin (Frommen et al. 2007; Spence & Smith 2007). It might seem redundant to base a study of grooming intensity in dependence of ectoparasite infestation on a formal model (Hawlena et al. 2007), or to use theoretical inference to predict whether parasite infestation affects signalling traits (Gilman et al. 2007). Once such relationship has been established, however, one might wish to know whether it is caused by a constraint that is overcome differently by individuals of varying quality, and whether this allows conspecifics to evaluate individuals by the quality of their signals (Hamilton & Zuk 1982), which would then require a tighter link to theoretical concepts. Such link might help, for example, when asking whether and how ornament size should correlate with other individual qualities (e.g. fighting...
success, Bokony et al. 2006), or whether and how females should adjust the sex ratio of offspring to the quality of their mate (e.g. Karino et al. 2006).

When studies are not based on a formal approach but rather on plausibility arguments or ‘common knowledge’ one may be caught by surprise when straightforward expectations are not met. For instance, dominance is usually assumed to raise the reproductive success of males, but this is not so in meadow voles (Spritzer et al. 2006). It is plausible to assume that increasing group size reduces neophobia, but ravens explore novel objects more readily when on their own instead of in groups (Stoewe et al. 2006). It is such discrepancies between general expectations and contradicting results where the issue becomes exciting and where a formal theoretical model may be particularly useful to guide further research.

When empiricists base their research firmly on a theoretical framework, most commonly either (1) predictions are derived from a general model and tested with a data set from a natural population providing the required information (Jennings et al. 2006; Porkert & Spinka 2006; Sergio et al. 2007; Stone 2007); or (2) predictions derived from heuristic models are tested by an experiment (Leitner et al. 2006; Mariette et al. 2006; Blanchet et al. 2007; Dyrzc & Halupka 2007). Sometimes, both approaches are applied in a study in parallel (Schwartz et al. 2007). These two approaches differ in an important way: the first one is correlative, so it is not suited to test for causal relationships. It affords an opportunity to generate appropriate questions and is important to direct research, especially early in an investigation and if the acquired information reflects natural conditions. The second approach is experimental, so in principle it is perfectly suited to uncover causal relationships. If precise predictions of a theoretical model can be tested by specifically designed experiments, this vulnerability to falsification is assertedly the strongest tool of scientific methodology (Popper 1965). Can we conclude then that the causal relation has been identified if the predictions of a model are confirmed by experimental data? We may hope so, but should not be tempted to conclude this prematurely. General models usually rest on extremely simple assumptions, by intention excluding a large number of potentially important variables and their interactions. If an experiment is to unravel causal relations, it must ensure that the assumptions of the model are met and that no significant factor or relationship between variables has gone unnoticed. Otherwise we might have located a potential mechanism of a pattern, but not identified the appropriate process behind it.

To unravel mechanisms underlying behaviour is not a small task. We may illustrate the difficulty by examining the research history of a well-studied paradigm within the realm of optimal foraging theory. The time animals are expected to spend collecting food in a patch should depend on its profitability and the effort required to find an alternative food patch of sufficient quality. The marginal value theorem (MVT) modelled by Charnov (1976) predicts that if net rate of energy intake is to be maximized, a patch should be left when the foraging rate obtained there falls below the average intake rate for the entire habitat. This prediction has been tested numerous times in a wide range of animals, and it revealed generally an excellent match between prediction and experimental results (Stephens & Krebs 1986; Cuthill et al. 1994; Cezilly & Benhamou 1996). However, many empirical studies test only qualitative predictions of the model (e.g. Naef-Daenzer 2000; Amano et al. 2006). When animal decisions in complex foraging situations are tested for quantitative predictions of the MVT, the fit is by no means convincing (Nonacs 2001; Carmel & Ben-Haim 2005; Searle et al. 2005; Wajnberg et al. 2006; see also Hancock & Milner-Gulland 2006). A survey of 26 studies revealed a consistent deviation of foragers’ decisions in more than three out of four cases (Nonacs 2001). This is a disturbing discrepancy, suggesting that in addition to the assumptions inherent in the MVT a realistic model of optimal patch choice decisions might need to consider intrinsic behavioural propensities, opportunity costs, simultaneous benefits of foraging patterns, nutritional states of animals, specific nutritional needs, limitations on information about patch and general habitat quality, costs of sampling, social forces and cognitive constraints, among others.

This may teach us that interpreting coincidence between qualitative predictions of a general model and empirical results as proof that the model explains the causal mechanisms (ultimate or proximate) underlying the trait is all too naive. Nevertheless, this is one of the most popular procedures in behavioural ecology. One might call it the ‘pseudo-proof fallacy’, because a failure to falsify the predictions of a model does not mean that the mechanism modelled is indeed underlying the pattern. This is nothing but the impossible proof of the null hypothesis. To be content with concordance between predictions and results at a qualitative level may be dangerous and misleading. Misleading because it makes us think the data are explained by the model from which the predictions were generated; dangerous because we may stop searching further when we have only detected one
potential explanation, without regarding alternatives or essential extension. Often different models make similar predictions, i.e., the same pattern may be generated by completely diverse processes (Carmel & Ben Haim 2005). Even worse, by changing small and seemingly unimportant assumptions the predictions of a model may reverse (Hamilton & Heg 2007). Biologists bear the insuperable burden of dealing with the formidable complexity inherent in all living matter. This is what earns biology the reputation of being a ‘soft science’. The study of behaviour, which represents the most complex expression of biological organization, suffers in particular from this problem. This is fine as long as we are aware that quick answers in behavioural biology are prone to be dirty.

Is there a way to avoid the pseudo-proof fallacy? Obviously, an important step is to abstain from over-interpreting a match between theoretical predictions and empirical results. However, the true remedy for the illusion of superficial understanding is to unravel underlying mechanisms, ultimate and proximate. This is notoriously demanding and unfit for quick answers. The function of heuristic theory in this process is to guide the research by helping to identify crucial factors and assess the importance of underlying assumptions. If empiricists and theoreticians are to benefit from each others’ approaches, modellers should strive for informing their models by realistic assumptions, and ethologists should use theory as guidance to ask the right questions and to generate falsifiable hypotheses (Lima & Bednekoff 1999; Gulbransen et al. 2006; Mirza et al. 2006; Slos & Stoks 2006). Ideally, both are combined in one study (e.g., Kacelnik 1984, Cant & Field 2001; Luttbeg & Langen 2004; Stamps et al. 2005). One possibility is to model conditions and relationships in a particular system, which can generate testable, quantitative predictions when parameterized with the respective data (Alonzo & Warner 2000; Skubic et al. 2004; Abe et al. 2007). Such approaches afford a most constructive interaction between theory and empiricism and are particularly suited to improve our understanding of mechanisms underlying behaviour.

Traditionally, studies of the ultimate causes of behaviour make more use of theory than those of proximate mechanisms. The firm theoretical basis of behavioural ecology is its major power. Empirical research in this field is based on formal evolutionary concepts, for instance on optimal foraging theory (Beauchamp 2006, 2007; Sanderson et al. 2006; Willisch & Ingold 2007) life history and contest theory (Jennings et al. 2006; Kemp et al. 2006; Takeuchi 2006; Brown et al. 2007), or sexual selection and sex allocation theory (Arnqvist & Anders 2006; Bellemain et al. 2006; Jawor & Breitwisch 2006; Fairn et al. 2007). However, we should be aware that the majority of empirical papers in behavioural ecology is not strongly rooted in formal theory. The standard procedure is rather a post hoc consolidation of collected data with seemingly adequate models and hypotheses. A critical appraisal of publications gives the impression that the strong theoretical basis of empirical behavioural ecology is more apparent than real.

It seems that behavioural physiology is even less rooted in theory, but there are excellent examples demonstrating that the study of mechanisms may benefit a lot also from a firm theoretical basis (Wingfield et al. 1990; Amdam & Omholt 2003; Endler & Mielke 2005; McNamara & Buchanan 2005; Ruina et al. 2005; Klerman & St Hilaire 2007). We may hope that ethologists studying behavioural mechanisms at the proximate level increasingly make use of theoretical models to guide their research (e.g., Bealor & Kerekian 2006; Winterrowd & Weigl 2006; Burt et al. 2007; Hobel & Gerhardt 2007; Vannoni & McElligott 2007). However, the limitation here is often that respective theory has not been developed yet. A closer collaboration between theoreticians and empiricists in the study of behavioural mechanisms is highly desirable.

In general, there is substantial scope for improvement regarding the use of theory in behavioural biology. If we consider publications in Ethology in the last 2 yr, for example, nearly 10% of original research papers reported on tests of predictions derived from a formal model as outlined above (e.g., Lister & Neff 2006; Goubault et al. 2007; Lopez-Rull et al. 2007; Robbins et al. 2007). Another 10% used theory in a more general form to derive expectations for their study (e.g., in the context of predation avoidance, Cooper 2006; Husak et al. 2006; Pecor & Hazlett 2006; parental investment, Ebensperger et al. 2006; Magee & Neff 2006; Foster & Burley 2007; and signalling, Friedl 2006; van Oort et al. 2006; Bertram 2007). Of these 45 empirical papers with more or less formal theoretical basis, 71% used an experimental approach to test predictions resulting from a theoretical model. I hasten to add that this approach is not the only stairway to heaven in science, but it is clearly a strong approach and hence always worth to be seriously considered when planning empirical research.

A frequent notion in behavioural articles is that the study set out to ‘test a model’ (e.g., Moller 1994; Hammond et al. 2006; Blanckenhorn et al. 2007; Sparling et al. 2007). This is an unfortunate phrasing because it implies that the model can be falsified by the study, i.e. it must be abandoned if the predictions are not
met by the empirical data. However, a discrepancy between a model’s predictions and empirical data may say nothing about the general quality or applicability of the model. Instead, it usually means that the respective system is not adequately represented by the model. Most often this is due to the fact that intentionally, theoretical models make simplifying assumptions, as I have outlined above. Rather than rejecting a model when data do not meet a prediction, extensions of the model or alternative explanatory concepts deserve consideration.

One might argue that journal editors are hypocritical when demanding that empirical research be based on theory, if neither a match between model predictions and data allows to conclude that the modelled process correctly describes the observed pattern, nor a mismatch allows to reject the model. Obviously I would not agree with such opinion. We must be aware that theory serves mainly one crucial purpose in our search for scientific explanations: it guides our thinking.

References


Jennings, D. J., Gammell, M. P., Carlin, C. M. & Hayden, T. J. 2006: Is difference in body weight, antler length, age or dominance rank related to the number of fights between fallow deer (Dama dama)? Ethology 112, 258–269.


