Spectacular phenomena and limits to rationality in genetic and cultural evolution*

Magnus Enquist[1,2], Anthony Arak[3], Stefano Ghirlanda[1], and Carl-Adam Wachtmesiter[1]

[1]Group for interdisciplinary cultural research, Stockholm University, Kräftriket 7B, 112 36 Stockholm
[2]Zoology Institution, Stockholm University, Svante Arrheniusvägen 14D, 112 36 Stockholm
[3]Archway Engineering (UK) Ltd, Elland, HX5 9JP, U.K.

Reprint of December 3, 2006

Abstract

In studies of both animal and human behaviour, game theory is used as a tool for understanding strategies that appear in interactions between individuals. Game theory focuses on adaptive behaviour, which can be attained only at evolutionary equilibrium. Here we suggest that behaviour appearing during interactions is often outside the scope of such analysis. In many types of interaction, conflicts of interest exist between players, fueling the evolution of manipulative strategies. Such strategies evolve out of equilibrium, commonly appearing as spectacular morphology or behaviour with obscure meaning, to which other players may react in non-adaptive, irrational ways. We present a simple model to show some limitations of the game theory

---

[*] First published in *Transactions of the Royal Society of London* **B357**, 1585–1594, 2002 (special issue on game theory and evolution). Permission to reproduce in any form is granted provided no alterations are made, no fees are requested, and notice of first publication is included. Corresponding author (as of December 3, 2006): Anthony Arak, arak@archway-engineering.com. Minor language differences may be present relative to the published version.

approach, and outline the conditions in which evolutionary equilibria cannot be maintained. Evidence from studies of biological interactions seems to support the view that behaviour is often not at equilibrium. This also appears to be the case for many human cultural traits, which have spread rapidly despite the fact that they have a negative influence on reproduction.

Introduction

The traditional way to analyze social strategies is through the application of game theory (Maynard Smith, 1982; Fudenberg & Tirole, 1992). The objective of game theory is to find and describe strategic equilibria, often referred to as Nash equilibria. At such equilibria no player can gain anything by using an alternative strategy. The evolutionary justification for this is that natural selection will favour change of strategies until such a solution is reached. Individuals using an equilibrium strategy appear to make adaptive or "rational choices" in order to maximize their reproductive success, selecting the best course of action from the set of possible strategies.

It follows that, if evolutionary processes are at equilibrium, predictions about behaviour can be obtained simply by asking what is the most profitable way to behave, without considering the dynamics of the evolutionary process (Parker & Maynard Smith, 1990; Grafen, 1991).

However, for evolutionary change to occur at all, strategies must exist out of equilibrium at least some of the time (Maynard Smith, 1978). The question remains open as to how persistent and wide-ranging such non-equilibrium conditions are in nature. It is commonly assumed that for much of the time strategies are at a stable equilibrium. New variants that arise are penalized by natural selection and the original situation is restored. Occasionally, a new variant does succeed in invading the population, resulting in a brief, transitional period of change, terminating in a new stable state. If this view is correct, we may be somewhat justified in ignoring the relatively short bursts of evolutionary change, and analyze behaviour purely in terms of optimization theory or game theory. Consequently, it then becomes possible to perceive almost every trait and behaviour as adaptive.

The idea of rationality or perfect adaptation has been criticised on the grounds that many constraints apply to behaviour mechanisms (Simon, 1955, 1956; Maynard Smith, 1978; Gould & Lewontin, 1979; Binmore, 1987; Rubinstein, 1998), that adaptation takes time (Maynard Smith, 1978), and that it is not always compatible with genetic mechanisms (Karlin, 1975) or evolutionary dynamics (Lande, 1981; Dieckmann & Law, 1996; Eshel, 1982). Here we consider a further factor that may prevent an equilibrium being reached: the nature of the interaction itself. When an advantageous trait evolves in one player this can be to the disadvantage of other players, and vice-versa. This scenario may result in endless cycles of adaptation and counteradaption among the different classes of player, with the result that evolution proceeds out of equilibrium for much of the time, with behavioural strategies in an almost continuous state of flux (Parker, 1979, 1983). Under these conditions strategies may emerge and persist that cannot be part of a game theoretical equilibrium. The players in such games seek to manipulate one another, and behaviour evolves that appears to be irrational when judged against optimization principles.

Given these two very different views of evolution it is important to understand which outcome is most likely, and under what conditions equilibria may be maintained. We explore this question by first presenting a simple game theory model of behaviour that generates equilibrium strategies, but is unrealistic in a number of respects. In particular, the number of possible strategies is severely restricted and individuals in the game are assumed to be omniscient about the environment. When more realistic assumptions are adopted, the equilibrium disappears and exploitative strategies dominate.

The results of our review challenge the notion that the analysis of behaviour can be achieved purely by the application of game theory. Indeed, many interactions appear to be outside the scope of game theory. We suggest that in many circumstances evolution is likely to proceed "out-of equilibrium" for much of the time. The strategies that emerge in such games are more appropriately viewed as staging posts on the road of an evolutionary race (Dawkins & Krebs, 1978, 1979), rather than as stable end-points predicted by game theory.

Although many of our examples come from animals, our conclusions are equally important to the study of human interactions. Because new innovations appear much more quickly in cultural than in genetic evolution, it can be argued that human behaviour is much more susceptible to invasion by manipulative strategies compared with the situation in other species. The evolution of traits "out-of-equilibrium" may therefore also account for much of the richness observed in human culture.

A problem for game theory

Some problems arising in game theory can be illustrated by a simple game between two players, an actor and a reactor. We call this the Game of Presence.

The actor is either present ($\nu = 1$) or absent ($\nu = 0$) but has no choice of actions. The reactor, based on whether or not the actor is present, decides upon an effort x ($x > 0$). In the presence of the actor the return on this investment to the reactor is first increasing and then decreasing with x. The benefit to the actor of the reactor's effort is ever-increasing with x. In the absence of the actor, providing an effort $x > 0$ returns a negative payoff to the reactor. Note that, as in most games, there is a conflict between the two players concerning the amount of effort x to be made by the reactor. The game models a widespread type of interaction. For example, a female bird (reactor) must decide how much to invest in reproduction depending on whether the male (actor) is present or absent in the territory. One example of such a game may be formally written as follows

$$F_A(x) = \begin{cases} 0 & \text{if } \nu = 0 \\ \\ x & \text{if } \nu = 1 \end{cases} \tag{1}$$

$$F_R(x) = \begin{cases} 0 & \text{if } \nu = 0 \\ \\ x - x^2 & \text{if } \nu = 1 \end{cases} \tag{2}$$

where F_A is the payoff to the actor and F_R is the payoff to the reactor.

Since we have (deliberately) assumed that the reactor is the only player that has a choice of actions, we solve the game simply by finding the effort that maximizes the reactor's payoff. If the actor is present the optimal effort is $x = 0.5$. Actor and reactor then receive in return 0.5 and 0.25 respectively. If the actor is absent, the optimal effort for the reactor is $x = 0$. This solution is a "Nash equilibrium" (no better response strategy exists) and it is also evolutionary stable because if the reactor's effort drifts away from the optima of 0.5 and 0, selection will return it to these levels.

This game may seem so trivial as to not warrant a formal analysis. However, it nicely illustrates some of the problems of applying game theory to reality, problems that pervade the entire game theory approach to behaviour. First, in reality, information about the presence or absence of the actor is not automatically provided to the reactor but must be inferred from sensory input (figure 1). The presence or absence of the actor is detected by a mechanism that reacts to the stimulation or physical energy (e.g. light or sound) that reaches the reactor. The actor must be recognized when present even when viewed from different distances and angles, in different light conditions and against different backgrounds.

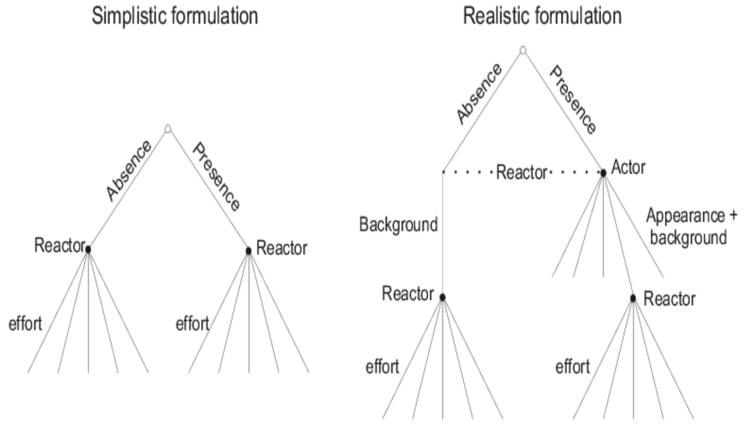

Figure 1: The Game of Presence. The figure compares two ways of formulating the Game of Presence. In the simplistic formulation (the traditional way of formulating games), it is assumed that the reactor knows whether the actor is present or absent. In the realistic formulation the reactor must infer this information using its sense organs, discriminating the actor on the background from the background alone. Only a few of the possible appearances are shown. In reality, the number of actor as well as background appearances is infinite.

Second, the actor may take on a variety of appearances, and because such appearances give rise to different stimulation, they may elicit quite different reactions from the reactor. In game theory, it is common to restrict, consciously or unconsciously, the number of strategies considered, e.g. by limiting the number of appearances to the minimum needed to convey the relevant information, or by considering variation along a single dimension only. Furthermore, strategies available to players are prescribed in advance while evolution is an unfolding process in which new strategies become possible as a result of evolution itself.

There are several reasons for making these simplifying assumptions. One is to make mathematical analysis possible. Another motive is to limit the number of solutions to the game or to eliminate those considered implausible (van Damme, 1987). However, as we shall see, these restrictions cannot always be justified from an evolutionary point of view.

A more realistic formulation

Let us now consider a more realistic formulation of the Game of Presence, in which the actor can change its appearance (by random mutation), and the reactor has one response to each such appearance (figure 1). Responses are also subject to change by mutation, so that over time reactors may increase or decrease their responsiveness to different appearances.

What is the outcome of such a game? In simulations, the game proved to be highly dynamic with rapid changes in actor appearances accompanied by reactor counter-adaptations (figure 2). In most simulations, appearances evolved that enabled actors to manipulate reactors into producing an effort $x > 0.5$. A stable equilibrium was found only when the number of appearances available to actors was restricted to one, corresponding to the simplistic version of the game presented above. As the number of possible appearances was allowed to increase, reactors were less able to resist manipulation, and thus produced a greater effort. Note also that actors typically do not use the most effective appearance and thus also behave suboptimally.

Why don't reactors evolve effective countermeasures against manipulation? This is not possible because reactor responses towards appearances not currently used by actors are neutral with respect to selection, and subject to mutations and subsequent drift. Thus, at any given time, there is a set of hypothetical appearances not currently in use that would stimulate reactors to produce efforts larger than the optimum. When such appearances arise by chance mutation in the population of actors, they are strongly favoured by selection and rapidly spread: this has been referred to as sensory exploitation (Basolo, 1990; Ryan, 1990). Selection acting on the population of reactors then tends to desensitize their mechanisms to these new appearances, restoring the effort back towards $x = 0.5$. However, at any time, there are likely to be appearances arising in the actor population, some of which can by-pass the reactors' defense mechanisms.

In conclusion, the realistic game has a very different outcome from the simplistic model. Stable equilibria do not exist and the game resembles an evolutionary race in which at least some players behave irrationally.

Theory versus reality

The traditional aim of game theory has been to find and describe equilibria. The problem of novel strategies has been addressed in several ways. In many signal-

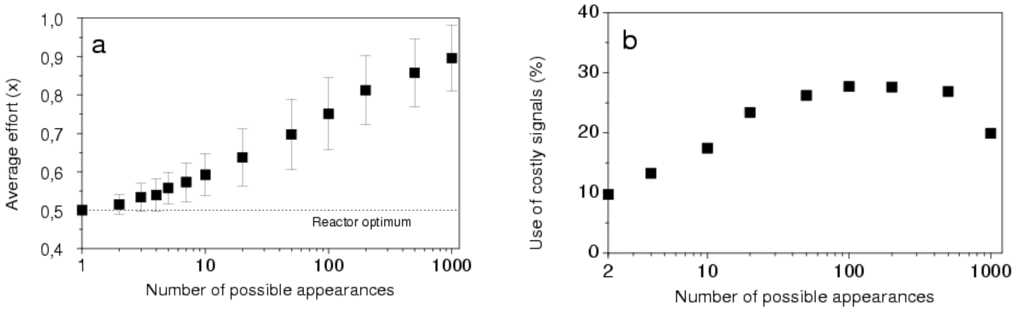

Figure 2: Simulations of the realistic formulation of the Game of Presence. **a**, The more strategies (appearances) available to actors, the greater effort they are able to elicit from reactors. Each simulation lasted 15,000 generations (average calculated on the last 10,000 generations only). The bars describe the fluctuation in effort (average SD from 10 simulations for each number of appearances). Simulations started at a Nash equilibrium: all actors use a particular appearance, and all reactors respond with the optimal effort $x = 0.5$ to any actor appearance. Such equilibrium is stable only with one possible appearance, corresponding to the simplistic version of the game (figure 1). With two or more appearances there is no stability. The effort fluctuates and signals replace each other (the rate of change of the most common signal varied between 0.004 to 0.014 per generation). **b**, The use of costly signals from a new set of ten simulations in which half of the appearances were costly to use. In all other respects the simulations were the same as above. Costly signals were used to a considerable extent. The degree of manipulation was somewhat lower in comparison with **a**. Description of the simulations: An actor strategy is to use one appearance a out of a set of appearances $\{1,2,\ldots,n\}$. A reactor strategy is a set $\{r_1,r_2,\ldots,r_n\}$ such that $r_i \in [0,1]$ is the response to appearance i. Responses are set independently. Actor and reactor populations consisted of 100 individuals. In each generation all actors interact with all reactors leading to an expected reproduction (fitness) for each individual, according to expressions (1) and (2) in the text. Due to random factors actual reproduction is not equal to expected reproduction. Specifically, at each generation we build a new population by picking at random 100 individuals from the old population, with the probability of individual i being picked proportional to its fitness i.e. $p_i = f_i / \sum_k f_k$. In addition, 1% of the individuals in the new populations are mutants rather than perfect copies of the parent. Actors mutated by selecting an appearance at random from the set of appearances. Reactors mutated by changing the response to one appearance, chosen at random. The new response was in turn a random number between 0 and 1 (uniform distribution). To produce **b** half of the appearances entailed a cost of 0.1 to the actor.

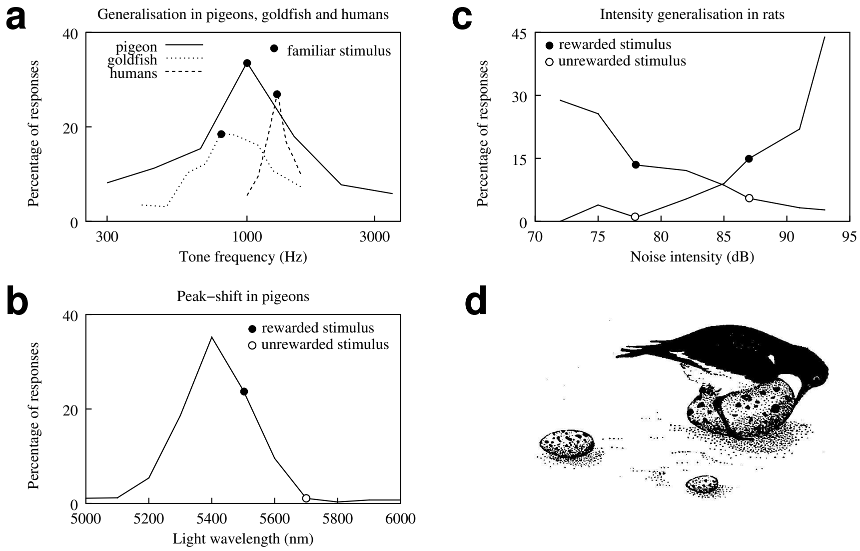

Figure 3: Examples of responses of animals to novel variations of familiar stimuli. **a**, The general finding is that modified stimuli often elicit similar responses as the familiar stimulus (Pavlov, 1927; Mackintosh, 1974). This phenomenon is known as generalization, and follows basic rules that are independent of species, context, sensory modality (Ghirlanda & Enquist, 2003). The modified stimuli are usually less effective than the familiar ones in eliciting a particular behaviour (response decreases when departing from the rewarded stimulus). Sometimes, however, novel stimuli can elicit even stronger responses (responding does not peak on the rewarded stimulus in **b** and**c**). This happens especially when intensity of stimulation is varied (**b**) or when animals must distinguish between two or more stimuli (**c**). This phenomenon has been found with respect to both innate and learned behaviour in several disciplines (Tinbergen, 1951; Hanson, 1959). Despite similarity in findings it has been referred to by a number of different names: supernormal stimulation in ethology, peak-shift and overgeneralization in psychology, and more recently receiver bias. As a whole, the data show that animals do not respond to novel stimuli in any of the idealized ways that have been considered in game theory modeling (no response, same response or a rational response). The empirical responses seem a “reasonable” way of dealing with novel stimuli, but at the same time such a general scheme of reaction (arising from the interaction of past experiences with the nervous system) cannot be optimal in any particular situation. An example of apparent biases is given in **d**: an oystercatcher, *Haematopus ostralegus*, is given a choice of what of three eggs to incubate. The smaller egg in the figure corresponds to the natural size of this species’ eggs (Tinbergen, 1951). (**a**) Illustrates a generalization in pigeons, goldfish and humans; (**b**) illustrates a density generalization in cats; and (**c**) illustrates a peak-shift in pigeons. Filled circles denote a positive stimulus, open circles denote a negative stimulus.

ing models explicit assumptions about reactions to novel stimuli are often made such as “no reaction” (Johnstone, 1995) and “same reaction as to the closest existing stimulus” (Grafen, 1990*a*). In classical game theory developing new equilibrium concepts has been a major issue, with the aim of limiting solutions to those considered more plausible. Generally, such refinements of the Nash equilibrium are based on small departures from the equilibrium (so-called “trembles” or “mistakes”) towards which players adapt (Selten, 1975, 1983; van Damme, 1987; Fudenberg & Tirole, 1992). This translates into “superrationality”, the notion that players behave rationally even in novel situations (Auman, 1986). However, neither superrationality nor the assumptions found in signaling games seem justifiable from knowledge of animals’ reactions to novelty (figure 3) or from evolutionary considerations. In particular, the idea that animals always evolve adaptive responses to “mistakes” seems untenable when the strategy space available to players is very large, or when certain strategies are used only at low frequency. By ignoring a fundamental property of real games, the possibility of manipulation occurring in games is denied.

By contrast, in reality manipulation may be the rule rather than the exception. Three conditions can be identified that promote instability and are necessary for manipulation to evolve:

1. The number of strategies available to players must not be too restricted, and new strategies must be allowed to arise from time to time. The number of strategies is often deliberately restricted in game theory models while in reality the potential strategy set is very large.

2. Direct interactions must occur between players. By this we mean that decisions must be taken after observations of other players’ actions, giving rise to “dynamic” games. Whenever such actions influence the reactor, manipulation is a possibility. In contrast, games do not suffer from instability when all decisions are made prior to any interactions: so-called “static” games (Fudenberg & Tirole, 1992). An example of a static game is the sex ratio game (Maynard Smith, 1982). Sex in most species is decided early in life before any interactions with other individuals have occurred, so there is no possibility for individuals to influence each other’s decisions. Such clear cases may be rare in nature, but are often studied theoretically.

3. Conflict must exist between players. Conversely, when players’ interests coincide exactly (i.e. the optimal response for the reactor is also optimal for

the actor), there are no incentives for manipulation. In addition to actor-reactor conflicts, evolution may be driven by conflicts between two or more actors competing for the attention of reactors (Dawkins & Krebs, 1978; Krebs & Dawkins, 1984; Arak & Enquist, 1995).

The importance of mechanism

A guiding principle in the study of behaviour has been that questions about ultimate causation (evolutionary explanations) and questions about proximate causation (behavioural mechanisms) are logically distinct and cannot replace one another (Tinbergen, 1951). If behaviour is at equilibrium, game theory alone can predict many details of behaviour. The study of mechanisms is then reduced to describing how adaptive behaviour is actually implemented. In contrast, if evolution proceeds out of equilibrium we need to consider details of mechanism in order to understand the form of behaviour that evolves. This is because the mechanism itself may act as an agent of selection upon behaviour (Staddon, 1975; Ryan, 1990; Guilford & Dawkins, 1991; Enquist & Arak, 1998).

Consider, for example, a signaling game. If the form of a signal is not at equilibrium (i.e. more efficient signals exist) the variants that are favoured depend on what mechanism is used by reactors for processing signals. A crucial issue is whether mechanisms can evolve that are not susceptible to manipulation, so that game theory predictions can be rescued.

As a starting point, we may first visualize the reactor's mechanism as a "look-up table" in which the response to every stimulus appears as an entry in one cell of the table. Modeling behaviour with such a mechanism is equivalent to the approach of game theory and optimization theory, that traditionally seek an independent, best response to each move by the opponent (Binmore, 1987; Fudenberg & Tirole, 1992). Note that we already used a simple look-up table in the simulation of the Game of Presence (figure 1). Can such an idealized mechanism evolve? The answer clearly is no. One reason is that an impossibly large memory would be required. If we try to imagine a look-up table based on all possible stimuli reaching the sense organs we immediately see a flaw in the idea. Suppose there are just 1000 receptors (a modest assumption in most cases) and each receptor can be in only one of two states, "on" or "off". Then there are 2^{1000} possible patterns of stimulation that could be experienced, each needing its own entry in the look-up table. The problem becomes more acute if, as it is often the case, a sequence of stimuli is relevant for optimal decision making. Although in game theory such sequences are explicitly described in game trees (Fudenberg & Tirole, 1992), it is difficult to imagine how they could all be stored in memory. Estimates of the storage capacity of the human brain vary between 2^{33} and 2^{52} bits (Crevier, 1993). The storage capacity of the genome is even more limited.

A second difficulty concerns the manner in which entries in a look-up table could be optimized. Evolution proceeds by trying out new solutions and retaining those innovations that are improvements. However, there are limits to the number of memory parameters that evolution can optimize, because the combinations to be tried out are too many even with a modest number of parameters. A third problem is that reactions towards novel stimuli are completely undetermined, since corresponding table values have not been under any selection.

The only escape from the above problems is to use a mechanism with a much coarser memory structure (fewer parameters) than a look-up table. In real mechanisms, responses depend on the architecture of the nervous system and the memory stored within it. It is conceivable that each memory parameter affects responses to many stimuli, and responding to a single stimulus is governed by many parameters. Now, reactions to novel stimuli are not undetermined, but depend on the architecture and the current content of memory. In addition, real mechanism suffers from constraints that make some responses more difficult to implement than others (Simon, 1956; Binmore, 1987; Maynard Smith, 1978; Dukas, 1998).

Evolution has favoured those mechanisms that show some “intelligence” vis-à-vis reality. For instance, stimuli that are similar to one another often share some causal relationship with events in the outside world. Animals detect and use such regularities, generalizing knowledge about familiar situations to novel ones (figure 3). In addition, animals use general methods to cope with novelty, including exploratory behaviour and avoidance behaviour (Russell, 1973).

Although generalization of this kind would seem to be a sensible way of dealing with novelty, it will always lack precision, independent of how intelligent the organism is. This is because novel situations can differ from familiar ones in many details that cannot always be anticipated from previous experiences. For example, if an animal that lives on a diet of berries encounters a new kind of berry (e.g. novel in colour or shape), there is usually no way of knowing whether it is edible or poisonous. In case the initial reaction is inappropriate, some evolutionary time is needed to establish the proper response. Alternatively, individual learning, which is another aspect of intelligence (look-up tables do not learn), may be used to establish appropriate behaviour towards the novel berry. In both cases, there will be a period of sub-optimal behaviour. In the case of learning, it will never disappear since naïve individuals appear continuously in each generation.

Furthermore, a real mechanism will not treat all novel stimuli alike, but will have biases. In our example, some new berries will be preferred over others, independent of their values as food. Some berries may be even preferred to the familiar ones (cf. figure 3d).

figure 4 shows the effect of using a more realistic mechanism, an artificial neural network, in the Game of Presence. Instability is still present, with reactors spending more effort than is optimal for them to do so, but manipulation is somewhat less than in simulations presented in figure 2. This is because the network is more "intelligent" towards novel stimuli than the look-up table used in the earlier simulation. At least two factors contribute to this "intelligence": the reluctance of the network to produce the same response to very different stimuli and the existence of a common output mechanism from which it is difficult to elicit responses that deviate substantially from the reactor's optimum.

However, the evolutionary consequences of intelligence are often mixed. To solve a problem with greater accuracy a more complex mechanism is usually needed. For example, eyes with more receptors allow for more accurate discrimination. How does such increased complexity affect reactions towards novel stimuli? Are humans less susceptible to manipulation than, say, a snail? One may think that if a more complex mechanism evolves, it will inevitably be more efficient and make less sub-optimal responses than the mechanism from which it evolved. For example, one may add a bias-correcting device to the original mechanism. But such a new mechanism will also have its own biases. Thus, while added complexity may function well in solving currently existing problems, it will not necessarily make the mechanism more efficient when faced with new problems. In fact the opposite may be true because, as complexity increases, the number of different stimuli the organism can perceive is also likely to increase, as well as the number of physiological and biochemical processes that are open to interference (figure 5). An increase in complexity may also allow qualitatively different stimulation to be effective. For example, the ability of a mechanism to recognize patterns in time may favour signals that are variable in time over monotonous repetitions of the same signal (Wachtmeister & Enquist, 2000).

Predictions and empirical support

Several predictions can be made about behaviour when conflicts of interest occur in interactions. These are as follows:

1. *Strategies are not at evolutionary equilibrium* (Dawkins & Krebs, 1978;

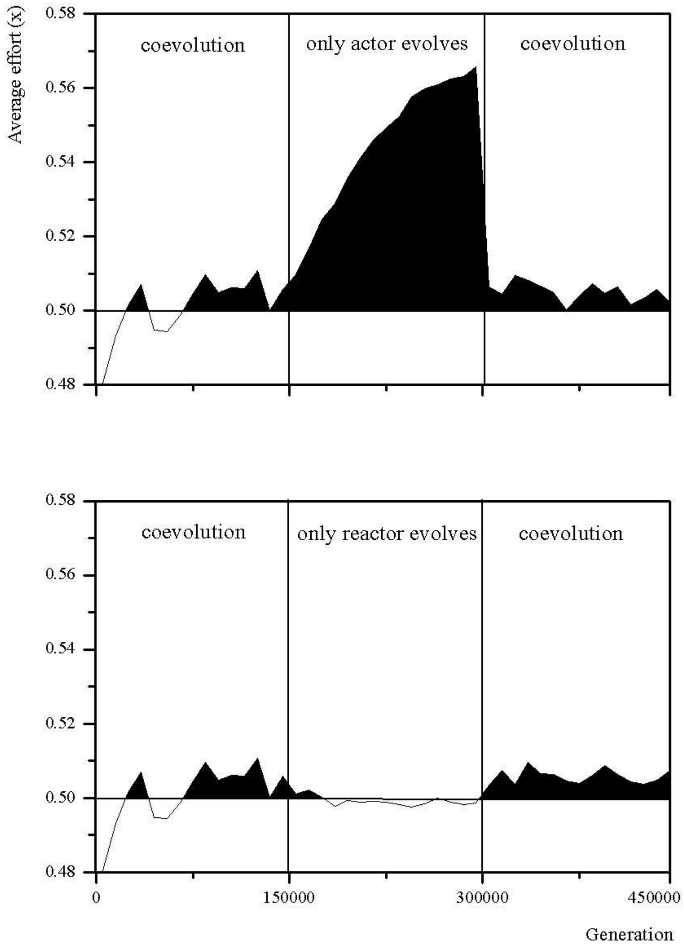

Figure 4: Simulation of the Game of Presence using a three-layer neural network as the reactor. Such networks are much more realistic than look up tables and generalize spontaneously. Diagram **a** and **b** both show how the average rector effort changes over time. The reactors optimal effort is 0.5. Shading shows when the reactor is manipulated (effort > 0.5). First actor and reactor were allowed to coevolve for 150,000 generations (both diagrams show this part). From this point the simulation was split in two. In one (a) only the actor was allowed to evolve. In the other (**b**) only the reactor evolved. After another 150,000 generations coevolution was again introduced. As shown some degree of manipulation is maintained during coevolution. When the reactor's evolution was stopped the actor became more efficient at manipulation. When instead evolution of the actor was stopped the reactor evolved an optimal effort.

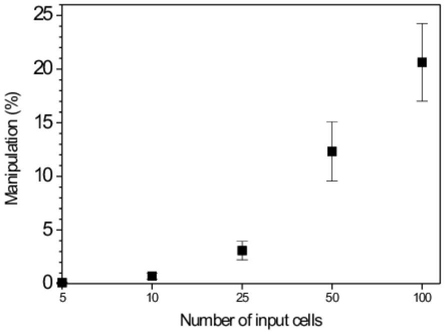

Figure 5: Effects of mechanism complexity on manipulation. The diagram demonstrates how varying the size of the reactor's artificial retina influences the opportunity for manipulation in an evolutionary process. On average the reactor is more susceptible to manipulation when the number of receptors increases. The results are from simulations of coevolution between actors and reactors similar to those described as in figure 4.

Arak & Enquist, 1995). The strongest evidence in support of this prediction comes from an experiment on the fruitfly, *Drosophila*, in which the evolutionary process was studied directly (Rice, 1996). In fruitflies, as in many other species, there is a conflict between males and females over mating opportunities, males preferring to mate more often than females do. In the experiment, males were mated with females for several generations, but while the male line was allowed to evolve, females were experimentally prevented from doing so. After only 30-40 generations the evolving males were better at obtaining matings with already mated females, they produced more male progeny and were better at securing paternity than control males. But these advances were made at the expense of the females, which had higher death rates when exposed to evolving males than to controls. The conclusion of this experiment is that there is an on-going conflict between the sexes, in which neither sex is at the optimum. But when evolution in one sex is experimentally arrested, the other sex rapidly wins an advantage.

Further evidence comes from comparative studies. In two species of dance flies of the genus *Rhamphomyia* females compete for males that carry a nuptial prey, which males trade for copulations (Funk & Tallamy, 2000). In both species males prefer females with large abdomens. In *R.f sociabilis*, thought to represent the ancestral condition, female abdomen size is a reliable indicator of the stage of egg development. In the long-tailed dance fly, *R. longicauda*, however, females have developed the ability to inflate their abdomens, which reduces the information content of the signal. Thus it appears that *R. longicauda* females with undeveloped eggs, use abdominal inflation to manipulate males into providing them with extra protein. Males that copulate with these females almost certainly suffer cuckoldry, because females usually mate several times before laying their eggs.

Comparative data from closely related species within certain genera of frogs (Ryan, 1990) and fish (Basolo, 1990), also suggests that current behaviour is not at equilibrium. In courtship behaviour, female congeners show similar preferences for certain male display traits, suggesting a common origin of the preference. Although these traits have evolved in males of some species they are absent in others, despite the existence of a female preference for them. This suggests that display used by males to attract females is not always optimal, and that display traits may sometimes evolve to exploit a pre-existing bias in females.

2. *Manipulation is common in nature*. The most obvious and undisputed cases of manipulation occur in conflicts between members of different species. Examples include brood parasitism in birds (Rothstein & Robinson, 1998) and fish (Sato, 1986), Batesian mimicry in butterflies (Brower, 1980), sexual mimicry in fireflies (Lloyd, 1980), prey mimicry in anglerfish (Pietsch & Grobecker, 1978), slave-making in ants (Froitzik et al., 2001) and many parasite-host interactions (Futuyma & Slatkin, 1983). In all these cases, there is an absolute conflict between players, in the sense that it would pay reactors to ignore the displays of actors altogether and so avoid becoming "victims" of manipulation. It is more difficult to determine whether manipulation is occurring in social interactions between members of the same species. Examples are begging behaviour and courtship display, in which it pays the reactor to respond to the actor's signals, but there is a conflict over the level of response. The problem here is that the optimum level of response is difficult to quantify. It is perhaps due to this difficulty that equilibrium explanations of behaviour are frequently emphasized (Grafen, 1990*b*; Godfray & Johnstone, 2000), even though the context of such interactions suggests that manipulation would be favoured. We propose that manipulation is always likely to be present to some degree in intraspecific social interactions when there is conflict, but departures from the optimum level of responding may sometimes be small and difficult to detect.

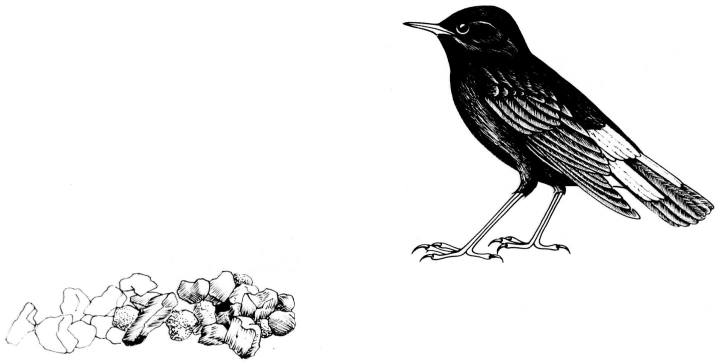

Figure 6: Stone carrying by black wheatears, *Oenanthe leucura*, a behaviour that is open to alternative interpretations. In this monogamous species, mainly the 40-gram male carries about 1-2 kilos of stones in the presence of the female preceding each reproductive attempt (Cramp, 1988; Moreno et al., 1994). The ritual is repeated up to 5 times per year and often with the same female, since black wheatears seem to pair for life. The amount of stones carried correlates with reproductive success. Explanations in terms of adaptive behaviour hold that the amount of stones carried is a reliable signal of mate quality that is used by females to assess their partners, and that females' reproductive investment is dependent on this assessment. An alternative explanation is that stone carrying is a ritualization of nest building behaviour. If it pays the male to reproduce earlier or at a higher rate than the female, males should try to influence females' reproductive decisions. A female is susceptible to manipulation, since she should only start reproduction in the presence of a willing male. Stone carrying may have developed as an exaggeration of a cue used by the females to detect the presence of such a male, if males originally signaled willingness by starting to construct a nest.

3. *The "meaning" of displays is often obscure*. When behaviour is at equilibrium, it is usually possible to describe traits in terms of their functions. Displays are often described as transmitting a particular kind of information, for example about species identity, fighting ability, mate quality, and so on. However, when evolution proceeds out of equilibrium it becomes difficult to attach a precise meaning to many traits that appear. For example, many birds possess plumage consisting of several distinct patches of colour, and multiple ornaments such as elongated crests and tail feathers often contribute to displays. Visual display may also be combined with songs incorporating many different components. Finally, several different

displays components may follow one another in sequence, forming bizarre rituals (figure 6, and Wachtmeister, 2001). Although it is possible that each of these different components of display conveys a different kind of information, there is little empirical support for this (Møller & Birkhead, 1993). In fact, many displays appear to be highly redundant for information transfer (Wiley, 1983). It is more prudent to suggest that each component has appeared simply because it contributes to the overall effectiveness of the display in eliciting the desired response from reactors. The precise form that displays take is governed much more by "receiver psychology" than by their information content (Enquist & Arak, 1998; Guilford & Dawkins, 1991). It is quite possible that displays contain a mixture of information and non-informative advertisement. For example, the peacock's tail might contain information about species and sex identity as well as stimulating the female to mate more directly.

(iv) *More costly and ornamented forms of display evolve as the degree of conflict increases* (Dawkins & Krebs, 1978; Arak & Enquist, 1995). With increasing conflict, there is a stronger selection pressure on reactors to resist manipulation. This, in turn, creates counter-selection on actors to produce more effective signals, even if this means taking on higher costs. Conversely, when there is complete common interest between actor and reactor, displays evolve to be efficient at information transfer. Because, under these conditions, there is no incentive for manipulation, selection will tend to minimize the cost of displays. Although there is much circumstantial support for this prediction (Dawkins & Krebs, 1978; Arak & Enquist, 1995), so far there have been no rigorous tests.

4. *Evolutionary change is rapid in traits used for manipulation.* Continuous selection on actors for more efficient display, and on reactors for resistance to display creates the conditions for rapid evolution of morphology and behaviour involved in interactions. If this is true, when comparing closely related species, there should be greater differences in the form of displays used for exploitation than in displays used to convey information. Observations on the aggressive behaviour of cichlid fish seem to support this prediction. Among closely related species, body colouration used in aggressive visual display is highly variable in form suggesting that, in the early stages of fights, multiple ways have evolved to influence opponents. However, in escalated fights, other displays such as tail beating and mouth wrestling become more important, and these displays vary little between species (En-

quist, unpublished observations). It is believed that the latter evolved because they allow accurate assessment of relative fighting ability.

Typically traits used in within-species conflicts evolve more rapidly than traits not involved in such conflicts. Examples of rapidly evolving traits are gamete proteins, reproductive organs, reproductive tracts, seminal fluid proteins, courtship display, threat displays, begging behavior, plant root behavior, pollen and ovary characteristics (reviewed by Rice & Holland, 1997). A consequence of rapid evolution in such traits is that systematicists frequently use them to discriminate between closely-related species (Arnqvist et al., 2000).

Human culture

The evolutionary scenarios described above may be very important for understanding human behaviour and culture. Advertisement and persuasion are phenomena that game theoreticians have had problems explaining (Rubinstein, 1998). It is of course rational to try to persuade through advertising, given that it works, but why does it work? According to the rationality paradigm we should not be persuaded to do something that is not to our advantage, such as paying a higher price for a product that can be bought cheaper from a less colourful supplier. In the light of the theory reviewed above, the existence of persuasion and advertisement can be predicted. Advertisement works because of the design of our brains. There are empirical psychological data showing that humans like other animals are sensitive to the form of signals and not just to the information they convey (Mowen & Minor, 1997). Some of the qualities found in ritualized biological signals, such as large size, frequent repetition, symmetry, elaborate ornamentation and mimicry, also appear in human advertisement. Furthermore, the same types of conflicts are relevant in advertisement and in biological games. Conflicts exist between different suppliers and between suppliers and consumers. The degree of ornamentation may also depend on conflict. Road signs can illustrate this. Some are set up by transport authorities to direct traffic. These signs are informative but not highly ornamented (little conflict). In contrast, signs set up by shops, restaurants and other commercial organizations are generally bigger and contain a high degree of ornamentation that is not necessarily indicative of the quality of products or services that are being advertised (high conflict). Also predicted is the decrease in efficiency of a particular advertisement as consumers counter-adapt to its persuading effect.

Ornamentation and rituals in humans are by no means limited to commercial contexts. In fact, major parts of all cultures include creation and enjoyment of stimulation such as music, gardens, decorated homes, films, paintings, poetry and so on. At times attempts have been made to explain these phenomena as adaptive responses, but these have been generally unconvincing. Many cultural expressions are of very recent origin. Records of artistic culture date back hardly more than 50,000 years and much of today's art has been shaped during the last few hundred years. Thus, it is not plausible that our reactions to art are biological adaptations. An alternative view that emerges from this paper is that human art, in its broadest sense, is a cultural exploration of the human mind. Artists have continuously invented new ways of stimulating our senses. This process has taken place not over an evolutionary time scale, but at a very rapid pace, often during the life span of individual human beings. It is therefore inconceivable that countermeasures to these highly attractive, persuasive stimuli can have evolved in humans. Attempts to seek adaptive explanations for many cultural traits will be futile.

This view may also offer an insight into findings about human demography that have troubled evolutionary biologists (Borgerhoff Mulder, 1998). One is that when general standards of living improve in a country the birth rate goes down. A related observation is that rich people tend to have fewer children. Obviously, this cannot be an adaptive behaviour in the sense of genetic evolution. An increase in material resources should lead to an increase in reproduction, not a decrease. From an abstract point of view cultural stimulation can be seen as a form of manipulation of gene-based cognitive programs that originally evolved because they increased their bearers' success in reproduction. When conditions improve not only do resources for reproduction increase but also the diversity and access to cultural stimulation. Similarly, rich families have greater access (more money) to such stimuli than poorer families. In practice, individuals may trade-off rewards from a diversity of cultural and biological activities resulting in less effort being allocated to reproduction.

In conclusion, the hypothesis that communication systems often exist "out of equilibrium" may go some way towards reconciling the existence of many aspects of human culture with evolutionary thinking. Spectacular cultural phenomena can evolve that convey little meaningful information, but still have strong impact on spectators.

Conclusions

In this review we have shown that to think of animals or humans as perfectly adapted, rational beings is sometimes difficult to justify based on empirical evidence and evolutionary theory. Real games with conflicts and interactions will always, as one element, contain continuous evolution of “salesmanship” in actors and “sale resistance” in reactors that may produce significant departure from rationality. The idea that manipulation is important in biological evolution is not new (Dawkins & Krebs, 1978; Futuyma & Slatkin, 1983; Rothstein & Robinson, 1998). However, while embracing the idea of manipulation occurring in coevolution between different species such as parasites and hosts, biologists have usually played down the possibility of manipulation within species, and have instead favoured adaptive explanations of behaviour. Likewise, game theory models within economics and other human disciplines have seldom considered the possibility of manipulation and persuasion. Actor-reactor coevolution has been studied in a number of different biological contexts, but often from different perspectives. Specific theories have been developed for sexual selection, intraspecific interactions, coevolution between different species, and host-parasite interaction at the biochemical or physiological level. We think that these evolutionary processes share significant similarities and that the differences should not be exaggerated.

Finally, although we have stressed that certain traits may often persist out-of-equilibrium, it would be wrong to conclude that adaptive behaviour can not evolve and that game theory and optimization theory should be abolished. In the special case when selection pressures are constant the use of optimization theory is justified. For example, the extraordinary flight of an albatross can be satisfactorily explained in terms of “good design” relative to the particular aerodynamic circumstances at sea. Furthermore, instability does not feature in games with little conflict, or where players cannot influence each other’s decisions. However, the majority of situations are more complex, involving some degree of conflict and interactions between players. Game theory has been successfully applied to a number of such situations, suggesting that instability does not always dominate actor-reactor evolution (Parker, 1983). Clearly more work is needed to evaluate precisely what outcomes are possible when evolution proceeds out of equilibrium.

Acknowledgments

We thank Göran Arnqvist, Carl Bergstrom, Innes Cuthill, Sasha Dall, Reuven Dukas, David Harper, Ben Kenward, Risa Rosenberg, John McNamara, John Maynard Smith and two anonymous reviewers for valuable comments on the manuscript. This research was supported by the Swedish National Research Council, Riddarhuset, and Marianne och Marcus Wallenbergs Stiftelse.

References

Arak, A. & Enquist, M. 1995 Conflict, receiver bias and the evolution of signal form. *Philosophical Transactions of the Royal Society, Series B* 349, 337–344.

Arnqvist, G., Edvardsson, M. & Friberg, U. 2000 Sexual conflict promotes speciation in insects. *Proceedings of the National Academy of Science of the USA* 97, 10460–10464.

Auman, R. 1986 Rationality and bounded rationality. *Games and Economic Behavior* 21, 2–14.

Basolo, A. 1990 Female preference predates the evolution of the sword in swordtail fish. *Science* 250, 808–810.

Binmore, K. 1987 Modeling rational players. Part I. *Economics and Philosophy* 3, 179–214.

Borgerhoff Mulder, M. 1998 The demographic transition: are we any closer to an evolutionary explanation? *Trends in Ecology and Evolution* 13(7), 266–270.

Brower, L. P., ed. 1980 *Mimicry and the evolutionary process*. Chicago, IL: University of Chicago Press.

Cramp, S. 1988 *The birds of the western palearctic*. No. V. Tyrant Flycatchers to Thrushes. Oxford: Oxford University Press.

Crevier, D. 1993 *AI: the tumultuous history of the search for artificial intelligence*, vol. 1. New York: BasicBooks.

Dawkins, R. & Krebs, J. 1979 Arms races between and within species. *Proceedings of the Royal Society, Series B* 205, 489–511.

Dawkins, R. & Krebs, J. R. 1978 Animal signals: information or manipulation? In *Behavioural ecology. An evolutionary approach* (eds. J. R. Krebs & N. B. Davies), pp. 282–309. Oxford: Blackwell Scientific Publications.

Dieckmann, U. & Law, R. 1996 The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology* 34, 579–612.

Dukas, R., ed. 1998 *Cognitive ecology. The evolutionary ecology of information processing and decision making*. Chicago: The University of Chicago Press.

Enquist, M. & Arak, A. 1998 Neural representation and the evolution of signal form. In Dukas (1998), pp. 1–420.

Eshel, I. 1982 Evolutionarily stable strategies and viability selection in Mendelian populations. *Theoretical Population Biology* 22, 204–217.

Froitzik, S., DeHeer, C. J., Hunjan, D. N. & Herbers, J. M. 2001 Coevolution in host-parasite systems: behavioural strategies of slave-making ants and their hosts. *Proceedings of the Royal Society of London* B268, 1139–1146.

Fudenberg, D. & Tirole, J. 1992 *Game theory*. Cambridge MA: The MIT Press.

Funk, D. & Tallamy, D. 2000 Courtship role reversal and deceptive signals in the long-tailed dance fly, Rhamphmyia longicauda. *Animal Behaviour* 59, 411–421.

Futuyma, D. & Slatkin, M., eds. 1983 *Coevolution*. Sunderland, Massachusetts: Sinauer Associates.

Ghirlanda, S. & Enquist, M. 2003 One century of generalization. In press on *Animal Behaviour*.

Godfray, H. J. C. & Johnstone, R. A. 2000 Begging and bleating: the evolution of parent offspring signalling. *Philosophical Transactions of the Royal Society of London* B355, 1581–1591.

Gould, S. & Lewontin, R. 1979 The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptionist programme. *Proceedings of the Royal Society, Series B* 205, 581–598.

Grafen, A. 1990*a* Biological signals as handicaps. *Journal of Theoretical Biology* 144, 517–546.

Grafen, A. 1990*b* Sexual selection unhandicapped by the Fisher process. *Journal of Theoretical Biology* 144, 473–516.

Grafen, A. 1991 Modelling in behavioural ecology. In *Behavioural ecology. An evolutionary approach* (eds. J. R. Krebs & N. B. Davies), pp. 5–31. Oxford: Blackwell Scientific Publications.

Guilford, T. & Dawkins, M. S. 1991 Receiver psychology and the evolution of animal signals. *Animal Behaviour* 42, 1–14.

Hanson, H. 1959 Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology* 58(5), 321–333.

Johnstone, R. 1995 Honest advertisement of multiple qualities using multiple signals. *Journal of Theoretical Biology* 177, 87–94.

Karlin, S. 1975 General two-locus selection models: some objectives, results and interpretations. *Theoretical Population Biology* 7, 364–398.

Krebs, J. & Dawkins, R. 1984 *Animal signals: mind-reading and manipulation*, vol. 2, pp. 380–402. Oxford: Blackwells.

Lande, R. 1981 Models of speciation by natural selection on polygenic traits. *Proceedings of the National Academy of Science USA* 78(6), 3721–3725.

Lloyd, J. E. 1980 Male *Photuris* firefiles mimic sexual signals of their females' prey. *Science* 210, 669–671.

Mackintosh, N. 1974 *The psychology of animal learning*. London: Academic Press.

Maynard Smith, J. 1978 Optimization theory in evolution. *Annual Review of Ecology and Systematics* 9, 31–56.

Maynard Smith, J. 1982 *Evolution and the theory of games*. Cambridge: Cambridge University Press.

Møller, A. & Birkhead, T. 1993 Cuckoldry and sociality: a comperative study of birds. *American Naturalist* 142, 118–140.

Moreno, J., Soler, M., Møller, A. & Linden, M. 1994 The function of stone carrying in the black wheatear, Oenanthe leucura. *Animal Behaviour* 47, 1297–1309.

Mowen, J. & Minor, M. 1997 *Consumer behavior*, vol. 5. Upper Saddle River, New Jersey: Prentice-Hall.

Parker, G. 1979 *Sexual selection and sexual conflict*, chap. 4, pp. 123–166. New York: Academic Press.

Parker, G. A. 1983 Arms races in evolution - an ESS to the opponent-independent cost game. *Journal of Theoretical Biology* 101, 619–648.

Parker, G. A. & Maynard Smith, J. 1990 Optimality theory in evolutionary biology. *Nature* 348, 27–33.

Pavlov, I. P. 1927 *Conditioned reflexes*. Oxford: Oxford University Press.

Pietsch, T. & Grobecker, D. 1978 The compleat angler: aggressive mimicry in an Antennariid anglerfish. *Nature* 201, 369–370.

Rice, W. 1996 Sexually antagonistic male adaptation triggered by experimantal arrest of femal evolution. *Nature* 381, 232–234.

Rice, W. R. & Holland, B. 1997 The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific Red Queen. *Behaviour, Ecology and Sociobiology* 41, 1–10.

Rothstein, S. & Robinson, S., eds. 1998 *Parasitic birds and their hosts*. Oxford: Oxford University Press.

Rubinstein, A. 1998 *Modelling bounded rationality*. Cambridge, MA: The MIT Press.

Russell, P. A. 1973 Relationships between exploratory behaviour and fear: a review. *British Journal of Psychology* 64, 417–433.

Ryan, M. 1990 Sexual selection for sensory exploitation in the frog Physalaemus pustulosus. *Nature* 343, 66–67.

Sato, T. 1986 A brood parasitic catfish of mouthbrooding cichlid fishes in lake tanganica. *Nature* 323, 58–59.

Selten, R. 1975 Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory* 4, 25–55.

Selten, R. 1983 Evolutionary stability in extensive two-person games. *Math Soc. Sciences* 5, 269–363.

Simon, H. 1955 A behavioural model of rational choice. *Quarterly Journal of Economics* 69, 99–118.

Simon, H. 1956 Rational choice and the structure of the environment. *Psychological Review* 63, 129–138.

Staddon, J. 1975 A note on the evolutionary significance of “supernormal stimuli. *American Naturalist* 109(969), 541–545.

Tinbergen, N. 1951 *The study of instinct*. New York and London: Oxford University Press.

van Damme, E. 1987 *Stability and perfection of Nash equilibria*. Berlin: Springer-Verlag.

Wachtmeister, C.-A. 2001 Display in monogamous pairs: a review of empirical data and evolutionary explanations. *Animal Behaviour* 61, 861–868.

Wachtmeister, C.-A. & Enquist, M. 2000 The evolution of courtship rituals in monogamous species. *Behavioral Ecology* 11(4), 405–410.

Wiley, R. H. 1983 the evolution of communication: information and manipulation. In *Communication* (eds. T. R. Hallidday & P. J. B. Slater), vol. 2. Oxford: Blackwell Scientific.