The death of costly signalling?

Michael Lachmann
Santa Fe Institute
Santa Fe, NM
USA

Carl T. Bergstrom
Department of Biology
Emory University
Atlanta, GA, USA

Szabolcs Számadó
Department of Plant Taxonomy and Ecology
Eötvös Lorand University
Budapest, Hungary

December 13, 2000

How do organisms communicate honestly despite conflicts of interest? Over the past quarter-century, the "costly signalling" hypothesis — that signal honesty can be ensured by appropriate signal cost — has emerged as the dominant explanation for this puzzle. First proposed by Zahavi [1, 2] and formalized by Grafen [3] and Godfray [4], this hypothesis has led to a proliferation of theoretical models [5, 6, 7, 8, 9, 10, 11, 12, 13, 14] and empirical tests (reviewed in [15, 16, 17, 18]). Unfortunately, these empirical studies suggest that honest signalling is not always accompanied by the predicted signal costs (reviewed in [12]), and some signalling systems (including human language) appear not to require signal cost at all. In response to these difficulties, researchers have attempted to identify special mechanisms by which signalling can be honest even with low or zero signal cost (also reviewed in [12]). Here, we show that no special mechanism is necessary. While the cost of out-of-equilibrium signals plays an important role in stabilizing honest signalling, the signals actually employed at equilibrium need not be costly. *Therefore, even unrelated individuals with conflicting interests can communicate honestly using cost-free or very cheap signals; contrary to the "handicap principle," waste is not required to ensure honest signals.* We illustrate this by constructing examples of cost-free signalling equilibria for the two paradigmatic signalling games of Grafen [3] and Godfray [4]. Our findings (1) significantly revise previous theoretical conclusions regarding the requirement for signal cost in honest signalling systems, (2) explain

the discrepancy between empirical signalling studies and theoretical predictions, (3) suggest why some animal signals use cost to ensure honesty while others do not, and (4) provide ways in which signalling theory can be used to address the "problem of deception" in the evolution of human language.

Communication — whether among organisms, cells within organisms, economics agents, humans in social groups, or any other types of individuals — is often viewed as an evolutionary game played between a signaller and a signal receiver [3, 4, 5, 19]. In its basic form, a signalling game can be described as follows: The signaller has private information, typically pertaining to her own condition (e.g. quality [3] or hunger level [4]). The signal receiver chooses an action, such as feeding the signaller, with a payoff that depends on this information. To convey the information to the receiver, the signaller sends a signal which may be costly to the signaller. A formal description is provided in the Methods.

Two types of signalling equilibria have been described in the literature: In a *separating equilibrium*, signallers of different qualities always send distinct signals, and thus the receiver can perfectly infer the quality of the signaller. In a *pooling equilibrium*, signallers of different qualities might send identical signals, and thus the receiver cannot completely infer the signaller quality from the signal. It has already been shown that certainly pooling equilibria can be arbitrarily cheap [12], allowing inexpensive transfer of limited information, but separating equilibria were thought to require substantial "strategic" cost. In the Methods of the present paper, we prove that separating equilibria can be arbitrarily cheap. This result obtains because signalling equilibria are stabilized by the costs of out-of-equilibrium signals, not by the costs of equilibrium signals; long as the former are sufficiently costly, the latter can be zero for all signallers. Therefore, signal cost is never necessary at equilibrium; contrary to the primary tenet of costly signalling theory [1, 2, 3, 4, 20, 21, 22], waste is *not* necessary for signal honesty.

To illustrate this point, we show that honest signalling is possible without signal cost at equilibrium in two classic models of "costly signalling". Figure 2 shows a cost function which allows arbitrarily-cheap signalling in Godfray's model of signalling among relatives [4]. Figure 3 shows a cost function allowing arbitrarily cheap signalling in a game analogous to Grafen's sexual selection model [3]. From these figures, we see that while the relation between signal cost and signaller quality can stabilize honest signalling, signals need not be expensive at equilibrium. Indeed, previous authors concluded that cost was necessary only because they imposed specific constraints on the form of the signal function. In Godfray's case, this constraint is the implicit assumption that all signallers pay the same cost to send any given signal, regardless of their own condition. In Grafen's model, this constraint is that signal cost takes the form q^s, where q is the quality of the signaller and s is the magnitude of the signal. Further details are provided in the Methods.

[Figure 1 about here.]

[Figure 2 about here.]

[Figure 3 about here.]

If the signal cost can be arbitrarily cheap at equilibrium — that is, if there is no *a priori* mathematical reason that signal honesty requires signal cost — should we expect that signal costs in nature will always be very close to zero? Not necessarily. Cost-free signalling will not be stable unless a very specific relationship obtains between signaller quality and the cost of signal production, i.e., unless the cost function $D(q, s)$ takes a very specific form. In some cases, it may be reasonable to assume constraints on the form of the signal cost function; for example, signal cost will be unrelated to signaller quality in some systems [4]. Under such circumstances, the models do predict that honest signals will be costly at equilibrium.

By contrast, cheap or cost-free signalling can occur when there is great flexibility in the relation D of cost to signal (i.e., a greater flexibility in the *signaling mechanism*). This flexibility is essential to impose the strategic necessity of sufficient cost on signallers who send 'wrong' signals, while allowing very low costs for those sending the 'right' signals.

When should we see sufficient flexibility in the form of cost function to allow cheap signalling? Sharp differences in signal costs may obtain more readily when costs are socially imposed by "punishment" of cheaters, than when costs are physiologically imposed the rigors of producing extravagant signals. For social enforcement to be effective, however, the information being conveyed must be independently verifiable by the signal receiver; otherwise detection and punishment of deceptive signallers is not possible. Therefore, we expect that animals will use cheap signals when the signal receiver has (1) coincident (or partially-coincident, for some signalling systems [12, 10]) interests with the signaller, (2) conflicting interests, but the ability to independently verify the information conveyed. When interests conflict and signals are not verifiable, signal honesty must be maintained by the lessflexible process of signal production and thus we would be more likely see classic "Zahavian" costly signalling. Hence many passerine species can signal verifiable traits — size, aggressiveness, and fighting ability — using forehead patches [8, 23], whereas peacocks have to use costly tails to signal genetic quality, a trait which is difficult for peahens to verify on the timescale of their mating decisions.

Human language provides a particularly striking example of how flexible cost structure can facilitate cost-free communication. In human verbal communication, the costs of lying are socially imposed by societal sanction against liars; once again there can be great flexibility in the cost-to-signal relationship. At the same time, social imposition of cost provides an important degree of rigidity to the signalling system. When the signaller physiology directly imposes the costs of signal production (as in the peacock), signallers are selected to evolve cheaper ways of sending equivalent signals. This process can ultimately bring about the decay of honest signalling [24]. When the cost signal production is imposed socially, by receivers, no such breakdown will occur: signallers *cannot* evolve cheaper ways of signalling and receivers lack the incentive to do so.

Recent theoretical work on human language evolution provides a number of important insights into the nature of this major evolutionary transition [25, 26, 27, 28]. However, a major unresolved problem remains: by what mechanisms could language evolve to be cheap but honest despite conflicting interests among communicating individuals [29]? For this reason, linguistic models are typically set in an Eden in which signaller and receiver have entirely coincident interests. Human language almost certainly did not evolve in such an Eden. We imagine that conflicting interests would have been frequent, and that the problem of honesty would have exerted a continual influence on the development of language. Here, we have shown that cost-free signalling systems such as human language can be addressed in a framework of non-coincident interests and socially imposed signal costs. As such, signalling theory can now be brought to bear on the language evolution problem. The results of models based on coincident interests (e.g., selection for greater precision and greater volume of information transmission, neither of which is necessarily favored when interests conflict [10]) can be reconsidered under more general evolutionary scenarios.

Our findings remove the discrepancy between models of discrete and continuous games; arbitrarily cheap signalling was previously known be possible in the former [11, 13], but thought to be impossible in the latter [1, 3, 4, 22]. Moreover, these findings have the potential to explain empirical studies which, by measuring negligible signal costs despite conflicting interests, appeared to contradict game-theoretic predictions (see ref. [12]). In addition, the realization that it is not the cost but instead the derivative (the cost differences) that is selected has important implications for future empirical studies of honest signalling. Namely, measurement of cost at equilibrium is not sufficient. Equilibrium cost tells us very little because many different costs (even zero cost) are possible at equilibrium under the revised theory. Instead, the theory's strong and unambiguous predictions involve the consequences of deviations from equilibrium; manipulation experiments in which signal magnitudes are altered may be necessary to test the "costly signalling" hypotheses.

By focusing on signal cost, rather than on the slope of signal cost (with respect to signal type and/or signaller quality), previous signalling models overlooked the existence of alternative separating equilibria — including cost-free separating equilibria — in signalling games. As a consequence, a misleading set of conclusions have gradually become entrenched in the study of communication: signal costs are typically assumed to be essential for honest signalling when signaller and signal receiver have conflicting interests. We have shown that contrary to the established view, honest communication does not require signal cost at equilibrium. Nonetheless, so-called costly signalling theory — perhaps more accurately termed "*honest* signalling theory" — still predicts a specific quantitative relation between a signal's cost, the signaller's condition or quality, and the signal's meaning, i.e., the response that it induces. Thus costly signalling theory is by no means dead. Our work should be seen as significantly refining rather than refuting prior work; indeed we believe that the methods and intuitions described here will lead to rapid progress in our understanding of signalling and the evolution of communication.

Methods

A signalling game can be described formally as follows: In each round, the signaller's condition, or quality, is q. Given this quality, she selects an signal $s = T(q)$ and pays a display cost $d = D(q, s)$. Upon receiving a signal s, the receiver takes an action $r = R(s)$. The payoff to the signaller, $H(q, r)$, depends on her quality and the action taken by the receiver. The payoff to the receiver, $G(q, r)$, depends on the same parameters. The "strategies" in this game are simply the signaller's choice of what signal to send, given by the function $T(q)$, and the receiver's choice of how to respond to each signal, given by the function $R(s)$. At equilibrium, a signaller of quality q will not benefit from sending a signal different than $T(q)$, paying the appropriate cost. The receiver will not benefit by responding to a signal s with an action other than $R(s)$.

Given the (continuously differentiable) payoffs for signaller and receiver, $H(q,r)$ and $G(q,r)$, is there a continuously differentiable cost-function $D(q,s)$ for which the equilibrium signals will be completely honest and yet cheap or free? At a separating equilibrium receivers can precisely infer the condition of any signaller ([3, 14]); therefore we shall identify signals with the qualities that they advertise. The receiver will have some optimal response $r = z_r(q)$ to a signaller in condition q. By definition, at a separating equilibrium, a signaller with quality q will advertise quality q, and will get the highest payoff from this choice of signal. To standardize the meaning of "cost", we define the cost to be 0 for sending some arbitrary signal q_0, no matter what the actual quality q of the signaller is. This can be seen as the cost of 'not signalling' [30]. All other signals have a non-negative cost. Then the net payoff to the signaller $H\Big(q, z_r(q')\Big) - D\Big(q, q'\Big)$ will have a maximum at $q' = q$, i.e., for any alternative signal $q' \neq q$,

$$H\Big(q, z_r(q)\Big) - D\Big(q, q\Big) \geq H\Big(q, z_r(q')\Big) - D\Big(q, q'\Big). \tag{1}$$

This implies that

$$\frac{\partial}{\partial q'} D(q, q')\Big|_{q'=q} = \frac{\partial}{\partial q'} H(q, r(q'))\Big|_{q'=q}. \tag{2}$$

Since it also holds that

$$D(q,q) = \int_{q_0}^{q} \frac{\partial D}{\partial q'}(q,q')dq', \tag{3}$$

we can choose D such that for all q, equation (2) is met but $\frac{d}{dq'}D(q',q')$ is arbitrarily small for $q' < z_r(q)$. Thus we can construct a smooth cost function D with arbitrarily small signal costs at equilibrium. We now provide two examples.

Godfray's model. In Godfray's model, an offspring of quality $q \in [0.5, 2.5]$ (where low q implies high hunger level) signals to elicit resource transfer from its parent. The offspring receives a personal fitness benefit $1 - e^{-rq}$ for response r; the parent receivers a personal fitness benefit of $1 - 0.08r$. To handle the relatedness between two players, Godfray employs Hamilton's rule: signaller and receiver each attempt to maximize their inclusive fitnesses, i.e., the sum of their personal fitnesses weighted by relatedness. Net payoffs are therefore $H(q,r) = (1+k) - e^{-rq} - 0.08rk$ and $G(q,r) = 2 - e^{-rq} - 0.08r$ (see [4]). When subsequent offspring will be full sibs, $k = 1/2$. Thus the parent's optimum response curve is given by $z_r(q) = (1/q)\log(q/0.08)$, and $\frac{\partial H}{\partial r} = -0.04 - qe^{-qr} = 0.04$ everywhere along the equilibrium path (Figure 1). By (2) we need to find a smooth cost function C with a partial derivative of cost with respect to induced response of $\frac{\partial C}{\partial r} = 0.04$ everywhere along equilibrium path and with $C(q, z_r(q)) = E(q) < \epsilon$ for all q. One such function has a partial derivative $\frac{\partial C}{\partial r}$ which is a smooth approximation to a step function and is given by $\frac{\partial C}{\partial r} = a + (2a/\pi)\text{Arctan}[b(\epsilon,q)r - z_r(q)]$ where $a = \frac{\partial H}{\partial r} = 0.04$ and b can be found numerically such that the necessary conditions are met. Such a cost function is shown in Figure 2 for $\epsilon = 0.0001$, i.e, for which signal cost never exceeds 0.0001 at equilibrium.

Grafen's model. In Grafen's model, a male signals his quality $q \in [0.2, 0.8]$ to a prospective mate, who selects a response (e.g. inclination to mate) r. The signaller's payoff is given by $w(s, r, q) = r^{0.3} q^s$. The signal receiver's optimal response curve is $z_r(q) = q$. Here, the benefit and cost to the signaller are not linearly independent. In this game, however, we can determine the optimal strategic behavior for the signaller by maximizing any monotone increasing transformation of payoff; here we choose to use $\log w(s, r, q) = 0.3 \log r + s \log q$. At this point, we can see that one of Grafen's seemingly innocuous assumption, that signal cost takes the form $F(q, s) = q^s$, is actually a strong and arbitrary constraint. To treat a more general class of signalling games (as Godfray did), Grafen could have used the payoff function $w(s, r, q) = r^{0.3} q^{F(q,s)}$, or equivalently, $\log w(s, r, q) = 0.3 \log r + F(q, s) \log q$. Grafen's payoff function is a special case of this function. Define $H(q, r)$ to be the first term and $D(q, r) = C(F(q, z_r^{-1}(r))$ to be the second (parameterized in terms of the equilibrium response rather than the signal). As above, we construct a cost function with partial derivative $\frac{\partial D}{\partial r} = a(q) + (2a(q)/\pi)\text{Arctan}[b(\epsilon, q)r - z_r(q)]$, where $a(q) = \frac{\partial}{\partial r} H(q, r)\Big|_{r=z_r(q)}$ and where b is determined numerically. Such a cost function is shown in Figure 3 for $\epsilon = 0.01$.

References

[1] Zahavi, A. *Journal of Theoretical Biology* **53**, 205–214 (1975).

[2] Zahavi, A. *Journal of Theoretical Biology* **67**, 603–605 (1977).

[3] Grafen, A. *Journal of Theoretical Biology* **144**, 517–546 (1990).

[4] Godfray, H. C. J. *Nature* **352**, 328–330 (1991).

[5] Maynard Smith, J. *Animal Behaviour* **42**, 1034–1035 (1991).

[6] Nur, N. and Hasson, O. *Journal of Theoretical Biology* **110**, 275–297 (1984).

[7] Enquist, M. *Animal Behaviour* **33**, 1152–1161 (1985).

[8] Maynard Smith, J. and Harper, D. G. C. *Philosophical Transactions of the Royal Society of London, Series B* **319**, 557–570 (1988).

[9] Johnstone, R. A. and Grafen, A. *Journal of Theoretical Biology* **156**, 215–234 (1992).

[10] Lachmann, M. and Bergstrom, C. T. *Theoretical Population Biology* **54**, 146–160 (1998).

[11] Hurd, P. L. *Journal of Theoretical Biology* **174**(2), 217–222 (1995).

[12] Bergstrom, C. T. and Lachmann, M. *Proceedings of the National Academy of Sciences USA* **95**, 5100–5105 (1998).

[13] Számadó, S. *Journal of Theoretical Biology* **198**, 593–602 (1999).

[14] Noldeke, G. and Samuelson, L. *Journal of Theoretical Biology* **97**, 527–539 (1999).

[15] Godfray, H. C. J. *Nature* **376**, 133–138 (1995).

[16] Johnstone, R. A. *Biological Reviews* **70**(1), 1 (1995).

[17] Caro, T. M. *Trends in Ecology and Evolution* **10**, 500–503 (1995).

[18] Kilner, R. and Johnstone, R. A. *Trends in Ecology and Evolution* **12**, 111–15 (1997).

[19] Spence, M. *Quarterly Journal of Economics* **87**, 355–374 (1973).

[20] Zahavi, A. *Animal Behaviour* **42**, 501–503 (1991).

[21] Zahavi, A. *Animal Behaviour* **54**, 228–229 (1997).

[22] Getty, T. *Animal Behaviour* **56**, 127–130 (1998).

[23] Rohwer, S. *American Zoologist* **22**, 531–546 (1982).

[24] Godfray, H. C. J. *American Naturalist* **146**(1), 1–24 (1995).

[25] Pinker, S. and Bloom, P. *Behavioral and Brain Sciences* **13**, 707–784 (1990).

[26] Maynard Smith, J. and Szathmáry, E. *The Major Transitions in Evolution.* Oxford University Press, Oxford, (1997).

[27] Nowak, M. A., Plotkin, J. B., and Jansen, V. A. A. *Nature* **404**, 495–498 (2000).

[28] Rubinstein, A. *Economics and Language.* Cambridge University Press, (2000).

[29] Knight, C. In *Approaches to the evolution of language: social and cognitive bases,* Hurford, J. R., Studdert-Kennedy, M., and Knight, C., editors. Cambridge University Press, Cambridge (1998).

[30] Bergstrom, C. T. and Lachmann, M. *Philosophical Transactions of the Royal Society of London, Series B* **352**, 609–617 (1997).

List of Figures

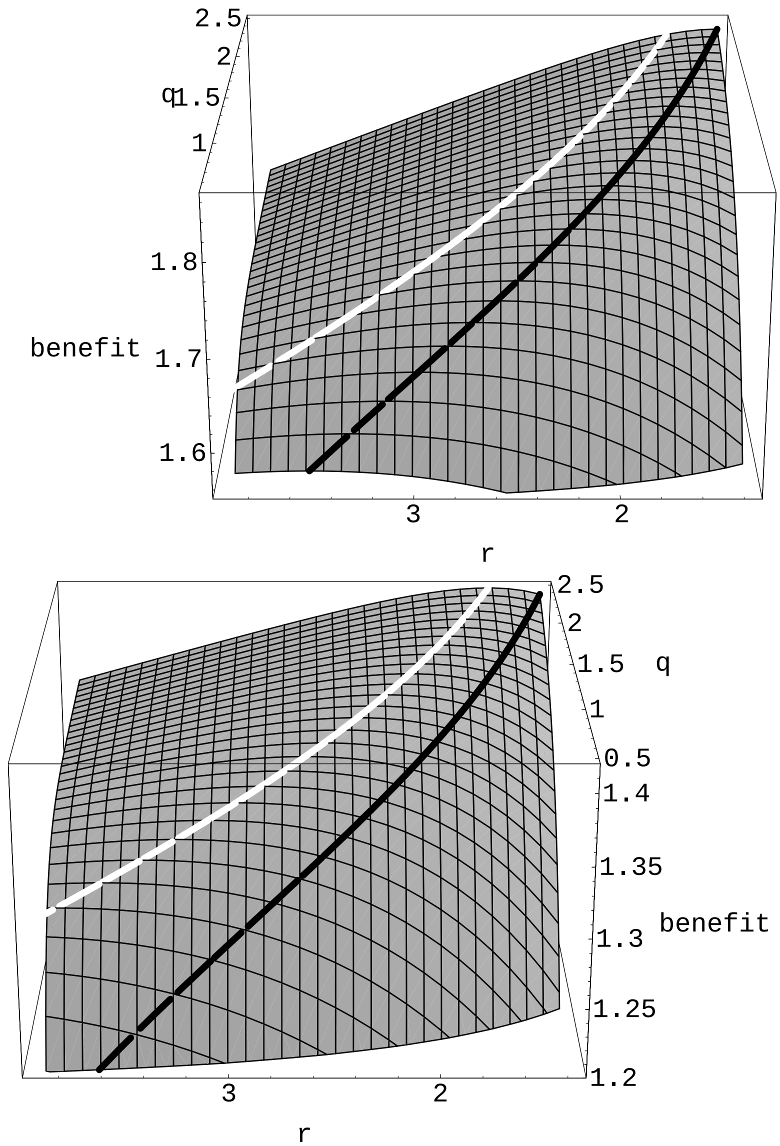

Figure 1: The benefit functions from response r in state q for Godfray's model (See Methods). The figures show (a) the parent's inclusive fitness benefit in the absence of signal cost and (b) the signaller's inclusive fitness benefit in the absence of signal cost. The parent's maximum benefit is shown by the black line and the signaller's maximum benefit is shown by the white line. At a separating equilibrium, the addition of signal cost to the signaller's benefit will move the signaller's maximum to be the same as the parent's maximum.

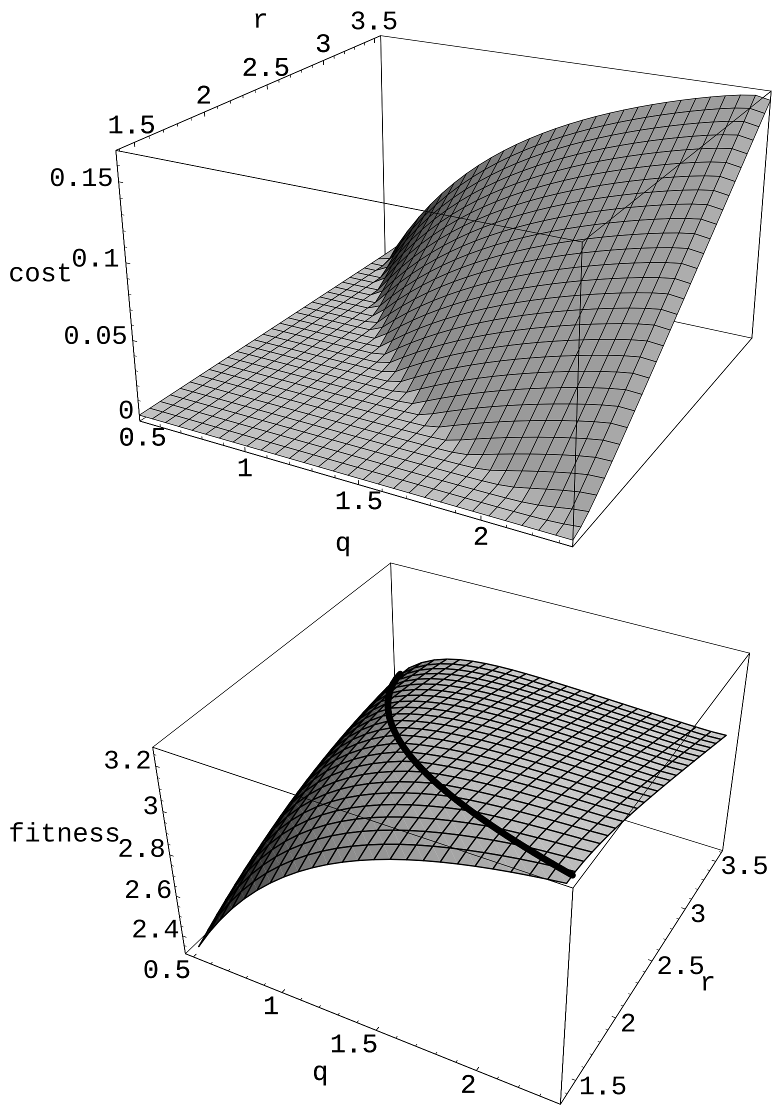

Figure 2: Cheap (cost< 0.001) signalling in Godfray's model: (a) signal cost and (b) signaller's net fitness when eliciting response r in state q. The receiver's equilibrium response is $z_r(q)$ shown by the solid line and indeed optimizes signaller fitness. (See Methods)

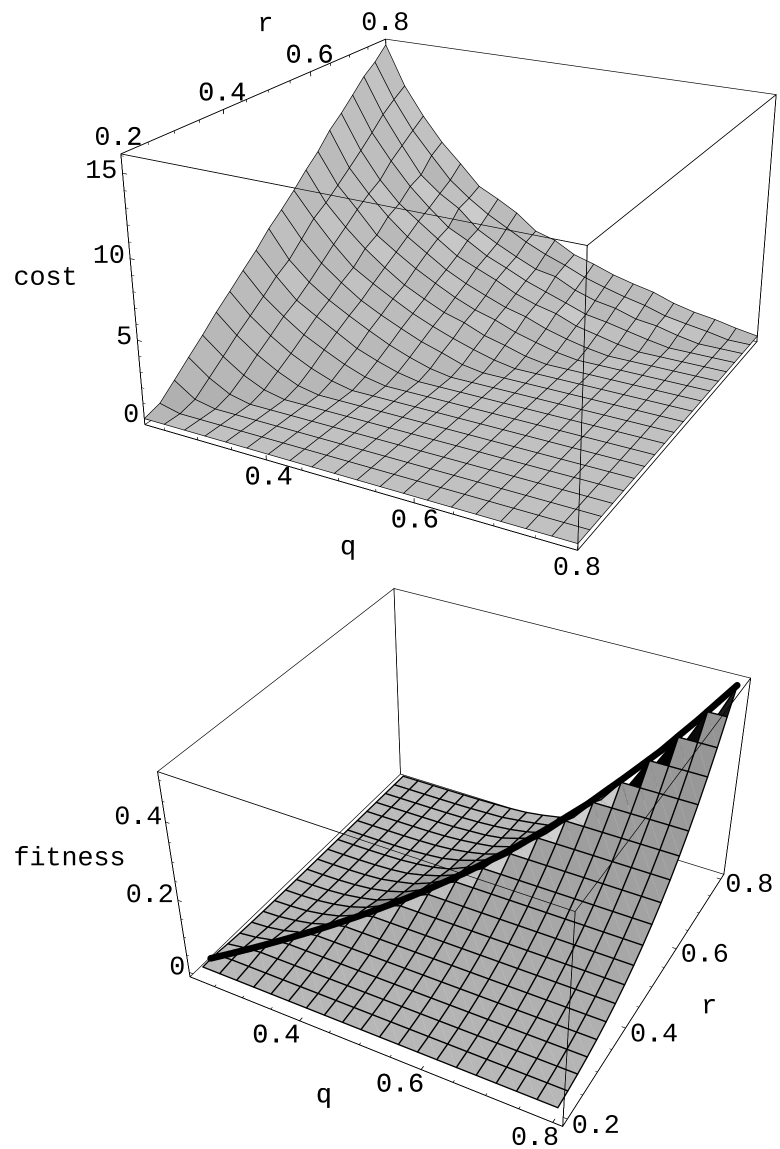

Figure 3: Cheap (cost< 0.01) signalling in Grafen's model: (a) signal cost and (b) signaller's net fitness when eliciting response r in state q. Equilibrium response $z_r(q)$ is shown by the solid line. (See Methods)