By Genes Alone:

A Model Selectionist Argument for Genetical Explanations of Cooperation in Non-

Human Organisms

Armin Schulz

Department of Philosophy

University of Kansas

3101 Wescoe Hall

Lawrence, KS 66045

awschulz@ku.edu

http://people.ku.edu/~a382s825/

Abstract

I distinguish two versions of kin selection theory—a purely genetic version (GKST) and a version that also appeals to cultural (i.e. non-genetically-derived) forms of cooperation (WKST)—and present an argument in favor of using the former when it comes to accounting for the evolution of cooperation in non-human organisms. Specifically, I first show that both GKST and WKST are equally mathematically coherent—they can both be derived from the Price equation—but not necessarily equally empirically plausible, as they are based on different assumptions about the inheritance system underlying the cooperative phenotype. Given this, I then, second, present a model selection theoretic argument in favor of GKST over WKST. This argument is based on the fact that, in non-human cases, the former theory is likely to be as empirically successful as WKST, while containing fewer degrees of freedom. I end by defending both the intrinsic importance of this argument and its relevance to the discussion surrounding the "gene's eye view of evolution."

By Genes Alone:

A Model Selectionist Argument for Genetical Explanations of Cooperation in Non-

Human Organisms

I.   Introduction

The evolution of cooperation—especially strongly altruistic cooperation—is a well-

known puzzle in evolutionary theory. Equally well-known (by now) is the fact that there

are several different ways of solving this puzzle (West et al., 2007; Sober & Wilson,

1998; Okasha, 2006). Among these solution strategies, one of the most widely accepted

ones is based on appealing to kin selection theory (KST)—i.e. the idea that natural

selection does not just target specific individuals, but also the "kin" (in a technical sense

made precise below) of these individuals. (Another widely accepted solution strategy is

based on appealing to multi-level selection theory. However, in this paper, the focus is

just on KST-based approaches towards the evolution of cooperation—though see also

note 18 below.)[1] However, as I show in this paper, the wide acceptance of KST masks the

fact that there are in fact several different versions of KST that can and have been

appealed to when it comes to accounting for the evolution of cooperation: in particular,

there is a narrower form of KST that only considers the degree of genetic similarity

among the cooperating organisms, and a wider form that also considers the non-genetic,

purely cultural degree of similarity among the cooperating organisms.

In this paper, I present and defend a model selection theoretic argument in favor of the

narrower, genetic form of KST when it comes to non-human cases of the evolution of

---

[1] For more details about KST and multi-level selection theory (MSLT), see e.g. Gardner et al. (2011);
Grafen (1985); Hamilton (1963) (for KST), and Damuth and Heisler (1988); Okasha (2006); Sober and
Wilson (1998) (for MLST).

cooperation. The argument is based on the ideas that (a) empirically, the resources of the narrower, genetic form of KST are sufficient to account for the evolution of cooperation in non-human organisms, (b) humans are known special cases that ought to be approached in a different framework from non-human organisms, and (c) that model selection theoretic considerations give preference to simpler over more complex theories. In other words, the goal in this paper is make precise why it is sometimes reasonable to favor a genetic form of KST, even though this theory is *nested in* a wider theory that includes appeal to cultural factors. This conclusion is noteworthy also, as it gets at an important general contrast in the study of (social) evolution concerning the epistemic value of the "gene's eye view of evolution."

The paper is structured as follows. In section II, I present and distinguish two forms of KST. In section III, I present a model selection theoretic argument in favor of the genetical form of KST. In section IV, I bring out the importance of this argument. I conclude in section V.

II.   Two Forms of KST

To understand KST—and the different forms the latter can take—it is best to start by assuming that we seek to determine the conditions under which heritable strongly altruistic cooperation—i.e. heritable traits that lead to the provision of costly help to others—can evolve (Wilson & Dugatkin, 1992; Sober & Wilson, 1998). As is well known, the evolution of altruistic cooperation is a very puzzling phenomenon, as it concerns cases where organisms have (behavioral or morphological) traits that provide fitness benefits to others and, at least in the case of strong altruism, fitness losses to

themselves. The evolution of these traits is thus very puzzling, as it seems to go against the very nature of natural selection: these are traits that it seems maladaptive to have. How can their evolution be explained?[2]

According to KST, the driver of the evolution of altruistic cooperation are kin relations (in what is potentially a broad sense of "kin"—see below for more on this). Specifically, the key idea behind KST is that the maladaptedness of altruistic cooperation is, in important regards, misleading. While altruistic cooperation can be disadvantageous to a given individual, it can still be advantageous to organisms that are similar to that individual—and what determines whether a trait is selected for is not just how it affects a given individual, but also how it affects other, similar individuals.

To make this idea more precise, note that it can be well captured by Hamilton's rule. In its most basic form, Hamilton's rule states that cooperation will be favored in a population if:

(1) $rb - c > 0$,

where r is a coefficient of relatedness among the interacting members of the population, b is the benefit the cooperation provides to the recipient, and c is the cost of the cooperation to the cooperator (see e.g. Gardner et al., 2011; Griffin & West, 2002; Hamilton, 1963; Queller, 2004). Two points need to be noted about Hamilton's rule (and KST in general).

---

[2] Of course, it is possible that they did not evolve by natural selection. However, given the fact that these traits are quite common (West et al., 2011), and that there are—perhaps despite initial appearances to the contrary—selective explanations available for them, I will not consider this further here. This is not to say that many such traits might not have evolved in ways that do not strongly depend on natural selection; it just means that this is not what is at stake here.

First, Hamilton's rule can be derived from the Price equation—a central, purely mathematical theorem of evolutionary biology (see e.g. Grafen, 1985; Sober & Wilson, 1998; Frank, 1998; Price, 1970, 1972; Rice, 2004; Frank, 2012; Luque, forthcoming). To see this, begin by recalling that the Price equation states that the change in the average value of a trait z with average fitness w is given by:

(2) $\Delta z = (1/w) [Cov(z_i, w_i) + E(w_i \Delta z_i)]$,

where z_i is the value of z of the ith member of the population, w_i is the fitness of z_i, and Cov(.) and E(.) are the covariance and expectation across all individuals in the population.[3] With this in mind, note that one way to derive Hamilton's rule from the Price equation is as follows (see Frank, 1998).

Start with a simple linear regression of individual i's fitness on its own genotype (g_i) and the (average) of that of its interactors (g_i'), with an independently distributed error term ε measuring the non-systematic (i.e. non-heritable) influences on the phenotype:

(3) $w_i = \alpha + \beta g_i + \gamma g_i' + \varepsilon$.

This can be substituted into the Price equation (2) (with g replacing z) to yield:

(4) $\Delta g = (1/w) [Cov(g_i, \alpha + \beta g_i + \gamma g_i' + \varepsilon) + E(w_i \Delta g_i)]$

---

[3] As noted by Frank (1998), there are many possible interpretations of z. The ones in text are just chosen for expositional clarity.

Expanding gives:

(5) $\Delta g = (1/w) [Cov(g_i, \alpha) + \beta\ Cov(g_i, g_i) + \gamma\ Cov(g_i, g_i') + Cov(g_i, \varepsilon) + E(w_i\Delta g_i)]$

Noting that $Cov(g_i, \alpha) = 0$ (as α is a constant), $Cov(g_i, \varepsilon) = 0$ (by definition of ε as the residual of the regression in (3)), $Cov(g_i, g_i) = Var(g_i)$, we can simplify (5) to get:

(6) $\Delta g = (1/w) [\beta\ Var(g_i) + \gamma\ Cov(g_i, g_i') + E(w_i\Delta g_i)]$

Some rearranging gives:

(7) $\Delta g = (1/w)\ Var(g_i) [\beta + \gamma\ Cov(g_i, g_i') / Var(g_i) + E(w_i\Delta g_i) / Var(g_i)]$.

Thus (assuming that $Var(g_i) \neq 0$), it will be true that

(8) $\Delta g > 0$ if and only if $\beta + \gamma\ Cov(g_i, g_i') / Var(g_i) > - E(w_i\Delta g_i) / Var(g_i)$.

Rewriting (8) by setting $\beta = c$, $\gamma = b$, $r = Cov(g_i, g_i') / Var(g_i)$ gives:

(9) $\Delta g > 0$ if and only if $-c + br > - E(w_i\Delta g_i) / Var(g_i)$.

Assuming that genetic transmission is faithful, i.e. that $E(w_i\Delta g_i) = 0$, (9) reduces to Hamilton's rule in its familiar form (1).

With this in the background, it now becomes possible to note the second key feature of KST: that it is not one theory, but a set of theories, tailored to specific evolutionary scenarios through the assumptions built into the derivation of Hamilton's rule (Gardner & West, 2010).[4] For present purposes, it is useful to distinguish two main versions of KST: a narrow, purely genetic form, and a wider, partially cultural form that allows for the cooperative trait to be genetically *or* purely culturally transmitted.

The genetic (and perhaps slightly more widely known) form of KST—GKST in what follows—is the one just derived. Slightly relabeled, it states:

(10) $r_g b - c > 0,$

where $r_g = \text{Cov}(g_i, g_{i'}) / \text{Var}(g_i)$, i.e. a measure of purely genotypic similarity among the interacting organisms. According GKST, then, cooperation will evolve if the ratio of net costs to benefits of cooperation (c/b) is less than the probability that the beneficiary of the cooperation is a genetically similar (at the relevant locus) organism: either because it is a relative or because it is a genetically similar non-relative (i.e. a "greenbeard"—see e.g. Gardner & West, 2010). This form of KST is thus built on the assumption that tracking genotypes is sufficient for tracking cooperative phenotypes.

By contrast, the wider form of KST—WKST in what follows—allows for (but does not require) the cooperative trait to also evolve without a genetic basis, i.e. purely by cultural evolution (in a broad sense). That is, according to WKST, what matters for the evolution of cooperation is whether phenotypically similar organisms—whether for

---

[4] Alternatively, one might conceive of KST as a family of models (Weisberg, 2013; Morrison, 2015). However, since settling this issue is not central for present purposes, I will not consider it further here.

genetic or cultural reasons—predominately interact with each other. Importantly, while perhaps less well known, this form of KST is just as mathematically coherent as the narrower GKST, and taken just as seriously in the literature (see e.g. Gardner et al., 2011).[5]

To understand this version of KST better, note that, in general, one can see an organism's heritable phenotype as stemming from two sources: its genetic makeup and its cultural background. (The environment is of course relevant as well; however, it is here taken to be part of the error term ε in (3) / (12)—i.e. "culture" is here assumed to include all systematic influences on an individual's phenotype that are not genetic in nature.)

Now, there can be no doubt that spelling out the difference between the genetic and the cultural sources of a phenotype is far from straightforward. On the one hand, it has turned out to be difficult to specify exactly what genes are—both on a molecular and an evolutionary level—and there is also increasing recognition of the importance of "gene-external" factors (such as epigenetic interactions and environmental inputs to gene expression) on the evolution and development of phenotypes (see e.g. Sarkar, 2005; Godfrey-Smith, 2009; Sterelny & Griffiths, 1999, chaps. 4-7; Oyama, 2000). On the other hand, the exact details of how cultural learning can and does work turn out to be tricky to specify (Henrich, 2015; Boyd & Richerson, 2005; Sterelny, 2012; Skyrms, 2010).

---

[5] Some authors (see e.g. Queller, 1985, 1992) have argued that, in phenotypic forms of Hamilton's rule, it is likely that there are "synergistic effects" that need to be taken into account: in particular, the cooperative phenotype may yield a benefit to other cooperators independently of the assortative effect measured by the correlation of genotypes g_i and g_i' (e.g. helping a helper may yield benefits that helping a non-helper does not). There are different ways of including this into Hamilton's rule; for example, Queller (1985) shows that, in some cases, these synergistic effects can be captured by adding an extra term to (1) like this: $\Delta p > 0$ if and only if $-c + br + dS > 0$, where S is a variable that captures the size of the synergistic (non-additive) effects of helping helpers. However, it is controversial to what extent this is unique to the phenotypic case (Queller, 1985, 1992, 2011; Birch & Marshall, 2014; Marshall, 2011; Sober, 2000). For this reason, I will not consider synergistic effects further here—noting just that, if synergistic effects *do* systematically affect WKST more than GKST, this would only favor my argument, as it adds degrees of freedom to the former. See also below.

However, this does not mean that we cannot maintain *that* there is a distinction between the genetic and the cultural evolution of phenotypes. While the details of exactly how to spell out this distinction are complex and controversial, the basic fact of there being such a distinction is not. This is all that is needed here: for present purposes, any (reasonable) way of drawing such a distinction will be adequate.

In turn, this means that we can set:

(11) $p_i = g_i + q_i,$

where g_i is, as before, the ith's individual's genotype, and q_i is a measure of i's cultural background—i.e. the heritable non-genetic sources of its phenotype p_i. Using (11) in place of g_i in regression equation (3) gives:

(12) $w_i = \alpha + \beta(g_i + q_i) + \gamma(g_i' + q_i') + \varepsilon.$

Plugging (12) into the Price equation (2) (with p replacing g), and following the same derivation as above gives:

(13) $\Delta p > 0$ if and only if

$\beta + \gamma[Cov(g_i, g_i')/Var(p_i) + Cov(g_i, q_{i'})/Var(p_i) + Cov(q_i, g_{i'})/Var(p_i) + Cov(q_i, q_{i'})/Var(p_i)]$

$> -E(w_i \Delta p_i)/Var(p_i).$

Note further that, by setting $a = Var(g_i) / [Var(g_i) + Var(q_i)]$, (13) can be rewritten as follows:

(14) $\Delta p > 0$ if and only if

$\beta + \gamma[aCov(g_i,g_{i'})/Var(g_i) + Cov(g_i,q_{i'})/Var(p_i) + Cov(q_i,g_{i'})/Var(p_i) + (1-a)Cov(q_i,q_{i'})/Var(q_i)] > -E(w_i\Delta p_i)/Var(p_i)$.

Rewriting (14) as above gives

(15) $\Delta p > 0$ if and only if $-c + b [a\, r_g + (1-a)\, (r_q + r_{gq} + r_{qg})] > -E(w_i\Delta p_i)/Var(p_i)$,

where $r_q = Cov(q_i, q_{i'}) / Var(q_i)$, i.e. a measure of the similarity in the cultural background among the relevant organisms, and $r_{gq} = Cov(q_i, g_{i'}) / Var(q_i) / r_{qg} = Cov(q_i, q_{i'}) / Var(q_i)$, i.e. measures of the degree to which genetic cooperators interact with purely cultural cooperators, and vice-versa.

   Finally, note that the right-hand side of (15) can further be rewritten so as to yield

(16) $\Delta p > 0$ if and only if $-c + b [a\, r_g + (1-a)\, (r_q + r_{gq} + r_{qg})] > -[E(w_i\Delta g_i)/Var(g_i)\, a+d]$,

where $d = E(w_i\Delta q_i)/[Var(p_i) + Var (g_i)]$. This is important to note, as in cases of cultural evolution, it is more plausible that there will some transmission bias across generations (Boyd & Richerson, 2005; Henrich, 2015; El Mouden et al., 2014; Morin, 2016; Lewens, 2015): for example, the presence of horizontal transmission will make the acquisition of the trait easier than it would be if it were purely vertically or genetically transmitted. If

so, then the ratio of the net costs to benefits of the cooperation (c/b) can be lower and cooperation can still be adaptive. Put differently: if we maintain the assumption that $E(w_i\Delta g_i)/Var(g_i) = 0$ (as above), so that there is no genetic transmission bias, it might still be true that $d \neq 0$—i.e. that there is a cultural transmission bias.

The structure of equation (16) is important, as it shows that GKST can be seen as a restricted version of WKST.[6] In particular, WKST contains elements not included in GKST: it allows that (i) $Var(q_i) \neq 0$ (implying that $a \neq 1$), (ii) $r_q \neq 0$, (iii) $r_{gq} \neq 0$, (iv) $r_{qg} \neq 0$, and (v) $E(w_i\Delta q_i) \neq 0$ (implying that $d \neq 0$). Put differently, to go from WKST to GKST, it needs to be assumed that (a) $E(w_i\Delta q_i) = 0$, and (b) $Var(q_i) = 0$ or (c) $r_q = r_{gq} = r_{qg} = 0$. Simplifying slightly: WKST states that cooperation will be favored if and only if the costs of cooperating relative to its benefits (c/b) are low relative to the probability that the beneficiary of the cooperation is another cooperator—either for genetic or for cultural reasons (see also Gardner et al., 2011; Grafen, 1985; Griffin & West, 2002; West et al., 2007, 2008; Queller, 2011; Frank, 1998).

In this way, we can derive the following two conclusions. First and most obviously, there are different forms of KST, depending on the exact assumptions that underlie Hamilton's rule—i.e. whether the cooperative trait is assumed to be purely genetically or also culturally transmitted. Second, KST, in either form, can be derived from the Price equation, and is thus at least mathematically coherent.

For what follows below, it is further important to note that the fact that the two forms of KST are equally mathematically coherent does not mean that they also need to be equally empirically successful. Just like evolution by natural selection can be seen as a

---

[6] Of course, it is also possible to define other restricted versions of WKST: for example, one could consider a form of KST that is restricted to cultural inheritance (i.e. q_i) only. However, the interest here is in GKST, as this is the theory that makes for an interesting empirical contrast to WKST. See also below.

mathematical triviality if the relevant conditions are satisfied (a set of heritable traits with fitness consequences—Lewontin, 1970; Godfrey-Smith, 2009), but whether the relevant conditions are satisfied in any individual case is an open empirical question (Sober, 2000), it can be the case that both GKST and WKST are mathematically coherent—but this does not mean that it is a meaningless question to ask which of these fit better to the empirical data. This is crucial for the rest of the discussion.

III.   A Model-Selectionist Argument for GKST

When comparing GKST and WKST, it may seem obvious that WKST is preferable to GKST: after all, it is more general—anything that can be achieved with GKST can be achieved with WKST, but the latter can, at least potentially, capture even more cases of the evolution of cooperation. What reason could there possibly be for *not* adopting the most general theory consistent with the data? However, as I try to show in this section, the situation is misleading: there are reasons to think that, in some cases, GKST should in fact be favored over WKST. This is a key point to emphasize: despite the fact that GKST is a special case of WKST, there are reasons to think that, in many cases, the former is a superior to the latter. Spelling out these reasons in detail is the aim of the rest of this paper. This can be done in three steps.

1. GKST is Simpler than WKST

There is a precise sense in which GKST is *simpler* relative to WKST. Specifically, GKST has *fewer degrees of freedom* than WKST: it is based on a restricted version of Hamilton's rule that does not allow $E(w_i \Delta q_i)$ to be non-zero, $Var(q_i)$ to be non-zero, or

any of r_q, r_{gq}, and r_{qg} to be non-zero.[7] Put differently, GKST tries to handle the same set of empirical phenomena as WKST, but with fewer theoretical tools—it aims to capture the evolution of cooperation (and other social traits) just by using measures of the genetic similarity among the relevant organisms, and not by also appealing to their cultural similarity. In this way, GKST can be straightforwardly seen to be the more parsimonious theory than WKST—in the precise sense of having fewer degrees of freedom.

2. GKST and WKST Are Empirically Equally Adequate—Except for a Known Special Case

However, despite the fact that GKST is simpler than WKST, there is good reason to think that, across all of the core biological applications of the two theories, they will be empirically largely equivalent to each other: they will agree on when we should expect the evolution of cooperation, and when not. This is due to the fact that, given the way the world happens to be, it is plausible that GKST will diverge from WKST only in a relatively small—and anyway anomalous—set of cases.[8]

Specifically, there is now a significant body of literature making clear that genuine cultural learning of the kind needed here—i.e. one which makes the cooperative trait sufficiently heritable for it to be able to evolve cumulatively—is adaptive only in small set of environments (Boyd & Richerson, 2005; Henrich, 2015; Richerson & Boyd, 2005). These are environments that change sufficiently fast to make using information received

---

[7] Note that WKST is a non-additive expansion of GKST. However, all that is relevant here is that WKST has more degrees of freedom than GKST; the nature and exact quantity of this increase are not so important—see also note 16.

[8] This marks an important contrast to other restricted forms of WKST: e.g. ones restricted to cultural factors only. The latter would *not* be empirically equivalent to WKST in non-human cases of the evolution of cooperation.

from the parental generation adaptive—also taking into account the costs of imitation learning (e.g. in terms of time or mistakes)—but not so fast as to make the information received from the parental generation outdated by the time it is received by the filial generation (Boyd & Richerson, 2005; Henrich, 2015).

Importantly furthermore, the major—and perhaps even only—example of organisms that live in such an environment are humans: humans occupy many different physical environments, lack major physical adaptations to these environments, are highly social, and profit from learning how to interact successfully with the particular local conditions they happen to find themselves in (Henrich, 2015; Richerson & Boyd, 2005; Boyd & Richerson, 2005; Sterelny, 2012, 2003). Indeed, the existence of a significant system of cumulative cultural inheritance has often been seen as the key driver of what makes humans unique in the biological world (Sterelny, 2012, 2003; El Mouden et al., 2014; Boyd & Richerson, 2005; Henrich, 2015).

Turning this around, the key point to note here is that, apart from humans, significant systems of cultural inheritance are hard to find. It is true that some have argued that there are cases of animal cultures (Laland & Janik, 2006; C. Heyes & Galef, 1996; Reader & Biro, 2010). However, the cultures involved in these cases are also widely seen to be spatio-temporally limited and not to affect the evolutionary dynamics of the traits in question (Henrich & McElreath, 2011; Creanza et al., 2012; Laland & Janik, 2006). In other words, there is good reason to think that *the kind of cultural learning* that might underwrite the purely cultural evolution of cooperation—unlike other forms of animal culture—is restricted to humans (Creanza et al., 2012).[9]

---

[9] The one widely studied case of non-human reciprocal altruism—concerning food sharing in vampire bats—is no exception to this either (Carter & Wilkinson, 2013; see also Hammerstein, 2003), as this does

This is important to note, for it is *already widely known* that human interactions have a number of unique features that make it useful to treat them separately from other cases of the evolution of cooperation: in particular, they are heavily based on symbol-driven interactions (Rosenberg, 2012). In fact, much of the literature in the human social sciences is testament to the fact that cooperation in the human realm plausibly needs to be studied, at least partly, in a different framework from other cases of evolution of the cooperation (Richerson & Boyd, 2005; Sterelny, 2012; C. M. Heyes & Frith, 2014; Rosenberg, 2012; El Mouden et al., 2014; Lewens, 2015). This is not to say that human cooperation needs to be seen to be completely divorced from biological factors; the point is just that it is widely accepted that the evolution of human cooperation needs to be seen as a special case that raises its own, unique issues.

Overall, therefore, it is plausible that in all the core biological, i.e. non-human cases, the simpler GKST will make it as easy to make sense of the evolution of cooperation (or its absence) as the more complex WKST.[10] More specifically, across all of the core biological applications of the two theories, they will agree on when we should expect the evolution of cooperation, and when not. The exceptions concern the special circumstances raised by the complex gene-culture coevolutionary relationships in the human realm; however, these special circumstances are anyway to be expected to raise different issues from the ones in the rest of biology. (I return to this point momentarily.)

---

not actually concern genuine evolutionary altruism, and is thus to be explained in a non-cooperative framework (Ramsey & Brandon, 2011).

[10] It is also worth noting that cases of mutualism do not obviously fall into the class of cases that cannot be handled well by genetic forms of GKST, as these do not obviously concern the evolution of altruistic cooperation (Wyatt et al., 2013; but see also Frank, 1994). Indeed, they may be better handled as cases where a population of organisms adapts to an environment partially constituted by other types of organisms (see also Gardner & West, 2010; Godfrey-Smith, 2009).

3. *Generalized Model Selection, GKST, and WKST*

So far, one might still think that WKST is to be preferred to GKST: all that the previous subsection has established is that there are reasons to think that Var $(q_i) = 0$ (and thus $r_q = r_{gq} = r_{qg} = 0$), and $E(w_i \Delta q_i) = 0$ in non-human cases. However, while this may be interesting, it may not be thought that this is a problem for WKST: it is just that this theory has elements in it that are not needed to capture the evolution of cooperation in non-human cases. What is wrong with a theory that is more complex than is needed in order to make sense of the data?

More specifically: why not go for the most general theory consistent with *all* of the data on cooperation—i.e. the one that applies to humans and non-human organisms? Why should we treat humans as a special case, and then formulate a separate theory—GKST— for non-human organisms only? In fact, it might seem that WKST is to be *favored* over GKST because of its greater generality. However, this is precisely where considerations from statistical model selection come in: for the latter speak against using more complex models where they can be avoided.[11] To understand this better, two points need to be noted about model selection theory.

First, there are many different kinds of model selection theory (in a broad sense), including the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), likelihood ratio tests, and cross-validation methods (Zucchini, 2000; Bretthorst, 1996; Forster & Sober, 2011; Burnham & Anderson, 2002; Schwarz, 1978; Goodman & Royall, 1988; Royall, 1997; Stone, 1974, 1977). These differ in the assumptions they are based on, their aims, and their concrete structure. While, in the present context, it is easiest to concentrate on likelihood ratio tests (Goodman & Royall, 1988; Royall, 1997;

---

[11] For a treatment of appeals to simplicity in science and philosophy more generally, see Sober (2015).

see also Abraham & Ledolter, 2006), the conclusions to follow do not depend on exact model selection framework used.

The second key point to note here is that all forms of model selection theory penalize more complex theories as compared to simpler theories (though the details for how they do this differ): more complex models need to fit the data considerably better to warrant being adopted (Hitchcock & Sober, 2004; Forster & Sober, 1994).[12] The justification for this depends on the specific model selection framework used: for example, the AIC approach is based on maximizing the expected predictive accuracy of the models under comparison, and more complex model are generally likely to be poorer at predicting new data (Forster & Sober, 1994; Zucchini, 2000; Hitchcock & Sober, 2004); by contrast, the BIC approach is based on maximizing the posterior probability of the models under comparison, and more complex models will generally have a lower posterior probability (Schwarz, 1978). In general, though, a key idea behind model selection theory is that, ceteris paribus, more complex theories are likely to get too close to the existing data points, which prevents them from successfully weeding out the noise from the signal in the data (Burnham & Anderson, 2002; Hitchcock & Sober, 2004; Forster & Sober, 2011, 1994).

While there is much more that can be said about the importance of and justifications for avoiding overly complex models (Forster & Sober, 1994; Hitchcock & Sober, 2004; Sober, 1988; Rochefort-Maranda, 2016), for present purposes, I shall just take it for granted. That is, my present argument is just premised on the fact *that* standard statistical approaches favor simple to complex models, holding goodness of fit fixed—I shall not

---

[12] Some only do so implicitly, though, by being asymptotically equivalent to AIC (Stone, 1977; Hitchcock & Sober, 2004).

consider *why* they do so. This means my argument is just as strong as the standard approaches towards statistical model selection theory; however, given the widespread acceptance of these approaches, this is only a weak restriction. I return to this point below in section IV.[13]

These points about model selection theory matter here, as they underwrite an argument for GKST over WKST. This argument can be stated very quickly.

1. GKST is simpler—has fewer degrees of freedom—than WKST.

2. However, its goodness of fit, in non-human cases, is about as high as that of WKST.

3. Model selection theoretic considerations require trading off the goodness of fit of a theory with its simplicity.

4. Overall, this therefore implies that, in non-human cases, a model selection perspective provides a reason to prefer GKST to WKST.

In other words: since the non-human data concerning the evolution of cooperation can be about equally well handled with a theory that does not allow for purely cultural clustering as a driver of this evolution, this kind of clustering should not be included in a theory of the evolution of cooperation in non-human organisms. In a bit more detail, this argument can be spelled out as follows.

---

[13] In this way, I also avoid Sober's (2002) charge that some model selectionist frameworks—like likelihoodism—lack an established epistemic foundation: this may be so, but all that I am claiming here is that, to the extent that likelihoodism is accepted, we have a reason to favor GKST over WKST. Given the wide acceptance of likelihoodism, this is still a sufficiently strong conclusion. At any rate, as will be made clearer below, the argument of this paper does not depend on the acceptance of likelihoodism. I thank xxxx for useful discussion of this issue.

Consider the currently known set of investigations about the non-human situations in which cooperation has evolved, and the situations it has not (or the situations in which it has evolved relatively easily and quickly, and the ones in which it has not): for example, concerning the "virulence" of bacteria (Brown et al., 2009), the formation of fruiting bodies in slime molds (Strassmann et al., 2011), or the frequency of warning calling in prairie dogs (Hoogland, 1983). Then ask: how many of these match what would be expected based on GKST, and how many match what would be expected based on WKST?[14] Now, by the above considerations, the answer to this question will be: about the same—that is, the success in accommodating the empirically observed cases of non-human social evolution using GKST will be as good as that using WKST. However, since model selection theory advises using the simpler of two equally empirically successful theories, this thus makes for an argument for GKST over WKST. Several further points concerning this argument need to be noted.

First, it is important to make clearer that a model selectionist framework is indeed applicable to the comparison between GKST and WKST. One concern one might have here is that since GKST and WKST can both be related to the Price equation—a mathematical theorem—they cannot be compared using a statistical framework like model selection theory.[15] However, this concern is unfounded: as noted earlier, the

---

[14] So, formally, one could define y_1=1 if and only if $(-c + b\ r_g) > 0$, y_1=0 otherwise (for GKST), and y_2=1 if and only if $-c + b\ [a\ r_g + (1\text{-}a)\ (r_q + r_{gq} + r_{qg})] > $ -d; y_2=0 otherwise (for WKST). Then one could generate a (large) data set containing all investigations of the evolution of cooperation, noting for each one the values (or at least estimates) of c, b, r_g, r_q, r_{gq}, r_{qg}, and d, and adding a variable z coded as 1 if cooperation did evolve in the case in question (or did so easily and quickly), and 0 if not. On this basis, one could then compare how well y_1 and y_2 match z. A more sophisticated version of such an approach would have y_i be graded, to model differences in the ease with which cooperation can evolve.

[15] Note also that, given the randomness inherent in evolutionary processes, there is no doubt that the comparison between GKST and WKST can be seen as a statistical inference problem to begin with. Furthermore, there is also little reason to think that the statistical properties of these evolutionary processes change from application to application (see also Forster & Sober, 1994).

derivability from the Price equation just means that the two theories are mathematically coherent—it does not speak to their empirical success. Just like we can compare, using model selection theory, an evolutionary model based just on selective pressures with one based on selective pressures and drift (Sober, 2008; Orzack & Sober, 1994), we can compare GKST to WKST: while all of these models can be seen to be mathematical trivialities if their assumptions are satisfied, *whether* their assumptions are satisfied in actual cases is an empirical question.[16]

The second point to note concerning the above argument for GKST over WKST is that it needs to be admitted that, in order to make this argument precise, we would need to compile a summary of the goodness of fit that can be achieved with GKST and WKST across a wide variety of different cases of the evolution of cooperation, and then assess this summary with a specific model selection framework (see note 14 above). This has not yet been done—nor is it likely that it will be done anytime soon, given the scale of the undertaking. However, this does not mean that the argument here laid out has no value whatsoever. In particular, I hope to have provided well supported reasons for thinking that such a model theoretic meta-analysis *would* favor GKST over WKST.

---

[16] Another concern one might have with placing the GKST / WKST comparison in a model selection framework is that these two theories contain no adjustable parameters, in the sense that is that there is no flexibility in how r_g, r_q, r_{gq}, r_{qg}, etc. are to be related to each other. However, this point does not raise major problems for the present argument either. On the one hand, likelihood ratio tests—for one—apply whether or not the two theories have adjustable parameters in this sense (Burnham & Anderson, 2002; Abraham & Ledolter, 2006). On the other hand, one can work around this issue by introducing adjustable parameters into the two theories and then applying AIC or BIC (for example). So, one could just compare (GKST*) Δg > 0 if and only if $s_1 + (- c + b\, r_g) > 0$ and (WKST*) $\Delta p > 0$ if and only if $s_1 + [-c + b\, r_g + s_2 [b [(a-1)\, r_g + (1- a)(s_3\, r_q + s_4\, r_{gq} + s_5\, r_{qg})]]] > -s_6 d$, where s_1 is a parameter capturing an overarching measurement error, and s_2, s_3, s_4, s_5, and s_6 are parameters capturing the weights that should be given to $Var(q_i)$, r_q, r_{gq} r_{qg}, and $E(w_i \Delta q_i)$ respectively. (One could further make s_2 to s_6 dummy parameters by restricting them to taking on the value of either 0 or 1.) It is then possible to let the data determine the best values of s_1 to s_6. The introduction of these parameters is furthermore made reasonable by the fact that, in actual applications of the two theories, we may need to replace population statistics (like variances and covariances) with their estimates (sample variances or covariances). If this is done, GKST* is nested in WKST*—it has $s_2 = s_3 = s_4 = s_5 = s_6 = 0$—and their comparison is a standard model selection problem (Burnham & Anderson, 2002).

While not fully resolving the debate between these two theories, therefore, I hope to have taken some steps towards such a resolution.

Third, as has been noted throughout the discussion, the above argument is restricted to non-human cases of the evolution of cooperation. In human cases, WKST may be more plausible: there, allowing for a < 1, $r_q \neq 0$, $r_{gq} \neq 0$, $r_{qg} \neq 0$, and $d \neq 0$ *is* likely to increase the fit to the data to a significant extent. However, as made clear earlier, these cases are known to raise unique issues anyway (Sterelny, 2012, 2003; Boyd & Richerson, 1985; Richerson & Boyd, 2005; Boyd & Richerson, 2005; Sober, 1992). In fact, this goes to the very heart of the present argument: when analyzing the evolution of cooperation, we have two choices. On the one hand, we can use WKST as our only theory for both non-human organisms and humans—that is, we can apply WKST across the board, even though parts of it are not needed in the non-human cases. On the other hand, we can use GKST for non-human organisms and WKST for humans. The point made here is that the latter is preferable to the former, as, in the non-human cases, GKST is likely to do a better job— statistically—than WKST, and, for independent reasons (e.g. concerning the fact that human social interactions are often symbolically driven) the human cases are known to require a different treatment anyway (El Mouden et al., 2014). The next question is what the importance is of this conclusion.

IV. The Importance of the Debate between GSKT and WKST

Why is comparing GKST and WKST important? There are two answers to this question. First and foremost, it is inherently interesting. As noted earlier, there are different ways of formulating KST (depending on which assumptions are appealed to when deriving or

otherwise setting out the theory), some of which are more complex—contain more degrees of freedom—than others. Given this divergence, there are theoretical, empirical, and practical reasons for caring about which version is actually to be adopted.

In particular, what the above discussion has shown is that (a) the simpler, genetics-focused form of KST (GKST) is mathematically just as valid as the more complex WKST (this was made clear in section II above), and (b) in core biological, non-human cases, GKST is empirically superior to WKST—in the sense that the empirical data, when combined with a model selectionist perspective, favor GKST over WKST. In turn, this also implies that, for practical purposes, in the KST-based investigation of the evolution of cooperation in non-human populations, we ought to see GKST as the major theoretical tool with which to approach these investigations. For these reasons, there can be no doubt that the comparison between GKST and WKST has much intrinsic importance.

However, there is also a second, even more wide-ranging reason for why the comparison between GKST and WKST is interesting. This reason concerns the fact that this comparison gets at an important general contrast in the study of (social) evolution: whether adopting a "gene's eye view" of (social) evolution is sufficient—or indeed preferable—to adopting a viewpoint that also includes, as major elements, non-genetic factors (Dawkins, 1989; Sterelny & Kitcher, 1988; Sober, 1990; see also Rubin, 2015).[17] Indeed, what makes WKST different from GKST is precisely the fact that the latter gives non-genetic elements only a negligible role in determining the evolution of cooperation (GKST allows non-genetic factors to underwrite the conditions for a high degree of

---

[17] This is sometimes expressed with the notion of "inclusive fitness"; however, for present purposes, the details of this are not so important. For more on this, see e.g. Maynard Smith (1976); Taylor and Frank (1996); West et al. (2007, 2008); Frank (1998); Hamilton (1964).

genetic similarity among interacting organisms, but does not directly take these non-genetic factors into account).[18]

Here it is important to note that, while it is true that the classical "gene's eye view" of evolution has been found to be overly simplified, this does not mean that a variant of this view is not still worth taking seriously. In particular, as noted earlier, it remains possible to see the difference between GKST and WKST as concerning the question of whether considering *broadly* genetic factors—however, exactly, this is to be spelled out (including reference to epigenetic factors such as DNA methylation and DNA activation that affect gene expression without altering DNA structure directly)—is sufficient to make sense of the evolution of cooperation, or whether non-genetic, purely cultural factors need to be taken into account as well. Formulated like this, this is still an important dispute that deserves to be taken seriously. For this reason, the fact that my argument provides considerations underwriting the gene's eye view of evolution (at least as far as accounting for the empirical data concerning the evolution of cooperation in non-human populations is concerned) is interesting.

However, it is also important to note that I do not want to claim that the argument presented in this paper shows the gene's eye view answers all of the questions surrounding the evolution of cooperation: after all, as has frequently been pointed out, the ability to empirically *account for* the evolution of cooperation using genetic factors alone

---

[18] In fact, one could see the comparison between GKST and WKST as related to the dispute as to whether KST, in general, or MLST, in general, should be seen as the key framework with which to approach the evolution of cooperation (for more on this dispute, see e.g. West et al., 2007, 2008; Wilson, 2008; Sober & Wilson, 1998; Okasha & Martens, 2016; Birch & Okasha, 2014; Okasha, 2015). This is due to the fact that MLST is often stated as emphasizing precisely the importance of non-genetic interactions in the explanation of the evolution of cooperation. So, for example Sober (2000, pp. 110-111), in laying out an MLST-based perspective towards the evolution of cooperation states: "the key to the evolution of altruism is population structure." In turn, this might lead one to think that, in spirit, MLST is very close to WKST (see e.g. Lehmann et al., 2007). However, laying out and justifying this argument calls for another paper.

is different from the ability to provide *an adequate causal explanation* of the evolution of cooperation (see e.g. Sober & Wilson, 1998; Okasha, 2006; Birch & Okasha, 2014; Okasha, 2015). Still, what I hope to have shown here is that, as far as statistical inference making is concerned, we should favor GKST to WKST in non-human cases. This is important, as it is just not obvious that it *must* be true that genetic factors are sufficient to account for the evolution of cooperation—the "bookkeeping argument" (Sterelny & Griffiths, 1999) should not be seen to be a trivial truth. The soundness of this argument needs to be *shown* whenever the argument is appealed to. So, while I do not want to claim that my argument has settled all the issues that need to be addressed here, I do think that it is sufficiently strong to be taken seriously.

V.   Conclusion

I have provided a model-selection based argument for a genetical version of KST— GKST—over a wider, not purely gene-focused form of KST—WKST. This argument is based on the claims that GKST is (i) simpler than WKST, and yet, (ii) in non-human populations equally empirically successful when compared to WKST. I have further shown that this argument is interesting, as, apart from the intrinsic importance of the comparison between GKST and WKST, it underwrites (without though fully establishing the truth of) a version of a "gene's eye of evolution."

Bibliography

Abraham, B., & Ledolter, J. (2006). *Introduction to Regression Modeling* (1st ed.). Independence, KY: Cengage.

Birch, J., & Marshall, J. A. (2014). Queller's Separation Condition Explained and Defended. *Am. Nat., 184*(4), 531-540.

Birch, J., & Okasha, S. (2014). Kin Selection and Its Critics. *BioScience*.

Boyd, R., & Richerson, P. (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Boyd, R., & Richerson, P. (2005). *The Origin and Evolution of Cultures*. Oxford: Oxford University Press.

Bretthorst, G. L. (1996). An Introduction to Model Selection Using Probability Theory as Logic. In G. Heidbreder (Ed.), *Maximum Entropy and Bayesian Methods* (Vol. 62, pp. 1-42): Springer Netherlands.

Brown, S. P., West, S. A., Diggle, S. P., & Griffin, A. S. (2009). Social evolution in micro-organisms and a Trojan horse approach to medical intervention strategies. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 364*(1533), 3157-3168.

Burnham, K. P., & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd ed.). New York: Springer.

Carter, G., & Wilkinson, G. (2013). Food Sharing in Vampire Bats: Reciprocal Help Predicts Donations More than Relatedness or Harassment. *Proceedings of the Royal Society B, 280*, 20122573.

Creanza, N., Fogarty, L., & Feldman, M. W. (2012). Models of Cultural Niche Construction with Selection and Assortive Matings. *PloS one, 7*(8), e42744.

Damuth, J., & Heisler, I. L. (1988). Alternative Formulations of Multi-Level Selection. *Biology and Philosophy, 3*, 407-430.

Dawkins, R. (1989). *The Selfish Gene* (2nd ed.). Oxford: Oxford University Press.

El Mouden, C., Andre, J. B., Morin, O., & Nettle, D. (2014). Cultural transmission and the evolution of human behaviour: a general approach based on the Price equation. *J Evol Biol, 27*(2), 231-241.

Forster, M., & Sober, E. (1994). How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. *British Journal for the Philosophy of Science, 45*(1), 1-35.

Forster, M., & Sober, E. (2011). AIC Scores as Evidence—A Bayesian Interpretation. In M. Forster & P. S. Bandyopadhyay (Eds.), *The Philosophy of Statistics* (pp. 535-549). Dordrecht: Kluwer.

Frank, S. A. (1994). Genetics of mutualism: the evolution of altruism between species. *J Theor Biol, 170*(4), 393-400.

Frank, S. A. (1998). *Foundations of Social Evolution*. Princeton: Princeton University Press.

Frank, S. A. (2012). Natural Selection. IV. The Price Equation. *J Evol Biol, 25*, 1002-1019.

Gardner, A., & West, S. A. (2010). Greenbeards. *Evolution, 64*(1), 25-38.

Gardner, A., West, S. A., & Wild, G. (2011). The genetical theory of kin selection. *J Evol Biol, 24*(5), 1020-1043.

Godfrey-Smith, P. (2009). *Darwinian Populations and Natural Selection* Oxford: Oxford University Press.

Goodman, S. N., & Royall, R. (1988). Evidence and Scientific Research. *American Journal of Public Health, 78*(12), 1568-1574.

Grafen, A. (1985). A Geometric View of Relatedness. *Oxford Surveys in Evolutionary Biology, 2*, 28-90.

Griffin, A. S., & West, S. A. (2002). Kin Selection: Fact and Fiction. *Trends in Ecology and Evolution, 17*, 15-21.

Hamilton, W. (1963). The evolution of altruistic behaviour. *American Naturalist, 97*, 354-356.

Hamilton, W. (1964). The Genetical Theory of Social Behavior. *J Theor Biol, 7*, 1-52.

Hammerstein, P. (2003). Why is Reciprocity so Rare in Social Animals? A Protestant Appeal. In P. Hammerstein (Ed.), *Genetic and Cultural Evolution of Cooperation* (pp. 83-94). Cambridge MA: MIT Press.

Henrich, J. (2015). *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton, NJ: Princeton University Press.

Henrich, J., & McElreath, R. (2011). The Evolution of Cultural Evolution. *Evolutionary Anthropology, 12*, 123-135.

Heyes, C., & Galef, B. G. (Eds.). (1996). *Social Learning in Animals: The Roots of Culture*. San Diego: Academic Press.

Heyes, C. M., & Frith, C. (2014). The cultural evolution of mind reading. *Science, 344*, 1243091.

Hitchcock, C., & Sober, E. (2004). Prediction Versus Accommodation and the Risk of Overfitting. *The British Journal for the Philosophy of Science, 55*(1), 1-34.

Hoogland, J. L. (1983). Nepotism and alarm calling in the black-tailed prairie dog (Cynomys ludovicianus). *Animal Behaviour, 31*(2), 472-479.

Laland, K. N., & Janik, V. M. (2006). The Animal Cultures Debate. *Trends in Ecology and Evolution, 21*, 542-547.

Lehmann, L., Keller, L., West, S. A., & Roze, D. (2007). Group selection and kin selection: two concepts but one process. *Proc. Natl. Acad. Sci. U.S.A., 104*, 6736-6739.

Lewens, T. (2015). *Cultural Evolution: Conceptual Challenges*. Oxford: Oxford University Press.

Lewontin, R. (1970). The units of selection. *Annual Review of Ecology and Systematics*, 1-18.

Luque, V. J. (forthcoming). One equation to rule them all: a philosophical analysis of the Price equation. *Biology and Philosophy*.

Marshall, J. A. R. (2011). Group selection and kin selection: formally equivalent approaches. *Trends Ecol Evol, 26*(7), 325-332.

Maynard Smith, J. (1976). Group Selection. *Q. Rev. Biol., 51*, 277-283.

Morin, O. (2016). *How Traditions Live and Die*. Oxford: Oxford University Press.

Morrison, M. (2015). *Reconstructing Reality: Models, Mathematics and Simulations*. Oxford: Oxford University Press.

Okasha, S. (2006). *Evolution and the Levels of Selection*. Oxford: Oxford University Press.

Okasha, S. (2015). The Relation Between Kin and Multi-Level Selection: An Approach Using Causal Graphs. *British Journal for the Philosophy of Science*.

Okasha, S., & Martens, J. (2016). The Causal Meaning of Hamilton's Rule. *Royal Society Open Science, 3*(160037).

Orzack, S. H., & Sober, E. (1994). Optimality Models and the Test of Adaptationism. *The American Naturalist, 143*(3), 361-380.

Oyama, S. (2000). *Evolution's Eye*. Durgham, NC: Duke University Press.

Price, G. R. (1970). Selection and covariance. *Nature, 227*, 520-521.

Price, G. R. (1972). Extension of covariance selection mathematics. *Ann Hum Genet, 35*, 485-490.

Queller, D. C. (1985). Kinship, Reciprocity and Synergism in the Evolution of Social Behavior. *Nature, 318*(28), 366-367.

Queller, D. C. (1992). Quantitative Genetics, Inclusive Fitness and Group Selection. *American Naturalist, 139*, 540-558.

Queller, D. C. (2004). Kinship is Relative. *Nature, 430*, 975-976.

Queller, D. C. (2011). Expanded Social Fitness and Hamilton's Rule for Kin, Kith, and Kind. *Proceedings of the National Academy of Sciences, 108*, 10792-10799.

Ramsey, G., & Brandon, R. (2011). Why reciprocal altruism is not a kind of group selection. *Biology and Philosophy, 26*, 385-400.

Reader, S. M., & Biro, D. (2010). Experimental Identification of Social Learning in Wild Animals. *Learning Behavior, 38*(3), 265-283.

Rice, S. H. (2004). *Evolutionary theory: mathematical and conceptual foundations*. Sunderland, MA: Sinauer Associates.

Richerson, P., & Boyd, R. (2005). *Not By Genes Alone*. Chicago: University of Chicago Press.

Rochefort-Maranda, G. (2016). Simplicity and model selection. *European Journal for Philosophy of Science, 6*(2), 261-279.

Rosenberg, A. (2012). *Philosophy of Social Science* (4th ed.). Boulder, CO: Westview Press.

Royall, R. (1997). *Statistical Evidence – A Likelihood Paradigm*. Boca Raton, FL: Chapman and Hall.

Rubin, H. (2015). Genetic Models in Evolutionary Game Theory: The Evolution of Altruism. *Erkenntnis, 80*(6), 1175-1189.

Sarkar, S. (2005). *Molecular Models of Life*. Cambridge, MA: MIT Press.

Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics, 6*, 461-465.

Skyrms, B. (2010). *Signals: Evolution, Learning, and Information*. Oxford: Oxford University Press.

Sober, E. (1988). *Reconstructing the Past: Parsimony, Evolution, and Inference*. Cambridge, MA: MIT Press.

Sober, E. (1990). The Poverty of Pluralism: A Reply to Sterelny and Kitcher. *The Journal of Philosophy, 87*(3), 151-158.

Sober, E. (1992). Models of Cultural Evolution. In P. Griffiths (Ed.), *Trees of Life* (Vol. 11, pp. 17-39): Springer Netherlands.

Sober, E. (2000). *Philosophy of Biology* (2nd ed.). Boulder, CO: Westview Press.

Sober, E. (2002). Bayesianism—Its Scope and Limits. In R. Swinburne (Ed.), *Bayes's Theorem* (pp. 21-38). Oxford: Oxford University Press.

Sober, E. (2008). *Evidence and Evolution*. Cambridge: Cambridge University Press.

Sober, E. (2015). *Ockham's Razors: A User's Manual*. Cambridge: Cambridge University Press.

Sober, E., & Wilson, D. S. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.

Sterelny, K. (2003). *Thought in a Hostile World: The Evolution of Human Cognition*. Oxford: Wiley-Blackwell.

Sterelny, K. (2012). *The Evolved Apprentice: How Evolution Made Humans Unique*. Cambridge, MA: MIT Press.

Sterelny, K., & Griffiths, P. (1999). *Sex and Death*. Chicago: University of Chicago Press.

Sterelny, K., & Kitcher, P. (1988). The Return of the Gene. *The Journal of Philosophy, 85*(7), 339-361.

Stone, M. (1974). Cross-Validictory Choice and Assessment of Statistical Predictions (with Diosciussioon) *Journal of the Royal Statistical Society B, 36*, 111-147.

Stone, M. (1977). An Asumptotic Equivalence of Choice of Model by Cross-Validation and Akaike's Criterion. *Journal of the Royal Statistical Society B, 39*, 44-47.

Strassmann, J. E., Gilbert, O. M., & Queller, D. C. (2011). Kin discrimination and cooperation in microbes. *Annu Rev Microbiol, 65*, 349-367.

Taylor, P. D., & Frank, S. A. (1996). How to Make a Kin Selection Model. *J. Theor. Biol., 180*, 27-37.

Weisberg, M. (2013). *Simulation and Similarity: Using Models to Understand the World*. Oxford: Oxford University Press.

West, S. A., El Mouden, C., & Gardner, A. (2011). Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and Human Behavior, 32*(4), 231-262.

West, S. A., Griffin, A. S., & Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J Evol Biol, 20*(2), 415-432.

West, S. A., Griffin, A. S., & Gardner, A. (2008). Social semantics: how useful has group selection been? *J Evol Biol, 21*(1), 374-385.

Wilson, D. S. (2008). Social semantics: toward a genuine pluralism in the study of social behaviour. *J Evol Biol, 21*(1), 368-373.

Wilson, D. S., & Dugatkin, L. (1992). Altruism: Contemporary Debates. In E. Fox Keller & E. A. Lloyd (Eds.), *Keywords in Evolutionary Biology* (pp. 29-33). Cambridge, MA: Harvard University Press.

Wyatt, G. A., West, S. A., & Gardner, A. (2013). Can natural selection favour altruism between species? *J Evol Biol, 26*(9), 1854-1865.

Zucchini, W. (2000). An Introduction to Model Selection. *Journal of Mathematical Psychology, 44*, 41-61.