This is a *preprint* version of a paper forthcoming in *Biology and Philosophy*. The citable version can be found here: http://link.springer.com/article/10.1007/s10539-016-9521-7

# Does Intragenomic conflict predict Intrapersonal Conflict?

*David Spurrett*

*Philosophy, UKZN, Howard College Campus, 4041, SOUTH AFRICA*
*e-mail: spurrett@ukzn.ac.za*

**Abstract**

Parts of the genome of a single individual can have conflicting interests, depending on which parent they were inherited from. One mechanism by which these conflicts are expressed in some taxa, including mammals, is genomic imprinting, which modulates the level of expression of some genes depending on their parent of origin. Imprinted gene expression is known to affect body size, brain size, and the relative development of various tissues in mammals. A high fraction of imprinted gene expression occurs in the brain.

Biologists including Hamilton, Trivers and Haig have proposed that this may explain some intrapersonal conflict in humans. This speculation amounts to an inference from conflict within the genome (which is well-established) to conflict within the brain or mind. This is a provocative proposal, which deserves serious attention. In this paper I assess aspects of Haig's version of the proposal. I argue, first, that the notion that intragenomic conflict predicts personal inconsistency should be rejected. Second, while it is unlikely that it credibly predicts sub-personal agents representing conflicting genetic interests, it is plausible that it predicts that the division of cognitive labour could be exploited to turn sub-systems into proxies for conflicting interests.

**Keywords:** Intragenomic conflict; Intrapersonal Conflict; Genomic Imprinting; Parent-specific gene-expression; Behaviour

**Acknowledgments:** A version of this paper was presented at the 2013 conference of the Australasian Association of Philosophy. I am grateful to the audience on that occasion, as well as to David Haig, Daniel Dennett, Don Ross, Mariam Thalos, JP Smit, John Collier and Blaize Kaye for feedback on various versions of the text. Two reviewers for this journal also provided very helpful comments and suggestions.

# 1. Introduction

In a striking passage in his post-script to a collection on kin recognition W.D. Hamilton identifies what he says is perhaps the 'most interesting' consequence of intragenomic conflict, a point that he calls 'philosophical':

> We see that we are not even in principle the consistent wholes that some schools of philosophy would have us be. Perhaps this is some comfort when we face agonizing decisions, when we cannot 'make sense' of the decisions we do make, when the bitterness of a civil war seems to be breaking out in our inmost heart. (1987, p. 426)

Hamilton's remark is passing and self-consciously speculative. He is not alone in making such a suggestion, though. Robert Trivers has said more than once that intragenomic conflict implies that we "literally have a paternal self and a maternal self and they are often in conflict" (2009, p. 163). Perhaps the most detailed development of the proposition that genomic conflicts "might be expressed within the mind" is to be found in David Haig's essay 'Intrapersonal Conflict' (2006, p. 20). Daniel Dennett has recognised this aspect of Haig's thought, and referred to his "lovely papers on intrapersonal conflicts" arguing that if maternally inherited and paternally inherited genes "get out of whack, serious imbalances can happen that show up as particular psychological anomalies" (Dennett 2013).

These remarks suggest a link between intragenomic conflict expressed through parent-

specific gene-expression and intrapersonal conflict. Parent-specific gene-expression occurs when the level of expression of a gene is contingent on whether it was inherited from a male or female parent. This non-Mendelian effect has been extensively confirmed in flowering plants and placental mammals. Its most well-documented manifestations in mammals are modulations of foetal development and brain development. In the case of foetal development, as with seed development, there is fairly obvious scope for resource competition between the developing young and the resource-providing parent. The leading explanation for the phenomenon of parent-specific gene-expression — the kinship theory of genomic imprinting — is an application of Trivers' account of parent-offspring (and inter-sibling) conflict (Trivers 1974). This explanation, described below, relates patterns of parent-specific gene expression to kinship-based asymmetries of interest between genes.

Perhaps the best known position relating intragenomic conflict to human psychology is Crespi and Badcock's (2008) proposal that autistic- and psychotic-spectrum conditions are respectively to be explained by reference to the excessive influence of paternal and maternal genetic contributions. The issues in play here are, however, for the most part distinct: Hamilton, Trivers and Haig aren't primarily concerned with autism or psychosis. Crespi and Badcock for their part aren't especially interested in intrapersonal conflict, even though they are interested in the psychological effects of intragenomic conflict. My own concern is rather narrow: I am specifically interested in testing the plausibility of the provocative suggestion, relatively neglected by philosophers, that intragenomic conflict predicts (some) intrapersonal conflict. This means that various candidate explanations of intrapersonal conflict in which intragenomic conflict plays no role are also not directly relevant here either. Nobody is advancing, nor am I assessing,

the claim that intragenomic conflict is the *only* possible explanation for intrapersonal conflict.

Because he has considered the question in the most detail, my treatment will focus on David Haig. I begin by selecting and clarifying some features of intrapersonal conflict that offer themselves as candidates for explanation by reference to intragenomic conflict (Section 2). Then I describe the phenomenon of parent-specific gene-expression, sketch the kinship theory of genomic imprinting, and say a little about the mechanism of genomic imprinting (Section 3). Finally I consider whether intragenomic conflict indeed predicts intrapersonal conflict (Section 4). I argue, first, that the notion that it predicts intrapersonal inconsistency should be rejected. Second, while it is unlikely that it credibly predicts sub-personal agents representing conflicting genetic interests, it is plausible that it predicts that the division of cognitive labour could be exploited to turn sub-systems into proxies for conflicting interests.

# 2. Intrapersonal conflict

Although Hamilton, Trivers and Haig suggest a *relationship* between intragenomic and intrapersonal conflict, they don't say quite the same things about intrapersonal conflict. They make a range of references to metaphorical 'civil wars', to 'factions', and to multiplied or divided selves, with differently imagined capacities for expressing conflict, and mostly in the absence of a determinate notion of a person. Their remarks are also often self-consciously tentative. I will not, consequently, attempt to treat their suggestions comprehensively. Rather, I extract a few recurring features of their remarks on the phenotype of intrapersonal conflict, and confine myself to assessing the case for

a relationship between intragenomic conflict and *two* of them:

First, it or its effects are often described as *introspectible*. Hamilton (1987) refers to 'agonizing decisions', the 'bitterness of civil war', and to being unable to 'make sense' of our decisions. Haig focuses on subjectively effortful choices (2006, pp. 8-9), arguing that they are symptomatic of intrapersonal conflict, and of there not being a common currency in which the values of options are represented. Elsewhere he refers to intrapersonal persuasion (2003). Trivers sometimes refers to the representatives of different parts of the genome as 'internal voices' (1997).

Second, they suggest that intrapersonal conflict manifests in a kind of *inconsistency*. We see this when Haig draws attention to the challenge of explaining why we "often find it hard to make decisions and stick to them" (2006, p. 9). He also rejects the suggestion that we would "evolve a consistent set of genetic biases" as presupposing that we are not "also subject to conflict among genes" (2006, p. 15). And Hamilton, as noted, contends that intragenomic conflict shows that we are not "consistent wholes" (1987, p. 426), as well as showing that different parts of the genome might contribute to "contradictory" strategies in the same individual (1987, p. 425).

Third, the intrapersonal conflict is described as taking place between reasonably enduring *sub-personal agents* the interests of which are distinct from the perspective of kinship, and which (perhaps) act strategically towards one another. Hamilton's references to civil war clearly suggest this, as does Trivers' repeated assertion that we "literally have a paternal self and a maternal self and they are often in conflict" (e.g. 2009, p. 163), and may indeed act deceptively towards one another (e.g. 2000, p. 115).

Haig has suggested in a variety of ways that intragenomic conflict provides a reason to take seriously the notion that "we can be at war with ourselves" (2006, p. 9), and that the 'self' is divided, or multiplied (2008b).

I won't say more here about the *experience* of intrapersonal conflict. Interesting though that topic might be, there's no reason to suppose that *all* intrapersonal conflict is introspectible and ample reason for caution regarding the reliability of introspection as a guide to sub-personal psychology. I focus instead on the second and third features, inconsistency and sub-personal agents. These are roughly complementary in so far as inconsistency in the behaviour or expressed preferences of a single person can be understood as a predictable product of the interaction of sub-personal agents with conflicting interests. This kind of explanation is, of course, a familiar one, and the catalogue of conflicting sub-personal agents that have been entertained is large, including 'reason' and 'passion', demonic stowaways, and split personalities. (Ainslie 1992, Chapter 2, critically discusses divided agent accounts of ambivalence.) A common theme in many divided agent models is that the whole person's choices are understood as the product of two (or more) sub-personal agents, with *inter*-individual differences partly explicable by reference to the relative 'strength' of each, and *intra*-individual differences, i.e. inconsistency, partly explicable by reference to fluctuations in which sub-personal agent has the upper hand at any particular time. My present point does not require my defending or repairing any particular divided self model, but merely noting their existence as a common template for explanations of inconsistency. What is distinctive about the claim I assess here is the notion that the interests of at least some sub-personal agents diverge *because of* the asymmetric relatedness between parts of one genome and those of other individuals.

# 3. Intragenomic conflict and genomic imprinting

I begin with some distinctions. *Parent-specific gene-expression* (PSGE) refers to the phenomenon that the level of expression of some genes is contingent on whether the gene was inherited from a male or female parent. *Genomic imprinting* is a mechanism that produces and explains the phenomenon of parent-specific gene-expression. This mechanism is incompletely understood, and appears to be heterogeneous, but the basic principle is that markers introduced in the gametes modify the expression of some genes in offspring, until being reset in their gametes. *Intragenomic conflict*, or the kinship theory of genomic imprinting, is a theory that predicts or explains PSGE from the perspective of the differential relatedness of, and hence conflicting interests among, *parts* of an individual genome to other individuals. Genomic imprinting, then, is a proximal mechanism for a phenomenon, PSGE, for which the kinship theory of genomic imprinting offers an ultimate rationale. These distinctions are not always drawn in quite the same way, even in important papers in the literature, and 'genomic imprinting' is sometimes used as an umbrella term for parent-specific gene expression, genomic imprinting, and the kinship theory.[1]

The 'fit' between the kinship theory of genomic imprinting, parent-specific gene expression and imprinting is good but not perfect. Some known instances of PSGE or imprinting currently lack plausible explanation in terms of fractional kinship (Burt and

[1] The important early paper by Haig and Westoby (1989) did not use the term 'genomic imprinting', but referred to 'parent-specific gene expression'. In a commentary on that paper Haig (2002) discusses some ambiguities in the term 'genomic imprinting'. The term 'kinship theory of genomic imprinting' is due to Trivers and Burt (1999).

Trivers 2006, p. 101ff) and it is possible that some of them never will be explained in those terms (Haig 2002, p. 19). It is also possible that not all PSGE depends on imprinting, as long as other mechanisms can achieve the same effect. Finally, the kinship theory of genomic imprinting may predict cases where there are gains to be had from parent-specific expression, but for some reason the process of selection cannot find (or has not yet found) means of achieving them that pay more than they cost. That the fit is imperfect does not matter for what follows. PSGE, and imprinting, are well-established in humans. A significant fraction of both is *better* explained by the kinship theory of genomic imprinting than available alternatives (Burt and Trivers 2006; Haig 2004). The kinship theory of genomic imprinting is not required to explain *all* parent-specific gene expression before we can ask whether it predicts personal inconsistency or conflicting sub-personal agents.

### 3.1. Intragenomic conflict

Consider reproduction from the perspective of autosomal genes in a placental mammal.[2] Typically genes are present in duplicate, with one allele inherited from each parent; and both alleles are transcribed equally, and indifferently to parental origin. (This is Mendelian received wisdom — even without knowledge of the mechanisms of inheritance, Mendel's research led him to maintain that the *action* of inherited 'factors' is independent of history, including parental origin.) Since autosomal genes have an equal chance of being transmitted to the haploid gametes, it is standardly supposed that their interests coincide entirely. On this view — what Trivers has called the 'phenotype

[2] I generally won't repeat 'autosomal' or 'mammal' but both are implicit throughout. The points made here don't apply, or apply only with modifications, in the cases of different taxa or genetic systems, or other genes, including mitochondria and sex chromosomes (Burt and Trivers 2006, p. 133ff).

view' (2009, p. 167) — genes are equal shareholders in whatever reproductive success the whole individual enjoys, and there is neither prospect of benefit from conflict between genes within a single genome, nor a mechanism for expressing conflict.

Against this standard view, although it is is undoubtedly *largely* correct (that is, for the majority of genes or, as an approximation, for the whole organism), it is easy enough to specify cases in which the *interests* of genes do not entirely coincide. To do this, instead of focusing on the inclusive fitness of the whole individual, we should compare the inclusive fitness of *parts* of the individual genome, as distinguished by parent of origin.

Consider a developing foetus, sharing half of its genes with the mother in whose body it resides, and half with its father. The genes that the foetus shares with its mother have a greater common interest in the mother's future reproductive opportunities (possibly with different fathers) than genes that it shares with its father. There is thus a *partial* asymmetry of interest in the level at which a foetus extracts nutritional resources from the mother, with paternally inherited genes favouring a greater burden than maternally inherited ones (Moore and Haig 1991; Haig 2002, p. 53). Similar considerations apply with respect to partial siblings. Haig (2003) asks us to consider Bob (named in honour of Robert Trivers) who has both a paternal and a maternal half-sibling (called Paddy and Maddy respectively). Bob's maternally derived genes have a one in two chance of being shared with Maddy and so would favour benefiting Maddy as long as the benefit to Maddy ($B$) exceeded twice the cost ($C$) to Bob ($B>2C$). Since Bob's paternally derived genes are absent from Maddy, no benefit to Maddy would justify *any* cost to Bob. So Bob's maternally derived and paternally derived genomes "are in conflict whenever $B > 2C > 0$." (Haig 2003, p. 419). Corresponding considerations apply to

Bob's dealings with Paddy. These asymmetries of interest within the genome of a single individual are reason enough to reject the simple 'equal shareholder' view described above. The kinship theory of genomic imprinting can be, and has been, further elaborated (e.g. Burt and Trivers 1998), but the points I have brought forward suffice for present purposes.

A likely rejoinder is to maintain out that this is all idle speculation as genes don't 'know' which parent they are from, and couldn't 'do' anything different even if they did. If genes didn't ever 'know' what parent they were inherited from, then they'd be operating behind a genetic "veil of ignorance" and, as Haig notes, regarding the case of Bob, "an uninformed gene has one chance in four of being present in Maddy and would favor transfer of the benefit if $B > 4C$" (Haig 2003, p. 419), which is to say that assuming uninformed genes leads to the 'phenotype view'. The analysis of conflicting interest provided by the kinship theory can be understood, though, as making a *conditional prediction*: if there are marginal gains to be had for fractions of the genome from expressing these conflicts, together with means to achieve them that are accessible to natural selection, then selection can be expected to exploit them.

### 3.2. Imprinting and parent-specific gene expression.

Some genes behave as though they 'know' what parent they're from, in the sense that how much the gene is expressed in an individual is conditional on whether it was inherited from a male or female parent. This biased expression depends on the molecular markers referred to as genomic imprints, and can take the form of mono-

allelic expression, where one gene at a particular locus is entirely silenced.[3] The differences in level of expression are often specific to the type of cell in which transcription occurs.[4] Genomic imprinting occurs during gametogenesis, and the marking persists in the somatic cells of the embryo and adult, but is reset in the gametes of the next generation.[5] Imprinting is a marker of the sex of the *most recent* parent of origin. It is not primarily related to the sex of the offspring, but some parent-specific gene expression is contingent on the sex of the individual (see, e.g., Gregg et al. (2010) on imprinted gene action in mice). Other kinds of marking are easy to imagine — for example, indicators of grandparental origin, parental age or number of previous births — and possibly occur, but are not relevant here. On the other hand, as I explain shortly, some known examples of imprinted genes display the very properties predicted by the analysis of asymmetric interest provided by the kinship theory.

One source of insight into the effects of imprinting and parent-specific gene-expression is provided by experiments in which pronuclei (genetic material from the gametes, *after* imprints would have been set) are transplanted into egg cells. It has been found that mouse embryos with entirely male-derived (androgenetic) or entirely female-derived (partheno-/gyno-genetic) pronuclei exhibit what is charmingly called 'embryonic lethality' (McGrath and Solter 1984). By contrast, chimeric embryos, made by aggregating cells containing pronuclei transplanted from male and female gametes in varying ratios with 'normal' embryonic cells, have been found to be viable. The

[3] Mono-allelic expression carries costs, including exposure to harmful recessive traits, and the effective absence of a 'backup' allele in cases of disadvantageous mutation on the other.
[4] The number of imprinted genes in humans is not yet definitively known, but recent estimates cluster around values between 100 and 1000. Just as regular gene expression can be contingent on factors including cell type and developmental stage, so can imprinted gene expression.
[5] This consistency of effect is accompanied by fairly heterogeneous mechanisms of imprinting. See Burt and Trivers (2006).

development of embryos with relatively more paternal or maternal genetic contribution can then be compared to each other, and to unbiased cases (Barton et al. 1984). When markers are added to the genetic material it is possible to determine the relative contribution of paternally and maternally derived cells to different tissues at different stages. Keverne et al. (1996) found that the more parthenogenetic (mother-biased) mouse chimeras had relatively smaller bodies than wildtype controls, and that androgenetic (father-biased) chimeras had relatively larger bodies. Despite their smaller body sizes parthenogenetic chimeras had larger brains than controls while androgenetic chimeras had smaller brains despite their larger body sizes. Imprinted genes were, furthermore, not uniformly expressed in the brain, with androgenetic cells relatively more represented in the hypothalamic structures but *not* in the cortex. A converse pattern was found for parthenogenetic cells, which made significant contributions to the cortex (especially frontal) and to the striatum and hippocampus but relatively little to the hypothalamus.[6]

As previously noted, paternal interests favour greater resource extraction from the mother than maternal interests. Relative body size is a clear indication of different levels of resource extraction. So the relatively larger bodies of androgenetic chimeras, and the relatively smaller ones of parthenogenetic chimeras straightforwardly fit the conflict hypothesis. The effects on brain development are less easy to interpret. We need a theory relating maternal and paternal genetic interest to brain size, or the relative sizes of specific neural tissues, before we can begin to make sense of the experimental

[6] For a more recent and detailed analysis of imprinted gene effects on mouse brain development see Gregg et al. (2010).

results.[7]

A paradigm case of imprinted gene expression in mice that is well understood (though not primarily involving the brain) concerns the genes *Igf2*, which is paternally expressed, and *Igf2r,* which is maternally expressed. (In primates *Igf2r* is not imprinted, and its expression is biallelic.) *Igf2* is active in most foetal tissues, and produces Insulin-like growth factor 2 (IGF2), which promotes growth by stimulating cell division. *Igf2r*, on the other hand, produces IGF2R which binds to and degrades IGF2. In cases where *Igf2* is inactive, body size is reduced. If, atypically, two copies of *Igf2* are active because the maternally inherited copy is silenced, body size is increased. Where both *Igf2* and *Igf2r* are active, their effects largely cancel each other out. As Burt and Trivers say, two "oppositely imprinted genes … with strong counteracting effects on growth that cancel out to give little or no net effect" is "precisely what is expected of internal genetic conflict based on evenly matched paternal and maternal alleles" (2006, p. 101).

Another striking instance, though more difficult to interpret, is provided by Prader-Willi syndrome (PWS) and Angelman syndrome. The syndromes are caused by reciprocal mutations in the same cluster of genes; these can occur in a variety of ways, the common consequences being that in PWS the relative expression of some maternally inherited genes is increased, while in Angelman the imbalance favours paternally inherited genes.[8] PWS infants show low birth weight and (before weaning) weak crying,

[7] There are some candidate theories in the human case, for example, Crespi and Badcock's (2008), which proposes that maternal genetic interests favour 'mentalising' cognition while paternal ones favour 'mechanising' cognition, and that these capacities asymmetrically depend on different brain tissues. See section (4.2) below.

[8] In the PWS case, deletion of the paternally inherited copy of a region of DNA, maternal uniparental disomy for the gene, or mutation that affects the imprinting mechanism responsible for gene silencing produce the same effects. Angelman syndrome is produced by the opposite

poor suckling, sleepiness and low activity, among other characteristics. *After* weaning PWS is associated with hyperphagia and obesity. Angelman infants on the other hand tend to show higher birth weight, and exhibit excessive smiling and laughter, hyperactivity, and frequent waking. Both syndromes are also associated with a variety of cognitive deficits. There is ongoing debate over *interpretation* of these phenotypes from the perspective of the kinship theory of genomic imprinting (see, *inter alia*, Úbeda and Haig 2003, Crespi and Badcock 2008, Haig 2009).

These are, from the perspective of the 'phenotype view', surprising results. Although much remains unknown about the machinery and consequences of genomic imprinting, the kinship theory of genomic imprinting itself, as well as the phenomena of parent-specific gene expression and genomic imprinting are well established. What follows does not depend upon how good the evidence in favour of them is, but rather on what their implications *might* be for intrapersonal conflict, and more specifically for the propositions that parent-specific gene expression predicts individual inconsistency and conflict among sub-personal agents along lines determined by fractional kinship.

# 4. Assessing the intragenomic-intrapersonal conflict connection

To recapitulate, in some cases the links between the kinship theory of genomic imprinting and the effects of parent-specific gene expression mediated by imprinting are fairly clear and direct. With respect to embryonic body size the phenotypic

---

conditions, i.e. maternal deletion, paternal uniparental disomy, or corresponding mutation of the imprinting mechanism (Cassidy et al. 2000).

consequences of parent-specific gene expression make sense from the standpoint of the kinship theory, *and* at least some of the genes responsible for variations in size are known to be imprinted. The results concerning brain development, though demonstrably involving parent-specific gene expression, do not straightforwardly suggest a conflict-based interpretation. Burt and Trivers say of Keverne et al.'s findings that the "meaning of these facts is mostly obscure" (2006, p. 130), and Haig that "The phenomenon and the genes involved are still too poorly known to make anything other than wild speculations" (2006, p. 19). I'm hoping, rather, to identify and assess disciplined speculations.

As noted above (in section 2), I'm restricting my attention to the suggestions that intragenomic conflict might explain inconsistency in the behaviour of a whole person, or point to the existence of warring sub-personal agents corresponding to the divergent interests of fractions of the genome. (In doing so, I'm setting aside questions about the introspectible aspects of intrapersonal conflict.) The questions corresponding to the suggestions at hand may be formulated as follows:

- Does the kinship theory of genomic imprinting provide an adaptive explanation for inconsistency?[9]
- Might discriminating gene expression lead to the creation, or explain the existence, of sub-personal agents representing divergent genetic interests?

The two questions are independent, in the sense that either could be answered while the other remains unresolved. We could have an adaptive explanation for inconsistency, but

[9] That is, of course, adaptive from the perspective of *parts* of a genome, even if carrying a net cost for the whole.

lack an account of how discriminating gene expression could produce it. (It's possible, that is, that selection hasn't, or can't, find a way to produce traits which we can determine would be adaptive if produced.) Conversely, an explanation of how discriminating gene expression might produce sub-personal agents would be of interest even if it didn't predict individual inconsistency.

Haig's expository starting point in his (2006) essay is not the kinship theory of genomic imprinting itself, but the phenomenon that some choices seem to require effort. After referring to some remarks of William James (1983[1890]) on the topic, Haig observes that effortful, or conflicted, choice can seem baffling from a simplistic evolutionary perspective:

> A fitness maximizing computer would simply calculate the expected utilities of the different alternatives and then choose the alternative with the highest score. Why should some kinds of decisions be more difficult to make than others? Is the subjective experience of effort merely a measure of the computational complexity of a problem, or is something else going on? (2006, p. 9)

Apparent inner conflict is *prima facie* maladaptive compared to the ideal of a fitness maximising computer. Haig suggests three (possibly) complementary explanations for inner conflict. He allows that some conflict may be apparent rather than real. In addition, some plausibly arises from "constraints on the perfection of adaptation", by which he means considerations including trade-offs with other design goals, and diminishing marginal returns from improvements to a trait. Finally, some conflict may be genuine, and reflect "disagreement over ultimate ends between different agents that

contribute to mental activity" (2006, pp. 9-10).

The critical foil that Haig describes here — the 'fitness maximizing computer' — has two important features. The first is a view about consistency in *behaviour*. If some agent is produced by natural selection, we might expect that its behaviour allocation would tend to track the contributions to the fitness of the whole agent of the various behaviours in its repertoire. (In an ideal limit case — assuming no counteracting factors or trade-offs — behaviour allocation would *optimise* contribution to fitness.) The view that natural selection favours a certain kind of consistency — in promoting individual fitness — in behaviour allocation is by no means novel. A well-known formulation of the project of behavioural ecology goes as follows:

> Any attempt to understand behavior in terms of the evolutionary advantage that it might confer has to find a "common currency" for comparing the costs and benefits of various alternative courses of action. (McNamara and Houston 1986, p. 358)

McNamara and Houston describe the methodological ideal that the behavioural ecologist should seek to convert the varied costs and benefits of all behaviours, actual and possible, into a representation of their consequences for fitness, in other words to interpret behaviour from the perspective of a fitness maximising computer.[10] This

[10] McNamara and Houston go on to argue that behavioural ecologists have mostly achieved more modest successes relating behaviour to domain-specific currencies such as rate of calorie intake in foraging. The notion of a unidimensional common currency follows fairly directly from the link to fitness. As Maynard-Smith put it: "Paradoxically it has turned out that game theory […] is more readily applied to biology than to the field of economic behavior for which it was originally designed … the theory requires that the value of different outcomes […] might be measured on a single scale. In human applications this measure is provided by 'utility' – a

commitment is not a view about *how* behaviour is produced or controlled; for many purposes behavioural ecologists are agnostic about that and are content to fit behaviour to some proxy for fitness, for example, in optimal foraging theory relating different strategies to rate of calorie intake.

Haig's critical foil, though, is *not* agnostic about the mechanisms by which behaviours are selected. This brings us to its second feature, which is the view that behaviour is selected by a process that computes "expected utilities". Claims about the representation of value are often taken to follow from claims that behaviour maximises some outcome, on the hypothesis that only a system representing values could achieve behaviour allocation sensitive to trade-offs between options with varying returns. Whether this really is a consequence is not directly at issue here (but see Spurrett 2014). That it *does* follow is the view defended in, for example, Shizgal and Conover:

> In natural settings, the goals competing for behavior are complex, multidimensional objects and outcomes. Yet, for orderly choice to be possible, the utility of all competing resources must be represented on a single, common dimension. (Shizgal and Conover 1996, pp. 37-38)

Shizgal and Conover's claim is that behaviour that is sensitive to tradeoffs between rewards in different modalities and of varying magnitudes — this is what they mean by 'orderly choice' — requires that the rewards be internally represented in some general and unified way. If, the thought goes, they are so represented, then behaviour allocation can be made orderly.

---

somewhat artificial and uncomfortable concept: In biology, Darwinian fitness provides a natural and genuinely one-dimensional scale" (1982, quoted in Glimcher 2002, p. 323).

The two commitments together — consistency in tracking fitness *plus* behaviour selection involving calculation of fitness returns — amount to a strong version of what Trivers calls the 'phenotype view' for organisms capable of behaviour, and so form an unsurprising target for Haig's purposes. Haig can be understood, then, as seeking to undermine *both* the consistency commitments found in McNamara and Houston, and the view of process described by Shizgal and Conover. In the first case intragenomic conflict undermines the homogeneity of interest supposed by standard behavioural ecology, and in the second place the expression of conflicting genetic interests undermines unity in the process of behaviour selection. I consider each line of thinking in turn.

### 4.1. An adaptive rationale for inconsistency?

Perhaps there is a quite simple argument from intragenomic conflict to inconsistency in behaviour. According to the 'phenotype view', genes are equal shareholders in the reproductive success of the individual they occupy, and so have coinciding interests in how that individual behaves. To the extent that behavioural dispositions have a genetic basis, the genes should concur on what those dispositions will be. And to the extent that these dispositions are under selection pressure, they will tend to track fitness. The overall set of inherited dispositions will be some sort of harmonious net effect of the contributions of all of the genes. But, the counterargument goes, the kinship theory of genomic imprinting shows that the interests of genes *don't* entirely coincide, and genomic imprinting provides a mechanism (in some species) by which at least some of these conflicts can be expressed. Rather than reflecting the net effect of a collection of

*coinciding* interests, therefore, inherited behavioural tendencies will be the product of partly *conflicting* interests, and so will tend, to some extent, to result in inconsistency. Where the phenotype view reasons from shared interests on the part of genes to a consistent set of behavioural dispositions, the conflict view reasons from contention among the genes to a *conflicted* set of behavioural dispositions. That genomic conflict *complicates* the standard picture is what Haig is driving at when he points out that although "behavioural ecologists recognize the possibility of conflicts between individuals, they usually assume that individuals have well-defined, unitary interests" (Haig 2008a, p. 209).

Three considerations favour Haig's line of argument. *First*, there's nothing special about behaviour, that makes it harder for selection to work on behavioural dispositions (to the extent that they have a genetic basis) than on other heritable characteristics. *Second,* and more specifically, behaviour can be a *more* promising target for the expression of some conflict than differential physical development. Conflict expressed inside the mother's body stops with birth, and only conveniently allows expression of conflict between embryo (or genetic interests within it) and mother. For most of the lifespan, and for conflict (or co-operation) with individuals other than the mother, discriminating behaviour is a more promising avenue of expression (Burt and Trivers 2006, pp. 124-5; Isles, Davies and Wilkinson 2006; Úbeda and Gardner 2010, 2011). Any behaviour favoured by inclusive fitness understood according to the 'phenotype view' is, accordingly, a potential candidate for *modulated* or *discriminating* expression sensitive to fractional relatedness. (Hamilton's remarks on 'civil war' quoted in my Introduction occur, one should recall, in the context of a post-script to a collection on kin recognition.)

The *third* point favouring the simple suggestion made above is even more specific. Imprinted gene activity *does* have some behavioural effects. As noted above, a substantial fraction of established PSGE in mammals is focussed on the brain. (Genes related to foetal development and genes related to brain development form the two largest groups of known imprinted genes.) As well as extreme cases such as Prader-Willi and Angelman syndromes (see section 3.2 above), there are credible examples of behavioural phenomena which evince imprinted activity, including aspects of maternal caring behaviour, infant suckling, and, perhaps, infant crying (Isles, Davies and Wilkinson 2006; Úbeda and Gardner 2010; Haig 2014). There are, furthermore, some predictions linking behaviour in adulthood to imprinted gene activity (e.g. Úbeda and Gardner 2011). For present purposes let us *assume* that there are indeed such cases.

These sources of encouragement are, however, insufficient to the task at hand. They are a justification for making the prediction (already partly confirmed) that *if* there are parent-of-origin specific gains to be had from modulations of behavioural dispositions, *then* selection will favour them as long as it finds a mechanism. Accordingly, if genomic imprinting enables parent-specific gene expression to achieve the modulations in question, then the genes in question are *prima facie* candidates for imprinting, including counter-imprinting or imprinting of genes with counteracting effects. That is interesting and probably important, but it doesn't get us to *inconsistency*.

Recall the chimeric mouse embryos constructed by Keverne *et al*. The body sizes of the chimeras aren't independent of the manipulations in the relative representation of maternally-derived and paternally-derived genes. One conclusion we should draw is that

the size of any particular embryo at a time isn't *simply* the effect of a genome given a nutritional budget and other relevant resources, but that the expression of the genome is *also* partly dependent on contingencies, expressed through imprinting, in the parent of origin for some genes. To the extent that it is meaningful to specify an embryonic body size at a time that represents the corporate interests of the whole genome, imprinting might result in an actual body size that is different. We might even say that the actual size is *inconsistent* with the 'corporate interest' size. For all that, the size of any particular embryo at a particular time is as determinate as body size ever is. It may be different to a supposed imprinting-free case, but it is what it is.

Similarly, the behavioural dispositions of an organism in which imprinting is active should reflect the *net effect* of the imprinted genome (plus development in interaction with the environment). Suppose that imprinting indeed acts on some genes that influence infant crying tendencies, and that they do so as expected, with paternally derived genes favouring marginally more crying as a means of extracting resources from the mother, and maternally derived ones conversely favouring marginally less crying (Haig 2014). On this supposition the crying tendency of any particular infant won't be independent of the parent of origin for those genes. There could be no fact of the matter about the crying tendency resulting from some portfolio of genes *without* information about parent of origin for parts of the portfolio. For all that, the crying tendency of any particular baby is determinate. It may be different to a supposed imprinting-free case, but it is what it is.

To think that intragenomic conflict *does* predict individual behavioural inconsistency is, perhaps, to equivocate two distinct claims. It is one thing for the heritable behavioural

dispositions of an organism to be consistent with a single set of genetic interests. The 'phenotype view' says that the behavioural dispositions of a living organism tend to do just that. The kinship theory shows that there can be conflict between the interests of different parts of a single genome, and imprinting provides one mechanism by which that conflict can be expressed. It follows from this that where imprinting is active, the behavioural dispositions of an organism may diverge from those that serve the unitary corporate interests of the genome. Given the dispositions that would serve the interests of the genome as a whole, the dispositions of an organism with imprinting may be inconsistent *with those dispositions*. This, though, is just not the same thing as a set of behavioural dispositions being *self-inconsistent*, in the sense of not conforming to *some* net set of dispositions or priorities. These two claims are not always carefully enough distinguished. At one point in his discussion, for example, Haig considers an objection to his line of thinking, and offers a response: "But surely we should evolve a consistent set of genetic biases. Not necessarily, if we are also subject to conflict among genes" (Haig 2006, p. 15). But, as I've just explained, a reason to think that the dispositions of an organism won't consistently serve the corporate interests of its genome isn't a reason to think that they won't be *self*-consistent *at all*.

**4.2. A mechanism for intra-personal conflict?**

The absence of an adaptive rationale for behavioural inconsistency leaves open the possibility of conflict in the choice-making process. Conflicting sub-personal agents might produce consistent output; alternatively, their conflict might explain deviations from consistency by reference to proximal rather than ultimate factors. As before, Haig's comment about the 'fitness maximizing computer' suggests the shape of the

inference from genomic conflict to motivational conflict. For we have, again, a simple ideal, and a complicating consideration arising from the hypothesis of intragenomic conflict. The simple ideal is that the process of natural selection will have built organisms that select behaviours by computing the fitness consequences of available behaviours, and that the consequences will be ranked according to the corporate interests of the organism's genes. The complicating consideration is that the interests of genes in an individual don't always coincide, and that there are mechanisms through which at least some of these conflicts can be expressed.

What we are looking for is not just *any* psychological effect of the action of imprinted genes. We are, rather, seeking, a connection between imprinted gene expression and conflicted motivation involving sub-personal representatives of conflicting interests. Trivers sometimes takes a single bound from genetic conflict to asserting that we 'can imagine' maternally and paternally derived genes 'urging' different things, or being the basis of conflicting 'internal voices' (1997, 2000). But it is the plausibility of the intermediate steps he omits that is at issue here. To begin with let us accept the critical foil Haig offers us, and consider the options available in it for the expression of intragenomic conflict. Suppose, then, that the 'phenotype view' points to a choice mechanism in the form of a fitness maximising computer. In an imprinting-free case, the genes relevant to motivation build a fitness maximizing computer and contribute to setting its *priorities* and mode of operation. Different genes contribute to constructing various preferences to their relative strengths, and the conditions under which they are activated. The total profile of the behavioural dispositions of a creature is the net effect of these various contributions.

Now consider how imprinted gene activity might influence or modulate the activity of such a contraption. If we idealise (away from the reality of polygenic traits) by speaking in terms of the options open to single genes, then the available 'moves' are to add or remove a preference, or to amplify or diminish existing ones. This fits with Haig's suggestion that imprinted genes "would be expected to influence broad behavioural tendencies and personality traits, rather than micro-managing every individual decision" (2006, p. 20). Perhaps *adding* a preference that would otherwise be absent is a less likely outcome of imprinted gene expression than the other options, but that doesn't matter here, because it's reasonable to be generous to the kinship theory of genomic imprinting when trying to work out its possible consequences.

Even being generous in this way, *none* of these moves, singly or in combination, lead to conflicting sub-personal agents, even though any of them could lead to different net preferences. This argument complements the 'net effect' argument in section (4.1) above, but differs in being focussed on mechanism rather than outcome. The processes in question could, as a result of imprinting, be wasteful in the same general way as the *Igf2/Igf2r* case. Suppose it is in paternal genetic interest to amplify some preference, say for infant crying, and in maternal interest to reduce it. Suppose furthermore that there is a gene product expressing each goal, and both genes are imprinted. Then one might see a wasteful state of affairs where one process negates the effect of the other, with no net behavioural effect, but at a metabolic cost. Even so, conflict expressed in the *production* of a motivational system doesn't mean that the resulting motivational system will *itself* be conflicted.

A natural worry here is that the notion of an internal 'fitness maximising computer' is

something of a straw man. This is in contrast to the behavioural ecology view considered in the preceding section. There it is close to a consensus position, routinely endorsed in textbooks, that selection tends to drive behaviour towards expressing the corporate interest of the whole genome. Even so, behavioural ecologists are often either agnostic about the mechanisms producing behaviour (being content to focus on relating behaviour to fitness) or are explicitly committed to the operation, at least in many species, of a wide range of distributed and semi-autonomous single-purpose control systems that produce appropriate behaviour *without* centralisation, computation or representation of value. Significant tendencies in evolutionary psychology, furthermore, while sharing the methodological ideal of the behavioural ecologists, are allergic to talk of centralised cognition, instead favouring the view that the human brain is a collection of specialised modules. Outside narrowly evolutionary approaches, there are many proposals regarding the architecture of cognition, including a variety of distributed, decentralised and parallel-process theories. There is not, then, a single default position that could be considered from the perspective of intragenomic conflict, with a view to looking for routes to the development of sub-personal agents. A comprehensive survey of available models and the opportunities they suggest for imprinted gene action would be intolerably unwieldy, so I offer a provisional and incomplete treatment in two steps.

The first step involves noting that the considerations raised above regarding the fitness maximising computer apply generally as long as there is a single system in which motivational competition plays out. Such a system might weigh options according to fitness contribution, or subjective utility, or reward. It might have kinks of various kinds, for example, valuing gains and losses asymmetrically, as in prospect theory, or

hyperbolically discounting delayed rewards as in some behavioural psychology.[11] Still, as long as a *unified* behaviour selection system is built by genes, conflicting genes might add, remove or modulate weights, and stretch or bend the kinks, but *whatever* changes they bring about by doing so, they won't bring about the existence of sub-personal agents. The resulting profile of preferences may well differ from one that serves the corporate interests of the genome (and differ because of intragenomic conflict) but the preferences themselves still won't be conflicted as a result of the genetic conflict. The bottom line is that even if one expands the range of single systems being contemplated, that still won't get you from conflicted genes to sub-personal agents.

The second, and longer, step is to consider fairly generally models of behaviour selection that are committed not to a single motivational system, but to two separate systems.[12] There is considerable variation among dual system, or dual process, models of cognition (Frankish and Evans 2009). In many cases the two systems are understood not in terms of conflict, but rather of division of labour. So 'habit based' systems and 'planning based' systems aren't generally supposed to have distinct and separate *interests*, even though they might at times produce different recommendations. The two systems are, rather, understood to operate in different ways, with distinctive costs, benefits and likely error types associated with each. So the predicament of an organism containing both involves efficient scheduling of the use of the more expensive and

[11] Both prospect theory (when combined with 'dual-system' approaches) and hyperbolic discounting have themselves been put to work in explanations of some kinds of behavioural inconsistency, and in models of intrapersonal conflict (e.g. Ainslie 1992). But in neither case is the conflict standardly connected with intragenomic interests. I emphasise again that the issue here isn't whether there are *any* explanations of inconsistency or conflicted choice, but whether intragenomic conflict provides one of the explanations.

[12] There could be more than two, without significant effect on the arguments here.

slower planning system so that it 'pays its way' by forestalling the errors of the quick and cheap habit system. Models of behaviour selection positing a division of labour between parallel systems can explain some behavioural inconsistency (for example, by reference to brute contingency regarding when or whether a planning system is called into action in time to prevent an inappropriate but habitual response) but they needn't, and typically don't, presuppose conflict in order to do so. Genuinely *conflicting* sub-personal agents should have incompatible preferences regarding the activity or allocation of resources by the whole agent *and* they should have a means of expressing their conflict.[13]

Haig seems to have a dual-system model of human motivation in mind, partly because of his sympathetic remarks, near the opening of his (2006) about William James's view on the interplay of 'reason' and 'passion'. (This at least *suggests* a way of thinking about the expression of conflicted genetic interests that differs from that of Trivers who refers, as we've seen, to a maternal and paternal 'self' but does not associate them with reason and passion.) Haig is careful, as noted above, to describe his suggestions as speculative; still, he entertains the thought that intragenomic conflict might be expressed through a dual system of motivation. This suggestion is, I want to argue, coherent, but one part of the burden of justification for taking it seriously needs to be made more explicit. (Whether the coherent suggestion is true is ultimately an empirical question.)

I've already noted that division of cognitive labour need not *always* be understood in

[13] Thaler and Shefrin's model involving a 'planner' and 'doer' is explicitly conflictual, even though the basis for understanding the conflict concerns time horizons rather than kinship (Thaler and Shefrin 1981).

terms of decomposition across agents with distinct interests. In addition, even overtly competitive cognitive processes need not reflect a *diversity* of interests. Many natural control systems, after all, involve opponent processes. Thermoreception, for example, typically involves separate transducers for temperatures above, and below, some threshold, rather than a single 'objective' encoding. This might make design sense because 'too hot' and 'too cold' receptors were initially separately discovered by natural selection. It might also be understood as a kludge, maintained in the absence of a solution discoverable by selection to the design problem of using firing rates to encode negative values. There's no need in this case to think in terms of one interest that 'wants' to be too cold, in a state of conflict with another that 'wants' to be too hot. Indeed, to suppose separate interests in the thermoreception case is obviously muddled. What there is, is a single whole-agent interest in staying within a certain thermal range, which is implemented by means of sub-systems that detect two opposed ways of being outside that range.

It can become meaningful to identify conflicting agents within a system, though, when parts of the system have different interests in the direction of overall system behaviour *and* means of expressing those differences. In the present case, the kinship theory offers a compelling analysis of such conflict (maternally derived and paternally derived genetic contributions really do have some incompatible interests). What we are looking for is a way of thinking about the means available to genes for expressing it, and, more specifically, whether those means get us to conflicting sub-personal agents, as remarks by Haig, Trivers and Hamilton sometimes suggest they do.

Since intragenomic conflict manifests itself by *modulation* in the level of expression of

some genes, it is highly implausible to expect it to lead the construction of organs or functional systems. So rather than looking for ways in which conflicting genetic interests might *build* sub-personal agents, I suggest that we should ask whether there are ways in which they could exploit the division of cognitive labour to express their conflict. Considering the question very generally, it seems clear that there are. If the contributions to overall functioning by some sub-systems are asymmetrically in the interest of maternally derived or paternally derived genes, *and* there are means by which discriminating gene expression can affect the balance between the systems, then sub-systems that needn't generally be regarded as agents, or as pursuing their own interests in preference to those of the whole agent, can serve as proxies for conflicting genetic interests.[14]

Crespi and Badcock (2008) argue, for example, that maternally derived genes should favour socially oriented cognition, whereas paternally derived ones should favour causal-mechanically oriented cognition. But these partly separate 'mentalising' and 'mechanising' capacities can be understood as *complementary* ways of making sense of the world, with healthy cognition being a matter of achieving an appropriate balance between them. So there's no suggestion here that competing genetic interests explain the *existence* of the capacities. Rather, Crespi and Badcock argue that the fact that mentalising and mechanising are to a significant extent implemented in different brain regions constitutes an opportunity for genetic competition to be expressed through the differential development of specific brain tissues. On their view, therefore, autism and schizophrenia ask to be seen as opposed disorders where the balance of genetic conflict

[14] I do not mean proxy in the sense of one agent with authority to act for another. That sense of 'proxy' implies too many agents. Rather, I mean proxy in the sense in which one value can be used to represent another in a calculation.

swings too far in one direction or the other. Their focus is, however, on explaining those conditions, not on inquiring into whether the individuals afflicted by them are subject to intrapersonal conflict.

Whether or not Crespi and Badcock's hypothesis — or any other that undertakes to relate genetic interests to brain regions — is correct in detail, the hypothesis itself is reasonable in principle. If different brain regions, or tissues, or cell types, make distinctive contributions to the overall behavioural strategy of an organism, and some genes have asymmetric interests in influencing that strategy *and* the means to act differentially upon the development of brain regions, tissues, etc., then such differentially developed regions, tissues, etc. can serve as proxies for intragenomic conflict. The extent to which they are elaborated or the ways they are developed in an individual might then depend partly on imprinted gene activity arising from the distinct interest some sections of the genome have in the whole individual's either possessing or lacking some disposition. In that case, our understanding of the balance of influence among different systems, and the resulting overall set of dispositions, would need, ideally, to include reference to competing genetic interests.

Haig speculates that 'reason' and 'passion' may map, at least roughly, onto the differential brain development discoveries in Keverne et al.'s chimeric mice. Recall that in those embryos androgenetic cells were relatively more represented in the hypothalamus and some other structures but *not* the cortex, while a converse pattern was found for parthenogenetic cells. Suppose that something like this is correct, *and* that the relative degree to which behaviour is controlled by cortical and libidinal areas depends on the sheer relative size of the respective tissues. Assume also that maternally

derived and paternally derived genetic interests each favour only one of the two systems, because of the behaviours the preferred system tends to promote. Then if imprinting could asymmetrically influence the relative rate of growth of some specific tissues (which we already know it does), the kinship theory of genomic imprinting would predict it. This wouldn't mean that the very existence of the hypothalamus, say, could in any *exclusive* way be explained by reference to paternally derived genetic interests, but it would mean that its properties, including its sheer size at a particular time, might depend in part on the contingencies of imprinted gene expression that were in conflict, ultimately, over how much influence the hypothalamus would have on the whole individual's behaviour.

The effect of imprinted gene activity in such a scenario would be a cognitive mechanism different from one that would serve the corporate interests of the whole genome. And it would achieve this departure by expressing conflict over what *kind* of cognitive mechanism the organism should have. This does not quite amount to sub-personal agents representing the distinctive interests of parts of the genome, unless those cognitive mechanisms functioned as sub-personal agents already. This scenario is better described, then, as one in which a division of cognitive labour is exploited to provide proxies for divergent genetic interests. As in section (4.1) above, this conclusion amounts to neither full endorsement nor full rejection of the suggested consequences of intragenomic conflict. That genetic factions could actually build functional representatives of their clashing interests is implausible. But it is not implausible to suppose, given the widespread division of cognitive labour, including partly independent and parallel systems for selecting behaviour, that intragenomic conflict could be expressed through modulation of the relative effectiveness of the

various cognitive sub-systems in ways that could not reasonably be understood without reference to the hypothesis of intragenomic conflict.

# 5. Conclusion

The kinship theory of genomic imprinting is a coherent application of the theory of inclusive fitness to portions of an individual genome. The analytical perspective that it offers predicts various conflicts. Not only that, a mechanism — genomic imprinting — conduces to the expression of (some of) these conflicts in some taxa. The most common known targets for imprinting in mammals (most of the evidence concerns mice and humans) are embryonic development, and the brain. The kinship theory of genomic imprinting, furthermore, predicts that imprinted gene activity will have behavioural effects. This prediction is borne out in some cases, although the research area is fairly new. A sizeable fraction of the known cases of imprinted gene expression fits the conflict hypothesis better than any other candidate explanation. Even if not all the remaining cases are found to fit that hypothesis (and there is no guarantee that they will), it still remains the leading explanation.

Some biologists have suggested, furthermore, that intragenomic conflict might explain or even predict intrapersonal conflict. This is the proposition which I have, in part, assessed here. My treatment has focussed specifically on the suggestions that the conflict perspective predicts individual behavioural inconsistency and that it predicts the existence of contending sub-personal agents representing divergent genetic interests. (I have not considered the view that intragenomic conflict might have introspectible effects.) There are good independent reasons, in any case, for doubting that natural

agents will be consistent in their behaviour, or that their behaviour selection processes will be highly unified. So neither behavioural inconsistency nor fragmented internal processes are really a puzzle. The issue is, rather, whether the hypothesis of intragenomic conflict can contribute anything towards explaining them.

My conclusions are mixed: while intragenomic conflict does not predict behavioural inconsistency, it does predict deviations in behavioural tendencies relative to the set of tendencies that best serve the corporate interests of the whole genome. Intragenomic conflict, moreover, does not plausibly predict the construction of sub-personal agents representing the interests of segments of the genome, even though it possibly predicts the differential elaboration of cognitive sub-systems in cases where the contribution those sub-systems make to overall behaviour enables them to act as proxies for the divergent interests of portions of the genome. It may be partly a matter of individual perspective whether this sounds like a qualified endorsement, or a qualified rejection, of the claim that intragenomic conflict predicts intrapersonal conflict. Either way, neither we nor the processes that shape our behavioural tendencies are entirely consistent wholes.

# References

Ainslie G (1992) *Picoeconomics*. Cambridge University Press, Cambridge.

Barton SC, Surani MA, Norris ML (1984) Role of paternal and maternal genomes in mouse development. *Nature* 311, 374–376.

Burt A, Trivers R (1998) Genetic conflicts in genomic imprinting. *P Roy Soc B-Biol Sci* 265: 2393-2397.

Burt A, Trivers R (2006) *Genes in Conflict: The biology of selfish genetic elements*

Bellknap/Harvard University Press, Cambridge, Mass.

Cassidy SB, Dykens E, Williams CA, (2000) Prader–Willi and Angelman syndromes: sister imprinted disorders. *Am J Med Genet* 97(2):136–146.

Crespi B, Badcock C (2008) Psychosis and autism as diametrical disorders of the social brain. *Behav Brain Sci* 31:241-320.

Dennett DC (2013) The normal well-tempered mind, Interview on *Edge.org*, transcript available at http://edge.org/conversation/the-normal-well-tempered-mind (accessed 10/8/2014).

Frankish K, Evans JStBT (2009) The duality of mind: An historical perspective. In: Evans JStBT, Frankish K (eds) *In two minds: Dual processes and beyond*. Oxford University Press, Oxford, pp 1-30

Glimcher P (2002) Decisions, Decisions, Decisions: Choosing a Biological Science of Choice. *Neuron* 36:323-332.

Gregg C, Zhang J, Weissbourd B, Luo S, Schroth GP, Haig D, Duluc C (2010) High-Resolution Analysis of Parent-of-Origin Allelic Expression in the Mouse Brain. *Science,* 329:643-648.

Haig D (2002) *Genomic Imprinting and Kinship.* Rutgers University Press, New Brunswick.

Haig D (2003) On intrapersonal reciprocity. *Evol Hum Behav* 24:418-425.

Haig D (2004) Genomic imprinting and kinship: How good is the evidence? *Annu Rev Genet* 38:553-85.

Haig D (2006) Intrapersonal conflict. In: Jones MK, Fabian AC (eds) *Conflict*. Cambridge University Press, Cambridge, pp 8-22

Haig D (2008a) Conflicting messages: genomic imprinting and internal communication. In d'Ettorre P, Hughes DP (eds) *Sociobiology of Communication.* Oxford University Press, Oxford, pp209-223

Haig D (2008b) Kinship asymmetries and the divided self. *Behav Brain Sci* 31:271-272.

Haig D (2009) Transfers and transitions: Parent–offspring conflict, genomic imprinting, and the evolution of human life history. *P Natl Acad Sci USA* 107:1731-1735.

Haig D (2014) Troubled Sleep: Night waking, breastfeeding and parent-offspring

conflict. *Evolution, Medicine and Public Health* 2014:32-39.

Haig D, Westoby M (1989) Parent specific gene expression and the triploid endosperm. *Am Nat* 134:147–55.

Hamilton WD (1987) Discriminating Nepotism: expectable, common and overlooked. In Fletcher DJC, Michener CD (eds) *Kin recognition in animals*. John Wiley and Sons, Chichester, pp 417-437

Isles AR, Davies W, Wilkinson LS (2006) Genomic Imprinting and the Social Brain. *Philos T Roy Soc B* 361:2229-2237.

James W (1983 [1890]) *The Principles of Psychology*. Harvard University Press, Cambridge, Mass.

Keverne EB, Fundele R, Narasimha M, Barton SC, Surani MA (1996) Genomic imprinting and the differential roles of parental genomes in brain development. *Dev Brain Res* 92:91-100.

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

McGrath J, Solter D. (1984) Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* 37:179–183.

McNamara JM, Houston AI (1986) The Common Currency for Behavioral Decisions. *Am Nat* 127:358-378.

Moore T, Haig D (1991) Genomic imprinting: a parental tug-of-war. *Trends Genet* 7:45-49.

Shizgal P, Conover K (1998) On the neural computation of utility. *Curr Dir Psychol Sci* 5:37-43.

Spurrett DS (2014) Philosophers Should Be Interested in 'Common Currency' Claims in the Cognitive and Behavioural Sciences. *S Afr J Philos* 33: 211-221.

Thaler RH, Shefrin HM (1981) An economic theory of self-control. *J Polit Econ* 89:392-406.

Trivers RL (1974) Parent-offspring conflict. *Am Zool* 14:249-264.

Trivers RL (1997) Genetic Basis of Intrapsychic Conflict. In: Segal N, Weisfeld GE, Weisfeld CC (eds) *Uniting Psychology and Biology: Integrative Perspectives on*

*Human Development*. American Psychological Association, Washington, DC, pp 385–395

Trivers R (2000) The Elements of a Scientific Theory of Self-Deception, *Ann NY Acad Sci* 907:114-131.

Trivers R (2009) Genetic conflict within the individual. Berlin: *Sonderdruck der Berliner-Brandenburgische Akademie der Wissenschschaften*. 14:149–199.

Trivers R, Burt A (1999) Kinship and genomic imprinting. In: Ohlsson R (ed) *Genomic Imprinting: An Interdisciplinary Approach*, Springer, Berlin. pp 1–21

Úbeda F, Gardner A (2010) A model for genomic imprinting in the social brain: Juveniles, *Evolution* 64:2587–2600.

Úbeda F, Gardner A (2011) A model for genomic imprinting in the social brain: Adults, *Evolution* 65:462–475.

Úbeda F, Haig D (2003) Dividing the Child. *Genetica* 117:103-110.