

Opinion

Embryo Selection and Mate Choice: Can 'Honest Signals' Be Trusted?

Dakota E. McCoy^{1,*,@} and David Haig¹

When a measure becomes a target, it often ceases to be a good measure – an effect familiar from the declining usefulness of standardized testing in schools. This economic principle also applies to mate choice and, perhaps surprisingly, pregnancy. Just as females screen potential mates under many metrics, human mothers unconsciously screen embryos for quality. 'Examinees' are under intense selection to improve test performance by exaggerating formerly 'honest' signals of quality. Examiners must change their screening criteria to maintain useful information (but cannot abandon old criteria unilaterally). By the resulting 'proxy treadmill', new honest indicators arise while old degraded indicators linger, resulting in trait elaboration and exaggeration. Hormone signals during pregnancy show extreme evolutionary escalation (akin to elaborate mating displays).

If the Stakes Are High, Measurement Is Hard

High-stakes testing is meant to measure performance in many contexts, but it also causes systemic, and unwanted behavior change. Often the things tested are easily measured proxies (e.g., standardized test scores at a middle school) for more difficult-to-measure attributes (e.g., how well the middle school teachers educate their pupils). Teachers 'teach to the test' rather than improving students' reading skills [1]. Poorly performing students are sometimes encouraged to stay home [2,3]. In a different context, test scores for university admissions are intended to provide objective measures of intellectual quality but come to also measure economic means as some families can afford expensive test-preparation courses that are unavailable to poorer families. Such tests convey some useful information but are distorted by confounding factors. When high stakes depend on an 'objective' measure, the measure may cease to provide reliable information because behavior changes to satisfy testing criteria, thus degrading the correlation between the proxies and the attributes that are the real target of evaluation. In economics and the social sciences, this is variously known as **Goodhart's law** (see [Glossary](#)) [4] or **Campbell's law** [5], succinctly phrased as 'When a measure becomes a target it ceases to be a good measure' [6]. In fields ranging from health care to academic publishing, many metrics become less correlated with quality over time ([Box 1](#)).

In economic and social contexts, Campbell's law predicts that simple indicators of quality can be corrupted; this may cause an arms race between the evaluating system and the agents being evaluated. As a result, both the indicator and the test become more complicated. For example, as Twitter bots spread to exploit automated measurements of impact, bot-detection algorithms must race to identify honest indicators of 'human-ness' [7]. As bots find ways to deceive each new indicator, both the algorithm and the bot are encumbered by the continued detection and production, respectively, of 'human-ness' indicators. The bot cannot afford to abandon old, degraded indicators of human-ness or it will be caught; the algorithm cannot afford to dally without finding new indicators of humanness or it will never catch the new and improved bots. The same

Highlights

Mate choice by honest signaling is a classic explanation for elaborate traits in nature. Many researchers have: (i) observed deceptive signaling, and (ii) wondered how honest signals relate to trait elaboration.

Honest signaling is analogous to high-stakes testing. Quality is hard to measure directly, so proxies (tests) are used. High-stakes testing causes 'teaching to the test' without improving educational outcomes.

Embryo choice is another high-stakes test. Mothers select healthy embryos and terminate sub-par embryos automatically. Embryos are selected to pass maternal tests without improving their quality. The resulting arms race causes extreme and elaborate signals during pregnancy.

We can better understand elaborate traits in nature if we interpret mate selection, and embryo choice, as a dynamic give-and-take between two parties with conflicting fitness interests.

¹Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

*Correspondence: dakotamccoy@g.harvard.edu (D.E. McCoy).
 @Twitter: [@WildlifeMcCoy](https://twitter.com/WildlifeMcCoy)

Box 1. When a Measure Becomes a Target, It Ceases to Be a Good Measure

In public schools in the USA, high-stakes testing has many unintended consequences. Teachers in Chicago guided poorly performing students away from subjects that counted towards performance ratings [2,3] and test scores in Texas skyrocketed after implementation of high-stakes testing but actual skills (e.g., reading) did not improve and ethnic performance gaps remained or even widened [1]. Educational outcomes are hard to measure, so we rely on test scores as a proxy of quality. Educational proxies degrade given sufficient incentive, as do proxies in many other fields [69]. Medicine is not immune to signal degradation. When hospitals were rated based on how long patients had to wait, emergency departments refused to admit patients from waiting ambulances until they could be seen, prompting ambulance companies to purchase makeshift tents to store patients in limbo [70]. In academic publishing, metrics such as impact factor over time become less informative of quality and depend heavily on field [71]. Likewise, researchers subset data or collect data selectively to achieve a significant *P* value, a practice known as ‘*P* hacking’ [72]. We want nutritional labels to convey food health, but food manufacturers game this system by labeling foods ‘low-fat’ but adding extra sugar [73]. These disparate examples show that metrics meant to improve performance often can be gamed. As metrics are gamed, regulatory agencies must establish more detailed and elaborate regulations, the result of an arms race resulting from high-stakes testing.

logic may explain why US tax law, an arena wherein people and corporations have strong incentives to deceive, is so cumbersome and elaborate. In contexts where honesty matters to the receiver but deception enriches the signaler, both the signal and the receiving apparatus become more elaborate over time.

Campbell’s law also applies to high-stakes examinations in biology. Mate choice and embryo selection are, respectively, indirect and direct means of selecting fit offspring. In both arenas, many candidates compete for few slots. Candidates’ intrinsic quality can be assessed through information that correlates with fitness (termed an **honest signal** in the domain of mate choice). But selection favors candidates who can exaggerate their own apparent quality. Indeed, in the well-studied domain of mate choice, substantial theoretical work [8–13] and many experimental results (Box 2) show that deception is widespread and suggest that antagonistic coevolution between the sexes may cause elaborate display traits. Signalers can find means to reduce the cost of, or invest disproportionately in, formerly honest signals – including by evolving enhancers to the signal. In response, receivers can raise the necessary signal threshold or adopt new, additional standards. By the ‘**proxy treadmill**’, receivers seek honesty and signalers seek to deceive, but receivers cannot unilaterally abandon a prior trait preference (e.g., if a female selects a male who does not conform to majority preference, her sons will be disadvantaged in the mating market).

We argue that the conflict of interest between examiners and examinees of lower quality promotes instability, escalation, and elaboration (of both signal and test). We hope to provide three useful contributions. First, the proxy treadmill arises from fundamental principles and should be expected in studies of honest signals. While it is uncontroversial that deception is widespread in mate choice, the idea of trait elaboration via arms races is not well-embedded in the field. In any study of honest signaling, we suggest that researchers should expect that the pressure to deceive begets pressure to evaluate better (causing signal elaboration). Second, research on deception in mate choice provides the framework to interpret the biology of human pregnancy: pregnancy is a high-stakes test of embryo quality. Just as males find deceptive means to enhance their apparent fitness to choosy females, we argue that embryos have evolved illusions of health and fitness as well as tools to coerce maternal bodies into continuing, rather than terminating, a pregnancy. Third, massive signal elaboration in embryo selection supports the theory of the proxy treadmill in other biological domains (such as mate choice). Mate choice is a complicated and controversial field involving multiple sensory systems, a chaotic environment, and competing (nonexclusive) selective pressures. Embryo selection occurs in a comparatively constrained

Glossary

Aesthetic residue: display trait which is no longer indicative of quality but is nonetheless retained in signaler–displayer examinations because it would be costly to abandon.

Campbell’s law: a principle in social sciences that states that quantitative proxies, or metrics, meant to measure complicated phenomena are subject to corruption when the indicator is used for decision-making. ‘When a measure becomes a target it ceases to be a good measure.’

Chorionic gonadotropin (CG): hormone produced by the placenta that is molecularly and functionally related to maternally derived luteinizing hormone; it maintains pregnancy by stimulating the corpus luteum to produce progesterone. It has evolved independently in equids and primates.

Honest signal: display trait thought to correlate with intrinsic quality (health or fitness) of the displayer, thus giving valid information to the observer.

Goodhart’s law: another name for Campbell’s law.

Luteinizing hormone (LH): hormone produced by the maternal pituitary that is required to maintain a pregnancy (it stimulates the corpus luteum to produce progesterone).

Parent–offspring conflict: an evolutionary theory based on differences between parents and offspring over the optimal investment a parent should give to a single offspring. This leads to many evolved signatures of conflict, such as genomic arms races.

Proxy treadmill: a proposed phenomenon by which signal traits become exaggerated as proxies of quality are continuously modified or replaced (because examinees are under strong selective pressure to inflate their apparent quality, thus devaluing any given proxy).

environment, the maternal body, and operates primarily via hormone signals for which the genetics are understood.

Elaborate mating displays demanded by the choosing sex, and elaborate hormone signals required for an embryo to be carried to term, are analogous to cumbersome regulations intended to detect bad behavior in many socio-economic arenas.

Choosing a Partner: The Displaying Sex Has Incentives to 'Cheat', but Some Forces Resist

In mate choice, members of one sex, usually females, assess the performance of the other sex in a mating display. Multiple selective pressures interact to shape female preference and favor particular male traits, including sensory bias (pre-existing preference for particular traits [14,15]), arbitrary aesthetic preferences [16–20], and the need to identify members of one's own species [21]; however, the most common explanation is that females use display traits to select a healthy and fit male ('honest signaling'). Under honest signaling theory, signals may indicate quality by incurring a cost [22–24] or by being causally linked to important physiological processes ('index' [25,26]).

Honest signals in mate choice are analogous to high-stakes tests: one sex uses proxies of quality to select mates of the opposite sex. Decades of research on mate choice have shown that signaling between two parties with conflicting interests is subject to deception [11,12,27–31] (Box 2). Signals evolve to manipulate or persuade receivers in a manner favorable to the sender, not merely to inform. The fundamental conflict of interest between chosen and choosing sexes [32] may cause an arms race where displayers exaggerate their appearance without improving their quality [11–13,33]. In sexual selection, the choosing sex (hereafter 'females') is selected to elicit honest information about displayers (hereafter 'males'), but males are selected to evolve deceptive strategies to appear higher-quality than they are. Low-quality males have a larger incentive to exaggerate. Deceptive enhancers may sweep through a population, reducing the value of that signal and raising the bar on how high-quality males must appear in order to mate [8,12,13,33–35].

Even as a signal becomes less valuable as an honest indicator, a female retains the preference for this signal because to unilaterally abandon her preference for a particular trait may disadvantage her sons (who, lacking that trait, would appear undesirable to the majority of females in the population). However, females are still selected to extract quality signals and will thus either adopt a more stringent version of the same test and/or adopt new tests of quality. In this manner signaling traits become more elaborate, as females demand new indicators of quality while retaining old, degraded indicators as **aesthetic residue** (for further discussion see following section). 'Honest' signaling is thus susceptible to bluffing displayers, which are widespread, and animals frequently evolve enhancers to their displays, associated traits with no obvious 'honest' content themselves that make the entire display more elaborate (Box 2). Deception in mate selection may be widespread [12] but frequently hidden because it is selected to be undetectable [36]. It likely is no coincidence that most deceptive signals have been described in small animals easy to observe, manipulate, and test.

Despite the pressure to deceive, many signals in nature are reliable. Some deceptive mating displays do not imply that signals convey no useful information. Many studies demonstrate correlations between displayer signal and quality. What selective forces resist the deterioration of honest assessment? First, certain tests of quality are challenging to manipulate. Index signals, which are tied to uncheatable physical processes rather than invoking a cost themselves, may be resistant to manipulation [8]. Courtship displays that are selected to reveal vigor are more

challenging to fake if they operate on multiple physical modalities (i.e., singing, drumming, moving one's whole body, and more). Finally, Campbell's and Goodhart's laws predict that honest assessment is possible as long as the examiner stays one step ahead of the examinee. Examiners can escalate the level of signal required, or add new required signals of quality, while retaining old signals as 'aesthetic residue'. We argue below that pregnancy is another clear case where signals between two individuals (mother and embryo) become more elaborate due to a conflict of interest. Mate choice is complex (due to the interaction of brains, sensory systems, and complex behaviors) but it should be subject to this same general phenomenon.

The Proxy Treadmill: New Honest Signals Evolve and Degraded Signals Linger

Observable proxies are chosen because they correlate with more difficult-to-observe qualities. However, this correlation tends to weaken because the proxy becomes a direct target of selection for test-taking ability. This can be thought of as 'sensory manipulation' of the examiner by the examinee. If examiners are to continue to select higher-quality examinees, they must seek new proxies that are better correlates of desired qualities. However, there is a con-

Box 2. Examples of Signal Elaboration Caused by Deception in Mate Choice

Deception and subterfuge abound in mate choice. When male fiddler crabs *Uca annulipes* lose their large display claw (Figure IA), 44% grow a large, but fragile, replacement ([36]; see also [74,75]). Male Trinidadian guppies *Poecilia reticulata* (Figure IB) prefer to display near relatively unattractive males to make themselves look better [76]. Further examples show that deception may cause trait elaboration.

Female long-tailed dance flies, *Rhamphomyia longicauda*, compete for protein-rich nuptial gifts from males (for which females exchange copulations). They fly in leks to attract male attention, and males prefer females that have distended abdomens [77]. Distended abdomens have been interpreted as an honest signal of egg development and thus of female mate quality. However, just before the lek, females swallow large quantities of air to inflate pouches on their abdomen and then further wrap their legs around the abdomen, creating an apparently huge abdomen, which completely masks egg development (Figure IC, [77]). In a related pouch-less species, *Rhamphomyia sociabilis*, abdomen size is indeed an honest correlate of egg development. This suggests that females evolved to deceive males after males were selected to prefer females at a later stage of egg development. Selection for deception caused elaboration.

Male nursery web spiders *Pisaura mirabilis* gather nuptial gifts to buy mating opportunities; these are nutritious substances often wrapped in silk. The 'worth' of male gifts varies, from nutritious prey to inedible popped red balloon [78,79]. It is thought that the gifts or the silk itself are honest signals of quality [78]. However, males offered deceptive, worthless gifts in 58% of field trials and 85% of laboratory trials [34]. Further, (i) worthless gifts were not connected to poor body condition; (ii) cheaters who offer worthless gifts in fact gain body mass; and (iii) males adaptively wrap more silk around worthless gifts, to conceal worthlessness and prolong female feeding time to allow mating [34]. The authors describe this as 'an antagonistic arms race between males under selection to deceive and females under selection to evolve resistance to deception' [34]. The arms race causes more elaborate male gifts.

Many male mammals are thought to scent-mark with urine to honestly convey their reproductive status, social rank [80], and, based on mark height, body size [81]. Indeed, dwarf mongooses (*Helogale parvula* [82]) spend longer investigating scent-marks that are higher from the ground. However, some individuals paint a larger-than-life portrait of themselves. Undersized male dwarf mongooses scent-mark higher than expected and, in fact, dwarf mongooses typically hand-stand to scent-mark (Figure ID, [82,83]). Small dogs (*Canis familiaris*) raise their leg at a higher angle than large dogs, thus exaggerating their relative mark height [84]. Males who could hand-stand and leg-cock could deceptively over-represent their own height; once these traits swept through the population, scent-marks would again honestly convey body size. The signal became more elaborate over evolutionary time.

In sea lamprey (*Petromyzon marinus*), the pheromone 3-keto petromyzonol sulfate (3kPZS) originated as a nonsexual, larvally produced migratory cue and was coopted by deceptive males to attract females [85]. This 'sensory trap' [86] could be reclaimed as an honest signal [87], whereby females assess male body size and condition via pheromone production [88]. Once 3kPZS became a proxy, selection favored males who could 'dramatically upregulate the 3kPZS biosynthetic pathway', causing males to have an 'up to 8000-fold increase' of production in their livers [89]. Further, small males have disproportionately larger livers and greater 3kPZS production [88], a potentially deceptive overinvestment. In response, female reception adaptively shifted to avoid costly confusion between larvally produced and sexually produced 3kPZS [89]. An originally deceptive signal was reclaimed as honest, was further corrupted, and promoted receiver adaptation to better discriminate.



Figure 1. Examples of Deceptive Signaling in Mate Choice. (A) When male fiddler crabs *Uca annulipes* lose their large claw (used for signaling and fighting), they regrow a huge but brittle claw. (Photo credit Rujuta Vinod; license CC BY-SA 4.0). (B) Female (bottom) and male (top) Trinidadian guppies (*Poecilia reticulata*); males preferentially display near comparatively unattractive males. (Photo credit Amy E. Deacon, Hideyasu Shimadzu, Maria Dornelas, Indar W. Ramnarine, and Anne E. Magurran; license CC BY 4.0). (C) A female long-tailed dance fly (*Rhizophomyia longicauda*) inflates sacs on her abdomen to appear as if she has well-developed eggs, a trait favored by observing males. (Photo credit Dave Funk.) (D) Male dwarf mongooses that are undersized scent-mark higher than expected and many individuals handstand. (Photo credit Lynda Sharpe.)

servative force, recognized by Fisher, that prevents a simple replacement of the degraded proxy by a better proxy [20]. This is the ‘desirability’ of the examiner’s offspring when they become examinees. A degraded proxy remains ‘sexy’ because an examiner who unilaterally drops the proxy will have sons who lose out in the mating market because of the established preferences of other females. Princeton University’s attempt to halt grade inflation is a useful analogy [37]. Amidst steadily increasing university grade point averages, Princeton University implemented grade deflation in 2004 but abandoned this policy in 2014 when no peer institutions followed suit [38]. Critics pointed out that low grades could harm students applying for jobs and that some students (perhaps concerned about their grades) chose to attend other schools [37], placing Princeton at a competitive disadvantage in university rankings. Unilateral changes by examiners invoke a serious selective cost.

This is the proxy treadmill. Proxies are initially adopted because they convey reliable information about male quality: females who adopt a reliable proxy are winners, as are the high-quality males they choose. But high-stakes examinations unwittingly select for males that appear better than their objective quality. Because ‘dissemblers’ are rewarded, high-quality males

who do not disassemble lose their advantage. Females who adopt entirely new examinations in response place their own sons at a selective disadvantage because they will be evaluated by the established criteria of other females. Therefore, old preferences are retained as aesthetic residue while new preferences are established to test quality. In this manner, ‘looking good on paper’ can sweep through a population, placing selective pressure on females to update their discriminatory abilities by adoption of new proxies while retaining older criteria as purely aesthetic preferences. This dynamic can potentially favor an extraordinary elaboration of desirable traits.

Imagine that female birds assess males based on how red their feathers are (redness, resulting from metabolized carotenoid pigments, is considered an honest index of metabolic function [25,26]). If some males evolve structural components to the feather that enhance redness by 25%, females can simply raise their preference bar by 25% to preserve an honest signal. Alternatively, females can develop a new test – say, the vigor of a male’s display dance – in addition to the redness test. While redness is no longer a reliable signal of quality, the female preference for redness may remain as aesthetic residue, while males now must also dance vigorously. As traits become more exaggerated or multiple traits stack together to resist signal erosion, sexually selected traits may become more elaborate. (We note that the proxy treadmill is not the only mechanism by which traits become more elaborate; for example, an animal’s fundamental sensory ecology can shape diversification [14]).

Honest signaling is not an equilibrium state because the information conveyed by signals is corrupted over time as males find ways to inflate their own appearance of quality (but see [Outstanding Questions](#)). But females are under constant selection to find new sources of useful information, whether this can be extracted from deceptive signals, or by adding new criteria of selection. This idea has long been described (e.g., [11,13,29,33]), but has not fully been adopted by researchers of sexual selection who investigate honest signals. This may be because sexual selection is extremely complex; mate choice involves multisensory perception in chaotic environments. It is challenging to document trait costs, assess whether honest signalers have higher fitness, and reconstruct the evolutionary history of traits, required to demonstrate the proxy treadmill in operation. A comparatively more constrained area of signaling, *in utero* embryo choice, provides a useful analogy.

Choosing a Child: Signal Elaboration by the Proxy Treadmill during Human Pregnancy

One rationale for ‘honest signaling’ or ‘good genes’ models of mate choice is that females select males based on the expected quality of their genetic contribution to offspring. This is indirect selection of offspring: a high-quality male may father some low-quality offspring because of the vagaries of genetic recombination and epistatic interactions between maternal and paternal genomes. A more direct way to improve the quality of offspring would be to examine offspring themselves before major commitment of resources [39]. Mothers, in species with postzygotic provisioning of offspring, are incentivized to terminate investment in subpar offspring early and redirect resources to better quality offspring [40,41]. The earlier a choice can be made, the greater the savings. Thus, the most intense period of mortality in human life-history occurs during the first month after conception: at least 22% of human embryos [42], perhaps more than 70% [43], are eliminated. Embryos audition for the role of a lifetime before unforgiving judges.

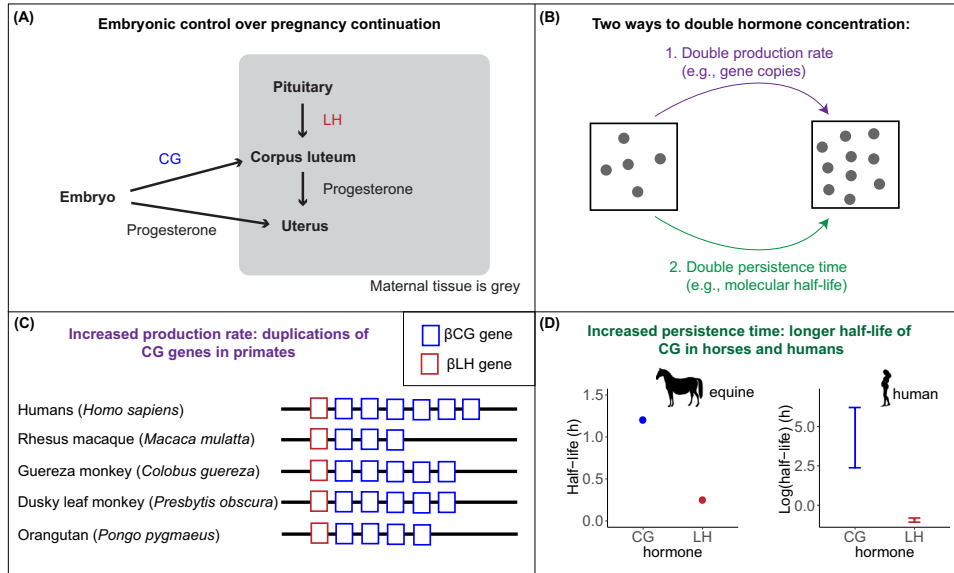
We focus on one criterion of maternal choice: the hormonal output of early embryos. Hormone signaling between embryo and mother is analogous to mating displays between males and females. Embryonic signaling is, in some ways, more evolutionarily tractable than sexual signaling. We understand the genetic underpinnings of many hormone signals and signaling occurs within the maternal body rather than in complex external environments. As such, embryo choice is a useful analogy to mate choice. Equally, substantial work on mate choice [11–13,33] underpins our investigation of pregnancy as a high-stakes examination.

Embryonic hormones may provide useful information about embryo quality, but this system has escalated via an intergenerational arms race due to differing evolutionary interests of mothers and embryos [44,45]. Researchers study this conflict of interest between mothers and embryos under the theoretical umbrella of **parent-offspring conflict**. Mothers are selected to gestate only high-quality offspring, but embryos are selected to exaggerate their perceived quality [46]. Akin to males exaggerating their appearance in front of choosy females, embryos hijack systems of maternal control and evolve amplified hormonal signals. As remarked by Roberts [47], under this high-stakes testing regime, the embryo represents its own interests ‘possibly even to the point of relaying misleading information regarding its own fitness and potential. The trophoblast-maternal interface thus represents a battleground that has shaped remarkable rates of evolutionary change’ [47].

Progesterone is essential for the maintenance of mammalian pregnancy. The ancestral source of progesterone was the corpus luteum formed from the ruptured follicle at ovulation. By 7 weeks of gestation, the human placenta becomes a sufficient source of progesterone to obviate the need for luteal progesterone [48]. However, prior to this ‘luteoplacental shift’, pregnancies miscarry if the corpus luteum is removed or ceases to produce progesterone. The corpus luteum initially produces progesterone in response to **luteinizing hormone (LH)** produced by the maternal pituitary, but the quantity of LH required to maintain the corpus luteum rapidly increases, resulting in menstruation (shedding the endometrial lining) 2 weeks after ovulation. Many early-stage human embryos are terminated through the process of menstruation, after they have embedded, at the regular menstrual time or after a slightly delayed period.

Early embryos of simian primates [49] and equids [50] have independently evolved a placental version of LH, **chorionic gonadotropin (CG)**, that is secreted into the maternal circulation to maintain progesterone production by the corpus luteum, thus preventing loss of the pregnancy (Figure 1A). The rapid increase of the LH requirement to maintain pregnancy is not observed in mammalian groups that do not produce CGs and is much greater than the ancestral levels of LH necessary to maintain progesterone production. Embryonic secretion of CGs can be conjectured to have originated as a simple form of ‘cheating’ in which embryos forged a statement they had passed uterine quality control and submitted it to the corpus luteum, rendering null and void whatever prior testing procedures had existed, akin to a student forging ‘A+’ on their report card [44].

However, even initially dishonest signals can be reclaimed as honest signals of quality. CG is a glycoprotein that reveals basic competences in protein synthesis and glycosylation [55]. The maternal evolutionary response was to rapidly increase LH requirements after ovulation and use levels of CG as a proxy of embryo quality. Embryos that produced abundant CG were healthier than embryos that produced lesser amounts. Tiny embryos came to be assessed on their relative prowess in production of CG and must now produce ‘heroic concentrations’ of CG [56] to avoid being terminated at menstruation. The extreme levels of CG produced by human embryos are the outcome of a long history of ‘grade inflation’: of selection on embryos



Trends in Ecology & Evolution

Figure 1. Hormone Signaling during Pregnancy Reflects an Evolutionary Arms Race between Embryos and Mothers, Causing Elaborate Signals. (A) In most mammals, pregnant females produce luteinizing hormone (LH) in their anterior pituitary, which causes the corpus luteum to produce progesterone (thus maintaining the endometrial lining rather than shedding it and the embryo). By withdrawing LH, she could terminate a pregnancy. But over evolutionary time, embryos seized control by producing a biochemical analog of LH, chorionic gonadotropin (CG), thus allowing pregnancies to continue even if the maternal body reduces or withdraws all LH. Embryos also became an independent source of progesterone. (B) The standing concentration of a hormone within the maternal body can be increased in two ways: (1) by increasing the production rate (e.g., through gene duplications); and (2) by increasing the persistence time (e.g., the molecular half-life). (C) Primates have had many gene duplications of β CG from an ancestral β LH gene [49]. (D) CG convergently evolved a much longer half-life in both equines and humans [51–54], allowing it to remain in circulation longer. Silhouette credits: horse (Openclipart, uncredited, public domain); woman (publicdomainpictures.net, credit Tanya Hall, public domain).

to perform better at maternal examinations, regardless of their intrinsic quality, and of corresponding escalation of the ‘passing grade’ to maintain quality control.

While in theory CG could remain honest, as an index of protein synthesis and glycosylation, embryos have found deceptive means to increase CG concentrations regardless of their quality. Primate embryos evolved at least two initially ‘dishonest’ test-taking strategies: duplications of β CG genes and extended half-lives of CG (Figure 1B,C). There are six copies of β CG in each haploid human genome and three in rhesus macaques [49]. Before an increased number of gene copies became fixed in an ancestral population, embryos with more gene copies had an ‘unfair’ advantage relative to embryos with fewer gene copies. Second, CG has evolved to have a much longer half-life than ancestral LH [51,57]. The half-life of LH is around 21–25 minutes, while that of hCG varies from 11 to 462 hours [51–53] because of a C terminal peptide of β CG that is absent in β LH [58–63]. An extended half-life increases the effective concentration of a hormone for the same rate of production. Therefore, the novel, longer-lived C terminal peptide would initially have given an ‘unfair’ advantage to embryos that possessed the peptide relative to those that retained the ancestral CG with a much shorter half-life. A similar C terminal peptide that extends half-life has evolved independently for the CG-LH of horses [50]. The half-life of equine CG is further extended relative to equine LH by differential glycosylation. As a result, equine LH is removed from

circulation 5.7 times more rapidly than equine CG (when equine LH and equine CG were injected into horse liver and kidney, 75% of CG but only 25% of LH remained in circulation after 30 minutes [54]; Figure 1D).

In summary, to pass maternal tests embryos found shortcuts that were decoupled from their intrinsic quality. The predicted maternal countermeasure, in evolutionary time, was to adjust upward the amount of CG required to maintain pregnancy which, in turn, favored increased embryonic production of CG. The predicted, and observed, outcome of this coevolution of proxy and test is an escalation to high levels of production of placental hormones with decreasing maternal responsiveness [45]. This coevolutionary arms race explains the ‘bewildering’ lack of coherence in placental structure and physiology across mammals [47].

Concluding Remarks: Why Are Placentas Like Birds-of-Paradise?

Placentas must achieve a simple task: passing goods between mother and fetus. It should be a piece of cake. But placentas are wildly diverse across mammals, ‘evolving and experimenting at a disturbingly rapid pace’ [47], reminiscent of the bizarre diversity of birds-of-paradise [64]. Pregnancy in placental mammals is a case of antagonistic coevolution between signaler and receiver. The sensory onslaughts of placental hormones in embryo choice, and fantastic ornaments in mate choice, are outcomes of arms races between examiners and examinees in the context of high-stakes testing.

Elaborate mating displays and hormone signals during pregnancy are analogous to burdensome socio-economic evaluations and regulations: in all cases, a fundamental conflict of interest promotes rapid escalation and elaboration between examiners and examinees. Many researchers of mate choice have described this arms race, but the framework has not been widely adopted. Researchers investigating honest signals should expect, and look for, deception and arms races. Embryo selection, described herein, supports the theory that one source of rapid elaboration and escalation in evolutionary biology is the proxy treadmill.

In addition, embryo–maternal relations are not typically considered to be conflictual, nor even are hormones produced during pregnancy typically considered as signals between two individuals with differing evolutionary interests. Research on mate choice provides a strong framework to interpret pregnancy as a high-stakes testing bed for embryos, rife with manipulation and deception.

The proxy treadmill framework offers several predictions. We predict that in lineages where placental hormones have evolved to manipulate maternal physiology (e.g., placental CG as a mimic of maternal LH to extend pregnancy and prevent termination), maternal receptors will demonstrate rapid evolution of the binding sites to resist embryonic manipulation. Likewise, we hypothesize that females who judge males on the quality of sensory displays will have evolved heightened sensory perception over relevant domains (e.g., discriminating yellow wavelengths if males have a yellow color display). We suggest that in lineages where the stakes are particularly high (e.g., mating success rates are very low), the choosing sex will come to rely on multiple independent quality signals as predicted by Campbell [5] in the socio-economic context. We also predict that enhancers should frequently evolve to ‘honest’ signals (those with some well-understood physiological link to quality). For example, carotenoid-colored feathers should be accompanied by structural modifications that enhance the color. Finally, we predict that illusions will tend to evolve in quality signaling pathways, such as placental hormones that mimic maternal hormones or optical illusions [90,91] in visual displays.

Outstanding Questions

How prevalent are deceptive signals in surveys of wild animal populations?

What sorts of signals are more resistant to deception?

Do successful tests maintain value by eliminating the bottom of the distribution or by discriminating at the top?

Are preferred signals with a long history in a clade (‘basal’ signals) more likely to be: (i) inherently honest and resistant to deception; or (ii) aesthetic residue (not condition-dependent), having uncoupled from true quality long ago by the proxy treadmill?

Do species with more exaggerated signaling traits also have more numerous signals of quality (in both mate choice and embryo selection)?

We suggest that Campbell's law applies to biology in many additional contexts related to embryo choice. For example, some marsupials give birth to more young than the pouch contains nipples, causing a high-stakes race to the teat. Northern quoll *Dasyurus hallucatus* have eight teats but up to 17 neonates immediately after the birth scramble; in one case, all losers in the race for a teat had disappeared 12 hours after birth [65]. Do marsupial neonates show Campbellian adaptations to this high-stakes test? Plants, too, frequently initiate many more seeds than will ultimately come to maturity [66]. The Australian plants *Banksia* spp. are a particularly dramatic example: in *Banksia elegans*, fewer than one in a thousand zygotes develop into a provisioned seed [67]. Is this a (very) high-stakes quality test? Indeed, in the related *Banksia spinulosa*, researchers found that the parent plant was selectively aborting lower-quality embryos (i.e., aborting self-fertilized embryos, which are inferior to outcrossed embryos [68]).

Both mate choice and embryo choice are selective sieves determining whose genes persist and whose disappear. The stakes could not be higher. And as every poker player will tell you, when the stakes are high, you cannot trust anyone.

Acknowledgments

We are very grateful to two anonymous reviewers for their feedback. We would like to thank John Christy for particularly useful feedback, including suggesting the terminology 'aesthetic residue' and suggesting multiple specific predictions and outstanding questions. We are grateful to Arvid Ågren, Pavitra Muralidhar, Carl Veller, and Inbar Mayaan for comments on this manuscript. D.E.M.'s research is conducted with Government support under and awarded by DoD, Air Force Office of Scientific Research, National Defense Science and Engineering Graduate (NDSEG) Fellowship, 32 CFR 168a. D.E.M. is also supported by a Theodore H. Ashford Graduate Fellowship.

References

- Goldstein, H. (2004) Education for all: the globalization of learning targets. *Comp. Educ.* 40, 7–14
- Jacob, B. (2005) Accountability, incentives and behavior: evidence from school reform in Chicago. *J. Public Econ.* 89, 761–796
- Macfadyen, L.P. et al. (2014) Embracing big data in complex educational systems: the learning analytics imperative and the policy challenge. *Res. Pract. Assess.* 9, 17–28
- Goodhart, C.A.E. (1983) *Monetary Theory and Practice: The UK-Experience*, Macmillan International Higher Education
- Campbell, D.T. (1979) Assessing the impact of planned social change. *Eval. Program Plann.* 2, 67–90
- Strathern, M. (1997) 'Improving ratings': audit in the British University system. *Eur. Rev.* 5, 305–321
- Haustein, S. et al. (2016) Tweets as impact indicators: examining the implications of automated "bot" accounts on Twitter. *J. Assoc. Inf. Sci. Technol.* 67, 232–238
- Maynard-Smith, J. and Harper, D. (2003) *Animal Signals*, Oxford University Press
- Johnstone, R.A. and Grafen, A. (1993) Dishonesty and the handicap principle. *Anim. Behav.* 46, 759–764
- Guilford, T. and Dawkins, M.S. (1991) Receiver psychology and the evolution of animal signals. *Anim. Behav.* 42, 1–14
- Krebs, J.R. and Dawkins, R. (1984) Animal signals: mindreading and manipulation. In *Behavioural Ecology: An Evolutionary Approach* (Krebs, J.R. and Davies, N., eds), pp. 380–402, Blackwell Scientific Publication
- Dawkins, M.S. and Guilford, T. (1991) The corruption of honest signalling. *Anim. Behav.* 41, 865–873
- Hill, G.E. (1994) Trait elaboration via adaptive mate choice: sexual conflict in the evolution of signals of male quality. *Ethol. Ecol. Evol.* 6, 351–370
- Endler, J.A. and Basolo, A.L. (1998) Sensory ecology, receiver biases and sexual selection. *Trends Ecol. Evol.* 13, 415–420
- Dawkins, M.S. and Guilford, T. (1996) Sensory bias and the adaptiveness of female choice. *Am. Nat.* 148, 937–942
- Prum, R.O. (2012) Aesthetic evolution by mate choice: Darwin's really dangerous idea. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 2253–2265
- Lande, R. (1981) Models of speciation by sexual selection on polygenic traits. *Proc. Natl. Acad. Sci. U. S. A.* 78, 3721–3725
- Kirkpatrick, M. (2006) Sexual selection and the evolution of female choice. *Evolution (N. Y.)* 36, 1–12
- Kirkpatrick, M. and Ryan, M.J. (1991) The evolution of mating preferences and the paradox of the lek. *Nature* 350, 33–38
- Fisher, R.A. (1999) *The Genetical Theory of Natural Selection: A Complete Variorum Edition*, Oxford University Press
- Hill, G.E. (2015) Sexiness, individual condition, and species identity: the information signaled by ornaments and assessed by choosing females. *Evol. Biol.* 42, 251–259
- Zahavi, A. (1975) Mate selection—a selection for a handicap. *J. Theor. Biol.* 53, 205–214
- Walker-Bolton, A.D. and Parga, J.A. (2017) "Stink flirting" in ring-tailed lemurs (*Lemur catta*): male olfactory displays to females as honest, costly signals. *Am. J. Primatol.* 79, e22724
- Olson, V.A. and Owens, I.P.F. (1998) Costly sexual signals: are carotenoids rare, risky or required? *Trends Ecol. Evol.* 13, 510–514
- Weaver, R.J. et al. (2018) Carotenoid metabolism strengthens the link between feather coloration and individual quality. *Nat. Commun.* 9, 73
- Weaver, R.J. et al. (2017) What maintains signal honesty in animal colour displays used in mate choice? *Philos. Trans. R. Soc. B Biol. Sci.* 372, 20160343
- Fisher, R.A. (1915) The evolution of sexual preference. *Eugen. Rev.* 7, 184–192
- Williams, G.C. (1966) Natural selection, the costs of reproduction, and a refinement of Lack's principle. *Am. Nat.* 100, 687–690
- Dawkins, R. and Krebs, J.R. (1978) Animal signals: information or manipulation. In *Behavioural Ecology: An Evolutionary Approach* (Krebs, J.R. and Davies, N., eds), pp. 282–309, Blackwell Scientific Publication
- Wickler, W. (1965) Mimicry and the evolution of animal communication. *Nature* 208, 519
- Wickler, W. (1968) *Mimicry in Plants and Animals*, Weidenfeld & Nicholson
- Parker, G.A. (2006) Sexual conflict over mating and fertilization: an overview. *Philos. Trans. R. Soc. London B Biol. Sci.* 361, 235–259
- Burk, T. (1988) Acoustic signals, arms races and the costs of honest signalling. *Florida Entomol.* 71, 400

34. Ghislandi, P.G. *et al.* (2017) Silk wrapping of nuptial gifts aids cheating behaviour in male spiders. *Behav. Ecol.* 28, 744–749
35. Burk, T. (1981) Signaling and sex in acalyptate flies. *Florida Entomol.* 64, 30–43
36. Backwell, P.R.Y. *et al.* (2000) Dishonest signalling in a fiddler crab. *Proc. R. Soc. London B Biol. Sci.* 267, 719–724
37. Finetter-Rosenbluh, I. and Levinson, M. (2015) What is wrong with grade inflation (if anything)? *Philos. Inq. Educ.* 23, 3–21
38. Stroebe, W. (2016) Why good teaching evaluations may reward bad teaching: on grade inflation and other unintended consequences of student evaluations. *Perspect. Psychol. Sci.* 11, 800–816
39. Macklon, N.S. and Brosens, J.J. (2014) The human endometrium as a sensor of embryo quality. *Biol. Reprod.* 91, 1–8
40. Buchholz, J.T. (1922) Developmental selection in vascular plants. *Bot. Gaz.* 73, 249–286
41. Haig, D. (1990) Brood reduction and optimal parental investment when offspring differ in quality. *Am. Nat.* 136, 550–556
42. Wilcox, A.J. *et al.* (1988) Incidence of early loss of pregnancy. *N. Engl. J. Med.* 319, 189–194
43. Roberts, C.J. and Lowe, C.R. (1975) Where have all the conceptions gone? In *Problems of Birth Defects* (Persaud, T.V.N., ed), pp. 148–150, Springer
44. Haig, D. (1993) Genetic conflicts in human pregnancy. *Q. Rev. Biol.* 68, 495–532
45. Haig, D. (1996) Altercation of generations: genetic conflicts of pregnancy. *Am. J. Reprod. Immunol.* 35, 226–232
46. Haig, D. (2019) Cooperation and conflict in human pregnancy. *Curr. Biol.* 29, R455–R458
47. Roberts, R.M. (1996) Interferon- τ and pregnancy. *J. Interf. Cytokine Res.* 16, 271–273
48. Csapo, A.I. *et al.* (1972) The significance of the human corpus luteum in pregnancy maintenance. I. Preliminary studies. *Am. J. Obstet. Gynecol.* 112, 1061–1067
49. Maston, G.A. and Ruvolo, M. (2002) Chorionic gonadotropin has a recent origin within primates and an evolutionary history of selection. *Mol. Biol. Evol.* 19, 320–335
50. Sherman, G.B. *et al.* (1992) A single gene encodes the beta-subunits of equine luteinizing hormone and chorionic gonadotropin. *Mol. Endocrinol.* 6, 951–959
51. Casarini, L. *et al.* (2018) Two hormones for one receptor: evolution, biochemistry, actions, and pathophysiology of LH and hCG. *Endocr. Rev.* 39, 549–592
52. Cole, L.A. (2012) hCG, five independent molecules. *Clin. Chim. Acta* 413, 48–65
53. Yen, S.S. *et al.* (1968) Disappearance rates of endogenous luteinizing hormone and chorionic gonadotropin in man. *J. Clin. Endocrinol. Metab.* 28, 1763–1767
54. Smith, P.L. *et al.* (1993) Equine lutropin and chorionic gonadotropin bear oligosaccharides terminating with SO₄-4-GalNAc and Sia alpha 2, 3Gal, respectively. *J. Biol. Chem.* 268, 795–802
55. Haig, D. (1999) Genetic conflicts of pregnancy and childhood. In *Evolution in Health and Disease* (Stearns, S.C., ed.), pp. 77–90, Oxford University Press
56. Zeleznik, A.J. (1998) In vivo responses of the primate corpus luteum to luteinizing hormone and chorionic gonadotropin. *Proc. Natl. Acad. Sci. U. S. A.* 95, 11002–11007
57. Henke, A. and Gromoll, J. (2008) New insights into the evolution of chorionic gonadotropin. *Mol. Cell. Endocrinol.* 291, 11–19
58. Furuhashi, M. *et al.* (1995) Fusing the carboxy-terminal peptide of the chorionic gonadotropin (CG) beta-subunit to the common alpha-subunit: retention of O-linked glycosylation and enhanced in vivo bioactivity of chimeric human CG. *Mol. Endocrinol.* 9, 54–63
59. Risma, K.A. *et al.* (1995) Targeted overexpression of luteinizing hormone in transgenic mice leads to infertility, polycystic ovaries, and ovarian tumors. *Proc. Natl. Acad. Sci. U. S. A.* 92, 1322–1326
60. Joshi, L. *et al.* (1995) Recombinant thyrotropin containing a β -subunit chimera with the human chorionic gonadotropin- β carboxy-terminus is biologically active, with a prolonged plasma half-life: role of carbohydrate in bioactivity and metabolic clearance. *Endocrinology* 136, 3839–3848
61. Lapolt, P.S. *et al.* (1992) Enhanced stimulation of follicle maturation and ovulatory potential by long acting follicle-stimulating hormone agonists with extended carboxyl-terminal peptides. *Endocrinology* 131, 2514–2520
62. Fares, F. *et al.* (2007) Development of a long-acting erythropoietin by fusing the carboxyl-terminal peptide of human chorionic gonadotropin β -subunit to the coding sequence of human erythropoietin. *Endocrinology* 148, 5081–5087
63. Fares, F.A. *et al.* (1992) Design of a long-acting follitropin agonist by fusing the C-terminal sequence of the chorionic gonadotropin beta subunit to the follitropin beta subunit. *Proc. Natl. Acad. Sci. U. S. A.* 89, 4304–4308
64. Laman, T. and Scholes, E. (2012) *Birds of Paradise: Revealing the World's Most Extraordinary Birds*, National Geographic Books
65. Nelson, J.E. and Gemmill, R.T. (2003) Birth in the northern quoll, *Dasyurus hallucatus* (Marsupialia: Dasyuridae). *Aust. J. Zool.* 51, 187–198
66. Stephenson, A.G. (1981) Flower and fruit abortion: proximate causes and ultimate functions. *Annu. Rev. Ecol. Syst.* 12, 253–279
67. Lamont, B.B. and Barrett, G.J. (1988) Constraints on seed production and storage in a root-suckering *Banksia*. *J. Ecol.* 76, 1069–1082
68. Vaughton, G. and Carthew, S.M. (1993) Evidence for selective fruit abortion in *Banksia spinulosa* (Proteaceae). *Biol. J. Linn. Soc.* 50, 35–46
69. McKee, M. (2004) Not everything that counts can be counted; not everything that can be counted counts. *BMJ* 328, 153
70. Braganza, O. (2018) Proxyeconomics, an agent based model of Campbell's law in competitive societal systems. *arXiv Prepr.* arXiv:1803.00345
71. Fire, M. and Guestrin, C. (2018) Over-Optimization of Academic Publishing Metrics: Observing Goodhart's Law in Action. *arXiv*, 1809.07841
72. Head, M.L. *et al.* (2015) The extent and consequences of p-hacking in science. *PLoS Biol.* 13, e1002106
73. Nguyen, P.K. *et al.* (2016) A systematic comparison of sugar content in low-fat vs regular versions of food. *Nutr. Diabetes* 6, e193
74. Clark, H.L. and Backwell, P.R.Y. (2016) Male mating success in a fiddler crab: a lesson in sample sizes. *J. Ethol.* 34, 119–126
75. Christy, J.H. and Rittschof, D. (2010) Deception in visual and chemical communication in crustaceans. In *Chemical Communication in Crustaceans*, pp. 313–333, Springer
76. Gasparini, C. *et al.* (2013) Do unattractive friends make you look better? Context-dependent male mating preferences in the guppy. *Proc. Biol. Sci.* 280, 20123072
77. Funk, D.H. and Tallamy, D.W. (2000) Courtship role reversal and deceptive signals in the long-tailed dance fly, *Rhaphomyia longicauda*. *Anim. Behav.* 59, 411–421
78. Albo, M.J. *et al.* (2011) Worthless donations: male deception and female counter play in a nuptial gift-giving spider. *BMC Evol. Biol.* 11, 329
79. Preston-Mafham, K.G. (1999) Courtship and mating in *Empis (Xanthempis) trigramma* Meig., *E. tessellata* F. and *E. (Polyblepharis) opaca* F. (Diptera: Empididae) and the possible implications of 'cheating' behaviour. *J. Zool.* 247, 239–246
80. Gosling, L.M. and Roberts, S.C. (2001) Scent-marking by male mammals: cheat-proof signals to competitors and mates. *Adv. Study Behav.* 30, 169–217
81. Alberts, A.C. (2002) Constraints on the design of chemical communication systems in terrestrial vertebrates. *Am. Nat.* 139, S62–S89
82. Sharpe, L.L. (2015) Handstand scent marking: height matters to dwarf mongooses. *Anim. Behav.* 105, 173–179
83. Sharpe, L.L. *et al.* (2012) Handstand scent marking in the dwarf mongoose (*Helogale parvula*). *Ethology* 118, 575–583
84. McGuire, B. *et al.* (2018) Urine marking in male domestic dogs: honest or dishonest? *J. Zool.* 306, 163–170
85. Buchinger, T.J. *et al.* (2013) Evidence for a receiver bias underlying female preference for a male mating pheromone in sea lamprey. *Proc. Biol. Sci.* 280, 20131966
86. Christy, J.H. (1995) Mimicry, mate choice, and the sensory trap hypothesis. *Am. Nat.* 146, 171–181