

UNIVERZITA KARLOVA

Přírodovědecká fakulta

Katedra filosofie a dějin přírodních věd

Teoretická a evoluční biologie



Mgr. Julie Nekola Nováková

Prediktory a koreláty altruismu u lidí z evolučního hlediska

Predictors and correlates of altruism in humans from the evolutionary viewpoint

Disertační práce

Vedoucí práce: prof. RNDr. Jaroslav Flegr, CSc.

Praha, 2025

Ráda bych poděkovala především svému školiteli, profesorovi Jaroslavu Flegrovi, za vedení a trpělivost, své rodině za vytrvalou podporu a všem kolegům za inspiraci a spolupráci. Bez nikoho z vás by tato práce nespátřila světlo světa.

Julie Nekola Nováková

Prohlašuji, že jsem svou závěrečnou práci zpracovala samostatně a že jsem řádně citovala všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu v rámci této ani jiné instituce. Studie uvedené v této práci byly podpořeny Grantovou agenturou Univerzity Karlovy (GAUK 1502018).

V Praze dne 31. 1. 2025



Julie Nekola Nováková

ANOTACE

Cílem této dizertační práce bylo experimentálně ověřit předpovědi vyplývající z několika hlavních teorií evoluce altruismu (velmi zjednodušeně řečeno, bezpodmínečného pomáhání u lidí i zvířat). Pro vysvětlení této „evoluční záhady“ bylo předloženo několik možných vysvětlení, včetně recipročního altruismu, pohlavního výběru, a dokonce i manipulace hostitele mikrobiomem. Právě to jsou hypotézy, které jsem testovala ve své práci. V ní nejprve podávám přehled o současném stavu poznání a stručný úvod k přiloženým publikacím. Následují čtyři práce, které jsem dokončila během svého doktorského studia a které se zabývají altruismem jako nákladnou signalizací v rámci pohlavního výběru, hormonálními vlivy na altruismus a jejich důsledky pro ultimátní (evoluční) příčiny tohoto chování, zdravotním stavem a dalšími potenciálními prediktory altruismu vyplývajících z teorie přímé reciprocity a nakonec novější hypotézou mikrobiální manipulace hostitele k altruismu.

SYNOPSIS

The goal of this thesis was to experimentally test predictions stemming from several major theories of evolution of altruism (put very simply, unconditional helping in humans as well as non-human animals). Multiple potential explanations have been put forth to explain this “evolutionary enigma”, including direct reciprocity, sexual selection, and even host manipulation by its microbiota. These were the hypotheses tested in my work. In the thesis, I first review the state of the art and provide an introduction to the enclosed publications. This is followed by the four papers that I completed during my doctoral studies, tackling altruism as costly signaling in sexual selection, hormonal influences of altruism and their implications for the ultimate (evolutionary) causes of this behavior, health and other potential predictors of altruism stemming from the direct reciprocity theory, and finally the relatively new hypothesis of host manipulation towards altruism.

ABSTRAKT

Altruismus již dlouho představuje svého druhu evoluční záhadu, jelikož toto chování – kdy jedinec pomáhá druhému na vlastní náklady – se z hlediska přírodního výběru jeví jako nevýhodné. Otázka zní: jak mohou jedinci (nebo jejich alely) altruismem maximalizovat svou fitness, aby se toto chování mohlo v populaci dlouhodobě udržet?

Evoluci altruismu se snaží vysvětlit více hypotéz, přičemž nejčastěji se uvažuje o příbuzenském výběru, reciprocitě, pohlavním výběru a interdemeckém/víceúrovňovém výběru. Každá z těchto hypotéz poskytuje různé předpovědi, co by mělo ovlivňovat výskyt a rozsah altruistického chování v dané populaci a u konkrétních jedinců. Přestože kolem těchto teorií existuje rozsáhlá literatura, jen málo prací se je pokusilo navzájem přímo porovnat prostřednictvím experimentálního testování.

Cílem této práce je zaplnit tuto mezeru a prostřednictvím několika studií provedených mnou ve spolupráci s kolegy(němi) otestovat různé předpovědi, které s sebou nesou teorie recipročního altruismu, nákladné signalizace v kontextu pohlavního výběru a manipulační hypotézy. Studie byly provedeny jako laboratorní experimenty i rozsáhlejší online průzkumy a převážně zahrnovaly experimentální hry (hry na diktátora a ultimátum; DG a UG) jako nepřímé ukazatele altruismu. V souladu s předchozími pracemi výsledky zahrnutých studií naznačují, že altruismus u lidí je komplexní jev, jehož evoluci mohlo pravděpodobně v různé míře formovat více faktorů současně.

Ve Studii 1 bylo mým cílem zjistit, zda priming fotografiemi (ne)atraktivních příslušníků opačného pohlaví ovlivňuje chování účastníků v DG a UG, které se často používají k hodnocení altruistických preferencí, a zda altruističtější se chovající lidé mají více romantických a sexuálních partnerů. Muži vystavení primingu atraktivity činili vyšší nabídky v UG a účastníci obou pohlaví, kteří byli náročnější ve své minimální přijatelné nabídce v UG, měli více reportovaných sexuálních partnerů, což lze s jistou opatrností interpretovat ve prospěch nákladné signalizace v rámci sexuálního výběru.

Studie 2 se zabývala vztahem salivárních hladin testosteronu, resp. kortizolu a altruistického chování v DG. Zjistili jsme, že muži s vyšší hladinou testosteronu poskytovali štedřejší příspěvky, což je také v souladu s předpověďmi vyplývajícími z pohlavního výběru.

Testování předpovědí vyplývajících z podmínek přímé reciprocity bylo cílem Studie 3. Na základě kognitivní náročnosti přímé reciprocity a také pravděpodobnosti budoucích interakcí s ostatními jedinci se dala očekávat pozitivní korelace altruistického chování s pamětí, kognitivními schopnostmi, nižším časovým diskontováním a lepším zdravím; nicméně jsme nezjistili žádnou významnou souvislost altruismu s pamětí a zdravím a převážně statisticky nevýznamné trendy v případech časového diskontování. Stejně tak míry kognitivní reflexe a inteligence neměly žádný vztah k dáváním v DG. V rozporu s předpoklady však negativně korelovaly s nabídkami a minimálními přijatelnými nabídkami v UG.

Nakonec cílem Studie 4 bylo ověřit předpovědi manipulační hypotézy (manipulace hostitele směrem k altruistickému chování); naše data ze dvou různých populací tuto hypotézu nepodpořila.

Tyto studie, ve vzájemném kontextu a v souladu s důkazy z dřívějších prací, naznačují, že ačkoli v evoluci altruismu u lidí pravděpodobně hrálo roli více faktorů, při testování předpovědí různých hypotéz bylo nejsnazší rozeznat vliv nákladné signalizace. To naznačuje, že pohlavní výběr by mohl být nejvýznamnějším selekčním faktorem altruismu u lidí v současnosti. Přesto by bylo přílišným zjednodušením mu automaticky přisuzovat největší celkový vliv a v průběhu evoluce člověka mohlo působit mnoho možných cest, které mohly působit souběžně a jejichž roli lze v současnosti jen obtížně rozlišit.

Klíčová slova

altruismus; kooperace; nákladná signalizace; evoluční psychologie; experimentální hry; manipulační hypotéza; reciprocita

ABSTRACT

Altruism has long presented an evolutionary enigma, as this behavior – where an individual helps another at its own cost – appeared disadvantageous from the viewpoint of natural selection. The question stood: how could individuals (or their alleles) still maximize their fitness through altruism, so that this behavior could persist in a population in the long term?

Multiple hypotheses have aimed to explain the evolution of altruism, the most widely considered being kin selection, reciprocity, sexual selection and interdemic/multilevel selection. Each of the hypotheses gives different predictions as to what should influence the occurrence and extent of altruistic behavior in a given population and specific individuals. However, while extensive literature exists on these theories, few works have attempted to directly compare them through testing.

The aim of this thesis is to fill the gap, and through several studies conducted by myself and my colleagues, to test different predictions yielded by the theories of reciprocal altruism, costly signaling in the sexual selection context, and host manipulation hypothesis. The studies were conducted as laboratory experiments as well as larger-scale online surveys and predominantly included experimental games (Dictator Game and Ultimatum Game; DG and UG) as proxies of altruism. In line with previous works, the results of the included studies suggest that altruism in humans is a complex, multifaceted phenomenon whose evolution may have been shaped by multiple factors simultaneously to a varying extent.

In Study 1, my aim was to investigate whether priming by photos of (un)attractive members of the opposite sex influences contributions in the DG and UG, oft-used to assess altruistic preferences, and whether more altruistically behaving people have more self-reported romantic and sexual partners. Men exposed to the attractiveness priming made larger UG offers and participants of both sexes who were more demanding in their minimum acceptable offer in the UG had more self-reported sexual partners, which can be with some caution interpreted in favor of costly signaling in sexual selection.

Study 2 looked at connections of salivary levels of testosterone and cortisol and altruistic behavior in the DG, and we found that high-testosterone men made more generous contributions, also in line with predictions stemming from sexual selection.

Testing predictions stemming from the conditions of direct reciprocity was the aim of Study 3. Based on the cognitive requirements of reciprocity and likelihood of future interactions with others, a positive correlation of altruistic behavior with memory, cognitive abilities, lower temporal discounting and better health could be expected; nevertheless, we found no notable connection of altruism with memory and health, and non-significant trends in case of temporal discounting. Likewise, measures of cognitive reflection and intelligence had no relationship with giving in the DG. However, they correlated negatively with offers and minimum acceptable offers in the UG, in contrast to predictions.

Finally, the aim of Study 4 was to test predictions of the hypothesis of host manipulation towards altruistic behavior; our data from two different samples did not support for this hypothesis.

The studies, synthesized together and in line with evidence from existing body of work, suggest that while multiple factors were likely at play in the evolution of altruism in humans, it was easiest to discern the effect of costly signaling in testing predictions of different hypotheses. This suggests that sexual selection might be the most prominent selection factor for altruism throughout the course of present-day human life. Nevertheless, it would be too much of a simplification to automatically attribute the greatest overall influence to it, and many possible pathways that could have acted simultaneously over the course of human evolution and may be difficult to discern in the present day.

Keywords: altruism; cooperation; costly signaling; evolutionary psychology; experimental games; host manipulation; reciprocity

TABLE OF CONTENTS

1.	Background: The “enigma” of altruism	11
1.1.	Why search for evolutionary roots of altruism?	12
1.2.	Defining altruism	14
1.3.	Explanations of altruism, and how to test them	15
1.3.1.	Interdemic selection	16
1.3.2.	Kin selection	18
1.3.3.	Reciprocity	20
1.3.4.	Sexual selection, or behavioral peacock tails	26
1.3.5.	Microbial manipulation	28
1.4.	Where do roots of altruism figure in other disciplines?	29
2.	Testing the foundations of altruism in humans	31
2.1.	Introduction	31
2.2.	Study 1: Sexual selection towards altruism	36
2.3.	Study 2: Testosterone and cortisol connection with altruism	36
2.4.	Study 3: Health and altruism – testing foundations of reciprocity	37
2.5.	Study 4: Host manipulation towards altruism	38
3.	Conclusions	39
4.	References	40
5.	Supplement	55
5.1.	Looking Like a Million Dollars: Does Attractiveness Priming Increase Altruistic Behavior in Experimental Games?	
5.2.	Generosity as a status signal: Higher-testosterone men exhibit greater altruism in the dictator game	
5.3.	Testing Trivers: Discerning the role of reciprocity in the evolution of altruism	
5.4.	Testing the microbiome-altruism hypothesis: Antibiotic intake doesn’t impair altruistic behavior	

LIST OF PAPERS INCLUDED IN THIS DOCTORAL THESIS

- **Novakova, J.,** Machová, K., Sýkorová, K., Zíka, V., & Flegr, J. (2021). Looking Like a Million Dollars: Does Attractiveness Priming Increase Altruistic Behavior in Experimental Games?. *Frontiers in Psychology*, 1974.
 - JNN, as the lead author, designed the experiment, participated in running the experimental sessions and data collection, led the data analysis and wrote the paper, with contributions and comments from other authors.
- **Novakova, J.,** Tureček, P., Machová, K., Sýkorová, K., Zíka, V., & Flegr, J. (2024). Generosity as a status signal: Higher-testosterone men exhibit greater altruism in the dictator game. *Evolution and Human Behavior*, 45(5), 106615.
 - JNN, as the lead author, designed the experiment, participated in running the experimental sessions and data collection, and wrote the paper apart from sections of the methods and supplement concerned with the statistical analysis (done by PT), with contributions and comments from other authors.
- **Novakova, J.,** Machová, K., Zíka, V., & Flegr, J. (under review). Testing Trivers: Discerning the role of direct reciprocity in the evolution of altruism. Under review in *Behavioural Processes*.
 - JNN, as the lead author, designed the experiment, participated in running the experimental sessions and data collection, led the data analysis and wrote the paper, with contributions and comments from other authors.
- **Novakova, J.,** & Flegr, J. (under review). Testing the microbiome-altruism hypothesis: Antibiotic intake doesn't impair altruistic behavior. Under review in *Evolution and Human Behavior*.
 - JNN, as the lead author, suggested the study, participated in the data collection, co-conducted the data analysis with JF and wrote the paper, with important contributions and comments from JF.

LIST OF THEMATICALLY RELEVANT PAPERS NOT INCLUDED IN THE THESIS

- Vranka, M., Frollová, N., Pour, M., **Novakova, J.**, & Houdek, P. (2019). Cheating customers in grocery stores: A field study on dishonesty. *Journal of Behavioral and Experimental Economics*, 83, 101484.
 - JNN participated in data collection and writing of the paper.
- Bouwmeester, S., Verkoeijen, P. P., Aczel, B., Barbosa, F., Bègue, L., Brañas-Garza, P., ... **Novakova, J.** ... & Wollbrant, C. E. (2017). Registered replication report: Rand, greene, and nowak (2012). *Perspectives on Psychological Science*, 12(3), 527-542.
 - JNN participated in data collection and analysis of the Czech-based study.
- Houdek, P., **Novakova, J.**, & Stastny, D. (2016). Ultrasociality: when institutions make a difference. *Behavioral & Brain Sciences*, 39.
 - JNN co-wrote the paper together with PH.
- Houdek, P., & **Novakova, J.** (2016). Frozen cultural plasticity. *Behavioral & Brain Sciences*, 39.
 - JNN co-wrote the paper together with PH.
- **Novakova, J.**, & Flegr, J. (2013). How much is our fairness worth? The effect of raising stakes on offers by proposers and minimum acceptable offers in dictator and ultimatum games. *PloS one*, 8(4), e60966.
 - JNN participated in data collection, analyzed the data together with her supervisor JF and wrote the paper, with JF's comments and supervision.

A full list of peer-reviewed papers as well as non-peer-reviewed conference contributions and popular science publications can be accessed in the candidate's CV submitted with this thesis, selected ones also in her [ORCID profile \(0000-0002-1984-9739\)](https://orcid.org/0000-0002-1984-9739).

1. BACKGROUD: THE “ENIGMA” OF ALTRUISM

Charles Darwin was puzzled. How could natural selection result in the production of sterile workers, helping another individual procreate, and moreover looking and behaving much differently from their fertile counterparts? It “at first appeared to me insuperable, and actually fatal to the whole theory” (Darwin 1859, 1872).

Eusociality, a very special case of altruism (put simply, helping another at one’s own expense), presented an enigma to Darwin – though the difference in appearance and behavior of workers in many eusocial species of insects unsettled him even more than the “mere” fact of helping the queen procreate. He returned to altruism, or as he put it, “moral sense”, at more length in *The Descent of Man* (Darwin 1871). It was a riddle: How could such behavior, costly to the altruist, be favored by natural selection?

The attention given to the problem was more than justified. Nature is not quite, in Tennyson’s words, “red in tooth and claw”; instead, we can observe a myriad of examples of both intraspecific and interspecies altruism and cooperation.

In some instances, such behavior can be explained by immediate mutual benefits, such as in the famous example of cleaner fish (though even this may include elements of reputation and reciprocity due to the temptation to cheat; see Bshary & Grutter (2006)). A mutualistic scenario, rendering advantages to be enjoyed by both parties without delay, is easy to explain. However, how do we explain cases where only one side seems to benefit, or the benefits are very much delayed in time? What roots shaped the evolution of altruism?

Darwin put forth a hypothesis we could classify as kin selection mixed with group selection (Darwin 1871); a century after him (and after much debate in the meantime, to which I’ll briefly return in subchapter 1.3.1), William Hamilton (1964a,b) and others formalized different hypotheses as to the selective advantages of altruism.

We shall investigate these influences specifically on altruism in humans in this work. However, a broader background is necessary to understand the potential underpinnings of altruistic behavior. How important was kin selection for the evolution of altruism among non-kin? How often did reciprocity play a role? What evidence points to sexual selection as altruism’s driving force? We are going to briefly look at what is considered altruism under the framework of evolutionary biology, how it may differ from the psychological view, what are the major hypotheses trying to explain the evolution of altruism and how we can empirically test them. Finally, four papers stemming from my doctoral studies at Charles University

follow this introduction, presenting specific cases of testing the proposed evolutionary underpinnings of altruistic behavior.

1.1 Why search for evolutionary roots of altruism?

One common source of confusion when approaching people from outside of evolutionary biology are the questions of whether altruism can really be beneficial to the altruist and what “explains” altruism.¹

Especially psychologists often resist the notion that altruistic behavior may ultimately benefit the altruist (“it’s not altruism, then!”, many would argue); some may claim that if we gain something, the behavior in question can no longer be considered an act of altruism. That is because of the conflation of different levels of explanation, proximate and ultimate – so let us have a look at these.

When met with the theories of reciprocity, kin selection, costly signaling and other ones, a psychologist could counter: “But we behave altruistically because it *makes us feel good*. That’s the reason.” In the same vein, a neurobiologist might object: “No, the reason behind it is increased activity in multiple limbic and cortical regions.”

All of these can be true at the same time; they are just different levels of explanation. To make sure there is no confusion about what level we are trying to explain here, let us make a very brief detour to Niko Tinbergen’s famous four questions (1963)².

In asking why some behavior exists, we can divide the answers into proximate (the “how”) and ultimate (the “why”), and further into historical and present-moment explanations: we are therefore left with four kinds of questions. Table 1 on the following page illustrates this important distinction on the example of altruism.

¹ This kind of confusion is not new; in fact, it has been with us ever since Darwin published his seminal theory, as shown by an interesting paper on “The Evolution of Morality” by James Seth (1889). Over half a century later, Keith (1948) writes in the British Medical Journal: “I believe that one of the central theses of the book – namely, that “altruism signifies a complete abandonment of the evolutionary outlook” (p. 72)-is not only immoral but false. It may have been true in palaeolithic times; it is not true now. Those who practise altruism in our own society accumulate less money than the egoists. But the poor breed quicker than the rich.”

² Or should we say Huxley-Tinbergen’s questions? Tinbergen (1963) himself wrote: “Huxley likes to speak of “the three major problems of Biology”: that of causation, that of survival value, and that of evolution – to which I should like to add a fourth, that of ontogeny.”

		Temporal perspective	
		Present-moment perspective	Historical perspective
Causal perspective	Proximate (How does it work?)	Mechanism (What neural, hormonal and other correlates and proximate psychological causes shape altruism?)	Ontogeny (How does altruism develop throughout the life of an individual, from the genetic level to environmental influences and psychological ontogeny?)
	Ultimate (Why does it work?)	Adaptation (What is the current adaptive value of altruism? Why is there selection towards it?)	Phylogeny (How did altruism evolve in different branches of life, and how did it change throughout time?)

Table 1. ALTRUISM IN LIGHT OF TINBERGEN’S FOUR QUESTIONS. The table summarizes different questions that can be asked about the origins of altruism, depending on one’s causal and temporal perspective, based on Tinbergen (1963).

Therefore, a psychologist may be entirely correct when ascribing altruism to the “warm glow” feelings (e.g. Grossman & Levy 2024, Hartmann et al. 2017), just the same as a neurobiologist may be correct when ascribing it to activity in the nucleus accumbens, anterior cingulate cortex, medial prefrontal cortex and other regions of the brain (Filkowski et al. 2017). But although these explanations are also products of evolutionary history, we do not dive deeper into them here; suffice to say that they exist. What we are chiefly interested in here is the *adaptive value* of altruism – its survival benefits.

That brings us back to the question of whether survival benefits of behavior make it non-altruistic. The short answer is no. On the proximate level, altruism doesn’t seem to yield any benefits to the altruist and absolutely can be motivated by genuinely wanting to help another individual, even at one’s own cost. On the ultimate level, it can still increase one’s chance of procreating, the progeny’s survival, and/or transmission of certain alleles.

To conclude this brief detour, altruistic motivation doesn’t preclude altruism from having an adaptive value. Any behavior that has persisted over evolutionarily meaningful time must have – or at least until recently have had (meaning that it would be a postadaptation; see Flegr 2023a) – one itself, unless we consider hypothetical special cases such as linkage to loci coding traits whose adaptive value outweighs the cost of the linked, now maladaptive behavior.

Having cleared the issue of different levels of explanation, we can move towards answering the question “what is altruism”, before embarking on the journey to the different hypotheses proposed to explain its existence.

1.2 Defining altruism

To quote the entry ‘Biological Altruism’ in the Stanford Encyclopedia of Philosophy (Okasha 2020), “in evolutionary biology, an organism is said to behave altruistically when its behaviour benefits other organisms, at a cost to itself”. That is the most widely used definition of altruism in biology (let us for the moment put aside related definitions from other fields, which are more applicable to their level of scientific reasoning – see the overview of Tinbergen’s questions above).

However, one can find competing definitions. West et al. (2007) state in their work attempting to clarify the terminology surrounding altruism and cooperation: “Altruism: behaviour which is costly to the actor and beneficial to the recipient; in this case and below, cost and benefit are defined on the basis of the lifetime direct fitness consequences of a behaviour.”

Note the “cost ... on the basis of the lifetime direct fitness consequences”. In other words, under West et al.’s definition, altruism is *always detrimental to the altruist*, even in the long view. I am personally puzzled by this definition, as it effectively precludes altruism from possibly having any survival value. Then, its existence would truly be an enigma! However, the authors really do clarify it this way: “if a cooperative behaviour was costly in the short term, but provided some long-term (future) benefit, which outweighed that, it would be mutually beneficial and not altruistic” (West et al. 2007, pg. 419). Under their point of view, “reciprocal altruism is not altruistic”. Elsewhere, they point out that the nonexistent benefit they mention is a *direct fitness benefit* (which really should have been spelled out everywhere it applied); therefore, in their view, kin selection is practically the sole plausible explanation for altruism (as they also discount interdemic and other forms of selection).

While I urge everyone interested in more definitions of altruism and related concepts to read West et al. (2007), their own definition is one I will *not* adhere to in this thesis. I thought it prudent to point out its existence, but it narrows altruism down far too much in my opinion and basically equates it to kin selection, which in the broader view constitutes one of the possible pathways towards altruism. Here, we will use a broader definition of altruism as behavior that benefits the recipient and puts a *varyingly long-term*, but *not ultimate* cost on the donor.

The definition of altruism itself is, unfortunately, not the only one the relevant literature doesn't quite agree on. We need not dwell on most other definitions here, but let us briefly stop at the debate over strong vs. weak altruism and whether to truly consider the latter altruistic. The distinction, introduced by Wilson (1979, 1980), is in the cost the altruist "pays": is the act of altruism costly only relative to other members of the group (ie. does it increase the "absolute" fitness of the altruist as well, but increases the fitness of his or her compatriots *more*, so that *relatively*, he or she is worse off against them)? Then it is a case of weak altruism – the altruist *gains*, but gains *less* than others. If the altruist "loses" in absolute terms, then the act is a case of strong altruism. Does it sound like mere semantics? It's actually not; please remember it well into the next section; you'll need it when we get to interdemic selection in subchapter 1.3.1.

Finally, we need to clear the issue of altruism vs. cooperation. Some researchers use the terms interchangeably, but I would advise against that; most commonly, cooperation is an umbrella term encompassing more types of behavior including altruism (ie. helping another at a cost to oneself), but also mutualism ((semi-)simultaneously mutually advantageous cooperation mostly within but also across species, represented likely by most cases of cooperative hunting, cleaning, some cases of grooming and other types of behavior). As West et al. (2007) point out, behavior classifiable as cooperation should be selected for because of its benefits for the recipient, ruling out "accidental benefits", microbial mutualism, etc.

With that, definitions are off the table as well, and we can move forward to the core of this work: different possible roots of altruism and ways to scientifically test them.

1.3 Explanations of altruism, and how to test them

A number of different explanations for the evolution of altruism, not always mutually exclusive, have been proposed over time. The subsections below serve as a very brief overview of different hypotheses and most importantly ways to test them (which are sometimes not very straightforward or easy to implement), concisely outlining interdemic selection, kin selection, reciprocity, sexual selection and microbial host manipulation.

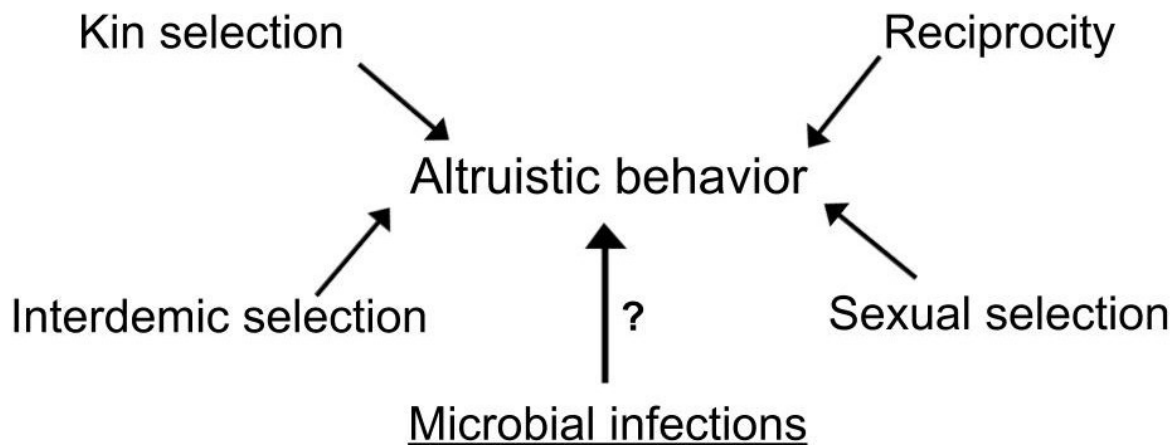


Fig 1. SUGGESTED EVOLUTIONARY PATHWAYS TO ALTRUISM. This diagram presents an overview of the likely processes contributing to altruism. The newest proposed possibility of those, host manipulation by microbial infections, includes a question mark, but although far more evidence supports the other hypotheses, the degree of their contribution is being widely debated. Reprinted from Novakova & Flegr (under review).

1.3.1 Interdemic selection

Interdemic selection, or rather its earlier sibling called group selection, was one of the earliest proposed explanations for the existence of altruism. This possibility was favored early on by Charles Darwin and many of his contemporaries. In Chapter V of *The Descent of Man*, Darwin (1871) supported this theory, but immediately added: “It is extremely doubtful whether the offspring of the more sympathetic and benevolent parents, or of those who were the most faithful to their comrades, would be reared in greater numbers than the children of selfish and treacherous parents belonging to the same tribe. He who was ready to sacrifice his life, as many a savage has been, rather than betray his comrades, would often leave no offspring to inherit his noble nature. ... Therefore, it hardly seems probable, that the number of men gifted with such virtues, or that the standard of their excellence, could be increased through natural selection, that is, by the survival of the fittest; for we are not here speaking of one tribe being victorious over another.” (Darwin 1871)

The seemingly missing individual benefits of unselfishness, which would make it prevalent through natural selection, were what puzzled Darwin. He subsequently made an early stab at reciprocity, in “In the first place, as the reasoning powers and foresight of the

members became improved, each man would soon learn that if he aided his fellow-men, he would commonly receive aid in return,” and at reputation, in “But another and much more powerful stimulus to the development of the social virtues, is afforded by the praise and the blame of our fellow-men.”

Since Darwin, group selection in various forms was invoked by Kropotkin, Allee, Wynne-Edwards and others (reviewed e.g. by Leigh Jr. (2009), Domondon (2013) and Harman (2014)). While many early “naïve” takes on group selection struggle to fit it into strictly darwinian, much less dawkinsian-type selection, it gradually developed into more complex form of group selection, oft-dubbed interdemic selection, focusing on how parameters like group size, relative fitness, fluctuation, migration or environmental stability change the influence of selection on the level of groups. A related higher-level concept that, similarly to the term interdemic selection, came to in the 1970s, is *multilevel selection*. The term encompasses selective pressures at different levels from an allele across the individual to groups, species etc. A good summary of different levels of selection can be found in Flegel (2023).

Proponents of interdemic selection argue that when we account for more complex relationships between the genotype and phenotype and for spatial effects, between-group selection can help explain the evolution of altruism alongside within-group selection (Wilson & Wilson 2007). Experimental evidence for the effectiveness of interdemic selection comes often from microbial growth studies (e.g. Rainey & Rainey 2003), but it has also been suggested for observed lioness’ cooperation in territorial defense (Heinsohn & Packer 1995) or emergence of institutional structures (Bowles et al. 2003).

One of the themes of interdemic selection is the role of “weak” (relative) altruism. Researchers opposing the role of group selection argue that since weakly altruistic acts are individually advantageous, they should be favored by simple individual selection, are not in fact altruistic and do not require any elements of selection between groups. However, since a help-receiving non-altruist is relatively (as well as absolutely) better off than the altruism in the social context, Wilson (1990) argues that it requires an interdemic selection component and models selection for weak and strong altruism on multiple levels. As he points out, “a single trait can be selected at both levels simultaneously”; the degree of influence of each level changes based on parameters of the trait, individual and group.

The debate over the extent of influence of group/interdemic selection in the evolution of altruism as well as in general is still ongoing, and given the clashes over terminology, may not be resolved anytime soon (Shavit & Millstein 2008). New models of

multilevel selection of human altruism are being continuously developed (e.g. Rosas 2008, Simon et al. 2013, Hermsen 2022). However, Kay et al. (2013) argue that most group selection models, supposedly discovering ways in which altruism can emerge and spread in spatially diverse groups without kin selection (see the next subsection), would not actually work without high degree of relatedness among individuals in the modeled population units, and they are to most extent only rediscovering and renaming kin selection.

With that, let us take a closer look on the phenomenon of kin selection itself.

1.3.2 Kin selection

Why do worker bees or ants completely sacrifice their own reproductive success in favor of that of their mother, the queen? Such altruism, along with the workers' markedly different appearance and instinct, puzzled Charles Darwin. But in *On The Origin of Species*, he wrote about the problem in Chapter VIII (Darwin 1872): "This difficulty, though appearing insuperable, is lessened, or, as I believe, disappears, when it is remembered that selection may be applied to the family, as well as to the individual, and may thus gain the desired end."

Kin selection: not formalized yet, but alluded to very specifically. The same questions of eusocial insects that initially puzzled Darwin and led him to this tentative conclusion drove William Hamilton (1964a,b), nearly a century later, to develop his theory of kin selection. At the heart of his theory lies the so-called Hamilton's rule:

$$r*b-c>0$$

Here, r means the degree of relatedness of two individuals (e.g. 0.5 for siblings in sexually reproducing diploid species), b is the benefit acquired by the helped individual (e.g. not starving because of donated food), and c the cost paid by the helping individual (e.g. the loss of the donated food). In our example in the brackets, the benefit is high (avoiding death by starvation), while the cost may not be (if the helping individual is sated and has enough food, sharing it doesn't deprive it of crucial resources). If the benefit times relatedness ratio outweigh the cost, it should be evolutionarily advantageous to help.

The most famous example is the existence of eusociality, where haplodiploidy (in the simplest case) leads to a higher degree of relatedness between the sisters than any of them would have with her own offspring, therefore making it advantageous to help the production of more sisters rather than investing into one's own reproduction, as Hamilton (1964a,b) argue. In reality, the situation is made slightly more complicated by the existence of some brothers (comparatively less related), the fact that in many haplodiploid species, the queen

takes multiple mates or multiple queens coexist in the same nest, and also the fact that the haplodiploidy hypothesis cannot explain eusociality in the diplo-diploid termites; therefore the ancestral lifelong monogamy has also been invoked to primarily explain eusociality, since it allows relatedness to never drop below 0.5, after which conditions – e.g. benefits of cooperative living and brood care – even slightly favoring the production of siblings rather than one’s own offspring can tip evolution in favor of eusociality (e.g. Heinze et al. 1994, Villesen et al., 2002, Boomsma 2007, Hughes et al. 2008, Boomsma et al. 2011, Boomsma et al. 2014); multilevel selection (see subsection 1.3.1 above) has also been invoked (Korb & Heinze 2004). On the other hand, kin selection seems to predict instances of “worker policing” very well (Wenseleers et al. 2006, Ratnieks & Helanterä 2009). For the only eusocial mammals, the naked mole rat and the Damaraland mole rat, kin selection (combined with environmental pressures) has been considered as the main pathway towards their eusociality (Jarvis et al. 1994).

This work is primarily concerned with altruism in humans, and although some authors called humans “ultrasocial” (Richerson & Boyd 1998, Gowdy & Krall 2015) – a term which encompasses species with a high level of social organization and building cities/states/their equivalents –, it is not clear how helpful it is to group humans together with most eusocial animals under this umbrella term. Humankind is definitely not a eusocial species, not anywhere close to transitioning to it; so, while the eusociality debate would itself fill multiple theses, let us abandon it here and zoom out to the broader picture. Some authors have gone as far as saying that “kin selection is the key to altruism” (Foster et al. 2005), even though their work referred primarily to the evolution of eusociality (again) and not altruism *sensu lato*. However, there is a broad body of evidence that kinship influences altruistic and cooperative – but also competitive – behavior in non-human animals (Griffin & West 2003, Smith 2014, Twyman & Gardner 2023), including primates (Silk 2009).

In humans, kinship appears to play a crucial role in altruistic behavior as well. Many observational studies found it of importance in forming alliances, providing support within social networks, forming stable cohabiting units, cooperating during disasters (see reviewed in Madsen et al. 2007), or make a greater social capital investment into their biological children as opposed no non-biological ones (Kuranchie 2021). In hypothetical life-or-death situations, people were more inclined to help closer kin and younger rather than older relatives (Burnstein et al. 1994). Likewise, they preferentially helped kin rather than non-kin of the same social distance in hypothetical economic experiments (Rachlin & Jones 2008).

Experimentally, people were shown for instance to invest more physical load when they could help closer relatives rather than distant ones (Madsen et al. 2007).

Before wrapping up kin selection, let me just for a moment take a slight detour to the related (no pun intended) “green-beard” model of altruism. Hinted at by Hamilton (1964b), it was more developed and popularized by Dawkins (1976). In its simplest form, the concept introduces an allele that serves to promote altruism and at the same time distinguish its bearers clearly (in Dawkins’ thought experiment, by giving them a green beard). Such a process seems to be at work in red fire ants (Ross & Keller 1998). However, green-beard models cannot be equated with kin selection, as a green-beard allele would only “help” other alleles at closely linked loci and overall degree of relatedness doesn’t play a role here; the greenbeard allele would only “care” about its own presence. I will not consider this model further here, but it is prudent to mention, and in rare cases such as the red fire ant, it seems to control altruism and spite.

To wrap up, kin selection is a relatively straightforward and sometimes well-measurable way to explain many cases of altruism across different branches of the animal kingdom. However, it has trouble explaining altruism among unrelated individuals. While kin selection doesn’t strictly require any reliable recognition mechanisms of kinship, much less its specific degree, and can be satisfied e.g. by acting altruistically towards spatially close individuals (which in many species tend to be related), how can it explain altruism towards members of clearly different social groups within the species? Ascribing such acts of altruism only to a “costly mistake” or “lack of adaptation to new social conditions” seems far-fetched. That is where other theories come into play.

1.3.3 Reciprocity

“An eye for an eye” is a succinct way of describing a reciprocal relationship, in case of this particular idiom a negative one. In case of altruism, the relationship is positive (or so it seems), and another idiom – “I scratch your back, you scratch mine” – is more apt. Reciprocity was suggested as an explanation for altruism among non-kin by Robert Trivers (1971). He argued that if members of a population are subject to the same risks (e.g. drowning, starving) and cheating (= refusing to reciprocate a previous altruistic act) has adverse consequences, reciprocal altruism could evolve.

Trivers also outlined conditions under which reciprocity in his model worked (ie. benefits outweighed the costs), that is: 1) many such altruistic situations occur over the lifetime, 2) a given individual repeatedly interacts with a small set of others, and 3) dyads are

exposed to altruistic situations relatively symmetrically. Several relevant biological parameters could help achieve this: 1) length of a lifetime, 2) low dispersal rate, 3) high mutual dependence (special cases of which are constituted by 4) parental care and 6) aid in combat), 5) lower dominance hierarchy. Others (Stevens & Hauser 2004, Melis & Raihani 2023) later argued that direct reciprocity as described by Trivers also demands good memory and cognition for remembering individuals and previous interactions, detecting cheaters and considering time-delayed payoffs; therefore they considered reciprocity of little importance in non-human animals.

However, we can indeed find some convincing evidence from other species, most famously blood-sharing among vampire bats (Wilkinson 1984, Wilkinson et al. 2016, Carter et al. 2020), but also cooperative nest defence in red-winged blackbirds (Olendorf et al. 2004), mobbing behavior in the pied flycatcher (Krams et al. 2007), grooming or food-for-grooming among primates (de Waal & Brosnan 2006, Schino & Aureli 2007, 2009, Majolo et al. 2012, Molesti & Majolo 2017), food sharing in rats (Kettler et al. 2021, Engelhardt & Taborsky 2022, Dolivo et al. 2023), allonursing reindeer mothers (Engelhardt et al. 2023) and possibly even interspecies reciprocal altruism among Amazonian parrots (Van Houtan et al. 2024), etc.

An interesting consequence of Trivers' hypothesis is that an individual's health, among other factors, should play a role in whether that individual engages in altruistic behavior or not: the higher the chances of repeated interactions, the more benefits can reciprocity yield. Thus, in theory, individuals in good health should be more likely to engage in reciprocal altruism. And, because of the need to remember other individuals and past interactions and to assess the future to some extent, better memory and cognition should also lead to greater altruism.

Therefore, we can make and test predictions about how health, memory and cognitive abilities should influence altruistic behavior. That is the core of the included Study 3, dubbed Testing Trivers. *However*, it's also a great simplification. To be clear, there is nothing wrong with simplifications *per se*; they are necessary in scientific research. But since Trivers, we have learned that more types of reciprocity than the initially proposed direct reciprocity are at play (not only in humans), and that the conditions outlined by Trivers and stemming from his theory don't strictly apply to other types of reciprocity, such as indirect/extended and generalized. Let us briefly go over them, as reviewed by Kroneisen & Bell (2022); quotes theirs.

Direct reciprocity (“I help you because you helped me.”) constitutes a gradual exchange between two individuals; put simply, “helping each other in turn”. For instance rat helping behavior (Engelhardt et al. 2024, 2023) or many cases of primate grooming were found to be consistent with direct reciprocity, but some were more complex as well, as we’ll see below (de Waal & Brosnan 2006). In their recent work, Rossetti & Hilbe (2023) found direct reciprocity models to fit most experimental human interactions well (with specifics depending on memory and other factors).

Indirect reciprocity (“I help you because you helped someone else.”) introduces an element of *reputation*. Interestingly, this has been hinted at by Darwin (1871) in *The Descent of Man* when pondering morality. In present-day evolutionary biology, a number of models have suggested that when reputation (image, social standing) is considered, discriminating strategies that cooperate with those known as cooperators can evolve under some conditions (Nowak & Sigmund 1998, 2005, Leimar & Hammerstein 2001, Panchanathan & Boyd 2003), even under – realistically – imperfect information (not publicly available, noisy, invented, etc.), if reputation is nuanced instead of a binary “good or bad” choice (Schmid et al. 2023). There has also been interesting work in formalizing the ingredients of reputation (e.g. Ohtsuki & Iwasa 2004, 2006; however, the ingredients resulted in a binary choice in these models), an approach that can yield further testable predictions.

How does indirect reciprocity (IR) stand against direct reciprocity (DR) in its evolutionary dynamics? Roberts (2007) as well as Suzuki & Akiyama (2007) found that when IR utilizes social context, not only the image of a cooperative/defective decision, it can exist stably with respect to DR. Schmid et al. (2021) employed a complex model comparing DR and IR, introducing mutations as well, and found that under low mutation rates and relatively reliable outside information, IR can evolve and persist. Pal et al. (2024) model DR, IR and generalized reciprocity, finding that while all three modes can coexist, generalized reciprocity hinders cooperation in case of many initial defections (and is vulnerable to the introduction of defecting strategies) and it tends to be outperformed by DR and IR.

Experimentally, IR has been supported e.g. by participant cooperation choices in Wedekind & Milinski (2000), manipulation of the Trust Game setup to include indirect payoffs (Dufwenberg et al. 2001) and in the field through effects of observability on cooperation (Yoeli et al. 2013) and influence of online profile history on received help (von Apeldoorn & Schram 2016). Recently, Oda & Hayashi (2024) showed that people felt more inclined to help strangers described as prosocial and in difficulty because of circumstances outside of their control – effectively because of their reputation.

Interestingly, a new model has shown that while reputation may promote cooperation in smaller groups, but impair it on larger scales (Schnell & Muthukrishna 2024); it also suggests a possible need for better integration of reciprocity with multilevel selection models. More such models that attempt to synthesize the local and global level of population and their dynamics are sorely needed, as well as experimental follow-ups of the abundant theoretical work (see review by Okada 2020).

Generalized reciprocity (“I help you because someone else helped me.”), in the context of evolutionary biology, means extending altruism towards previously unknown individuals. Models aimed at studying its plausibility supported it (e.g. Pfeiffer et al. 2005), especially under assortment of encounters instead of a well-mixed population (Rankin & Taborsky 2009); but see the comparison of the three key modes of reciprocity by Pal et al. (2024) outlined above.

Interestingly, in laboratory experiments, not only humans but also rats were shown to exhibit generalized reciprocity (Rutte & Taborsky 2007), helping unknown conspecifics after they have been helped by another; however, this was disputed by Schweinfurth et al. (2019). In any case, they do appear capable of direct reciprocity (Kettler et al. 2021, Engelhardt & Taborsky 2022, Dolivo et al. 2023).

Strong reciprocity (Gintis 2000; not outlined by Kroneisen & Bell (2022), but important to mention here) relates to cooperating with cooperators and punishing defectors, even at the punisher’s own cost. Gintis (2000) came up with a model of its emergence. Its existence is supported by data from the field as well as lab experiments (e.g. Fehr et al. 2002, but see the call for more data from cooperation “in the wild” by Guala 2012).

Attitudinal reciprocity (de Waal 2000, de Waal & Brosnan 2006, Suchak & de Waal 2012), last but not least, is sometimes used as a separate category, denoting sequential interactions stemming from “mirroring” of the partner’s attitudes and ultimately from empathy and prosociality. As opponents of Triversian direct reciprocity often argue that for it to succeed, some altruistic tendencies have to already exist in order to be repaid (note: nobody claimed there would have been none), attitudinal reciprocity seems to offer a solution to this problem.

Note that this division is typical of evolutionary biology and ethology; cultural anthropologists typically divide reciprocity into generalized, balanced and negative (Sahlins, 1972). In this context, generalized reciprocity means something different than in our evolutionary outlook. For anthropology, it is a relationship between two people who don’t “keep an exact ledger”, but reciprocally help each other (e.g. one invites the other for lunch,

the other later invites the first one for coffee, theatre, etc., and they don't attempt to keep the value of each exchange comparable and don't stress over one being unable to reciprocate). Balanced reciprocity would be basically direct reciprocity "with a ledger", and as the exchange mostly occurs simultaneously, would rather fall under mutualism in the biology context. Finally, negative reciprocity means attempting to selfishly (in the psychological sense of the word) "get something for nothing/very little"; in the biology context, this would be rather dubbed theft or harassment. I won't be using the social psychology definitions of reciprocity in the thesis; it is just important not to confuse them with the categories of reciprocity used in the evolutionary and ethological contexts. Perhaps they won't clash in the future; recently, an interesting analytical attempt at unification of different definitions and categories of reciprocity has been made by Guevara Beltran et al. (2023). Such attempts are important, since under some definitions it has been argued that reciprocity is fundamentally unable to explain the evolution of altruism and cooperation, since its stabilizing effect relies on punishing (at least by not cooperating with repeatedly) defectors, and a sufficiently costly punishment could stabilize any kind of behavior. I personally find this conclusion a little puzzling; I agree with the last part, but why would that mean that reciprocity is insufficient for explaining altruistic behavior?

In any case, the theory can and should be tested to dispel or support these doubts. Testing whether the direct reciprocity conditions described above (e.g. health status, memory, cognitive ability) influence altruism is a way to relatively simply, in terms of the complex landscape we are traversing here, test the importance of direct reciprocity. Finding out for instance that health has no impact on altruistic behavior, however, would not refute the whole reciprocity theory, as it should by now be clear that its reality is far more complex than the simplest Triversian model.

Memory and cognitive load: Previous studies on humans as well as non-human animals (e.g. Larose & Dubois 2010) revealed that impaired working memory indeed interferes with reciprocal cooperation and altruism (Soutschek & Schubert 2015, Duffy & Smith 2014, Milinski & Wedekind 1998) or choice of cooperative partners (Murty et al. 2016). However, others found no effect of memorization tasks on generosity (Hauge et al. 2015, Cappelletti et al. 2011), and yet others found a detrimental effect of concurrent memory load on altruism, but no effect of cumulative memory load (Mitkidis et al. 2022); the varying approaches across different studies make generalizations difficult, but there is at least tentative evidence for the role of working memory.

Some authors suggested that we may have memory mechanisms primed to memorize cheaters, especially lower-status and unexpected ones (Mealey et al. 1996, Mieth et al. 2016); however, the question whether we have enhanced memory for faces of cheaters yields mixed answers (Mehl & Buchner 2008, Buchner et al. 2009, Bell et al. 2014).

Cognitive abilities: High mental load has been experimentally shown to increase generosity (Schulz et al. 2014), decrease it in women but not men (Grossman et al. 2014), but also to have no notable effect (Tinghög et al. 2013). It also led to greater acceptance rates of unfair offers, but only modulated by the status of the proposer (Harris et al. 2020). Empathy may be a moderating factor of cognitive load on prosocial decisions (Gamble et al. 2023). But interpretation of studies with even slightly different designs can be tricky, as shown by Kessler & Meier (2014), who ran an identical study on mental load and charitable giving five times, but in different order with other tasks within the same lab session. When the study was ran first, cognitive load led to *decreased* generosity, but when it was ran last, participants in the cognitive load treatment actually gave *more*; tiredness or an overall load from previous tasks might be behind the difference, which shows us the importance of replications. Unaccounted-for factors might be partly responsible for studies' results.

With that in mind, we might want to explore the role of cognitive abilities, not only present-moment load, in altruism. However, they may not be the best predictor of altruism; students with better math scores gave less in the Dictator Game in a study by Benjamin et al. (2013), and several studies (Ben-Ner et al. 2004, Ponti & Rodriguez-Lara 2015, Corgnet et al. 2015, Czerwinka et al. 2018) found people who scored higher in the Cognitive Reflection Test (CRT) to behave more selfishly in the experimental Dictator Game. The study by Chen et al. (2013) suggests that results associated with one type of scores may not be transferrable to other scores, as participants with higher math scores in the SAT tests were found to be more generous, while those with higher grade point averages behaved more selfishly; we must therefore tread carefully in interpreting data from different studies. In addition, different modes of cognitive abilities appear to predict altruistic behavior to a different degree (and direction) during ontogeny (Rose et al. 2024); something to consider in future studies in humans as well as non-human animals. During a person's life, the feedback direction might change as well; perhaps, while in children cognition might predict altruism (and it would be hard to argue for the opposite direction of the relationship), in old people it might be vice versa, with altruistic dispositions possibly slowing cognitive decline, as interpreted e.g. by Corrêa et al. (2018). However, we could also conjecture that people whose mind remains

sharper up to an advanced age behave more prosocially; that presents another open question to be answered.

But perhaps the cognitively demanding effective “scorekeeping” is not necessary for mutually successful reciprocal interactions, if there are already present prosocial tendencies and attitudinal mirroring (Suchak & de Waal 2012); in that case, though, it’s moving the question of altruism’s origins somewhat away from reciprocity. This vein of reasoning would be supported by the above-reviewed evidence that greater cognitive reflection and deliberation does not yield more altruistic behavior. Moreover, a survey of “extreme altruists” (people who actively and willingly risked their life to save others) suggests that deliberation is usually not part of extremely altruistic acts (Rand & Epstein 2014).

Health: A multitude of studies have found a positive connection between health and altruistic behavior. *However*, they typically don’t mention reciprocity and its preconditions at all and instead focus on the other possible direction of influence: altruism as a factor promoting health, and not vice versa. It makes perfect sense from the point of view of gaining larger social support networks and a good reputation (see Indirect reciprocity above). A review by Post (2005) found a positive connection between health and altruism in most studies, interpreted exactly as an effect of helping on well-being. Similarly, others also found altruists to be healthier people (Figueredo et al. 2006, Weziak-Bialowolska et al. 2024), at least in the case of mental health and with the relationship possibly moderated by life satisfaction (Cheng et al. 2024, Mayers et al. 2025).

Perhaps, though, reciprocal altruism can arise simpler still; a new study by Da Silva and Bonini (2024) suggests that it could emerge from random encounters classifiable as theft. It’s not a new concept, although I haven’t seen it formalized in a model before this work. Nevertheless, Gurven (2004) pointed out that “‘reciprocal’ TS [tolerated scrounging] ... is essentially identical to RA [reciprocal altruism]”, and in a commentary to his paper, Stevens & Cushman (2004) argued in favor of tolerated scrounging because of the immediacy of payoffs (to the thief) and no need for complex cognitive mechanisms required by reciprocal altruism. Literature on seemingly altruistic phenomena in non-human animals also abounds with cases of likely harassment/theft (e.g. Stevens 2004, Zapata et al. 2009, Gloneková et al. 2016, 2021, Paul & Bhadra 2017).

But there may be an even simpler pathway towards altruism than building from tolerated theft or another mechanism to reciprocity, and it has to do with partner choice.

1.3.4 Sexual selection, or behavioral peacock tails

We all know how the story goes: Boy meets girl, she becomes infatuated with his extravagant colorful tail and mates with him. Their offspring carry the alleles for extravagant tails and preference towards it, and the most colorful and most extravagant sons of the new generation leave most offspring in the following one... We are, of course, talking about peacock plumage, a textbook example of Fisherian runaway sexual selection (Flegr 2023c). We start with an indicator of fitness and preference towards it – and may end up in a positive feedback loop inexorably leading to the trap of a sexually selected, but otherwise maladaptive trait.

Could altruism, at least in some cases, be such an example of runaway sexual selection and constitute a “behavioral peacock tail”? Such as it is, it could explain widespread altruism in one-off interactions and towards socially distant individuals. Or could it constitute a handicap (Zahavi 1995) of the trait bearer, suggesting to potential partners that the individual can exist well enough even with this handicap and is therefore a worthy prospective mate? Or is it an honest costly signal? In fact, these possibilities do not exclude one another (under most definitions, that is).

Perhaps surprisingly, sexual selection towards altruism has been studied relatively sparsely in non-human animals (see the call by Covas & Doutrelant (2018) for more research in this overlooked area), despite its groundwork being laid by ornithologists (Shultziner & Dattner (2016) argue that Zahavi’s model species, the Arabian babbler, is a good example of sexual selection towards altruism). In humans, though, evidence suggests a connection between altruism, mating success and attractiveness.

Biases in favor of attractive people, including greater helping efforts directed towards them, run in many areas of the human society, among other things the labor market, everyday social interactions, and even experimental situations (Maestripieri et al. 2017), and in many cases may be explained by sexual selection (Roberts 2017). Waitresses with more attractive hair color or make-up were given larger tips by male customers (Guéguen 2012, Jacob et al. 2010).

In experimental games, men as well as women cooperated more with more attractive people of the opposite sex, and men were in general more altruistic towards women than towards other men (Farrelly et al. 2007, Bhogal et al. 2016). Lucas & Koff (2013) found people of both sexes to be more generous towards more attractive members of the opposite sex (with further effects of fertility phase in women). Other similar studies, though, found men to be more generous to women, but with no modulating effect of attractiveness (Saad & Gill 2001, Bhogal et al. 2017). A more recent study of online helping behavior did find a positive effect of attractiveness (Schwarz & Baßfeld 2018).

Being altruistic towards attractive potential mates is just one facet of sexual selection pathways towards altruism. Models such as Gintis et al. (2001) or Pilot (2005) suggest that under some conditions, even unconditional altruism as an honest signal of partner quality could evolve by sexual and social selection. More widespread altruistic behavior, not directed solely to attractive potential mates, could therefore be expected, even to the point of runaway competitive altruism (Roberts 1998, Van Vugt et al. 2012), and there should be preference for altruistic mates in partner choice.

This was observed by e.g. Barclay (2010) and Phillips et al. (2015), especially in women. For details of further evidence, including studies diving into more specific facets of prosociality or into phenotypic/genetic quality signalization by altruism, see a review by Bhogal et al. (2019). Recent research also demonstrated that altruistic partners are desirable regardless of sexual orientation (Duxbury & Bhogal 2024).

Studying the mating and reproductive success of altruists is a logical next step. Arnocky et al. (2017) found more altruistic people (with the trait self-reported as well as measured in an experimental game) to have more self-reported sexual partners. In addition, Allen et al. (2024) observed that Australian people who did more altruistic (charity, volunteering) work during a time span of 10 years also had more children (in total as well as conceived during this time period), suggesting some role of altruism in reproductive success.

Therefore, there is quite compelling evidence for greater altruism directed towards more attractive people, for altruism as a signal for the opposite sex, preference for altruistic mates, and altruists' actual reproductive success. The studies mentioned above are key works, but the list is by no means exhaustive. I would like to direct a reader interested in a longer in-depth take to a thorough new review by Boog et al. (2024), which outlines evidence coming from different disciplines and methodologies, and also their limitations and still open questions about altruism, mate choice and mating success. The authors make a point that while the overall breadth of evidence for altruism as a product of sexual selection is great, only the psychological evidence so far has solid foundations, and argue for a greater variety of research methods and number of studies in other disciplines; that is a wish I certainly stand behind.

1.3.5 Microbial manipulation

Interestingly, the manipulation hypothesis (see review Thomas et al. 2005) has been applied to altruism as well by Lewin-Epstein et al. (2017). Put simply, the hypothesis suggests that microbes could biochemically manipulate their hosts toward altruistic behavior, bringing the

microbes an evolutionary advantage due to an easier transmission from one host to another. In this case, altruism would function as a xenoadaptation (Flegr 2023b).

It is an elegant, simple hypothesis in that it doesn't require any complex cognitive abilities or social structures. On the other hand, it may be *too* simple to provide practical real-world predictions, and its take-home message seems to be mainly that horizontally transferred alleles (ie. ones present in transmitted microbes) can produce behaviors that are detrimental, costly from the “vertically transmitted alleles’ point of view”.

In principle, the hypothesis is testable by looking whether individuals treated with antibiotics (or antivirals or antiparasitics in case of other infections) manifest any change in altruistic behavior (whether in the course of their lifetime, or on the population level compared to those who had not received such treatment). Studying this possibility was the aim of my Study 4 included in this thesis.

1.4 Where do roots of altruism figure in other disciplines?

Knowing what factors contributed to the evolution of human altruism (and to what extent, possibly in what lineages of life, which is not explored experimentally within the scope of this work) is not only an interesting research question in itself, but trickles into other important topics, whose brief non-exhaustive overview is given below.

Convergent evolution of behavior: Altruism is unlikely to share completely common roots in all lineages where it is present. Knowing its roots can help us map its possible independent occurrence in different lineages. Comparative studies and meta-analyses can be especially useful here, when possible.

Insights into related fields: Mapping the selection value of altruism and thus constraining its ultimate origins can be helpful in synthesis with other-level causes of altruism (proximate mechanistic, ontogeny, phylogeny) – for instance, if we learned that reciprocity were most prominent in humans, we might want to ask in what age do cognitive systems for reciprocal “transactions” develop and what these systems are (which brain regions, hormonal regulation etc. are involved). This, in turn, can lead to greater understanding of their connection with other systems and evolutionary history.

Cultural evolution: Cultural evolution almost certainly played a major role in the advancement of human altruism, overlaying the biological influences and acting in synergy with them. Synthesis of biological and cultural evolution models (where especially multilevel selection concepts can be helpful) is much-needed for understanding human as well as non-human animal societies.

Philosophy: Ever since Darwin, philosophers have pondered the implications of the evolutionary theory for human society and morality. Darwin himself didn't shy away from these issues in *The Descent of Man* (Darwin 1871), and they have been the fodder of philosophy papers and books ever since. While I do not discuss the philosophical aspects of altruism and its evolution here, there is an ongoing discussion in the philosophical circles.

SETI: In a more speculative, as-of-yet impossible to test, but potentially civilization-altering question, multiple researchers asked the question whether intelligent species that might evolve on other worlds would also exhibit altruism, and to what extent. A decade ago, Douglas A. Vakoch (2014) edited a volume that ultimately asked this question and its implications (would another civilization's propensity for altruism mean a greater chance of sending radio or optical signals to potential-others?, would it mean that members of the civilization would extend altruistic preferences to species from other biospheres?). While we still have an N=1 as to known inhabited planets, it's an intriguing question for the future.

We could imagine even more questions the roots of altruism could help us answer. However, the aims of this work are not as ambitious as to answer these bigger questions, highly interdisciplinary and out of the scope of the conducted studies. It can, however, hopefully lay important groundwork for achieving these more ambitious goals step by step – it is important to keep them in sight.

2. TESTING THE FOUNDATIONS OF ALTRUISM IN HUMANS

“I have always been amazed, and I must admit annoyed as well, when I met, among fellow-zoologists, with the implied or stated opinion that the study of survival value must necessarily be guesswork, and that exact experimentation on the problem is in principle not possible.”

- Niko Tinbergen (1963)

Altruism is especially prevalent in human society, and evolutionary biologists since Charles Darwin (1871) have been particularly interested in its roots. At the same time, humans constitute somewhat special study participants: they can read written instructions, speak to experimenters... At the same time, they can be primed by a myriad of potential confounding variables that can be controlled for with arguably more difficulty than in most non-human animals (consider our cultural backgrounds, entertainment we consume, the huge variety of food that we literally consume...), however rich personal histories they may have.

Kurzban et al. (2015) review the literature on human altruism evolution and list more solid as well as tentative evidence for the role of kin selection, direct as well as indirect reciprocity and signaling. As they importantly point out, different explanations are not mutually exclusive. In fact, their synthesis is being increasingly called for, and e.g. Efferson et al. (2024) found that neither reciprocity or intergroup competition alone can explain human altruism well, but mended together – as cooperative reciprocity in-group and uncooperative reciprocity out-group – fit observed and experimental human behavior well.

However, this complexity – though it cannot be any other way, I’m afraid – makes ascertaining the degree of influence of different roots of the behavior a Gargantuan task. So, before we embark on a summary of the studies included in this thesis, in which I boldly attempted to add a little contribution to the said task, we must make one last detour: In order to understand different kinds of studies that can be conducted, an important note must be given to measuring altruism, in humans as well as nonhuman animals. Under the coarsest gaze, we can divide ways to measure altruistic behavior into four basic categories:

1. *Observations in the wild.* These can constitute observing animals in their natural habitats, making note of their social ties and exchanges, sharing resources or otherwise behaving in an altruistic way, and seeing whether altruistic behavior is “rewarded” by future interactions, be they reciprocal food sharing, defense, copulations, etc. For instance, studying whether sentinels are getting any benefits from others based on their patrolling behavior would fall under this category. In

humans, a brilliant example is presented by Gurven (2004), who reviewed cross-cultural food sharing patterns in order to discern between kin selection, reciprocal altruism, tolerated scrounging, and costly signaling.

2. *Field experiments.* These studies also take place “in the wild”, but certain conditions have been experimentally manipulated: for instance the occurrence of sentinel behavior, to return to the observational example above, by feeding selected individuals *ad libitum*, and then studying whether sentinels are rewarded for their behavior (Ostreiher et al. 2023). In humans, it could be changing the hair color (wig) or make-up of a waitress to see whether there are any differences in tipping from customers (Guéguen 2012, Jacob et al. 2010).
3. *Laboratory studies.* Experiments in the lab have the great advantage of (in the ideal case) carefully mapped and maintained conditions, well-known and well-selected sample of a certain size, and a great degree of freedom in manipulating the experimental conditions. They also have the (potentially great) disadvantage of artificial conditions and situations. Experimental games (e.g. Güth et al. 1982, Forsythe et al. 1994, Berg et al. 1995) fall under this category and, by including a set of parameters (such as anonymity, social distance, interaction asymmetry...) that can be easily manipulated, present a veritable cornucopia of potential experiments. My Studies 1–3 fall under this category.
4. *Online studies.* This category, applicable only to human participants, could be in some cases grouped with laboratory experiments, but I prefer setting it apart, since it doesn’t always include an *experimental* element and may only include surveys. Most of my Study 4 falls under this category.

Some researchers would prefer to add a special category for theoretical modeling studies, but these don’t *measure* altruism and study it by different means instead; therefore I have not included them in the list above. The studies in my thesis fall under points 3 and 4, with all the advantages as well as disadvantages that stem from that fact.

In the studies included in my doctoral thesis, altruism was typically measured through a proxy – behavior in experimental games, most importantly the Dictator Game, in some cases the Ultimatum Game and/or the Trust Game. I will very briefly describe the games below before embarking on summarizing the studies themselves.

The Dictator Game (further abbreviated DG) (Forsythe 1995) involves two players (typically in a fully anonymous, one-off setting), the first of whom is endowed by the

experimenter with an amount of money they can share with Player 2, who is passive. While, to repeat, the setting is typically anonymous and nothing forces Player 1 to share, people usually do share something – on average ~28% (Engel 2011).

The Ultimatum Game (UG) (Güth et al. 1982), while introduced earlier, is more complex than the DG. In the UG, Player 2 does have agency, and they can either accept or refuse the offer made by Player 1. If Player 2 accepts, each gets their share. If Player 2 refuses, neither gets anything. While, from the purely game-theoretical perspective, it would be most advantageous for Player 2 to accept any non-zero amount in a one-off setting, this is not what happens in reality. Most people in the position of Player 2 refuse low offers, and most in the position of Player 1 offer 50% of the pie, with ~42% being the average offer (Tisserand 2014).

Finally, the Trust Game (TG), also dubbed the Investment Game (Berg et al. 1995), involves one more step: Player 1, who is endowed with money by the experimenter, can send some to Player 2. The sent amount is doubled by the experimenter (which both players know it would be). In the final step, Player 2 can send some amount back to Player 1. Again, nothing compels them to, but most people do; and while Player 1 “should” send nothing, expecting Player 2 to keep everything they would receive, that is also not how real people behave. Player 1 usually transfers approx. 50% to Player 2, who on average returns about 37% of their received amount (Johnson & Mislin 2011).

There is a large body of altruism research – under the hats of experimental economics, psychology, ethology, evolutionary biology and more – connected to these games. They have been adapted for various animal studies as well (see Jensen et al. 2007, Kaiser et al. 2012, Proctor et al. 2013, Milinski 2013, Henrich & Silk 2013, Bueno-Guerra et al. 2019, Sánchez-Amaro et al. 2021, 2024), and cross-cultural studies (e.g. Henrich et al. 2001, 2006) and meta-analyses (Engel 2011, Cooper & Dutcher 2011, Larney et al. 2019, Cochard et al. 2021) exist to help us ascertain the robustness of the findings.

Typically, DG contributions are interpreted as altruistic, UG offers as part-altruistic, part-strategic concern, and refusals of low UG offers as altruistic punishment (Fehr & Gächter 2002, Fehr & Fischbacher 2004, Casal et al. 2019), although this interpretation has been disputed by some scholars (Brethel-Haurwitz et al. 2016); the initial allocation in the TG is considered to be a proxy of trust, as the name suggests, and the second allocation (amount sent back to Player 1) mainly as trustworthiness (Camerer 2003, Houser et al. 2010, Johnson & Mislin 2011), but both can also be interpreted as signs of altruism, as suggested by the designs of Cox (2004). All of this, of course, applies to one-shot games. Since I did not use

repeated games, I will not dive deeper into the strategies emerging from repeated DG, UG or TG (with the same/different partners, different time spans and (un)known number of rounds...), but will refer the interested reader to several pertinent works (e.g. Slembeck et al. 1999, Engle-Warnick & Slonim 2004, Avrahami et al. 2013, Dreber et al. 2014, Sass et al. 2015).

While the experimental games' interpretation is sometimes debated, self-reported altruism bears its own problems, though the typically used scales appear to have a good correlation with real-life traits (e.g. Rushton et al. 1981, Dee et al. 2003), and measuring actual real-life altruistic behavior may not be easy (although some studies were able to utilize e.g. identifiable samples of "extreme altruists", such as people who donated a kidney to a stranger or risked their life to save another; see Rand & Epstein 2014, Rusch 2022, Amormino et al. 2022). The DG, UG and TG were used as the sole proxies of altruism in some of my study designs, and in synergy with self-reported altruistic behavior in others. The goal was always to test the connection of altruism in the games (and in some cases self-reported real-life altruism) with factors hypothesized to influence altruistic tendencies.

In Table 2 below, I summarize the predictions stemming from major evolutionary theories of altruism, means to test them and whether they have been tested in any of the studies I have included in this work.

Theory	Selected predictions	Examples of testing
Interdemic selection	Altruism persists if altruists exhibit positive association, whether the population is continuous or forms isolated groups. Altruistic subpopulations could reproduce more frequently and drive the less altruistic ones to extinction. Altruism coded by multiple genes should be able to better persist in a meta-population. (Predictions of the rate of subpopulation – deme – origination and extinction, migration and mutation rates differ more across models.)	E.g. manipulating the frequency of altruism-coding alleles (where these are known – a caveat) in different subpopulations of animal species, and tracking their interdemic competition. Comparisons of altruistic norms in different societies through history and the societies' spread and stability (caveat: focuses on one specific in a world of complexity and randomness). Not tested by myself.

Kin selection	More help will be given to relatives than non-relatives. More help will be given to closer relatives.	E.g. experimental testing of willingness to help and extent of helping towards others of different levels of genetic (and, to eliminate this potential confounder, social) relatedness. Can be conducted across many different species in many settings, in humans also through surveys and games. Not tested by myself.
Reciprocity	Better memory, health and cognition will correlate with greater direct reciprocity. (Different predictions are given by indirect and generalized reciprocity – concerning reputation, noise, starting conditions’ influence on stability of strategies, etc.)	E.g. memory and cognitive tests, health checkups and/or self-reported health condition can be correlated with altruism measured via a proxy of an experimental game (DG, UG, TG...), a more lifelike experimental situation, or self-reported. Population-level studies can be conducted via surveys of health and altruism (charity work, volunteering...). More specifically regarding factors influencing health and memory, hormone measurement and/or administration studies can be conducted, via altruism measured as outlined above. Tested in Study 3.
Sexual selection	More altruism will be directed towards more desirable prospective partners. Altruists will be deemed more sympathetic, attractive and desirable potential mates. More altruistic individuals will have a greater reproductive success.	E.g. animal studies where the apparent altruism of a given individual is experimentally increased/decreased and tracking its mating success. Testing whether people primed by photographs of less/more attractive people of their preferred sex behave more altruistically, or whether people who think they are paired with someone (un)attractive exhibit greater altruism.

		Field experiments including manipulating attractiveness (by using wigs, make-up, special clothing...) to see whether the level of altruism directed towards the focal person changes. “Counting babies” (measuring reproductive success) studies including altruism information. Tested in Studies 1 and 2.
Host manipulation	.	E.g. studying whether individuals who have recently taken antibiotics exhibit lower altruism. Tested in Study 4.

TABLE 2. PREDICTIONS FROM DIFFERENT THEORIES. The table summarizes major theories trying to explain the evolution of altruism, their predictions and means of testing.

2.1 Study 1: Sexual selection towards altruism

In the paper “Looking Like a Million Dollars: Does Attractiveness Priming Increase Altruistic Behavior in Experimental Games?” (further referred to as Study 1), our aim was to investigate whether priming by photos of (un)attractive members of the opposite sex influences contributions in the DG and UG, oft-used to assess altruistic preferences, and whether more altruistically behaving people have more self-reported romantic and sexual partners. No significant effects were observed in behavior in the DG. As to the UG, men exposed to the attractiveness priming made larger UG offers, while women exposed to the same treatment made lower ones. Participants of both sexes who were more demanding in their minimum acceptable offer (MAO) in the UG had more self-reported sexual partners. With some caution, these results can be interpreted in favor of the costly signaling hypothesis, although the lower UG offers made by women in the attractiveness treatment were puzzling (and can be tentatively attributed to sex differences in partner seeking strategies, but that should be taken with a grain of salt in the absence of further supporting evidence from the same sample).

2.2. Study 2: Testosterone and cortisol connection with altruism

If sexual selection and health (in connection to reciprocity or otherwise) play a role in regulating altruistic behavior, we should expect hormonal levels to influence it and be able to make predictions about their impact on altruism. In “Generosity as a status signal: Higher-testosterone men exhibit greater altruism in the dictator game” (further Study 2), we studied the effect of testosterone and cortisol (and their interaction) on behavior in the DG.

Purely from the perspective of reciprocity, we should expect higher cortisol levels (as potential signifiers of stress and worse health condition) and to some extent higher testosterone levels (since testosterone dysregulates the immune system) to correspond negatively to altruism, since health would be an important factor of future interaction chances. On the other hand, from the perspective of sexual selection, one would expect testosterone to be connected to higher altruism primarily in men, if altruism served as an honest costly signal. Of course, these are highly simplified predictions that exist in a highly complex reality, so any effects would likely be small and not easily extracted from the influence of other factors. In addition, the (joint) influence of testosterone and cortisol on human behavior is itself complex, especially on traits associated with dominance and aggression (see Mehta & Josephs 2010, the dual hormone hypothesis).

We found that high-testosterone men made more generous contributions in the DG (which could be understood as status seeking and/or an honest signal of resources), in line with predictions stemming from sexual selection.

Since the preparation of the paper, a new related study has been published: no connection was found between prosociality and markers of chronic stress among the Hadza (Fedurek et al. 2024). More such studies across different societies with varying social hierarchies would be highly desirable to test especially the direct reciprocity and costly signaling hypotheses.

2.3 Study 3: Health and altruism – testing foundations of reciprocity

In Study 3, we aimed to test predictions stemming from the preconditions of Triversian direct reciprocity. Because of the cognitive requirements of reciprocity, widely expected in the literature, and the likelihood of future interactions with others, we hypothesized a positive correlation of altruistic behavior with memory, cognitive abilities, lower temporal discounting and better health.

However, in our sample of participants playing the DG and UG (as well as reporting on their altruistic behavior in more real-life context), taking part in several tests (Cognitive Reflection Test, Cattell’s intelligence test – A version, a temporal discounting task) and

answering questions about their health, we found no notable connection of altruism with memory and health, and predominantly non-significant trends in case of temporal discounting. These were more in line with the life history theory than direct reciprocity. Likewise, measures of cognitive reflection and intelligence had no relationship with giving in the DG. However, they correlated negatively with offers and minimum acceptable offers in the UG, in contrast to predictions.

It would be premature to minimize the role of (direct) reciprocity based only on these findings on one small sample, but it appears that at the very least, effects of the preconditions of Triversian direct reciprocity are small or very small if they occur in present-day humans.

2.4 Study 4: Host manipulation towards altruism

Lewin-Epstein et al. (2017) hypothesized that microbial infections could be an important driving force of evolution of altruism by manipulating their hosts toward altruistic behavior (thus increasing the chances of transmission from one host to another); in light of the many findings of our microbiome's importance and also in light of the manipulation hypothesis, such a claim was not surprising. The authors themselves proposed to test the hypothesis by seeing whether antibiotic (ATB) intake, potentially wiping the responsible microbes, had any effect on altruistic behavior.

We tested this in two related studies: first on data from 604 student participants who played a hypothetical Dictator Game and provided data about their antibiotics intake, and subsequently in an online study containing 8,427 valid responses from the general public. We found no relationship of antibiotics intake (either recently or in childhood development) and altruistic behavior in experimental games in either study. Instead, we observed an opposite-than-expected relationship of self-reported antibiotics intake in childhood with real-life helping behavior in the second study.

These results suggest that at least in humans, microbial host manipulation towards altruism is playing (and has played?) a negligible role, if any at all.

3. CONCLUSIONS

“The history of the altruism-morality debate in biology has provided many insights both about ourselves and about the ways in which we come to know ourselves. How we will re-write this history one hundred years hence remains unknown.”

- Oren Harman (2014)

Charles Darwin appeared to be puzzled by altruism, and to some extent, so are we up to this day – with multitudes more data and a deeper understanding of the workings of biological evolution, but uncovering ever more complexity as we dig even deeper. It would almost certainly be a mistake to think that a single simple cause could explain such a complex phenomenon as altruism, spread well over the animal kingdom – or to explain it even in a single species, such as humans.

The widespread prevalence of altruism in humans is striking and enables our unusually large and complex societies to exist and thrive. It’s a continued source of fascination, old and new research questions across disciplines and much debate.

I would love to say that my work has been pivotal in understanding this phenomenon better and that it ushered the field into a new, brighter age. Alas, it has not fared that well; I can only hope that it has provided some incremental advances towards greater understanding of the roots of altruism in humans.

In my work, I attempted to experimentally test predictions stemming from three different hypotheses proposed as potential roots of altruism: reciprocity, sexual selection and microbial host manipulation. The evidence collected and analyzed in my studies comprising the core of this thesis tentatively points towards costly signaling (within sexual selection) as an important factor in the evolution of altruism.

A greater synthesis of field and lab studies and theoretical models from the points of view of different fields is needed to map the origins of human altruistic behavior to whatever extent possible. My work has, hopefully, contributed to this greater picture – if only through incremental steps. Aptly, I hope that it has been *helpful*.

4. REFERENCES

- Allen, M. S., Robson, D. A., Mishra, M., & Laborde, S. (2024). A prospective and retrospective 10-year study of altruism and reproductive success. *Evolutionary Behavioral Sciences*, 18(4), 299–307. <https://doi.org/10.1037/ebs0000320>
- Amormino, P., Ploe, M. L., & Marsh, A. A. (2022). Moral foundations, values, and judgments in extraordinary altruists. *Scientific Reports*, 12(1), 22111. <https://doi.org/10.1038/s41598-022-26418-1>
- Apeldoorn, J. van, & Schram, A. (2016). Indirect Reciprocity; A Field Experiment. *PLOS ONE*, 11(4), e0152076. <https://doi.org/10.1371/journal.pone.0152076>
- Arnocky, S., Piché, T., Albert, G., Ouellette, D., & Barclay, P. (2017). Altruism predicts mating success in humans. *British Journal of Psychology*, 108(2), 416–435. <https://doi.org/10.1111/bjop.12208>
- Avrahami, J., Güth, W., Hertwig, R., Kareev, Y., & Otsubo, H. (2013). Learning (not) to yield: An experimental study of evolving ultimatum game behavior. *The Journal of Socio-Economics*, 47, 47–54. <https://doi.org/10.1016/j.socec.2013.08.009>
- Barclay, P., & Willer, R. (2006). Partner choice creates competitive altruism in humans. *Proceedings of the Royal Society B: Biological Sciences*, 274(1610), 749–753. <https://doi.org/10.1098/rspb.2006.0209>
- Bell, R., Schain, C., & Echterhoff, G. (2014). How selfish is memory for cheaters? Evidence for moral and egoistic biases. *Cognition*, 132(3), 437–442.
- Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2013). Who is ‘Behavioral’? Cognitive Ability and Anomalous Preferences. *Journal of the European Economic Association*, 11(6), 1231–1255. <https://doi.org/10.1111/jeea.12055>
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior*, 10(1), 122–142.
- Bhogal, M. S., Farrelly, D., & Galbraith, N. (2019). The role of prosocial behaviors in mate choice: A critical review of the literature. *Current Psychology*, 38(4), 1062–1075. <https://doi.org/10.1007/s12144-019-00308-8>
- Bhogal, M. S., Galbraith, N., & Manktelow, K. (2016). Physical Attractiveness and Altruism in Two Modified Dictator Games. *Basic and Applied Social Psychology*. <https://www.tandfonline.com/doi/abs/10.1080/01973533.2016.1199382>
- Bhogal, M. S., Galbraith, N., & Manktelow, K. (2017). Physical Attractiveness, Altruism and Cooperation in an Ultimatum Game. *Current Psychology*, 36(3), 549–555. <https://doi.org/10.1007/s12144-016-9443-1>
- Ben-Ner, A., Kong, F., & Putterman, L. (2004). Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving. *Journal of Economic Psychology*, 25(5), 581–589. [https://doi.org/10.1016/S0167-4870\(03\)00065-5](https://doi.org/10.1016/S0167-4870(03)00065-5)
- Boog, D. A., Egas, M., & Ploeger, A. (2024). Displaying Altruism as a Sexual Signal in Human Mate Choice is an Adaptation—An Interdisciplinary Overview of the Evidence. *Evolutionary Psychological Science*, 1–19.
- Boomsma, J. J. (2007). Kin Selection versus Sexual Selection: Why the Ends Do Not Meet. *Current Biology*, 17(16), R673–R683. <https://doi.org/10.1016/j.cub.2007.06.033>

- Boomsma, J. J. (2009). Lifetime monogamy and the evolution of eusociality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1533), 3191–3207. <https://doi.org/10.1098/rstb.2009.0101>
- Boomsma, J. J., Beekman, M., Cornwallis, C. K., Griffin, A. S., Holman, L., Hughes, W. O. H., Keller, L., Oldroyd, B. P., & Ratnieks, F. L. W. (2011). Only full-sibling families evolved eusociality. *Nature*, 471(7339), E4–E5. <https://doi.org/10.1038/nature09832>
- Boomsma, J. J., Huszár, D. B., & Pedersen, J. S. (2014). The evolution of multiqueen breeding in eusocial lineages with permanent physically differentiated castes. *Animal Behaviour*, 92, 241–252. <https://doi.org/10.1016/j.anbehav.2014.03.005>
- Bowles, S., Choi, J.-K., & Hopfensitz, A. (2003). The co-evolution of individual behaviors and social institutions. *Journal of Theoretical Biology*, 223(2), 135–147. [https://doi.org/10.1016/S0022-5193\(03\)00060-2](https://doi.org/10.1016/S0022-5193(03)00060-2)
- Bshary, R., & Grutter, A. S. (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature*, 441(7096), 975–978. <https://doi.org/10.1038/nature04755>
- Brethel-Haurwitz, K. M., Stoycos, S. A., Cardinale, E. M., Huebner, B., & Marsh, A. A. (2016). Is costly punishment altruistic? Exploring rejection of unfair offers in the Ultimatum Game in real-world altruists. *Scientific reports*, 6(1), 18974.
- Buchner, A., Bell, R., Mehl, B., & Musch, J. (2009). No enhanced recognition memory, but better source memory for faces of cheaters. *Evolution and Human Behavior*, 30(3), 212–224. <https://doi.org/10.1016/j.evolhumbehav.2009.01.004>
- Bueno-Guerra, N., Völter, C. J., de las Heras, Á., Colell, M., & Call, J. (2019). Bargaining in chimpanzees (Pan troglodytes): The effect of cost, amount of gift, reciprocity, and communication. *Journal of Comparative Psychology*, 133(4), 542–550. <https://doi.org/10.1037/com0000189>
- Burnstein, E., Crandall, C., & Kitayama, S. (1994). Some neo-Darwinian decision rules for altruism: Weighing cues for inclusive fitness as a function of the biological importance of the decision. *Journal of Personality and Social Psychology*, 67(5), 773–789. <https://doi.org/10.1037/0022-3514.67.5.773>
- Cappelletti, D., Güth, W., & Ploner, M. (2011). Being of two minds: Ultimatum offers under cognitive constraints. *Journal of Economic Psychology*, 32(6), 940–950. <https://doi.org/10.1016/j.joep.2011.08.001>
- Camerer, C. (2003). Behavioral game theory: Experiments in strategic interaction. Princeton University Press.
- Carter, G. G., Farine, D. R., Crisp, R. J., Vrtillek, J. K., Ripperger, S. P., & Page, R. A. (2020). Development of New Food-Sharing Relationships in Vampire Bats. *Current Biology*, 30(7), 1275–1279.e3. <https://doi.org/10.1016/j.cub.2020.01.055>
- Carter, G., & Wilkinson, G. (2013). Does food sharing in vampire bats demonstrate reciprocity? *Communicative & Integrative Biology*, 6(6), e25783. <https://doi.org/10.4161/cib.25783>
- Casal, S., Fallucchi, F., & Quercia, S. (2019). The role of morals in three-player ultimatum games. *Journal of Economic Psychology*, 70, 67–79. <https://doi.org/10.1016/j.joep.2018.11.003>
- Chen, C.-C., Chiu, I.-M., Smith, J., & Yamada, T. (2013). Too smart to be selfish? Measures of cognitive ability, social preferences, and consistency. *Journal of Economic Behavior & Organization*, 90, 112–122. <https://doi.org/10.1016/j.jebo.2013.03.032>

- Cheng, Y., Wei, Y., & Tang, S.-L. (2024). Does Helping Others Always Benefit Health? Longitudinal Evidence on the Relationship between Helping Behavior and Depression: The Mediating Role of Life Satisfaction and the Moderating Effect of IADL. *Depression and Anxiety*, 2024(1), 2304723. <https://doi.org/10.1155/2024/2304723>
- Clavien, C., & Chapuisat, M. (2013). Altruism across disciplines: One word, multiple meanings. *Biology & Philosophy*, 28(1), 125–140. <https://doi.org/10.1007/s10539-012-9317-3>
- Clutton-Brock, T. (2002). Breeding Together: Kin Selection and Mutualism in Cooperative Vertebrates. *Science*, 296(5565), 69–72. <https://doi.org/10.1126/science.296.5565.69>
- Cochard, F., Le Gallo, J., Georgantzis, N., & Tisserand, J.-C. (2021). Social preferences across different populations: Meta-analyses on the ultimatum game and dictator game. *Journal of Behavioral and Experimental Economics*, 90, 101613. <https://doi.org/10.1016/j.socec.2020.101613>
- Cooper, D. J., & Dutcher, E. G. (2011). The dynamics of responder behavior in ultimatum games: A meta-study. *Experimental Economics*, 14(4), 519–546. <https://doi.org/10.1007/s10683-011-9280-x>
- Corgnet, B., Espín, A. M., & Hernán-González, R. (2015). The cognitive basis of social behavior: Cognitive reflection overrides antisocial but not always prosocial motives. *Frontiers in Behavioral Neuroscience*, 9. <https://doi.org/10.3389/fnbeh.2015.00287>
- Corrêa, J. C., Ávila, M. P. W., Lucchetti, A. L. G., & Lucchetti, G. (2019). Altruistic behaviour, but not volunteering, has been associated with cognitive performance in community-dwelling older persons. *Psychogeriatrics*, 19(2), 117–125. <https://doi.org/10.1111/psyg.12372>
- Covas, R., & Doutrelant, C. (2019). Testing the Sexual and Social Benefits of Cooperation in Animals. *Trends in Ecology & Evolution*, 34(2), 112–120. <https://doi.org/10.1016/j.tree.2018.11.006>
- Cox, J. C. (2004). How to identify trust and reciprocity. *Games and Economic Behavior*, 46(2), 260–281. [https://doi.org/10.1016/S0899-8256\(03\)00119-2](https://doi.org/10.1016/S0899-8256(03)00119-2)
- Czerwinka, M., Staniszevska, A., & Kompa, K. (2018). Cognitive Reflection Test in Predicting Rational Behavior in the Dictator Game. In K. Nermend & M. Łatuszyńska (Eds.), *Problems, Methods and Tools in Experimental and Behavioral Economics* (pp. 301–312). Springer International Publishing. https://doi.org/10.1007/978-3-319-99187-0_22
- Darwin, C. (1859, def. ed. 1872) *On the Origin of Species*. Retrieved from Project Gutenberg Dec. 30, 2024, from <https://www.gutenberg.org/files/2009/2009-h/2009-h.htm>
- Darwin, C. (1871, sec. ed. 1874) *The Descent of Man*. Retrieved from Project Gutenberg Dec. 30, 2024, from <https://www.gutenberg.org/cache/epub/2300/pg2300-images.html>
- Dawkins, R. (1976). *The selfish gene*. Oxford University Press. ISBN 0-19-857519-X.
- De Waal, F. B. M. (2008). Putting the Altruism Back into Altruism: The Evolution of Empathy. *Annual Review of Psychology*, 59(1), 279–300. <https://doi.org/10.1146/annurev.psych.59.103006.093625>
- Dee, D. Y., Lee, J. Y., & Kang, C. H. (2003). Development and Validation of an Altruism Scale for Adults. *Psychological Reports*, 92(2), 555–561. <https://doi.org/10.2466/pr0.2003.92.2.555>
- Denny, A. J., Franks, N. R., Powell, S., & Edwards, K. J. (2004). Exceptionally high levels of multiple mating in an army ant. *Naturwissenschaften*, 91(8), 396–399. <https://doi.org/10.1007/s00114-004-0546-4>

- Dolivo, V., Engqvist, L., & Taborsky, M. (2023). When deciding to cooperate by direct reciprocity, Norway rats sometimes benefit from olfactory competence and seem not impaired by insufficient cognitive abilities. *Animal Cognition*, 26(4), 1119–1130. <https://doi.org/10.1007/s10071-023-01754-5>
- Domondon, A. T. (2013). A history of altruism focusing on Darwin, Allee and E.O. Wilson. *Endeavour*, 37(2), 94–103. <https://doi.org/10.1016/j.endeavour.2012.12.001>
- Dreber, A., Fudenberg, D., & Rand, D. G. (2014). Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics. *Journal of Economic Behavior & Organization*, 98, 41–55. <https://doi.org/10.1016/j.jebo.2013.12.007>
- Duffy, S., & Smith, J. (2014). Cognitive load in the multi-player prisoner's dilemma game: Are there brains in games? *Journal of Behavioral and Experimental Economics*, 51, 47–56. <https://doi.org/10.1016/j.socec.2014.01.006>
- Dufwenberg, M., Gneezy, U., Güth, W., & van Damme, E. E. C. (2001). Direct versus indirect reciprocity: An experiment. *Homo oeconomicus*, 18(1/2), 19–30.
- Duxbury, A., & Bhogal, M. S. (2024). The Effect of Sexual Orientation on Preferences for Altruism in Human Mate Choice. *Evolutionary Psychological Science*, 10(4), 383–387. <https://doi.org/10.1007/s40806-024-00419-x>
- Efferson, C., Bernhard, H., Fischbacher, U., & Fehr, E. (2024). Super-additive cooperation. *Nature*, 626(8001), 1034–1041. <https://doi.org/10.1038/s41586-024-07077-w>
- Eibl-Eibesfeldt, I., & Salter, F. K. (1998). *Indoctrinability, Ideology, and Warfare: Evolutionary Perspectives*. Berghahn Books.
- Engelhardt, S. C., Paulsson, N. I., & Taborsky, M. (2024). Norway rats recruit cooperation partners based on previous receipt of help while disregarding kinship. *iScience*, 27(12). <https://doi.org/10.1016/j.isci.2024.111314>
- Engelhardt, S. C., & Taborsky, M. (2022). Food-exchanging Norway rats apply the direct reciprocity decision rule rather than copying by imitation. *Animal Behaviour*, 194, 265–274. <https://doi.org/10.1016/j.anbehav.2022.09.005>
- Engelhardt, S. C., Weladji, R. B., Holand, Ø., Røed, K. H., & Nieminen, M. (2023). Evidence suggesting that reindeer mothers allonurse according to the direct reciprocity and generalized reciprocity decision rules. *PLOS ONE*, 18(12), e0295497. <https://doi.org/10.1371/journal.pone.0295497>
- Engelhardt, S. C., Paulsson, N. I., & Taborsky, M. (2023). Assessment of help value affects reciprocation in Norway rats. *Royal Society Open Science*, 10(10), 231253. <https://doi.org/10.1098/rsos.231253>
- Engle-Warnick, J., & Slonim, R. L. (2004). The evolution of strategies in a repeated trust game. *Journal of Economic Behavior & Organization*, 55(4), 553–573. <https://doi.org/10.1016/j.jebo.2003.11.008>
- Farrelly, D., Lazarus, J., & Roberts, G. (2007). Altruists Attract. *Evolutionary Psychology*, 5(2), 147470490700500205. <https://doi.org/10.1177/147470490700500205>
- Fedurek, P., Aktipis, A., Cronk, L., Danel, D., Lacroix, L., Lehmann, J., Mabulla, I., Makambi, J. E., & Berbesque, J. C. (2024). Prosocial reputation and stress among contemporary hunter-gatherers: The Hadza case. *Scientific Reports*, 14(1), 21403. <https://doi.org/10.1038/s41598-024-72238-w>
- Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425(6960), 785–791. <https://doi.org/10.1038/nature02043>

- Fehr, E., Fischbacher, U., & Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, 13(1), 1–25. <https://doi.org/10.1007/s12110-002-1012-7>
- Filkowski, M. M., Cochran, R. N., & Haas, B. W. (2016). Altruistic behavior: Mapping responses in the brain. *Neuroscience and Neuroeconomics*, 5, 65–75. <https://doi.org/10.2147/NAN.S87718>
- Flegr, J. (2023a). Postadaptations. In T. K. Shackelford (Ed.), *Encyclopedia of Sexual Psychology and Behavior* (pp. 1–9). Springer International Publishing. https://doi.org/10.1007/978-3-031-08956-5_385-1
- Flegr, J. (2023b). Xenoadaptations. In T. K. Shackelford (Ed.), *Encyclopedia of Sexual Psychology and Behavior* (pp. 1–5). Springer International Publishing. https://doi.org/10.1007/978-3-031-08956-5_389-1
- Flegr, J. (2023c). Runaway Sexual Selection Model. In T. K. Shackelford (Ed.), *Encyclopedia of Sexual Psychology and Behavior* (pp. 1–10). Springer International Publishing. https://doi.org/10.1007/978-3-031-08956-5_2063-1
- Foster, K., Wenseleers, T., & Ratnieks, F. (2006). Kin selection is the key to altruism. *Trends in Ecology & Evolution*, 21(2), 57–60. <https://doi.org/10.1016/j.tree.2005.11.020>
- Gamble, R. S., Henry, J. D., & Vanman, E. J. (2023). Empathy moderates the relationship between cognitive load and prosocial behaviour. *Scientific Reports*, 13(1), 824. <https://doi.org/10.1038/s41598-023-28098-x>
- García, J., & van den Bergh, J. C. J. M. (2011). Evolution of parochial altruism by multilevel selection. *Evolution and Human Behavior*, 32(4), 277–287. <https://doi.org/10.1016/j.evolhumbehav.2010.07.007>
- Gardner, A., West, S. A., & Wild, G. (2011). The genetical theory of kin selection. *Journal of Evolutionary Biology*, 24(5), 1020–1043. <https://doi.org/10.1111/j.1420-9101.2011.02236.x>
- Gintis, H. (2000). Strong Reciprocity and Human Sociality. *Journal of Theoretical Biology*, 206(2), 169–179. <https://doi.org/10.1006/jtbi.2000.2111>
- Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly Signaling and Cooperation. *Journal of Theoretical Biology*, 213(1), 103–119. <https://doi.org/10.1006/jtbi.2001.2406>
- Gloneková, M., Brandlová, K., & Pluháček, J. (2016). Stealing milk by young and reciprocal mothers: High incidence of allonursing in giraffes, *Giraffa camelopardalis*. *Animal Behaviour*, 113, 113–123. <https://doi.org/10.1016/j.anbehav.2015.11.026>
- Gloneková, M., Brandlová, K., & Pluháček, J. (2021). Further behavioural parameters support reciprocity and milk theft as explanations for giraffe allonursing. *Scientific Reports*, 11(1), 7024. <https://doi.org/10.1038/s41598-021-86499-2>
- Glynatsi, N. E., Akin, E., Nowak, M. A., & Hilbe, C. (2024). Conditional cooperation with longer memory. *Proceedings of the National Academy of Sciences*, 121(50), e2420125121. <https://doi.org/10.1073/pnas.2420125121>
- Glynatsi, N. E., McAvoy, A., & Hilbe, C. (2024). Evolution of reciprocity with limited payoff memory. *Proceedings of the Royal Society B: Biological Sciences*, 291(2025), 20232493. <https://doi.org/10.1098/rspb.2023.2493>
- Gowdy, J., & Krall, L. (2016). The economic origins of ultrasociality. *Behavioral and Brain Sciences*, 39, e92. <https://doi.org/10.1017/S0140525X1500059X>

- Griffin, A. S., & West, S. A. (2002). Kin selection: Fact and fiction. *Trends in Ecology & Evolution*, 17(1), 15–21. [https://doi.org/10.1016/S0169-5347\(01\)02355-2](https://doi.org/10.1016/S0169-5347(01)02355-2)
- Griffin, A. S., & West, S. A. (2003). Kin Discrimination and the Benefit of Helping in Cooperatively Breeding Vertebrates. *Science*, 302(5645), 634–636. <https://doi.org/10.1126/science.1089402>
- Grossman, Z., van der Weele, J., & Andrijevich, A. (2014). *A Test of Dual-Process Reasoning in Charitable Giving*. <https://escholarship.org/uc/item/4tm617f7>
- Grossman, P. J., & Levy, J. (2024). It's not you (well, it is a bit you), it's me: Self- versus social image in warm-glow giving. *PLOS ONE*, 19(3), e0300868. <https://doi.org/10.1371/journal.pone.0300868>
- Guala, F. (2012). Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences*, 35(1), 1–15. <https://doi.org/10.1017/S0140525X11000069>
- Guéguen, N. (2012). Hair color and wages: Waitresses with blond hair have more fun. *The Journal of Socio-Economics*, 41(4), 370–372. <https://doi.org/10.1016/j.socec.2012.04.012>
- Guevara Beltran, D., Ayers, J. D., Munoz, A., Cronk, L., & Aktipis, A. (2023). What is reciprocity? A review and expert-based classification of cooperative transfers. *Evolution and Human Behavior*, 44(4), 384–393. <https://doi.org/10.1016/j.evolhumbehav.2023.05.003>
- Gurven, M. (2004). To give and to give not: The behavioral ecology of human food transfers. *Behavioral and Brain Sciences*, 27(4), 543–560.
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of economic behavior & organization*, 3(4), 367–388.
- Halperin, T., & Levy, A. (2024). What, If Anything, Is Biological Altruism? *The British Journal for the Philosophy of Science*, 75(2), 465–486. <https://doi.org/10.1086/716097>
- Hamilton, W. D. (1964a). The genetical evolution of social behaviour. I. *Journal of theoretical biology*, 7(1), 1–16.
- Hamilton, W. D. (1964b). The genetical evolution of social behaviour. II. *Journal of theoretical biology*, 7(1), 17–52.
- Harman, O. (2014). A history of the altruism–morality debate in biology. *Behaviour*, 151(2/3), 147–165.
- Harris, A., Young, A., Hughson, L., Green, D., Doan, S. N., Hughson, E., & Reed, C. L. (2020). Perceived relative social status and cognitive load influence acceptance of unfair offers in the Ultimatum Game. *PLOS ONE*, 15(1), e0227717. <https://doi.org/10.1371/journal.pone.0227717>
- Hartmann, P., Eisend, M., Apaolaza, V., & D'Souza, C. (2017). Warm glow vs. altruistic values: How important is intrinsic emotional reward in proenvironmental behavior? *Journal of Environmental Psychology*, 52, 43–55. <https://doi.org/10.1016/j.jenvp.2017.05.006>
- Hauge, K. E., Brekke, K. A., Johansson, L.-O., Johansson-Stenman, O., & Svedsäter, H. (2016). Keeping others in our mind or in our heart? Distribution games under cognitive load. *Experimental Economics*, 19(3), 562–576. <https://doi.org/10.1007/s10683-015-9454-z>
- Heinsohn, R., & Packer, C. (1995). Complex Cooperative Strategies in Group-Territorial African Lions. *Science*, 269(5228), 1260–1262. <https://doi.org/10.1126/science.7652573>
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *American Economic Review*, 91(2), 73–78. <https://doi.org/10.1257/aer.91.2.73>

- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2006). Costly Punishment Across Human Societies. *Science*, 312(5781), 1767–1770. <https://doi.org/10.1126/science.1127333>
- Henrich, J., & Silk, J. B. (2013). Interpretative problems with chimpanzee ultimatum game. *Proceedings of the National Academy of Sciences*, 110(33), E3049–E3049. <https://doi.org/10.1073/pnas.1307007110>
- Hermesen, R. (2022). Emergent multilevel selection in a simple spatial model of the evolution of altruism. *PLOS Computational Biology*, 18(10), e1010612. <https://doi.org/10.1371/journal.pcbi.1010612>
- Houser, D., Schunk, D., & Winter, J. (2010). Distinguishing trust from risk: An anatomy of the investment game. *Journal of Economic Behavior & Organization*, 74(1), 72–81. <https://doi.org/10.1016/j.jebo.2010.01.002>
- Houtan, K. S. V., Rojas-Moscoso, J.-I., Houtan, H. N. V., & Gonzalez, O. (2024). *Cooperation, status, and altruism in a mixed society of Amazonian parrots*. <https://ecoevorxiv.org/repository/view/7868/>
- Hughes, W. O. H., Oldroyd, B. P., Beekman, M., & Ratnieks, F. L. W. (2008). Ancestral Monogamy Shows Kin Selection Is Key to the Evolution of Eusociality. *Science*, 320(5880), 1213–1216. <https://doi.org/10.1126/science.1156108>
- Jacob, C., Guéguen, N., Boulbry, G., & Ardiccioni, R. (2010). Waitresses' facial cosmetics and tipping: A field experiment. *International Journal of Hospitality Management*, 29(1), 188–190. <https://doi.org/10.1016/j.ijhm.2009.04.003>
- Jarvis, J. U. M., O'Riain, M. J., Bennett, N. C., & Sherman, P. W. (1994). Mammalian eusociality: A family affair. *Trends in Ecology & Evolution*, 9(2), 47–51. [https://doi.org/10.1016/0169-5347\(94\)90267-4](https://doi.org/10.1016/0169-5347(94)90267-4)
- Jensen, K., Call, J., & Tomasello, M. (2007). Chimpanzees are rational maximizers in an ultimatum game. *science*, 318(5847), 107-109.
- Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5), 865–889. <https://doi.org/10.1016/j.joep.2011.05.007>
- Johnstone, R. A., & Bshary, R. (2004). Evolution of spite through indirect reciprocity. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1551), 1917–1922. <https://doi.org/10.1098/rspb.2003.2581>
- Jonason, P. K., Baughman, H. M., Carter, G. L., & Parker, P. (2015). Dorian Gray without his portrait: Psychological, social, and physical health costs associated with the Dark Triad. *Personality and Individual Differences*, 78, 5-13.
- Kaiser, I., Jensen, K., Call, J., & Tomasello, M. (2012). Theft in an ultimatum game: Chimpanzees and bonobos are insensitive to unfairness. *Biology Letters*, 8(6), 942-945. <https://doi/full/10.1098/rsbl.2012.0519>
- Kay, T., Keller, L., & Lehmann, L. (2020). The evolution of altruism and the serial rediscovery of the role of relatedness. *Proceedings of the National Academy of Sciences*, 117(46), 28894–28898. <https://doi.org/10.1073/pnas.2013596117>
- Kessler, J. B., & Meier, S. (2014). Learning from (failed) replications: Cognitive load manipulations and charitable giving. *Journal of Economic Behavior & Organization*, 102, 10–13. <https://doi.org/10.1016/j.jebo.2014.02.005>
- Kerr, B., Godfrey-Smith, P., & Feldman, M. W. (2004). What is altruism? *Trends in Ecology & Evolution*, 19(3), 135–140. <https://doi.org/10.1016/j.tree.2003.10.004>

- Kettler, N., Schweinfurth, M. K., & Taborsky, M. (2021). Rats show direct reciprocity when interacting with multiple partners. *Scientific Reports*, 11(1), 3228. <https://doi.org/10.1038/s41598-021-82526-4>
- Korb, J., & Heinze, J. (2004). Multilevel selection and social evolution of insect societies. *Naturwissenschaften*, 91(6), 291–304. <https://doi.org/10.1007/s00114-004-0529-5>
- Krams, I., Krama, T., Igaune, K., & Mänd, R. (2008). Experimental evidence of reciprocal altruism in the pied flycatcher. *Behavioral Ecology and Sociobiology*, 62(4), 599–605. <https://doi.org/10.1007/s00265-007-0484-1>
- Kroneisen, M., & Bell, R. (2022). Memory as a cognitive requirement for reciprocal cooperation. *Current Opinion in Psychology*, 43, 271–277. <https://doi.org/10.1016/j.copsyc.2021.08.008>
- Kurokawa, S. (2017). The extended reciprocity: Strong belief outperforms persistence. *Journal of Theoretical Biology*, 421, 16–27. <https://doi.org/10.1016/j.jtbi.2017.03.021>
- Kurzban, R., Burton-Chellew, M. N., & West, S. A. (2015). The evolution of altruism in humans. *Annual review of psychology*, 66(1), 575–599.
- Larney, A., Rotella, A., & Barclay, P. (2019). Stake size effects in ultimatum game and dictator game offers: A meta-analysis. *Organizational Behavior and Human Decision Processes*, 151, 61–72. <https://doi.org/10.1016/j.obhdp.2019.01.002>
- Larose, K., & Dubois, F. (2011). Constraints on the Evolution of Reciprocity: An Experimental Test with Zebra Finches. *Ethology*, 117(2), 115–123. <https://doi.org/10.1111/j.1439-0310.2010.01850.x>
- Leigh Jr, E. G. (2010). The group selection controversy. *Journal of Evolutionary Biology*, 23(1), 6–19. <https://doi.org/10.1111/j.1420-9101.2009.01876.x>
- Leimar, O., & Hammerstein, P. (2001a). Evolution of cooperation through indirect reciprocity. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1468), 745–753. <https://doi.org/10.1098/rspb.2000.1573>
- Leimar, O., & Hammerstein, P. (2001b). Evolution of cooperation through indirect reciprocity. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1468), 745–753. <https://doi.org/10.1098/rspb.2000.1573>
- Lewin-Epstein, O., Aharonov, R., & Hadany, L. (2017). Microbes can help explain the evolution of host altruism. *Nature communications*, 8(1), 14040.
- Lucas, M., & Koff, E. (2013). How conception risk affects competition and cooperation with attractive women and men. *Evolution and Human Behavior*, 34(1), 16–22. <https://doi.org/10.1016/j.evolhumbehav.2012.08.001>
- Madsen, E. A., Tunney, R. J., Fieldman, G., Plotkin, H. C., Dunbar, R. I., Richardson, J. M., & McFarland, D. (2007). Kinship and altruism: A cross-cultural experimental study. *British journal of psychology*, 98(2), 339–359.
- Maestriperi, D., Henry, A., & Nickels, N. (2017). Explaining financial and prosocial biases in favor of attractive people: Interdisciplinary perspectives from economics, social psychology, and evolutionary psychology. *The Behavioral and Brain Sciences*, 40, e19. <https://doi.org/10.1017/S0140525X16000340>
- Majolo, B., Schino, G., & Aureli, F. (2012). The relative prevalence of direct, indirect and generalized reciprocity in macaque grooming exchanges. *Animal Behaviour*, 83(3), 763–771. <https://doi.org/10.1016/j.anbehav.2011.12.026>

- Marsh, A. A. (2016). Neural, cognitive, and evolutionary foundations of human altruism. *WIREs Cognitive Science*, 7(1), 59–71. <https://doi.org/10.1002/wcs.1377>
- Mayers, T., Eto, S., Maki, N., Araki, A., & Matsuda, H. (2024). Volunteering and Its Association with Depression, Loneliness, and Lifestyle of Older Adults: Insights from a Japanese Cross-Sectional Study. *Healthcare*, 12(21), 2187. <https://doi.org/10.3390/healthcare12212187>
- Mealey, L., Daoood, C., & Krage, M. (1996). Enhanced memory for faces of cheaters. *Ethology and Sociobiology*, 17(2), 119–128. [https://doi.org/10.1016/0162-3095\(95\)00131-X](https://doi.org/10.1016/0162-3095(95)00131-X)
- Mehl, B., & Buchner, A. (2008). No enhanced memory for faces of cheaters. *Evolution and Human Behavior*, 29(1), 35–41.
- Mehta, P. H., & Josephs, R. A. (2010). Testosterone and cortisol jointly regulate dominance: Evidence for a dual-hormone hypothesis. *Hormones and behavior*, 58(5), 898–906.
- Melis, A. P., & Raihani, N. J. (2023). The cognitive challenges of cooperation in human and nonhuman animals. *Nature Reviews Psychology*, 2(9), 523–536. <https://doi.org/10.1038/s44159-023-00207-7>
- Mieth, L., Bell, R., & Buchner, A. (2016). Cognitive Load Does Not Affect the Behavioral and Cognitive Foundations of Social Cooperation. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.01312>
- Milinski, M., & Wedekind, C. (1998). Working memory constrains human cooperation in the Prisoner's Dilemma. *Proceedings of the National Academy of Sciences*, 95(23), 13755–13758.
- Milinski, M. (2013). Chimps play fair in the ultimatum game. *Proceedings of the National Academy of Sciences*, 110(6), 1978–1979. <https://doi.org/10.1073/pnas.1222132110>
- Mitkidis, P., Lindeløv, J. K., Elbaek, C. T., Porubanova, M., Grzymala-Moszczynska, J., & Ariely, D. (2022). Morality in the time of cognitive famine: The effects of memory load on cooperation and honesty. *Acta Psychologica*, 228, 103664. <https://doi.org/10.1016/j.actpsy.2022.103664>
- Molesti, S., & Majolo, B. (2017). Evidence of direct reciprocity, but not of indirect and generalized reciprocity, in the grooming exchanges of wild Barbary macaques (*Macaca sylvanus*). *American Journal of Primatology*, 79(9), e22679. <https://doi.org/10.1002/ajp.22679>
- Morris, S. G. (2005). Identifying the Explanatory Weakness of Strong Altruism: The Needle in the 'Haystack Model.' *Philosophy of Science*, 72(5), 1124–1134. <https://doi.org/10.1086/508121>
- Murty, V. P., FeldmanHall, O., Hunter, L. E., Phelps, E. A., & Davachi, L. (2016). Episodic memories predict adaptive value-based decision-making. *Journal of Experimental Psychology: General*, 145(5), 548–558. <https://doi.org/10.1037/xge0000158>
- Nesse, R. M. (2007a). Runaway Social Selection for Displays of Partner Value and Altruism. *Biological Theory*, 2(2), 143–155. <https://doi.org/10.1162/biot.2007.2.2.143>
- Nesse, R. M. (2007b). Runaway Social Selection for Displays of Partner Value and Altruism. *Biological Theory*, 2(2), 143–155. <https://doi.org/10.1162/biot.2007.2.2.143>
- Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science (New York, N.y.)*, 314(5805), 1560–1563. <https://doi.org/10.1126/science.1133755>
- Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, 393(6685), 573–577. <https://doi.org/10.1038/31225>

- Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, 437(7063), 1291–1298. <https://doi.org/10.1038/nature04131>
- Nowak, M. A., Tarnita, C. E., & Wilson, E. O. (2010). The evolution of eusociality. *Nature*, 466(7310), 1057–1062. <https://doi.org/10.1038/nature09205>
- Oda, R., & Hayashi, N. (2024). Deciding Who Is Worthy of Help: Effect of the Probability of Reciprocity on Individuals' Willingness to Help Others. *Evolutionary Psychology*, 22(2), 14747049241254725. <https://doi.org/10.1177/14747049241254725>
- Ohtsuki, H., & Iwasa, Y. (2004). How should we define goodness?—Reputation dynamics in indirect reciprocity. *Journal of Theoretical Biology*, 231(1), 107–120. <https://doi.org/10.1016/j.jtbi.2004.06.005>
- Ohtsuki, H., & Iwasa, Y. (2006). The leading eight: Social norms that can maintain cooperation by indirect reciprocity. *Journal of Theoretical Biology*, 239(4), 435–444. <https://doi.org/10.1016/j.jtbi.2005.08.008>
- Okada, I. (2020). A Review of Theoretical Studies on Indirect Reciprocity. *Games*, 11(3), 27. <https://doi.org/10.3390/g11030027>
- Okasha, S. (2020). Biological Altruism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2020). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/sum2020/entries/altruism-biological/>
- Olendorf, R., Getty, T., & Scribner, K. (2004). Cooperative nest defence in red-winged blackbirds: Reciprocal altruism, kinship or by-product mutualism? *Proceedings of the Royal Society B: Biological Sciences*, 271(1535), 177–182.
- Osiński, J., Karbowski, A., Rusek, J., & Reinholz, A. (2021). The effect of delay and social distance in the dictator and ultimatum games. *Behavioural Processes*, 193, 104513. <https://doi.org/10.1016/j.beproc.2021.104513>
- Ostreiher, R., Mundry, R., & Heifetz, A. (2023). Are sentinels rewarded by their groupmates? A study in a cooperative breeder. *Behavioral Ecology and Sociobiology*, 77(11), 125.
- Pal, S., Hilbe, C., & Glynatsi, N. E. (2024). *The co-evolution of direct, indirect and generalized reciprocity* (No. arXiv:2411.03488). arXiv. <https://doi.org/10.48550/arXiv.2411.03488>
- Paul, M., & Bhadra, A. (2017). Selfish Pups: Weaning Conflict and Milk Theft in Free-Ranging Dogs. *PLOS ONE*, 12(2), e0170590. <https://doi.org/10.1371/journal.pone.0170590>
- Panchanathan, K., & Boyd, R. (2003). A tale of two defectors: The importance of standing for evolution of indirect reciprocity. *Journal of Theoretical Biology*, 224(1), 115–126. [https://doi.org/10.1016/S0022-5193\(03\)00154-1](https://doi.org/10.1016/S0022-5193(03)00154-1)
- Pfeiffer, T., Rutte, C., Killingback, T., Taborsky, M., & Bonhoeffer, S. (2005a). Evolution of cooperation by generalized reciprocity. *Proceedings of the Royal Society B: Biological Sciences*. <https://doi.org/10.1098/rspb.2004.2988>
- Pfeiffer, T., Rutte, C., Killingback, T., Taborsky, M., & Bonhoeffer, S. (2005b). Evolution of cooperation by generalized reciprocity. *Proceedings of the Royal Society B: Biological Sciences*, 272(1568), 1115–1120. <https://doi.org/10.1098/rspb.2004.2988>
- Phillips, T., Barnard, C., Ferguson, E., & Reader, T. (2008). Do humans prefer altruistic mates? Testing a link between sexual selection and altruism towards non-relatives. *British Journal of Psychology*, 99(4), 555–572. <https://doi.org/10.1348/000712608X298467>

- Pilot, M. (2005). Altruism as an advertisement – a model of the evolution of cooperation based on Zahavi's handicap principle. *Ethology Ecology & Evolution*, 17(3), 217–231. <https://doi.org/10.1080/08927014.2005.9522593>
- Ponti, G., & Rodriguez-Lara, I. (2015). Social preferences and cognitive reflection: Evidence from a dictator game experiment. *Frontiers in Behavioral Neuroscience*, 9. <https://doi.org/10.3389/fnbeh.2015.00146>
- Post, S. G. (2014). Altruism, happiness, and health: It's good to be good. *An Exploration of the Health Benefits of Factors That Help Us to Thrive*, 66-76.
- Proctor, D., Williamson, R. A., de Waal, F. B., & Brosnan, S. F. (2013). Chimpanzees play the ultimatum game. *Proceedings of the National Academy of Sciences*, 110(6), 2070-2075. <https://www.pnas.org/doi/abs/10.1073/pnas.1220806110>
- Rachlin, H., & Jones, B. A. (2008). Altruism among relatives and non-relatives. *Behavioural Processes*, 79(2), 120–123. <https://doi.org/10.1016/j.beproc.2008.06.002>
- Rachlin, H. (2019). Group selection in behavioral evolution. *Behavioural Processes*, 161, 65-72.
- Rainey, P. B., & Rainey, K. (2003). Evolution of cooperation and conflict in experimental bacterial populations. *Nature*, 425(6953), 72–74. <https://doi.org/10.1038/nature01906>
- Ramsey, G. (2016). Can altruism be unified? *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 56, 32–38. <https://doi.org/10.1016/j.shpsc.2015.10.007>
- Rand, D. G., & Epstein, Z. G. (2014). Risking Your Life without a Second Thought: Intuitive Decision-Making and Extreme Altruism. *PLOS ONE*, 9(10), e109687. <https://doi.org/10.1371/journal.pone.0109687>
- Rankin, D. J., & Taborsky, M. (2009). Assortment and the evolution of generalized reciprocity. *Evolution*, 63(7), 1913–1922. <https://doi.org/10.1111/j.1558-5646.2009.00656.x>
- Ratnieks, F. L. W. (2006). The evolution of cooperation and altruism: The basic conditions are simple and well known. *Journal of Evolutionary Biology*, 19(5), 1413–1414. <https://doi.org/10.1111/j.1420-9101.2006.01172.x>
- Ratnieks, F. L. W., & Helanterä, H. (2009). The evolution of extreme altruism and inequality in insect societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1533), 3169–3179. <https://doi.org/10.1098/rstb.2009.0129>
- Richerson, P. J., & Boyd, R. (1998). The evolution of human ultra-sociality. *Indoctrinability, ideology, and warfare: Evolutionary perspectives*, 71-95.
- Roberts, G. (1998). Competitive altruism: From reciprocity to the handicap principle. *Proceedings of the Royal Society B: Biological Sciences*, 265(1394), 427–431. <https://doi.org/10.1098/rspb.1998.0312>
- Roberts, G. (2007). Evolution of direct and indirect reciprocity. *Proceedings of the Royal Society B: Biological Sciences*, 275(1631), 173–179. <https://doi.org/10.1098/rspb.2007.1134>
- Roberts, G. (2017). Prosocial behavior as sexual signaling. *Behavioral and Brain Sciences*, 40.
- Rosas, A. (2008). Multilevel selection and human altruism. *Biology & Philosophy*, 23(2), 205–215. <https://doi.org/10.1007/s10539-007-9083-9>
- Rose, L., Kovarski, K., Caetta, F., Makowski, D., & Chokron, S. (2024). Beyond empathy: Cognitive capabilities increase or curb altruism in middle childhood. *Journal of Experimental Child Psychology*, 239, 105810. <https://doi.org/10.1016/j.jecp.2023.105810>

- Ross, K. G., & Keller, L. (1998). Genetic control of social organization in an ant. *Proceedings of the National Academy of Sciences*, 95(24), 14232–14237. <https://doi.org/10.1073/pnas.95.24.14232>
- Rossetti, C. S., & Hilbe, C. (2024). Direct reciprocity among humans. *Ethology*, 130(4), e13407.
- Rusch, H. (2022). Heroic behavior: A review of the literature on high-stakes altruism in the wild. *Current Opinion in Psychology*, 43, 238–243. <https://doi.org/10.1016/j.copsyc.2021.07.024>
- Rushton, J., Chrisjohn, R. D., & Fekken, G. (1981). The altruistic personality and the self-report altruism scale. *Personality and Individual Differences*, 2(4), 293–302. [https://doi.org/10.1016/0191-8869\(81\)90084-2](https://doi.org/10.1016/0191-8869(81)90084-2)
- Rutte, C., & Taborsky, M. (2007). Generalized Reciprocity in Rats. *PLOS Biology*, 5(7), e196. <https://doi.org/10.1371/journal.pbio.0050196>
- Saad, G., & Gill, T. (2001). Sex Differences in the Ultimatum Game: An Evolutionary Psychology Perspective. *Journal of Bioeconomics*, 3(2), 171–193. <https://doi.org/10.1023/A:1020583425623>
- Sahlins, M. (1972). *Stone Age Economics*. Chicago: Aldine-Atherton. ISBN 0-202-01099-6.
- Sánchez-Amaro, A., Maurits, L., & Haun, D. B. M. (2024). Chimpanzees engage in competitive altruism in a triadic ultimatum game. *Scientific Reports*, 14(1), 3393. <https://doi.org/10.1038/s41598-024-53973-6>
- Sánchez-Amaro, A., & Rossano, F. (2021). Chimpanzees and bonobos use social leverage in an ultimatum game. *Proceedings of the Royal Society B: Biological Sciences*, 288(1962), 20211937. <https://doi.org/10.1098/rspb.2021.1937>
- Sass, M., Timme, F., & Weimann, J. (2015). *The Dynamics of Dictator Behavior* (SSRN Scholarly Paper No. 2607954). Social Science Research Network. <https://doi.org/10.2139/ssrn.2607954>
- Schino, G., & Aureli, F. (2007). Grooming reciprocation among female primates: A meta-analysis. *Biology Letters*, 4(1), 9–11. <https://doi.org/10.1098/rsbl.2007.0506>
- Schino, G., & Aureli, F. (2010a). Primate reciprocity and its cognitive requirements. *Evolutionary Anthropology: Issues, News, and Reviews*, 19(4), 130–135. <https://doi.org/10.1002/evan.20270>
- Schino, G., & Aureli, F. (2010b). The relative roles of kinship and reciprocity in explaining primate altruism. *Ecology Letters*, 13(1), 45–50. <https://doi.org/10.1111/j.1461-0248.2009.01396.x>
- Schmid, L., Chatterjee, K., Hilbe, C., & Nowak, M. A. (2021). A unified framework of direct and indirect reciprocity. *Nature Human Behaviour*, 5(10), 1292–1302. <https://doi.org/10.1038/s41562-021-01114-8>
- Schmid, L., Ekbatani, F., Hilbe, C., & Chatterjee, K. (2023). Quantitative assessment can stabilize indirect reciprocity under imperfect information. *Nature Communications*, 14(1), 2086. <https://doi.org/10.1038/s41467-023-37817-x>
- Schnell, E., & Muthukrishna, M. (2024). Indirect reciprocity undermines indirect reciprocity destabilizing large-scale cooperation. *Proceedings of the National Academy of Sciences*, 121(19), e2322072121.
- Schulz, J. F., Fischbacher, U., Thöni, C., & Utikal, V. (2014). Affect and fairness: Dictator games under cognitive load. *Journal of Economic Psychology*, 41, 77–87. <https://doi.org/10.1016/j.joep.2012.08.007>
- Schwarz, S., & Baßfeld, L. (2019). Do men help only beautiful women in social networks? *Current Psychology*, 38(4), 965–976. <https://doi.org/10.1007/s12144-018-0086-2>
- Schweinfurth, M. K., Aeschbacher, J., Santi, M., & Taborsky, M. (2019). Male Norway rats cooperate according to direct but not generalized reciprocity rules. *Animal Behaviour*, 152, 93–101. <https://doi.org/10.1016/j.anbehav.2019.03.015>
- Seth, J. (1889a). The Evolution of Morality. *Mind*, 14(53), 27–49.

- Seth, J. (1889b). The Evolution of Morality. *Mind*, 14(53), 27–49.
- Sethi, R., & Somanathan, E. (2003). Understanding reciprocity. *Journal of Economic Behavior & Organization*, 50(1), 1–27. [https://doi.org/10.1016/S0167-2681\(02\)00032-X](https://doi.org/10.1016/S0167-2681(02)00032-X)
- Shavit, A., & Millstein, R. L. (2008). Group Selection Is Dead! Long Live Group Selection? *BioScience*, 58(7), 574–575. <https://doi.org/10.1641/B580702>
- Shultziner, D., & Dattner, A. (2016). The merits of self-handicap: The handicap principle as an explanation of altruism compared to reciprocal altruism. *Cogent Biology*. <https://www.tandfonline.com/doi/abs/10.1080/23312025.2016.1140857>
- Silk, J. B. (2009). Nepotistic cooperation in non-human primate groups. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1533), 3243–3254. <https://doi.org/10.1098/rstb.2009.0118>
- Simon, B., Fletcher, J. A., & Doebeli, M. (2013). Towards a general theory of group selection. *Evolution*, 67(6), 1561–1572. <https://doi.org/10.1111/j.1558-5646.2012.01835.x>
- Slater, P. J. B., Rosenblatt, J. S., Snowden, C. T., Roper, T. J., Brockmann, H. J., & Naguib, M. (2005). *Advances in the Study of Behavior*. Elsevier.
- Slembeck, T. (1999). Reputations and Fairness in Bargaining—Experimental Evidence from a Repeated Ultimatum Game With Fixed Opponents. *Experimental*, Article 9905002. <https://ideas.repec.org/p/wpa/wuwpx/9905002.html>
- Sloan Wilson, D. (1997). Altruism And Organism: Disentangling The Themes Of Multilevel Selection Theory. *The American Naturalist*. <https://doi.org/10.1086/286053>
- Smith, J. E. (2014). Hamilton’s legacy: Kinship, cooperation and social tolerance in mammalian groups. *Animal Behaviour*, 92, 291–304. <https://doi.org/10.1016/j.anbehav.2014.02.029>
- Sober, E., & Wilson, D. S. (1998). *Unto others: The evolution and psychology of unselfish behavior* (p. 394). Harvard University Press.
- Soutschek, A., & Schubert, T. (2016). The importance of working memory updating in the Prisoner’s dilemma. *Psychological Research*, 80, 172–180.
- Stevens, J. R. (2004). The selfish nature of generosity: Harassment and food sharing in primates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1538), 451–456. <https://doi.org/10.1098/rspb.2003.2625>
- Stevens, J. R., & Cushman, F. A. (2004). Cognitive Constraints on Reciprocity and Tolerated Scrounging. *Behavioral and Brain Sciences*, 27(4), 569–570. <https://doi.org/10.1017/s0140525x04320121>
- Stewart-Williams, S. (2007). Altruism among kin vs. nonkin: Effects of cost of help and reciprocal exchange. *Evolution and Human Behavior*, 28(3), 193–198. <https://doi.org/10.1016/j.evolhumbehav.2007.01.002>
- Suchak, M., & de Waal, F. B. M. (2012). Monkeys benefit from reciprocity without the cognitive burden. *Proceedings of the National Academy of Sciences*, 109(38), 15191–15196. <https://doi.org/10.1073/pnas.1213173109>
- Suzuki, S., & Akiyama, E. (2007). Evolution of indirect reciprocity in groups of various sizes and comparison with direct reciprocity. *Journal of Theoretical Biology*, 245(3), 539–552. <https://doi.org/10.1016/j.jtbi.2006.11.002>
- Thomas, F., Adamo, S. S., & Moore, J. (2005). Parasitic manipulation: where are we and where should we go?. *Behavioural processes*, 68(3), 185–199.

- Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift Für Tierpsychologie*, 20(4), 410–433. <https://doi.org/10.1111/j.1439-0310.1963.tb01161.x>
- Tinghög, G., Andersson, D., Bonn, C., Böttiger, H., Josephson, C., Lundgren, G., ... & Johannesson, M. (2013). Intuition and cooperation reconsidered. *Nature*, 498(7452), E1-E2.
- Tisserand, J.-C. (2014). Ultimatum game: A meta-analysis of the past three decades of experimental research. *Proceedings of International Academic Conferences*, Article 0802032. <https://ideas.repec.org/p/sek/iacpro/0802032.html>
- Turner, R., & Vallée-Tourangeau, F. (2020). Fiction effects on social cognition: Varying narrative engagement with cognitive load. *Scientific Study of Literature*, 10(1), 64–97. <https://doi.org/10.1075/ssol.19008.tur>
- Twyman, K. Z., & Gardner, A. (2023). Kin selection of time travel: The social evolutionary causes and consequences of dormancy. *Proceedings of the Royal Society B: Biological Sciences*, 290(2006), 20231247. <https://doi.org/10.1098/rspb.2023.1247>
- Vakoch, D. A. (Ed.). (2014). *Extraterrestrial Altruism: Evolution and Ethics in the Cosmos*. Springer. <https://doi.org/10.1007/978-3-642-37750-1>
- Van Houtan, K. S., Rojas-Moscoco, J. I., Van Houtan, H. N., & Gonzalez, O. (2024). Cooperation, status, and altruism in a mixed society of Amazonian parrots. (ecoevorxiv.org preprint)
- Van Vugt, M., Roberts, G., & Hardy, C. (2007). Competitive altruism: A theory of reputation-based cooperation in groups. In L. Barrett & R. Dunbar (Eds.), *Oxford Handbook of Evolutionary Psychology* (p. 0). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198568308.013.0036>
- Viana, D. S., Gordo, I., Sucena, É., & Moita, M. A. P. (2010). Cognitive and Motivational Requirements for the Emergence of Cooperation in a Rat Social Game. *PLOS ONE*, 5(1), e8483. <https://doi.org/10.1371/journal.pone.0008483>
- Villesen, P., Murakami, T., Schultz, T. R., & Boomsma, obus J. (2002). Identifying the transition between single and multiple mating of queens in fungus-growing ants. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1500), 1541–1548. <https://doi.org/10.1098/rspb.2002.2044>
- Warneken, F., & Tomasello, M. (2009). The roots of human altruism. *British Journal of Psychology*, 100(3), 455–471. <https://doi.org/10.1348/000712608X379061>
- Wedekind, C., & Milinski, M. (2000). Cooperation Through Image Scoring in Humans. *Science*, 288(5467), 850–852. <https://doi.org/10.1126/science.288.5467.850>
- Wenseleers, T., & L. W. Ratnieks, F. (2006). Comparative Analysis of Worker Reproduction and Policing in Eusocial Hymenoptera Supports Relatedness Theory. *The American Naturalist*. <https://doi.org/10.1086/508619>
- West, S. A., Cooper, G. A., Ghoul, M. B., & Griffin, A. S. (2021). Ten recent insights for our understanding of cooperation. *Nature Ecology & Evolution*, 5(4), 419–430. <https://doi.org/10.1038/s41559-020-01384-x>
- West, S. A., & Gardner, A. (2010). Altruism, Spite, and Greenbeards. *Science*, 327(5971), 1341–1344. <https://doi.org/10.1126/science.1178332>
- West, S. A., Griffin, A. S., & Gardner, A. (2007a). Evolutionary Explanations for Cooperation. *Current Biology*, 17(16), R661–R672. <https://doi.org/10.1016/j.cub.2007.06.004>

- West, S. A., Griffin, A. S., & Gardner, A. (2007b). Social semantics: Altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology*, 20(2), 415–432. <https://doi.org/10.1111/j.1420-9101.2006.01258.x>
- West, S. A., Griffin, A. S., & Gardner, A. (2008). Social semantics: How useful has group selection been? *Journal of Evolutionary Biology*, 21(1), 374–385. <https://doi.org/10.1111/j.1420-9101.2007.01458.x>
- Weziak-Bialowolska, D., Skiba, R., & Bialowski, P. (2024). Longitudinal reciprocal associations between volunteering, health and well-being: Evidence for middle-aged and older adults in Europe. *European Journal of Public Health*, 34(3), 473–481. <https://doi.org/10.1093/eurpub/ckae014>
- Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature*, 308(5955), 181–184. <https://doi.org/10.1038/308181a0>
- Wilkinson, G. S., Carter, G. G., Bohn, K. M., & Adams, D. M. (2016). Non-kin cooperation in bats. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1687), 20150095. <https://doi.org/10.1098/rstb.2015.0095>
- Wilson, D. S. (1992). On the relationship between evolutionary and psychological definitions of altruism and selfishness. *Biology & Philosophy*, 7(1), 61–68. <https://doi.org/10.1007/BF00130164>
- Wilson, D. S., & Wilson, E. O. (2007). Rethinking the Theoretical Foundation of Sociobiology. *The Quarterly Review of Biology*, 82(4), 327–348. <https://doi.org/10.1086/522809>
- Woodford, P. J. (2019). The many meanings of “cost” and “benefit:” biological altruism, biological agency, and the identification of social behaviours. *Biology & Philosophy*, 34(1), 4. <https://doi.org/10.1007/s10539-018-9667-6>
- Yoeli, E., Hoffman, M., Rand, D. G., & Nowak, M. A. (2013). Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences*, 110(supplement_2), 10424–10429. <https://doi.org/10.1073/pnas.1301210110>
- Zapata, B., González, B. A., & Ebensperger, L. A. (2009). Allonursing in Captive Guanacos, Lama guanicoe: Milk Theft or Misdirected Parental Care? *Ethology*, 115(8), 731–737. <https://doi.org/10.1111/j.1439-0310.2009.01660.x>
- Zwick, M., & Fletcher, J. A. (2014). Levels of Altruism. *Biological Theory*, 9(1), 100–107. <https://doi.org/10.1007/s13752-013-0145-8>

5. SUPPLEMENT

This section contains the following published and submitted papers by the author:

- Novakova, J., Machová, K., Sýkorová, K., Zíka, V., & Flegr, J. (2021). Looking Like a Million Dollars: Does Attractiveness Priming Increase Altruistic Behavior in Experimental Games?. *Frontiers in Psychology*, 1974.
- Novakova, J., Tureček, P., Machová, K., Sýkorová, K., Zíka, V., & Flegr, J. (2024). Generosity as a status signal: Higher-testosterone men exhibit greater altruism in the dictator game. *Evolution and Human Behavior*, 45(5), 106615.
- Novakova, J., Machová, K., Sýkorová, K., Zíka, V., & Flegr, J. (under review). Testing Trivers: Discerning the role of reciprocity in the evolution of altruism. Under review in *Behavioural Processes*.
- Novakova, J., & Flegr, J. (under review). Testing the microbiome-altruism hypothesis: Antibiotic intake doesn't impair altruistic behavior. Under review in *Evolution and Human Behavior*.



Looking Like a Million Dollars: Does Attractiveness Priming Increase Altruistic Behavior in Experimental Games?

Julie Novakova^{1*}, Kamila Machová¹, Katerina Sýkorová¹, Vojtěch Zíka^{1,2} and Jaroslav Flegr^{1,3}

¹ Laboratory of Evolutionary Biology, Department of Philosophy and History of Science, Faculty of Science, Charles University, Prague, Czechia, ² Center for Behavioral Experiments (CEBEX), Prague, Czechia, ³ Applied Neuroscience and Neuroimaging, National Institute of Mental Health, Klecany, Czechia

OPEN ACCESS

Edited by:

Danielle Sulikowski,
Charles Sturt University, Australia

Reviewed by:

Johannes Hewig,
Julius Maximilian University of
Würzburg, Germany
Farid Pazhoohi,
University of British Columbia, Canada

*Correspondence:

Julie Novakova
julie.novakova@natur.cuni.cz

Specialty section:

This article was submitted to
Evolutionary Psychology,
a section of the journal
Frontiers in Psychology

Received: 25 January 2021

Accepted: 03 May 2021

Published: 20 July 2021

Citation:

Novakova J, Machová K, Sýkorová K,
Zíka V and Flegr J (2021) Looking Like
a Million Dollars: Does Attractiveness
Priming Increase Altruistic Behavior in
Experimental Games?
Front. Psychol. 12:658466.
doi: 10.3389/fpsyg.2021.658466

The emergence of altruistic behavior constitutes one of the most widely studied problems in evolutionary biology and behavioral science. Multiple explanations have been proposed, most importantly including kin selection, reciprocity, and costly signaling in sexual selection. In order to test the latter, this study investigated whether people behave more altruistically when primed by photographs of attractive faces and whether more or less altruistic people differ in the number of sexual and romantic partners. Participants in the general population ($N = 158$, 84 F, 74 M) first rated the attractiveness of photographs of 20 faces of the opposite (sexually preferred) sex and then played the Dictator and Ultimatum Games (DG and UG). The photograph rating acted as priming; half the participants received photographs of people rated as more attractive than average in an earlier study, and the other half received photographs previously rated as less attractive. The attractiveness-primed participants, especially men, were expected to behave more altruistically—signaling that they are desirable, resource-possessing partners. We also expected altruists to self-report more sexual and romantic partners. The observed difference between altruistic behaviors in the attractiveness- and unattractiveness-primed groups occurred in UG offers, however, in the opposite than expected direction in women. The number of sexual partners was positively correlated to minimum acceptable offers (MAOs) in the UG, in line with expectations based on the theory of costly signaling.

Keywords: altruism, attractiveness, costly signaling, experimental games, reputation, sexual selection

INTRODUCTION

“To ignore the questions of survival value and evolution... makes it impossible to arrive at an understanding of behavioral problem.”

– Niko Tinbergen (1963).

Evolutionary explanations for altruistic behavior, fairness, and cooperation continue to garner debate even after decades of research studies. Theories ranging from mutualism, kin selection, interdemic selection, reciprocal altruism (direct or indirect, based on individual reputation), and

“greenbeard” altruism to costly signalization and sexual selection have been proposed (Fehr and Fischbacher, 2003; West et al., 2007; Clutton-Brock, 2009). Most are not mutually exclusive and could have contributed to the evolution of altruistic behavior to some extent. Cooperation among humans may also have been shaped by factors that contributed less in other primate species, such as cooperative breeding and cultural evolution (Silk and House, 2016).

Explanations of the adaptive value of altruism and its evolutionary origins do not undermine the psychological mechanisms that produce altruism, because function and mechanism are each different levels of question (Tinbergen, 1963). That altruism may be of adaptive benefit to the altruist and, thus, ultimately “selfish,” does not diminish the psychological motivations for selflessly helping others: experiencing the oft-cited “warm glow,” emphatic concern, adherence to social norms, or displaying an intrinsic helping preference. Those are merely different level facets of the phenomenon that should not be confused (Clavien and Chapuisat, 2013; Kurzban et al., 2015). In the same vein, we should not “hijack” the mechanistic terminology of motivation when deriving evolutionary explanations (De Waal, 2008).

Altruism as a costly signal (driven by sexual selection; e.g., Zahavi, 1995) represents one of the major evolutionary theories of altruism and has received a lot of attention among evolutionary biologists, psychologists, and behavioral scientists studying humans. It is supported by experimental and theoretical data, including models of cooperation as an honest costly signal (e.g., Gintis et al., 2001). It may, however, be difficult to discern the various selection pressures contributing to altruistic behavior. Roberts (1998) suggested a basic criterion for distinguishing reciprocity from competitive altruism (i.e., sexual selection, along with general reputation and alliance building): If the altruistic act is returned later in a similar manner, reciprocity can be invoked (e.g., in allogrooming impalas and allofeeding vampire bats). If the act is not returned, competitive altruism may apply (possibly in primate allogrooming, bird allopreening, or group cooperation). Moreover, Shultziner and Dattner (2016) reasoned that the criteria for indirect reciprocity, an extension of the original reciprocal altruism of Trivers (1971), have become so loose that they make the distinction practically untestable. It blends into competitive altruism through reputation concerns, which are considered by both theories. In human, we can rarely observe a clear reciprocal “transaction,” even across longer timescales, and one-shot altruism is commonplace. This suggests that either altruism used to be based on reciprocity in people earlier but is now a maladaptation of the “paleolithic mind” (Trivers in Buchanan, 2005; Phillips, 2015) or that competitive altruism may be a more important driving force for altruism in humans.

Mate choice based on sexual preference for altruists is supported by a number of studies. In the presence of observers (or even observer cues such as a printed image of human eyes), people are more inclined toward generosity (Haley and Fessler, 2005; Bateson et al., 2006; Burnham and Hare, 2007; Bourrat et al., 2011; Ernest-Jones et al., 2011; Nettle et al., 2013; although the effect may be limited to short cue exposure:

Sparks and Barclay, 2013), and they donate more under non-anonymous settings (Burnham, 2003). Competitive altruism, in general, seems common among humans (Van Vugt et al., 2007). These findings may be attributed to general social selection and reputation seeking (which includes sexual selection as the reputation of an altruist would be advantageous in mate choice).

The effect is even more pronounced in the presence of people of sexually desired sex of an individual, especially in men. Men are more likely to give money to beggars in the company of a woman, particularly in the early stages of a relationship (Latane, 1970), and also to give money to female beggars (Goldberg, 1995). Laboratory results indicate the same. Altruistic behavior is often measured through the Dictator Game (DG) (Forsythe et al., 1994) and the Ultimatum Game (UG) (Güth et al., 1982), both widely used in behavioral sciences. Both games have two players, who in typical laboratory settings cannot see each other or interact directly. In the DG, the first player is given an incentive (usually monetary) by the experimenter and can share it with the second player, who has no power over the outcome of the game. In the UG, the second player can decide whether to accept or reject the proposed offer. In the former, the second player receives the offered portion of the pie and the first player keeps the rest; in the latter, neither receives anything. Modified versions of both games have also been utilized in various animal cooperation studies (e.g., Jensen et al., 2007; Rutte and Taborsky, 2007, 2008; Kaiser et al., 2012; Proctor et al., 2013). DG giving is more clearly a sign of altruism, while the UG involves an element of strategic reasoning, making an offer that is unlikely to be rejected. Humans do not act as “pure gain-maximizers” without other motives, so not just any positive offer gets accepted. In most societies, an even split is the “surest bet,” though there is great variance both within and across societies (Henrich et al., 2006), and offers and rejection rates can be modulated by the stakes (Novakova and Flegr, 2013).

Both games have been used in previous studies that focused on altruistic behavior and physical attractiveness. Men tend to be more generous to the players of the opposite sex in the experimental games (Farrelly et al., 2007; Bhogal et al., 2016) or in the presence of third-party female observers (Iredale et al., 2008). Competitive altruism among men can also be observed in tasks that require some degree of self-sacrifice in order to obtain a reward for co-players (McAndrew and Perilloux, 2012). Even though the effect is more pronounced in men, it is not limited to them. Both sexes cooperated more with more attractive people of the opposite sex in several experimental games (Farrelly et al., 2007; Bhogal et al., 2016).

This pattern seems to occur across human societies. Foraging sex differences among hunter-gatherers may be caused by signaling benefits: Relatively inefficient, but difficult foraging strategies may persist chiefly among men due to costly signaling (Bird et al., 2001). Tognetti et al. (2012) studied cooperation in the Public Goods Game (PGG), played for rice and altruistic giving (*via* a donation test) in the presence of observers of either the same or opposite sex in rural Senegal. Men contributed more in the PGG when observed by either sex, but their contributions in the game gradually increased if observed by young women.

Women contributed more in the presence of other women, but their contributions gradually increased when observed by men, although they did not exceed the overall altruistic behavior shown in the presence of other women. In the donation test, both sexes were more likely to donate in the presence of female observers. The existence and extent of the donations did not influence the perception of observers on the attractiveness of players. The results suggest that sexual selection may be driving the altruistic and cooperative behavior of men, whereas other types of social selection play a greater role in the altruistic behavior of women.

However, Saad and Gill (2001) did not detect any effect of the physical attractiveness of recipients on UG offers but only detected that men were more generous toward women. Similarly, Bhogal et al. (2017) found no relationship between the attractiveness of people and the level of altruistic behavior of others directly toward them in a face-to-face DG/UG. But their games were played for chocolate coins, and food, unlike money, acts as an immediate reinforcer in humans and other animals (Mazur, 1994; Stephens et al., 2006; Epstein et al., 2010), which might account for the null results. Finally, in terms of non-financial altruism, Schwarz and Baßfeld (2019) and Bhogal et al. (2019) found a positive effect of attractiveness in both sexes on willingness to help a stranger on a social network and in imaginary credit sharing, respectively.

There is also evidence that altruistic people are perceived as more attractive mates, especially in female choice. Phillips et al. (2008) found mate preference for altruism in both sexes, more strongly in women. Their twin study (Phillips et al., 2010) suggests there may be a genetic component in preference for altruism. Barclay (2010) presented volunteers with descriptions of people with varying traits, including altruism, and found that more altruistic people were considered to be more desired partners by both sexes. Arnocky et al. (2017) showed that more altruistic people (with this trait both self-reported and measured in a modified DG) had more self-reported sexual partners, suggesting the role of sexual selection in human altruistic behavior.

We aimed to investigate the effect of attractiveness in both ways, i.e., expecting subjects (especially men) to behave more altruistically in the DG/UG (both used in previous studies) after viewing photographs of attractive people of the opposite sex and expecting altruists to have more self-reported sexual and romantic partners. As potential confounders, we also measured the Big Five personality traits, where especially agreeableness was previously associated with altruism (Ashton et al., 1998). In addition, we investigated the effect of the self-perceived attractiveness of participants on their behavior in the experimental games. In distinguishing these effects in a single study, we are filling an important gap in the existing literature studies. Based on the theory of costly selection, attractiveness-primed people and people with more partners and self-perceived as more attractive should be more altruistic in giving and offering but also more demanding from others (= less willing to accept low UG offers) in order to keep or increase their relative social status.

MATERIALS AND METHODS

Subjects and Course of the Experiment

The data were collected at the Charles University in Prague in June and July 2018/2019. Participants were invited *via* online recruitment on Facebook, with a maximum of 18 participants invited for each session, to take part in several tasks for a set of studies. Each session consisted of either men or women only in order to eliminate stochastic noise (the presence/absence of attractive opposite-sex individuals among the participants of an experimental session). Only self-reported heterosexuals were recruited for the study in order to clearly discern relationships between the measured variables on a sample of the target size. In total, 158 people participated in the experiments (74 men: mean age = 29, median of age = 29; 84 women: mean age = 31, median of age = 28).

Participants were greeted at the reception and led to the computer laboratory with seats separated by cardboard screens so that the participants could not see each other or otherwise interact during the experiment. They read and signed their informed consent and were able to ask questions about it before the commencement of the experiments. At the beginning of the computer-based survey, participants viewed 20 neutral-expression frontal portrait color photographs of people of the opposite sex. The priming part of the session was always overseen by an experimenter of the same sex as the participants.

The participants were asked to rate the attractiveness of the faces on an 8-point Likert scale [same as in the source study, Flegr et al. (2019), employed in order to avoid the “middle” response]. One half was randomly allocated photographs rated as less attractive on the same Likert scale by an independent sample in a previous questionnaire (Machová, 2018), and the other half was shown the 20 faces rated as more attractive. The unattractive sample of male photographs was originally rated 2.73 on average and the attractive was rated 3.97 ($t = 7.98$, $df = 37.67$, $p < 0.0001$) on an 8-point scale. The unattractive sample of female photographs had an average rating of 3.01 and the attractive was rated 5.21 ($t = 9.64$, $df = 33.05$, $p < 0.0001$) on the same scale. Differences of the priming sets in other characteristics are shown in **Supplementary Table 1**. The photographs used in the study of Machová (2018) and rated there for attractiveness belonged to materials from Flegr et al. (2019): randomly selected freely available photographs on the internet, where the selection criteria were estimated with ages 18–70, frontal view, neutral expression, resolution of at least 1,165×1,476 pixels, background, and hair possibly covered with a gray mask in postproduction.

Afterward, the subjects played first the DG and then the UG, each for 400 tokens (=40 CZK, approx. 2 USD), with no intermission. They were presented with the rules and truthfully told that they would be randomly matched with another participant in the room, and their earnings at the end of the experiment would reflect the outcome of both games. Each participant was asked how much they wanted to give and offer the other player, what is their minimum acceptable offer (MAO) in the UG, and also whether they would hypothetically accept or reject an offer of 120 tokens (30% of the pie). The survey contained control questions to assess whether the participants

understood the rules of the games before they played and also a question about previous knowledge of the games.

The participants subsequently filled in a questionnaire asking about the number of sexual and romantic partners during the past year and their life so far. Three questions were aimed at self-perceived attractiveness of an individual in terms of both physical appearance (of the face and body of an individual) and personality on a 5-point scale. Big Five personality factors (agreeableness, extraversion, openness, neuroticism, and conscientiousness) were also measured *via* Ten-Item Personality Inventory (TIPI) (Gosling et al., 2003), as previous research studies found a connection for each of them with some aspects of altruism and cooperation, and they might modulate the responses of the participants (Clark et al., 2014; Oda et al., 2014).

Finally, in order to avoid the possible confounding effects of the status of an individual, the participants were also asked about their economic and social-family life satisfaction. Other data were collected for related studies on conditions of Triversian reciprocity and hormonal effects, respectively: health, temporal discounting, cognitive reflection, intelligence, memory, risk-taking propensity, facial photograph, 2D:4D, and salivary cortisol and testosterone levels. At the end of the testing, subjects were informed about the goals of the experiment and were paid their reward.

Statistical Analysis

The analysis was conducted in R (R 3.6.1 for 64-bit Linux; R Core Team, 2013). Responses were checked for their validity (completion of the survey and answering control questions about understanding the principles of the games correctly), and the aggregate variables were computed (self-perceived attractiveness as added points from the perceived body, face, and personality attractiveness of an individual; personality traits from TIPI questions). Descriptive statistics were calculated first.

Afterward, a linear regression model was built for each of the studied dependent variables: allocation to another player in the DG, offer made to another player in the UG, the MAO in the UG, and reaction to the hypothetical offer of 120 tokens (30% of the pie) in the UG. In each starting model, the following predictors were used: participant sex, age, previous knowledge of the games, priming (and its interaction with sex), Big Five personality traits (and their interactions with sex), number of romantic partners in the previous year (and its interaction with sex), number of sexual partners in the previous year (and its interaction with sex), number of romantic partners in total (and its interaction with sex and age), number of sexual partners in total (and its interaction with sex and age), and self-perceived attractiveness of an individual (and its interaction with sex and age). Backward elimination was then applied to reach the best-fitting model.

With alpha set to 0.05, $N = 158$, 14 independent variables, and expected small effect size [estimated 0.14 for the purposes of the power analysis (PA)], the power of the test was 0.91, which was calculated using the R package “pwr” (Champely et al., 2018).

Finally, values of DG allocations and UG offers that are statistical outliers were kept in the analyses, because while DG and UG responses tend to have a strong central tendency to the mode, they also exhibit great range within and across different

samples (Henrich et al., 2006), and outlier exclusion in this case would potentially cause a great distortion from reality.

The dataset used in our analyses has been published on FigShare (10.6084/m9.figshare.12067407).

RESULTS

Descriptive Statistics

All 158 responses were valid, meaning that the participants finished the experiment and correctly answered the control questions. The mean DG allocation was 154.4 out of 400 tokens (~38.6%), median was 200 (50%), and SD was 76.2. The mean UG offer was 202 tokens (50.5%), median was 200 (50%), and SD was 51.8. One-sided Pearson's correlation test (expecting a positive correlation) showed both offers to be weakly correlated ($r = 0.152$, $t = 1.924$, $df = 156$, $p = 0.028$).

Mean MAO in the UG was 140.4 tokens (35.1%), median was 150 (37.5%), and SD was 69.6. Gender differences in allocations, offers, and MAOs were observed, but none were statistically significant (see **Table 1**). Finally, 56% (54% of women and 58% of men) of participants would accept a hypothetical offer of 120 tokens, while the rest would reject it. Previous knowledge of the games had no significant effect on the responses of the participants.

Confirmatory Statistics

Dictator Game Allocations

The only significant predictor of DG offer that remained after performing backward elimination on the starting full linear regression model was the trait agreeableness ($t = 2.269$, $SE = 5.291$, $p = 0.025$): the higher an individual scored in agreeableness, the more they allocated to others in the DG.

Ultimatum Game Offers

UG offers were most strongly predicted by priming ($t = -2.940$, $SE = 11.098$, $p = 0.004$), followed by its interaction with sex ($t = 2.862$, $SE = 16.184$, $p = 0.005$), extraversion–sex interaction ($t = -2.802$, $SE = 4.559$, $p = 0.006$), neuroticism–sex interaction ($t = 2.284$, $SE = 4.580$, $p = 0.024$), trends with extraversion ($t = 1.955$, $SE = 3.369$, $p = 0.052$), and neuroticism ($t = -1.807$, $SE = 3.584$, $p = 0.073$) alone. Priming by attractiveness, therefore, led to lower UG offers overall and in women but increased offers by male participants (see **Figure 1**). Extraversion predicted higher offers in women but lower offers in men. An opposite effect was observed for neuroticism.

When the analysis was subsequently performed for each of the sexes separately, the significant negative effect of attractiveness priming was retained in women ($t = -2.795$, $SE = 11.346$, $p = 0.007$), while the effects of extraversion ($p = 0.060$) and neuroticism ($p = 0.103$) were not significant in this limited sample. None of the effects were significant in a subsample of men alone (priming $p = 0.403$, extraversion $p = 0.143$, and neuroticism $p = 0.407$).

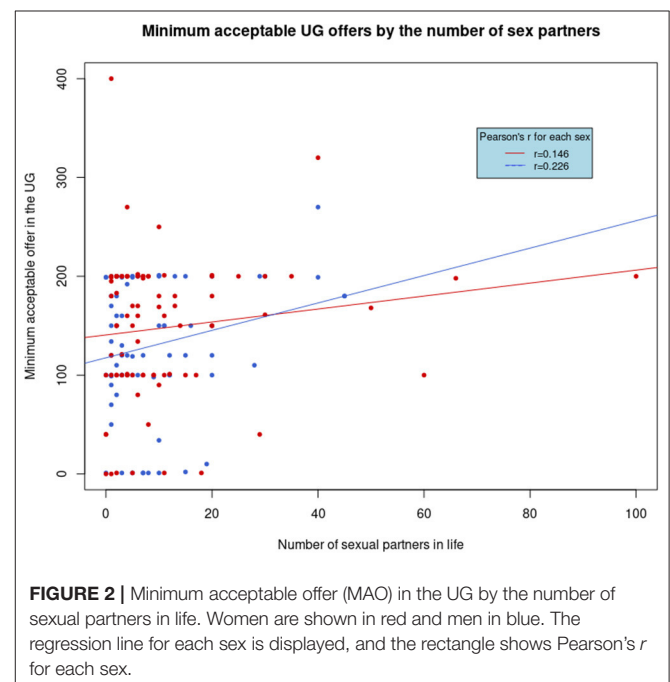
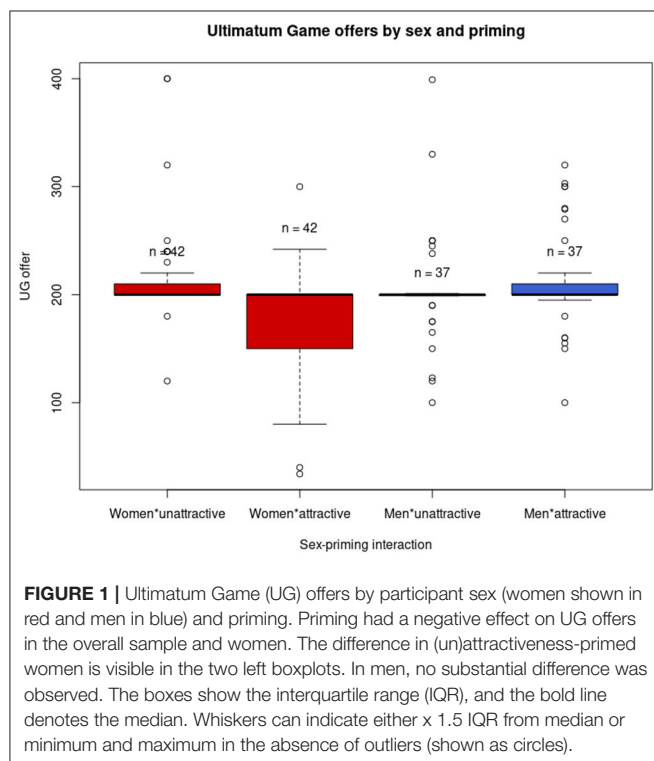
Ultimatum Game Responses

The minimum acceptable UG offer was significantly (and positively) predicted only by the number of lifetime sexual

TABLE 1 | Gender differences in DG and UG responses.

	All participants		Women		Men	
	Mean	Median	Mean	Median	Mean	Median
DG allocation	154.4 (38.6%)	200 (50%)	165 (41.25%)	200 (50%)	142.42 (35.61%)	200 (50%)
Statistical difference between sexes	–		$t = 1.830, df = 125.6, p = 0.070$			
UG offer	202 (50.5%)	200 (50%)	197.4 (49.35%)	200 (50%)	207.9 (51.98%)	200 (50%)
Statistical difference between sexes	–		$t = -1.284, df = 155.9, p = 0.201$			
UG MAO	140.4 (35.1%)	150 (37.5%)	148.2 (37.05%)	160.5 (40.13%)	131.5 (32.88%)	134 (33.5%)
Statistical difference between sexes	–		$t = 1.509, df = 154.3, p = 0.133$			

The numbers in parentheses show the fraction of the maximum possible sum. The basic difference of response between sexes was measured by two-sampled t-tests.



partners ($t = 2.269, SE = 0.399, p = 0.025$; shown graphically in **Figure 2**), with no significant interactions or trends, and no predictors proved significant for the rejection or acceptance of the hypothetical offer of 120 tokens.

DISCUSSION AND CONCLUSIONS

The behavior of participants in this study in both experimental games was roughly comparable to previous studies, where the average DG allocation is typically 25–30% of the pie, and the modal and median offers in the UG tend to be a fair split, with the average offer slightly above 40%. Offers of 20% or lower are more often rejected than those which are not, and offers of 30%

also tend to be frequently rejected (e.g., Camerer, 2003; Engel, 2011; Tisserand et al., 2015). Our subject pool seems moderately more generous than the average, but well within the norm, and roughly conformed to previous findings in the same cultural setting (Lindová et al., 2010; Novakova and Flegr, 2013; Veselý, 2014).

We found no effect of attractiveness priming on DG allocations or responses to UG offers, but it had the expected effect on UG offers made by men and the opposite effect on offers made by women. In addition, we found that subjects of both sexes who were more demanding and who sought fairness in the UG had more self-reported sexual partners. This is in line with the theory of costly signaling. An UG offer rejection is a costly act that is commonly interpreted as altruistic punishment; however,

we suggest that its role in establishing social rank and reputation is more probable.

The study finding that attractiveness priming positively influenced offers by men in the experimental situation of the UG is in line with the majority of studies so far (Farrelly et al., 2007; Bhogal et al., 2016, 2019; Schwarz and Baßfeld, 2019), whereas the observed opposite effect in women stands in contrast to them. While some previous research studies found no connection between attractiveness and altruistic behavior (Saad and Gill, 2001; Bhogal et al., 2017), this is to our knowledge the first occurrence of the opposite relationship in either sex. When we divided the sample by sex, the negative effect of priming remained significant in women ($n = 84$), while the positive effect in the smaller sample of men ($n = 74$) did not retain significance.

The curious negative effect of attractiveness on UG offers by women calls for explanation, and we can offer several speculations of such. It is worth pointing out that no effect manifested in the DG, which represents a clearer case of altruism, and its existence in the more complex UG may, therefore, be linked to the strategic element of the experimental situation.

While the participants were specifically primed for attractiveness by rating this trait of the faces on photographs, we cannot exclude the possibility that some other factors of the shown images may have influenced the decision-making of the participants. The original study of Machová (2018) included data not only for attractiveness but also for a number of other characteristics rated by volunteers online. The attractiveness of male faces rated by women correlated positively with most traits or descriptions (nice, an ideal partner for a one-night stand, an ideal life partner, sexy, charismatic, intelligent, fun, rich, generous, altruistic, trustworthy, healthy, mature, good skin quality, and symmetrical), weakly but positively with perceived faithfulness and maturity, and negatively with perceived masculinity and dominance. The attractiveness of female faces rated by men correlated positively with all traits but dominance and faithfulness. We can speculate that the halo effect of beauty (e.g., Zebrowitz and Franklin, 2014), i.e., the positive correlation of attractiveness with the majority of desirable traits and descriptions, could have enhanced the perception of another characteristic in our subjects, although attractiveness was made salient by the rating. **Supplementary Table 1** summarizes the differences in various characteristics by the attractive/unattractive sample. Dominance, however, which may have seemed a likely potential cause of the results, did not significantly differ in either men or women. It is possible, though, that another trait or emotion could have been attributed to the faces, since neutral-expression faces are often ascribed emotional states by observers (Hester, 2019). On the contrary, it does not explain the negative effect exhibiting in female participants only.

We can also theorize that we observed no significant effect of attractiveness priming in men because of the same-sex experimental sessions (chosen in order to exclude the confounding effect of more or less attractive opposite-sex co-players), and in short, there were no “real” members of the opposite sex to impress with the generosity of an individual. However, running mixed-sex sessions could introduce stochastic noise (and potentially even systemic bias if predominantly

less/more attractive women or men attended) into the results, and viewing photographs of members of the opposite sex directly before playing the experimental games should suffice to elicit an unconscious response to the presence of opposite sex (akin to the effect of “watching eyes” cues, reported by Haley and Fessler, 2005 and others). The immediacy of the games following the priming should also increase the effect, since long-exposure cues can cause habituation and weaker effects (Sparks and Barclay, 2013). A careful comparison of the effect of real-life (un)attractive observers and comparably (un)attractive images is needed in the future and constitutes an interesting line of research studies, akin to discerning nuances in the effects of observer cues brought up above (such as printed eyes) and real-life observers.

Allocations of DG in this study correlated positively with agreeableness (in line with Ben-Ner et al., 2004), and UG offers in the full sample were predicted by extraversion and neuroticism in interaction with sex, not unprecedented in previous literature (Ashton et al., 1998; Oda et al., 2014; Nakavachara, 2018), but these facts bear little relevance for costly signaling. However, the number of sexual partners in life was correlated to the MAOs in the UG: the higher the number, the higher was the MAO, as shown in **Figure 2**. A preference for an equal, or at least close-to-equal, split can indicate fairness (or inequity aversion), one of the psychological mechanisms of altruism, where rejecting an unfair offer acts as costly punishment. However, it is not necessarily altruistic, as it could also serve as a mechanism for avoiding the imposition of inferior status, rather than a prosocial act of punishing the responders so that they would behave more equitably in future interactions. While a cross-cultural study by Henrich et al. (2006) found a correlation between rejection rates in the UG and altruistic behavior in the DG, several other studies failed to find a link between costly punishment in the UG and prosocial behavior in other experimental games and situations (Ohmura and Yamagishi, 2005; Yamagishi et al., 2012; Brethel-Haurwitz et al., 2016) or failed to find meaningful third-party punishment, which could be considered altruistic (Pedersen et al., 2013) from an evolutionary, if not psychological, point of view.

These findings are in line with the current study, since sexual selection should lead to avoidance of being ascribed inferior social status. A correlation between DG allocation and MAO or between rejection of an unfair offer and MAO was not found, suggesting that UG rejections were more due to reputation-seeking rather than altruistic punishment. It is also in accordance with the positive relationship of self-perceived attractiveness and MAO in our sample. However, it could also be suggested that another mediating variable (such as self-esteem, which would likely be higher on average in people, especially men, with a greater number of sexual partners) was responsible for modulating the UG rejections and MAO. Self-esteem had been previously linked to altruistic behavior (but not necessarily altruistic motivations, see Batson et al., 1986; on the other hand, studies in different contexts did not find a common pattern, and the question of the hen and egg—whether self-esteem precedes altruistic behavior or the other way around, or the relationship of these traits goes in both directions—is not entirely clear; see e.g., Hessing and Elffers, 1985; Jiang et al., 2017).

In conclusion, we did not find evidence of attractiveness cues increasing altruistic behavior as predicted by the costly signaling theory. Women who had participated in this study behaved in the opposite direction; in men, the effect theoretically ran in the expected direction but was not significant. However, MAOs of UG correlated with the number of self-reported sexual partners, which is in line with the costly signaling theory. The findings highlight the need for replication of “well-known” effects (Ioannidis, 2005; Loken and Gelman, 2017). Further research studies are needed to address the relationships between the attractiveness of sexual cues with altruistic behavior and to shed more light on the role of sexual selection in the evolution of human altruism.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: FigShare: doi.org/10.6084/m9.figshare.12067407.

ETHICS STATEMENT

The study was approved by the research ethics board of the Faculty of Science, Charles University. The study was performed in accordance with the ethical standards as laid down in the

1964 Declaration of Helsinki and its later amendments. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JN: study conception, participant invitation, data collection, analysis, and writing the paper. KM, KS, and VZ: data collection and comments on the analysis and paper. JF: study conception and assistance with analysis and writing the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the grant agency of Charles University (grant no. GA UK 1502018) and the Czech Science Foundation (GACR) (grant no. 18-13692S).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.658466/full#supplementary-material>

REFERENCES

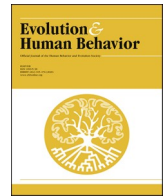
- Arnocky, S., Piché, T., Albert, G., Ouellette, D., and Barclay, P. (2017). Altruism predicts mating success in humans. *Br. J. Psychol.* 108, 416–435. doi: 10.1111/bjop.12208
- Ashton, M. C., Paunonen, S. V., Helmes, E., and Jackson, D. N. (1998). Kin altruism, reciprocal altruism, and the Big Five personality factors. *Evol. Hum. Behav.* 19, 243–255. doi: 10.1016/S1090-5138(98)00009-9
- Barclay, P. (2010). Altruism as a courtship display: some effects of third-party generosity on audience perceptions. *Br. J. Psychol.* 101, 123–135. doi: 10.1348/000712609X435733
- Bateson, M., Nettle, D., and Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biol. Lett.* 2, 412–414. doi: 10.1098/rsbl.2006.0509
- Batson, C. D., Bolen, M. H., Cross, J. A., and Neuringer-Benefiel, H. E. (1986). Where is the altruism in the altruistic personality? *J. Pers. Soc. Psychol.* 50:212. doi: 10.1037/0022-3514.50.1.212
- Ben-Ner, A., Kong, F., and Putterman, L. (2004). Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving. *J. Econ. Psychol.* 25, 581–589. doi: 10.1016/S0167-4870(03)00065-5
- Bhagal, M. S., Bartlett, J. E., and Farrelly, D. (2019). The influence of mate choice motivation on non-financial altruism. *Curr. Psychol.* 38, 959–964. doi: 10.1007/s12144-018-0070-x
- Bhagal, M. S., Galbraith, N., and Manktelow, K. (2016). Physical attractiveness and altruism in two modified dictator games. *Basic Appl. Soc. Psychol.* 38, 212–222. doi: 10.1080/01973533.2016.1199382
- Bhagal, M. S., Galbraith, N., and Manktelow, K. (2017). Physical attractiveness, altruism and cooperation in an ultimatum game. *Curr. Psychol.* 36, 549–555. doi: 10.1007/s12144-016-9443-1
- Bird, R. B., Smith, E., and Bird, D. W. (2001). The hunting handicap: costly signaling in human foraging strategies. *Behav. Ecol. Sociobiol.* 50, 9–19. doi: 10.1007/s002650100338
- Bourrat, P., Baumard, N., and McKay, R. (2011). Surveillance cues enhance moral condemnation. *Evol. Psychol.* 9:147470491100900206. doi: 10.1177/147470491100900206
- Brethel-Haurwitz, K. M., Stoycos, S. A., Cardinale, E. M., Huebner, B., and Marsh, A. A. (2016). Is costly punishment altruistic? Exploring rejection of unfair offers in the Ultimatum Game in real-world altruists. *Sci. Rep.* 6:18974. doi: 10.1038/srep18974
- Buchanan, M. (2005). Charity begins at homo sapiens. *New Sci.* 185, 33–37.
- Burnham, T. C. (2003). Engineering altruism: a theoretical and experimental investigation of anonymity and gift giving. *J. Econ. Behav. Org.* 50, 133–144. doi: 10.1016/S0167-2681(02)00044-6
- Burnham, T. C., and Hare, B. (2007). Engineering human cooperation. *Hum. Nat.* 18, 88–108. doi: 10.1007/s12110-007-9012-2
- Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.
- Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., et al. (2018). *Package 'pwr'*. R Package Version 1.
- Clark, C. B., Thorne, C. B., Vann, J., and Cropsey, K. L. (2014). Five-factor personality and cooperative behavior. *North Am. J. Psychol.* 16, 481–494.
- Clavien, C., and Chapuisat, M. (2013). Altruism across disciplines: one word, multiple meanings. *Biol. Philos.* 28, 125–140. doi: 10.1007/s10539-012-9317-3
- Clutton-Brock, T. (2009). Cooperation between non-kin in animal societies. *Nature* 462, 51–57. doi: 10.1038/nature08366
- De Waal, F. B. (2008). Putting the altruism back into altruism: the evolution of empathy. *Annu. Rev. Psychol.* 59, 279–300. doi: 10.1146/annurev.psych.59.103006.093625
- Engel, C. (2011). Dictator games: a meta study. *Exp. Econ.* 14, 583–610. doi: 10.1007/s10683-011-9283-7
- Epstein, L. H., Salvy, S. J., Carr, K. A., Dearing, K. K., and Bickel, W. K. (2010). Food reinforcement, delay discounting and obesity. *Physiol. Behav.* 100, 438–445. doi: 10.1016/j.physbeh.2010.04.029

- Ernest-Jones, M., Nettle, D., and Bateson, M. (2011). Effects of eye images on everyday cooperative behavior: a field experiment. *Evol. Hum. Behav.* 32, 172–178. doi: 10.1016/j.evolhumbehav.2010.10.006
- Farrelly, D., Lazarus, J., and Roberts, G. (2007). Altruists attract. *Evol. Psychol.* 5:147470490700500205. doi: 10.1177/147470490700500205
- Fehr, E., and Fischbacher, U. (2003). The nature of human altruism. *Nature* 425, 785–791. doi: 10.1038/nature02043
- Flegr, J., Blum, A. E., Nekola, O., and Kroupa, Š. (2019). What people prefer and what they think they prefer in short- and long-term partners. The effects of the phase of the menstrual cycle, hormonal contraception, pregnancy, and the marital and the parenthood status on partner preferences. *Evol. Hum. Behav.* 40, 112–125. doi: 10.1016/j.evolhumbehav.2018.09.003
- Forsythe, R., Horowitz, J. L., Savin, N. E., and Sefton, M. (1994). Fairness in simple bargaining experiments. *Games Econ. Behav.* 6, 347–369. doi: 10.1006/game.1994.1021
- Gintis, H., Smith, E. A., and Bowles, S. (2001). Costly signaling and cooperation. *J. Theor. Biol.* 213, 103–119. doi: 10.1006/jtbi.2001.2406
- Goldberg, T. L. (1995). Altruism towards panhandlers: who gives? *Hum. Nat.* 6, 79–89. doi: 10.1007/BF02734137
- Gosling, S. D., Rentfrow, P. J., and Swann, W. B. Jr. (2003). A very brief measure of the big five personality domains. *J. Res. Pers.* 37, 504–528. doi: 10.1016/S0092-6566(03)00046-1
- Güth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Org.* 3, 367–388. doi: 10.1016/0167-2681(82)90011-7
- Haley, K. J., and Fessler, D. M. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evol. Hum. Behav.* 26, 245–256. doi: 10.1016/j.evolhumbehav.2005.01.002
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., et al. (2006). Costly punishment across human societies. *Science* 312, 1767–1770. doi: 10.1126/science.1127333
- Hessing, D. J., and Elffers, H. (1985). General and physical self-esteem and altruistic behavior. *Psychol. Rep.* 56:930. doi: 10.2466/pr0.1985.56.3.930
- Hester, N. (2019). Perceived negative emotion in neutral faces: gender-dependent effects on attractiveness and threat. *Emotion* 19:1490. doi: 10.1037/emo0000525
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med.* 2:e124. doi: 10.1371/journal.pmed.0020124
- Iredale, W., Van Vugt, M., and Dunbar, R. (2008). Showing off in humans: Male generosity as a mating signal. *Evol. Psychol.* 6:147470490800600302. doi: 10.1177/147470490800600302
- Jensen, K., Call, J., and Tomasello, M. (2007). Chimpanzees are rational maximizers in an ultimatum game. *Science* 318, 107–109. doi: 10.1126/science.1145850
- Jiang, H., Chen, G., and Wang, T. (2017). Relationship between belief in a just world and Internet altruistic behavior in a sample of Chinese undergraduates: multiple mediating roles of gratitude and self-esteem. *Pers. Individ. Differ.* 104, 493–498. doi: 10.1016/j.paid.2016.09.005
- Kaiser, J., Jensen, K., Call, J., and Tomasello, M. (2012). Theft in an ultimatum game: Chimpanzees and bonobos are insensitive to unfairness. *Biol. Lett.* 8, 942–945. doi: 10.1098/rsbl.2012.0519
- Kurzban, R., Burton-Chellaw, M. N., and West, S. A. (2015). The evolution of altruism in humans. *Ann. Rev. Psychol.* 66, 575–599. doi: 10.1146/annurev-psych-010814-015355
- Latane, B. (1970). Field studies of altruistic compliance. *Represent. Res. Soc. Psychol.* 1, 49–61. doi: 10.1016/0022-1031(69)90006-7
- Lindová, J., Kubena, A. A., Štůrcová, A., Krivohlavá, R., Novotná, M., Rubešová, A., et al. (2010). Pattern of money allocation in experimental games supports the stress hypothesis of gender differences in *Toxoplasma gondii*-induced behavioural changes. *Folia Parasitol.* 57, 136–142. doi: 10.14411/fp.2010.017
- Loken, E., and Gelman, A. (2017). Measurement error and the replication crisis. *Science* 355, 584–585. doi: 10.1126/science.aal3618
- Machová, K. (2018). *Vzájemné vztahy vlastností vyčtených z fotografií obličejů* (thesis). Charles University, Prague. Available online at: <https://dspace.cuni.cz/bitstream/handle/20.500.11956/98093/120296471.pdf> (accessed March 1, 2020).
- Mazur, J. E. (1994). Effects of intertrial reinforcers on self-control choice. *J. Exp. Anal. Behav.* 61, 83–96. doi: 10.1901/jeab.1994.61-83
- McAndrew, F. T., and Perilloux, C. (2012). Is self-sacrificial competitive altruism primarily a male activity? *Evol. Psychol.* 10:147470491201000107. doi: 10.1177/147470491201000107
- Nakavachara, V. (2018). The economics of altruism: the old, the rich, the female. *J. Hum. Behav. Soc. Environ.* 28, 28–43. doi: 10.1080/10911359.2017.1393367
- Nettle, D., Harper, Z., Kidson, A., Stone, R., Penton-Voak, I. S., and Bateson, M. (2013). The watching eyes effect in the Dictator Game: it's not how much you give, it's being seen to give something. *Evol. Hum. Behav.* 34, 35–40. doi: 10.1016/j.evolhumbehav.2012.08.004
- Novakova, J., and Flegr, J. (2013). How much is our fairness worth? The effect of raising stakes on offers by proposers and minimum acceptable offers in dictator and ultimatum games. *PLoS ONE* 8:e60966. doi: 10.1371/journal.pone.0060966
- Oda, R., Machii, W., Takagi, S., Kato, Y., Takeda, M., Kiyonari, T., et al. (2014). Personality and altruism in daily life. *Pers. Individ. Differ.* 56, 206–209. doi: 10.1016/j.paid.2013.09.017
- Ohmura, Y., and Yamagishi, T. (2005). Why do people reject unintended inequity? Responders' rejection in a truncated ultimatum game. *Psychol. Rep.* 96, 533–541. doi: 10.2466/pr0.96.2.533-541
- Pedersen, E. J., Kurzban, R., and McCullough, M. E. (2013). Do humans really punish altruistically? A closer look. *Proc. R. Soc. B Biol. Sci.* 280:20122723. doi: 10.1098/rspb.2012.2723
- Phillips, T. (2015). Human altruism and cooperation explainable as adaptations to past environments no longer fully evident in the modern world. *Q. Rev. Biol.* 90, 295–314. doi: 10.1086/682589
- Phillips, T., Barnard, C., Ferguson, E., and Reader, T. (2008). Do humans prefer altruistic mates? Testing a link between sexual selection and altruism towards non-relatives. *Br. J. Psychol.* 99, 555–572. doi: 10.1348/000712608X298467
- Phillips, T., Ferguson, E., and Rijdsdijk, F. (2010). A link between altruism and sexual selection: genetic influence on altruistic behaviour and mate preference towards it. *Br. J. Psychol.* 101, 809–819. doi: 10.1348/000712610X493494
- Proctor, D., Williamson, R. A., de Waal, F. B., and Brosnan, S. F. (2013). Chimpanzees play the ultimatum game. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2070–2075. doi: 10.1073/pnas.1220806110
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*.
- Roberts, G. (1998). Competitive altruism: from reciprocity to the handicap principle. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 265, 427–431. doi: 10.1098/rspb.1998.0312
- Rutte, C., and Taborsky, M. (2007). Generalized reciprocity in rats. *PLoS Biol.* 5:e196. doi: 10.1371/journal.pbio.0050196
- Rutte, C., and Taborsky, M. (2008). The influence of social experience on cooperative behaviour of rats (*Rattus norvegicus*): direct vs generalised reciprocity. *Behav. Ecol. Sociobiol.* 62, 499–505. doi: 10.1007/s00265-007-0474-3
- Saad, G., and Gill, T. (2001). Sex differences in the ultimatum game: an evolutionary psychology perspective. *J. Bioecon.* 3, 171–193. doi: 10.1023/A:1020583425623
- Schwarz, S., and Baßfeld, L. (2019). Do men help only beautiful women in social networks? *Curr. Psychol.* 38, 965–976. doi: 10.1007/s12144-018-0086-2
- Shultziner, D., and Dattner, A. (2016). The merits of self-handicap: The handicap principle as an explanation of altruism compared to reciprocal altruism. *Cogent Biol.* 2:1140857. doi: 10.1080/23312025.2016.1140857
- Silk, J. B., and House, B. R. (2016). The evolution of altruistic social preferences in human groups. *Philos. Transac. R. Soc. B Biol. Sci.* 371:20150097. doi: 10.1098/rstb.2015.0097
- Sparks, A., and Barclay, P. (2013). Eye images increase generosity, but not for long: the limited effect of a false cue. *Evol. Hum. Behav.* 34, 317–322. doi: 10.1016/j.evolhumbehav.2013.05.001
- Stephens, D. W., McLinn, C. M., and Stevens, J. R. (2006). Effects of temporal clumping and payoff accumulation on impulsiveness and cooperation. *Behav. Proc.* 71, 29–40. doi: 10.1016/j.beproc.2005.09.003
- Tinbergen, N. (1963). On aims and methods of ethology. *Z. Tierpsychol.* 20, 410–433. doi: 10.1111/j.1439-0310.1963.tb01161.x
- Tisserand, J. C., Cochard, F., and Le Gallo, J. (2015). *Altruistic or Strategic Considerations: A Meta-Analysis on the Ultimatum and Dictator Games*. Besançon: CRESE, Université de Franche-Comté.
- Tognetti, A., Berticat, C., Raymond, M., and Faurie, C. (2012). Sexual selection of human cooperative behaviour: an experimental study

- in rural Senegal. *PLoS ONE* 7:e44403. doi: 10.1371/journal.pone.0044403
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57. doi: 10.1086/406755
- Van Vugt, M., Roberts, G., and Hardy, C. (2007). “Competitive altruism: development of reputation-based cooperation in groups,” in *The Oxford Handbook of Evolutionary Psychology*, eds R. I. M. Dunbar, and L. Barrett (Oxford: Oxford University Press), 531–540. doi: 10.1093/oxfordhb/9780198568308.013.0036
- Vesely, S. (2014). Ultimatum game with asymmetric information: a study of deception and fairness. *Stud. Psychol.* 56:53. doi: 10.21909/sp.2014.01.650
- West, S. A., Griffin, A. S., and Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evol. Biol.* 20, 415–432. doi: 10.1111/j.1420-9101.2006.01258.x
- Yamagishi, T., Horita, Y., Mifune, N., Hashimoto, H., Li, Y., Shinada, M., et al. (2012). Rejection of unfair offers in the ultimatum game is no evidence of strong reciprocity. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20364–20368. doi: 10.1073/pnas.1212126109
- Zahavi, A. (1995). Altruism as a handicap: the limitations of kin selection and reciprocity. *J. Avian Biol.* 26, 1–3. doi: 10.2307/3677205
- Zebrowitz, L. A., and Franklin, R. G. Jr. (2014). The attractiveness halo effect and the babyface stereotype in older and younger adults: similarities, own-age accentuation, and older adult positivity effects. *Exp. Aging Res.* 40, 375–393. doi: 10.1080/0361073X.2014.897151

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Novakova, Machová, Sýkorová, Zíka and Flegr. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Generosity as a status signal: Higher-testosterone men exhibit greater altruism in the dictator game

Julie Novakova^{a,*}, Petr Tureček^{a,b}, Kamila Machová^a, Kateřina Sýkorová^a, Vojtěch Zíka^{a,c}, Jaroslav Flegl^a

^a Laboratory of Evolutionary Biology, Department of Philosophy and History of Science, Faculty of Science, Charles University, Prague, Czech Republic

^b Center for Theoretical Study, Prague, Czech Republic

^c Center for Behavioral Experiments (CEBEX), Prague, Czech Republic

ARTICLE INFO

Keywords:

Androgen
Dictator game
Dominance
Dual hormone hypothesis
Evolution of altruism
Evolutionary psychology
Social decision making

ABSTRACT

Altruistic behavior can be modulated by many factors including hormonal levels, but their reported effects remain mixed. Understanding the proximate mechanisms of altruism such as these can help test predictions of ultimate, evolutionary explanations. We investigated the relationship of the endogenous salivary levels of testosterone and cortisol with Dictator Game (DG) offers as a proxy of altruism on a sample of general-population participants ($N = 158$, 84 F, 74 M). Bayesian data analysis and model comparison showed both testosterone and cortisol were negatively correlated with DG offers in women, while higher testosterone levels were associated with greater generosity in men. These results suggest that high testosterone may promote altruistic behavior in the service of status-seeking among men.

1. Introduction

1.1. Background

Altruism, meaning apparently selfless helping behavior toward others, has long constituted an evolutionary enigma. Hypotheses such as kin selection (Hamilton, 1964a, 1964b), reciprocity (Trivers, 1971) or costly signaling, dubbed ‘handicap hypothesis’ (Zahavi, 1995), have been proposed, and it seems that each could apply under some circumstances (West et al., 2011).

Proximate, mechanistic causes were also investigated, both of their own accord and tied to the ultimate evolutionary causes. Among the proximate mechanisms of altruistic behavior, hormonal levels had often been the focus in both humans and animals. But despite the growing attention toward this piece of the altruism puzzle, the specific hormonal influences and their connections to the ultimate causes remain largely unresolved.

Here we focus on the influence of testosterone (T) and cortisol (C) levels on altruism, especially in terms of the dual hormone hypothesis suggesting that basal T and C jointly influence behavioral traits associated with aggression and dominance (Mehta & Josefs, 2010). Conclusions insofar have been mixed, as summarized below. The aim of this

study is to shed more light on this complex question and help integrate the proximate hormonal influences with ultimate evolutionary explanations, such as dominance displays in order to attract mates and cement one’s social status.

1.2. Testosterone as an altruism modulator

A number of studies looked into the effect of sex hormones on pro-social behavior, with ambiguous results. Various experimental games are typically used as altruism proxies, among them the Dictator Game (DG; Forsythe et al., 1994), where one player is allocated a sum of money and may divide it between self and another player, who is passive in this case. People on average allocate around 28%, with women being usually more generous than men (Engel, 2011). In the earlier-devised Ultimatum Game (UG; Güth et al., 1982), the second player has agency and may accept or reject the first player’s offer. In the latter case, neither gets anything. In the Public Goods Game (Marwell & Ames, 1981), players can contribute to a common pot where the contributions are multiplied and divided equally (not based on the initial contribution) between players. Out of these games, the DG is the closest proxy for “pure” altruism unaffected by other social considerations.

High-T men were found to offer non-significantly more in the UG

* Corresponding author at: Viničná 7, 128 00 Prague, Czech Republic.

E-mail address: julie.novakova@natur.cuni.cz (J. Novakova).

<https://doi.org/10.1016/j.evolhumbehav.2024.106615>

Received 23 November 2023; Received in revised form 2 August 2024; Accepted 6 August 2024

Available online 23 August 2024

1090-5138/© 2024 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

than low-T men by [Burnham \(2007\)](#), and women with artificially raised T offered significantly more in the UG than those who'd received placebo (but at the same time, women who *believed* that they'd received testosterone offered less than those who believed they'd ingested placebo). [Zethraeus et al. \(2009\)](#) administered estrogen, testosterone or placebo to postmenopausal women, and found no effect of sex hormones on altruism, fairness and trust in several economic experiments, while [Zak et al. \(2009\)](#) found that men with artificially raised testosterone levels behaved *less* generously in the UG. Low-offer rejections in the UG, often interpreted in terms of fairness preferences but also as aggression, appeared to be increased by T ([Burnham, 2007](#); [Mehta & Beer, 2010](#)), decreased by it ([Kopsida et al., 2016](#)), as well as independent of it ([Cueva et al., 2017](#)).

Such mixed evidence would suggest that the effects of T may be context-dependent and influenced by other variables. This conclusion is supported by [Inoue et al. \(2017\)](#), who found that high-T men with a higher social rank (as maintained in a university rugby team hierarchy) behaved more dominantly – offered less and demanded more – in the UG, while high-T men of a lower rank behaved submissively. The authors concluded that high T likely promotes strategic, rank-dependent social behavior rather than straightforward dominance. A follow-up study ([Inoue et al., 2023](#)) found that high-T men whose status has risen recently do in fact behave more dominantly. In addition, high T was also found to correlate with greater parochialism in experimental games ([Diekhof et al., 2014](#); [Reimers et al., 2019](#); [Reimers & Diekhof, 2015](#)).

Proxies of earlier androgen exposure can also be measured to elucidate the effects of T on behavior. The second to fourth digit (2D:4D) ratio is sexually dimorphic and thought to be dependent on prenatal T exposure (e.g. [Manning et al., 1998](#), [Hönekopp et al., 2007](#); experimentally in rats [Talarovičová et al., 2009](#)), hence potentially serving as a proxy of androgen levels in intrauterine development. Lower digit ratios are thought to indicate higher intrauterine androgen exposure. However, the ratio may simply reflect shifts along the allometric line – larger people have a lower 2D:4D ratio ([Kratovich & Flegr, 2009](#)). Or, as [Knickmeyer et al. \(2011\)](#) suggested, it might at least in early childhood indicate neonatal T levels rather than prenatal androgen exposure.

Again, some studies found that participants with lower digit ratios (potentially higher T exposure) were more generous, such as [Millet and Dewitte \(2009\)](#) in the DG, but exposure to aggression cues inverted the relationship, and partly [Ronay and Galinsky \(2011\)](#) in a rare face-to-face UG setting. Other studies, in contrast, found people with lower digit ratios more selfish, such as [Brañas-Garza and Kovářík \(2013\)](#) or [Buser \(2012\)](#) across several experimental games. However, the latter study didn't measure the ratio and only asked the participants whether their index or ring finger on each hand is longer or of equal length, which elicited criticism. Others found no effect ([Brañas-Garza et al., 2019](#); [Neyse et al., 2020](#); [Parslow et al., 2019](#)). The potential interplay of digit ratio and actual T levels is also unclear. [Van Honk et al. \(2012\)](#) investigated the influence of testosterone on behavior in the Public Goods Game, and found no main effect in either men or women, but there was a significant effect of interaction of testosterone with the 2D:4D ratio: Participants with low prenatal T exposure behaved more prosocially after T administration, while those with high prenatal exposure exhibited no significant change. The potential effect may not persist across ethnic groups: In [Galizzi and Nieboer \(2015\)](#), an inverted U-shaped relationship between DG giving and digit ratio was observed in European-origin participants, but not across the whole participant pool or in other single ethnic groups. To add to the complexity of the problem, [Senci et al. \(2020\)](#) observed a U-shaped relationship of digit ratio and DG giving in a control sample not primed by social norms, and no 2D:4D ratio effect in their experimental sample primed by prescriptive norms on generosity, concluding that biological foundation of altruistic behavior exists, but can be overridden by social norms.

The mixed conclusions and vastly varying experimental designs using only male or female participants, or both sexes, natural salivary or

blood T, T administration or indirect cues of T exposure, different experimental games, situations or questions to assess participants' altruism, and often specific primings or cues make it nigh-impossible to reliably generalize the overall effects of T on altruistic behavior. [Stanton \(2017\)](#) reviewed the effects of T on decision-making in various contexts (economic, social, ethical, consumer behavior). The reported effects of T were mixed (mostly either promoting competition and risk-taking, or null). With caution, it can be concluded that the effects of T depend on context (such as the respective social rank, ingroup/outgroup status and sexual or aggression cues, of which little has been investigated thoroughly up to date).

1.3. Cortisol: stress response and altruistic decision-making

As a 'stress hormone', cortisol (C) can indicate impaired mental and/or physical health, and has been associated with impaired memory function, though the duration of the effect, influence of timing and relationship with acute/basal cortisol levels are not completely clear ([Bermejo et al., 2022](#); [Brunner et al., 2006](#); [Het et al., 2005](#); [Staufenbiel et al., 2013](#)). Its overall relationship with cooperative and other behavior appears across a range of vertebrate species appears to be complex, in some cases inverse U-shaped ([Raulo & Dantzer, 2018](#)).

Human studies also don't show a clear linear pattern. [Starcke et al. \(2011\)](#) compared control and stress-exposed participants' responses to hypothetical realistic moral dilemmas with an altruistic and egotistical reaction option, and while there was no significant difference between both groups, participants whose C was more elevated after the stress task exhibited more egotistical behavior.

[Vinkers et al. \(2013\)](#) studied the time-dependent effect of stress – immediately and with a 75-min latency – on decision-making in the DG and UG, and found the stressed participants to be less generous (time-independently) and more prone to altruistic punishment immediately after stress exposure, but less prone to it than controls after the latency period. Measured salivary C levels did not correlate with altruism in the games at any point. A study by [Schulreich et al. \(2022\)](#) corroborated these results, showing a negative effect of cortisol increase on charitable donations in the control group as well as the experimental group who were subjected to the Trier Social Stress Test.

Contrary to those findings, [Singer et al. \(2017\)](#) found that young men exposed to acute stress via the Trier Social Stress Test behaved more altruistically in experimental situations. Their C levels correlated positively with altruism. [Sparrow et al. \(2019\)](#) also found young adults with higher stress reactivity to be more generous, in this case in the DG. Interestingly, older adults showed a (non-significant) negative association of cortisol and generosity (while in general, older adults were still more altruistic than young adults, as shown in previous studies ([Bailey et al., 2013](#), [Matsumoto et al., 2016](#), [Nakavachara, 2018](#), [Ogawa et al., 2018](#), [Pornpattananangkul et al., 2019](#)). Finally, [Zhang et al. \(2019\)](#) found no effect of stress task-induced salivary cortisol on DG or UG behavior in women, but found men to behave more altruistically in the DG with more elevated C.

[Schulreich et al. \(2022\)](#) speculated that the apparent discrepancy between studies finding an increase or decrease of altruism, respectively, in relation to cortisol levels or their changes might come down to competing effects of acute stress: impairing mentalizing (comprehension of one's and others' intentions and behavior), while increasing empathy. The salience of various factors (such as the perceived need of the donation recipient) might be decisive in which of these opposite effects becomes more pronounced.

All these experiments, though, concerned cortisol levels after exposure to some form of acute stress, while basal C levels could be more relevant in the long term. These were studied by [Pfattheicher and Keller \(2014\)](#), who found that basal C correlated negatively with costly punishment of free-riders in the Public Goods Game (PGG). Costly punishment is generally associated with altruistic concerns. However, PGG contributions did not correlate with C levels. The study had male

participants only.

Self-reported chronic stress, which would be associated with increased basal C, was found to have no effect of donations in an anonymous Dictator Game (Ceccato, Kettner, Kudielka, Schwieren, & Voss, 2018). The role of basal cortisol – more interesting from the perspective of Triversian reciprocity, as chronic stress negatively affects memory and expected lifespan – in altruism has received little attention as opposed to cortisol changes under acute stress, and it remains to be investigated in more detail. The present study aims to be a part of that.

We are not aware of any C administration studies using the framework of experimental games up to this point, making clearly distinguishing its effects less simple. However, since high C, especially in the long term, is linked to stress and potentially impaired health and chances of long-term survival, we should expect it to correlate negatively with altruism, viewed through the lens of the reciprocal altruism theory that requires high chances of future interactions and good memory of previous ones.

1.4. Testosterone-cortisol relationship and the dual hormone hypothesis

The dual hormone hypothesis (DHH), posited by Mehta and Josephs (2010), suggests that basal T and C jointly influence behavioral traits associated with aggression and dominance, making high-T, low-C individuals more dominant and status-seeking, while high-T, high-C individuals would exhibit more status loss avoidance. Montoya et al. (2012) reviewed the literature on the role of T and C in reactive (impulsive) and proactive (instrumental) aggression, including social dilemmas and games, and found the findings from experimental games-utilizing studies with endogenous T and C alone or exogenous T suggestive of the above-described interplay of T and C. T seems to promote aggression and social status-seeking, while C is linked to social withdrawal, and its high levels can nullify or override the effects of testosterone.

Denson et al. (2013) investigated the effect of T*C in women on reactive (retaliative, impulsive) aggression, and observed women with both high T and C reacted more aggressively when insulted, while low levels of both hormones predicted more submissive behavior. More in line with the DHH predictions, Pfattheicher et al. (2014) found that high-T, low-C corresponded to destructive (antisocial), but not altruistic punishment in the PGG. In other words, high-T men with low C levels significantly more punished other players who'd contributed just as much as they or more in the PGG, but they did not punish free-riders any more. Neither of the hormones correlated with costly punishment. The study had male participants only. Behavior that could be interpreted as antisocial was similarly found in high-T, low-C individuals with chronic dominance in the Pfattheicher (2017) study, where male participants played a "reversed DG", where they could increase their earnings by decreasing another (dummy) player's earnings.

Zilioli et al. (2015) studied the effects of T and C on empathy, and found that high-T individuals with low basal C had lower empathy scores (in accordance with the folk perception of T as an 'anti-empathy' hormone), while the effect was reversed in individuals with high C levels. Sex differences were observed by Sýkora (2018), who found support for the dual hormone hypothesis in risk-taking in women, but not men, and not in competitiveness in either sex.

Prasad et al. (2019) found support for DHH when investigating the interplay of testosterone and acute stress cortisol response on DG offers; high-T individuals with low C response exhibited most dominant behavior, keeping most of the endowment for themselves. Men and women in the study showed similar patterns. In an earlier study using the UG and the interplay of acute stress-induced cortisol change with testosterone, though, sex differences were observed in the relationship of acute stress and "retaliation" (rejecting an unfair UG offer): women in the high-stress condition were less likely to retaliate than in the low-stress condition, while men were more likely to do so (Prasad et al., 2017).

Both basal and acute cortisol levels may therefore play a key role in modulating the dominance-seeking effects of testosterone. Moreover, the dual hormone hypothesis might explain the inconsistency in reported effects of T and C in studies investigating only one of these hormones. The DHH-framework studies outlined above also show the importance of differentiating between various behaviors of a single category (impulsive/reactive v. instrumental/proactive aggression in Denson et al. (2013) and Pfattheicher et al. (2014)) and highlight the role of sex.

A review and series of studies by Grebe et al. (2019) suggest that the majority of published studies are underpowered to detect a T-C relationship with behavioral traits. Their seven independent studies showed no T*C effects, and only a weak effect of T on status-seeking. They point out, however, that social environment context should be investigated with regard to the dual hormone hypothesis (where single-sex or mixed-sex settings could play a role, and opponent status cues may influence the realized behavior (Knight et al., 2022)). Dekkers et al. (2019) performed a meta-analysis of 30 papers investigating the DHH, yielding tentative support to the hypothesis (at least where social status is concerned). The effects seemed to be stronger in men. However, further studies are needed to elucidate the subject, which we're hoping to contribute to.

In the present study, we investigated the relationship of basal T and C (measured from salivary samples) with altruism in the DG. The majority of previous studies suggest that high-T, low-C individuals exhibit more dominant and status-seeking behavior, but the T*C connection with altruism remains far less clear, therefore we aimed to fill this gap.

2. Methods

2.1. Participants and experimental procedure

The experiments were conducted at the Charles University in Prague in June and July 2018 (for male participants) and 2019 (for female participants). Participants were invited through Facebook. A maximum of 18 enlisted participants could attend one session, and in total, 158 people participated (all central-European, 74 men, 84 women, mean age = 29, median of age = 28, SD = 9.3). The sessions took place in late afternoon only, always with same-sex heterosexual participants and always at the same time and the same season, in order to minimize the noise of diurnal and seasonal fluctuation of hormonal levels.

On site, the participants were greeted at the reception and led to the computer lab (each to their own station divided by partitions), where they read and signed their informed consent and experiment guidelines. Their saliva was collected into salivettes by passive salivation (rolling a swab in their mouth for three minutes). They were aware that C and T levels will be measured in the samples. After the salivettes were recovered by the experimenter, they could begin the online questionnaire including the Dictator Game and general demographic questions (sex, age, education, socioeconomic standing). Other variables (self-reported number of sexual and romantic partners, short personality survey, health, temporal discounting and risk-taking, Ultimatum Game behavior) were measured for a related study (Novakova et al., 2021); optional participant photographs (to be used in a later study) were also taken for a future study after the questionnaire was completed.

The DG was played for 400 tokens translatable into 40 CZK (roughly 2 USD). The participants were first presented with the rules and truthfully told that they would be randomly matched with another person in the room and that their earnings at the end of the experiment would reflect the game outcome. Before playing, the participants answered a control question to assess whether they understood the rules of the game correctly, and a question about having any previous knowledge of either DG or UG (from lectures, literature etc.). Finally, the participants were individually informed about the goal of the experiment and paid their reward gained in the experiment, with the opportunity to gamble in a die roll for a greater reward.

2.2. Saliva treatment and analysis

All saliva samples were stored in salivettes in a freezer with -20°C until radioimmunoassayed for T and C levels at the Institute of Endocrinology in Prague. For the full methodology see Flegel et al. (2008). In case of insufficient sample volume, only T assays were done, since T effects were expected to play a more substantial role in altruism, based on literature review. Insufficient material for analysis of either hormone was found in four women and ten men. Only T was analyzed two women and eleven men, and only C (where the amount of saliva for T analysis was found insufficient at a later stage) in two women (no men).

2.3. Statistical analysis

We have conducted a Bayesian data analysis and model comparison to assess the relationship between hormonal levels and altruism.

Sex (as an index variable with two levels), Testosterone, Cortisol, and the interaction between Testosterone and Cortisol (as a T/C ratio) were considered as potential predictors of the amount given by the participants in the DG. Logarithms of hormonal levels and the logarithm of T/C ratio were used to avoid the measurement's arbitrariness and other issues connected to raw T/C ratios raised by Del Giudice & Gangestad (2022). The logarithmic scale emphasizes the importance of the relative increase (i.e. 10% increase in the T level) over the absolute increase (i.e. 1 particle/mol). Base 1.1 was used for all logarithms to create an intuitive predictor scale with 1 unit representing a 10% increase. Logarithms were standardised before the analysis to allow for easy assignment of priors to model parameters. Still, despite being made on logarithmic scale, some predictions are visualised on the original scale to allow for a straightforward interpretation. In addition, because some studies that aim to evaluate the dual hormone hypothesis work with a more conventional interaction term, the T*C product, we include results of alternative analysis that works with the product instead of the ratio in Supplement S11.

To avoid the omission of non-trivial relationships, we fitted a broad set of models on the data. Each continuous predictor could enter the analysis in 6 levels of increasing complexity: 0: No association between the predictor and the altruism, 1: Linear relationship, 2–5: Spline model of a given order (order 2 utilised linear, order 3 quadratic, 4–5 cubic B-splines with nodes aligned with appropriate quantiles; see Supplement S1 for spline visualisation and further introduction of the method). Splines were constructed using the `bs()` function from the “splines” package (RCoreTeam, 2019). We used splines instead of more common polynomials, because spline-defined functions cover more versatile set of curves in the acceptable minimum-maximum boundaries for the same number of parameters as compared to polynomials, and because all parameters corresponding to spline weights are advised to have the same prior distribution, unlike slopes multiplying higher order polynomials (see also McElreath, 2020 for further discussion on Bayesian spline models).

A similar approach was recently brought forward as specification curve analyses by Simonsohn et al. (2020) who, however, work in the frequentist framework and explicit hypothesis testing. Our approach combines multi-model inference common in the Bayesian framework with model comparison and spline curve specification.

Sex could have a null relationship with the altruism (i.e. there was a common intercept parameter for men and women), or influence the altruism baseline (i.e. there was a separate intercept by sex), or interact with continuous variables (i.e. each slope or spline weight had a separate value for men and women). Thus 647 models were created ($6 \times 6 \times 6 \times 3-1$; because the by-sex intercept and slope models without hormonal predictors represent the same structure). Models' out-of-sample predictive accuracy was compared using WAIC (Widely Applicable Information Criterion) and the resulting weights.

$$w_i = e^{-0.5\Delta_i} / \sum_j e^{-0.5\Delta_j}$$

where Δ_i represents the difference between *i*th model's WAIC and the lowest WAIC in the set (McElreath, 2020).

It is well known that the dependent variable (donation in the DG) has a challenging distribution. There is an overrepresentation of people donating exactly half of the entrusted money (200 tokens = 20 CZK), and the distribution is constrained by the budget (0–400 here). We have selected a mixture distribution reflecting this fact, a middle-inflated beta-binomial distribution. There are two multiple logistic regressions with the same predictors and model terms in each model. Parameter values in these regressions are, however, independent. First regression, linked to a probability term p in a simple binomial distribution, decides the “middle inflation”; whether or not the participant donates the exact half of the money regardless of other regression terms. The second regression uses the standard beta-binomial distribution with $N = 400$, $\alpha = \mu \times \theta$, $\beta = (1 - \mu) \times \theta$ as a likelihood function. (The expected proportion of the budget donated (μ) and the overdispersion (θ) around this mean. These parameters are sufficient to model discrete values between 0 and 400.)

The regression terms are linked to the log-odds of p and μ . Parameter θ is assumed to be constant and independent of sex.

We used weakly regularised unbiased priors in all models. Prior distributions of regression parameters, both intercepts and slopes, were characterised by normal distributions $N(\text{mean} = 0, \text{sd} = 2)$. The posterior distributions were extracted using the Hamiltonian Monte Carlo sampler Stan (StanDevelopmentTeam, 2021) called by the rethinking stan() infrastructure (McElreath, 2020). All compared Stan models were built by a bespoke compiler. All data and computer code used in the analysis are available at <https://osf.io/pwztlm/>.

3. Results

3.1. Descriptive statistics

3.1.1. Dictator game

All 158 responses were valid: the participants finished the experiment and correctly answered the control question for assessing whether they understood the game. The mean amount allocated to another player in the DG was 154.4 out of 400 tokens (~38.6%), median 200 (50%), sd 76.2. Women gave on average 165 tokens, men 142.4 tokens. Previous knowledge of the DG or UG had no statistically significant effect on participants' behavior.

3.1.2. Testosterone

For 16 of the 158 participants, testosterone levels could not be reliably measured due to the small volume of extracted saliva. In the 142 remaining ones, the mean testosterone level was 0.566 nmol/l for men (sd = 0.341), and 0.667 nmol/l for women (sd = 0.222). Two men had very low testosterone levels (together with one woman they were marked as potential outliers, i.e. influential points, on the logarithmic scale), therefore we decided to re-run the analysis without the three data points. The results did not change (see Supplement S10). The two men were, however, not enough to drag the whole male mean down; the whole distribution of testosterone values in men resembles the distribution of T values in women, just shifted by approx. 0.1 (see raw data distributions in Supplement S9). We discuss the observed difference further in Discussion.

3.1.3. Cortisol

Cortisol levels were available for fewer participants ($N = 131$), due to the insufficient sample volume for both T and C assays in some cases. The mean C level was 2.250 nmol/l for men (sd = 1.550), and 2.900 nmol/l for women (sd = 1.250).

3.1.4. Testosterone-cortisol ratio

For participants with both T and C levels measured, T/C ratios were computed. The mean ratio level was 0.33420 for men (sd = 0.383), and

0.28625 for women (sd = 0.231). (Also see the comparison of raw hormonal levels, interaction terms, and their logarithms in Supplement S9.)

3.2. Models of T/C association with altruism

Biological sex and hormonal levels together influenced the expected donation in the DG. Models that contained separate intercepts and slopes for each sex offered a much better out-of-sample predictive accuracy (total weight = 0.79) than models with common slopes (total weight = 0.21) and models independent of sex (total weight = 0.00).

Out of the 100 best models, 77 were models with the interaction between sex and hormonal levels, 23 had at least by-sex intercepts. 100 worst models were exclusively models without the effect of sex. This can be interpreted as participant sex playing a key role in predicting altruistic behavior.

The best model (slo_051) combined the complicated relationship between log(C) and DG donation (fifth-order spline structure) with a straightforward linear effect of log(T/C) ratio (Fig. 1, see Supplement S3 for the posterior distribution of model parameters). Although non-trivial, C-related changes in the mean donation seem to be represented faithfully by this curve. In the first 39 models, the relationship between

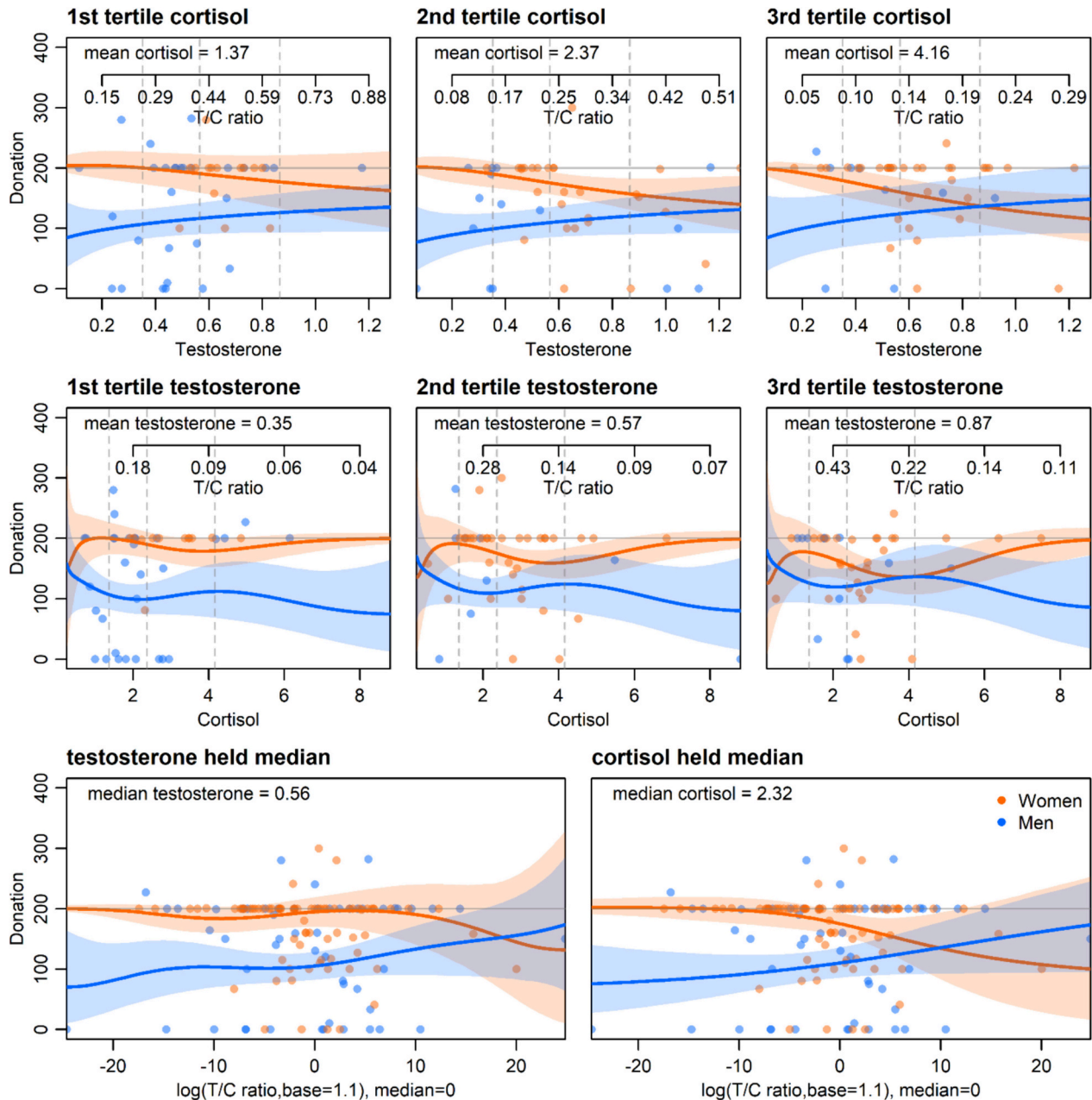


Fig. 1. Predictions of the best model. The dashed vertical lines in the two upper rows indicate positions of the respective tertile means in complementary panels. The raw data in these rows are limited to tertiles in panel titles. In the predictions from the perspective of T/C ratio, all raw data points are included (women shown in orange, men in blue). The counterfactual value in the top left corner was held constant to produce predictions of change in donation with the x-axis in each panel (e.g. if the T was held median, only C values varied to produce desired T/C ratios). The additional T/C ratio axes by tertiles in top two rows are added for easier interpretation of the interaction term, which changes with both T and C. They allow to read all panels as having T/C ratio on the x-axis (the mean value for each respective tertile is used to calculate corresponding T/C values for each x-axis tick). The solid lines represent mean predictions, and the semi-transparent corridors span 89% compatibility intervals based on 10,000 samples from the posterior distribution. To see the inverse visualisation with hormonal levels logarithms and the ratio on the original scale, see Fig. S4. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

C and donation is either the above-mentioned fifth-order spline (25 models), or none (14 models). The best simpler model can be found on position 40 (slo_220), and even the best model using just cortisol to predict altruism suggests the more complicated relationship (slo_050, 29th best model). Although the null C models fall second (total weight 0.28), it is improbable that the relationship is non-existent and we identified a false pattern. It just means that no simple curve can faithfully represent the relationship over the whole recorded cortisol range. The null model minimises the risk of overfitting, but at the expense of overall low accuracy. In overall, the models with by-sex slopes and fifth-order B-spline structure for log(C) predictor have an accumulated weight of 0.47. Given the predictive accuracy and the comparison to other models, the best model can be interpreted as realistically capturing the complex hormonal influences.

The cortisol-altruism relationship can be described as U-shaped, with U-reversal at the very beginning of the detected C range (i.e. incomplete M-shaped in women and incomplete W-shaped in men; see Fig. 1, and also Figs. S17, and S18, with the results of models that fall second and third in the comparison). Both men and women with very low C levels are predicted to donate similar amounts. Then, the donations rise sharply in women and decrease in men to decrease, respectively rise again to a similar amount around the mean of the third tertile. The lines reach plateaus of high donations in women and low donations in men for extremely high C values. Due to the limited sample size, however, the compatibility intervals around the extreme ends are very wide, and one shall be cautious in their interpretation. The concave profile of the curve characterising women donations up to the third tertile and the convex profile of donations by men do seem well-grounded.

With increasing T (either relative to C or absolute, depending on the model, because, to some extent, these variables are interchangeable), men's donations increase, while women's donations decrease (Fig. 2). Parcellation of the model into two logistic regressions brings little new information. If the donation is expected to rise according to the beta-binomial model's central tendency μ , so does the probability of donating the exact half of the budget p (see Supplementary Figs. S9–S18). That is because the fair share represents a very generous offer. This conclusion was not hard-wired in our model; the parameters of the regressions were independent. The fact that the general patterns agree is a good sanity check. It suggests that despite a complicated mixture

distribution, decisions whether or not to split the budget equally and how much to donate otherwise are not governed by fundamentally different underlying processes in participants' minds.

T performs better as an isolated predictor than C (models with only T have a cumulative weight of 0.08, C-only models just 0.03, models using only T/C ratio have a cumulative weight of 0.07). The best model using a single hormonal predictor employs two T splines (model slo_200, overall 5th best model; see Supplementary Fig. S19 for the visualisation: the function is growing, respectively declining over the whole T range, the first spline only moderates the increase, respectively decrease of the donation).

If testosterone is given enough space to interact with cortisol in men, an interesting pattern emerges. The positive effect of T on DG offers is most pronounced if C is sufficiently low. For higher C values in men, the relationship between T and generosity gets slightly flattened or, possibly even reversed (see, for instance, the testosterone-perspective plots in the alternative models with log(T*C) product in Supplementary Fig. S32. The best log(T*C) model WAIC = 934.53 was even slightly better than the best log(T/C) model WAIC = 936.48). Caution, however, is advisable when interpreting this interaction, because of the relatively low number of men with high C and T values (notice wide 89% CIs of posterior mean – shaded rectangles – in high-C men in Fig. 2).

The overall best model with T/C ratio suggests that the interaction may be non-trivially present in the effect of the choice to split the budget in half in women (see parameter estimates in Supplementary Fig. S6). Women with higher log(T/C) are expected to show lower probability of splitting the budget in half, while the continuous effect on the donation otherwise does not show conclusive difference from 0. Further studies should aim to replicate our effort to faithfully reproduce the mixture distribution present in DG offers to support or revise such subtleties in how hormonal levels are translated to realized signs of generosity.

Outside of extreme values, the set of models agree well that between means of the first and the third tertile of T levels (see the dashed vertical lines in Fig. 1 that indicate by-tertile means), the expected donation declines with increasing T in women and rises with increasing T in men. The effect of C between the respective tertiles is similar in women; they donate conclusively less with higher C, unlike men.

The aggregated donation predictions for tertile-mean combinations based on the set of 10 best models (Fig. 2) is in accord with this finding (see Fig. S5 for their overview). As T increases, men donate more and women less. Women's donations tend to decrease with increasing C. In contrast, men's donations do not respond strongly to this hormone's changes, perhaps except for the moderating effect of C on the effect of T (which is most pronounced when C is low as described above). These effects are preserved even if we use the 100 best models instead of 10 (see Supplementary Fig. S20, see Fig. S21 if you prefer plot alignment with alternating T levels).

These reported effects of hormones are robust, even though the model omitting hormones' effects scored relatively high (see Supplementary Fig. S4, look for model int_000, 10th best model overall, the relative weights of the best model slo_051 and this one are 0.68 and 0.32).

If we disrupt the relationship between hormones and donation by randomly shuffling the vectors holding T and C values, we arrive at a completely different picture (see Supplement S6). Models using separate intercepts and slopes for each sex no longer cluster at the beginning of the models' weight comparison; they become interchangeable with models predicting donations using common slopes. Furthermore, models with a low number of splines are much favoured over the complex models, which is not the case in the actual-data comparison in Fig. S4 (this difference between the empirical and the scrambled dataset is too large to be attributable to the sampling variation, see Supplement S7).

The difference in donation between men and women, regardless of the hormonal levels, is well established. The null model (no_000) is at the end of the overall comparison. The weights in a simulation, where

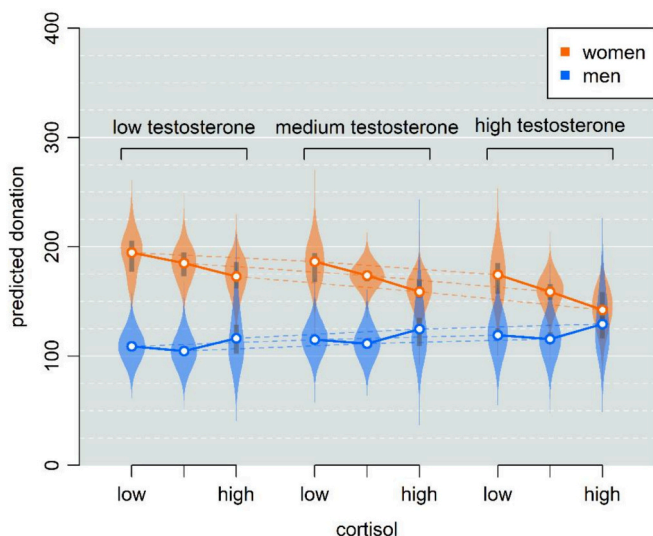


Fig. 2. Aggregated prediction based on posterior distributions of the 10 best models. The density plots depict the complete distribution of mean donation prediction, including sample variation within each model. The shaded rectangles span 89% of the posterior distribution means. White points indicate grand means of the posterior distribution means. Low, medium, and high levels correspond to mean values per 1st, 2nd, and 3rd tertile of data distributions.

hormonal levels and sex labels are randomly permuted (Supplement S6), suggest the reversal of the predicted out-of-sample accuracy than reported in Fig. S4 (see Fig. S23).

Still, there is a lot of unexplained variation in altruism even after sex and hormonal balance are accounted for. The θ characterising the overdispersion of donation around the predicted budget allocation is 1.93 [89% CI 1.42–2.52] according to the best model, which suggests relatively evenly spread beta distribution (higher θ correspond to the tighter distribution of probability density). The WAIC of the best model is 936.49, and the WAIC of the worst model is 957.05, while the estimated standard error of the best model's WAIC is 66.50. That signals that even the best model does not perform much better than the worst model in the set. It is the huge overdispersion that causes it (see Supplement S8), indicating there is still a lot of variation to explain. Put simply, the best model likely captures the real complexity of T and C influences on DG donations, but explains only a small variance in the DG donation distribution. Future studies should also investigate the mutual relationships of other physiological or neurological correlates. Nevertheless, our results suggest that the interaction between hormonal levels and sex is an integral part of any model aiming to predict donation in the DG precisely.

4. Discussion

Consistently with previous studies (see the Engel, 2011 meta-study), women in our sample were more generous in the DG than men. We observed that DG allocations consistently declined with increasing T in women and rose, albeit less steeply, with higher T in men. The relationship between altruism and C was more complicated, roughly N-shaped in women and H-shaped in men, again across all tertiles of T levels; but it can be approximated that women contributed less with rising C levels, while men exhibited a weaker opposite pattern (see Fig. 2) – this is what happens between the mean value of the first and third tertile.

According to the dual hormone hypothesis, behavioral patterns commonly associated with high T tend to be more pronounced in individuals with low basal C. The results we observed do not neatly fit this interpretation due to the nontrivial shape of the relationship of C with altruism; high-T and low-C, but to some extent also high-C men as opposed to median-C men behaved more generously in the DG.

While T, especially in men, is often thought to promote aggression and selfishness (more or less suggested also by the studies of Zak et al., 2009, Diekhof et al., 2014, Inoue et al., 2017, Reimers et al., 2019), the actual reality appears more complicated. In fact, high-T men in our sample donated more than low-T men (especially if they had low C levels), more in line with Burnham (2007). The dual hormone hypothesis suggests that high T (combined with low C) should be linked to social status-seeking and aggression, and high C to social withdrawal (modulating the effect of T on status-seeking).

In our study, greater altruism in high-T men could be interpreted in terms of status-seeking and status display, which corresponds to Durkee et al. (2020) findings that benefit-generation predicts status allocation cross-culturally, whilst cost-infliction only predicts status weakly or not at all. Interestingly, this interpretation is at odds with Inoue et al. (2017, 2023), who found high-T men to offer less and demand more in the UG, and Pfattheicher (2017), where high-T, low-C dominant men took more in a reversed DG. While the UG and reverse-DG can be compared to the DG only with caution, these findings and their interpretation invite the question: Is altruism a signal of dominance, or submission?

We suspect that the answer would be very much context-dependent, and that the question itself, asked without adding “under conditions...”, may be much too simplified to elicit any truly useful answers. It would be useful to view both the hormonal influences on altruism and its ultimate benefits more in the view of life histories (and thus varying costs and benefits), both within- and across studies (Roney, 2016).

If we interpret our results in light of broadly generalized mate choice

and seeking strategies between the sexes – where men are more likely to showcase the abundance of resources and women to look for these signals (e.g. Todd et al., 2007; Waynforth & Dunbar, 1995; Wiederman, 1993), it makes sense that the proximate mechanisms of hormonal influences on altruistic behavior should differ by sex, shaped by different ultimate causes, here status display (relevant for mate attraction and intra-sexual competition) for men.

A single study, of course, cannot bring any strong conclusions about the evolutionary origins of human altruism. It can, however, represent one important piece of the whole puzzle of factors contributing to the evolution of altruistic behavior. The data-driven approach that we employed revealed an interesting pattern that invites both further theoretical development and empirical replication.

There were several potential limitations to the study. In our sample, puzzlingly, women had higher mean T levels than men. A plausible explanation is that salivary T correlates closely with serum T in men, but not necessarily in women (Shirtcliff et al., 2002), making this a possible limitation. Higher salivary T for women may reflect relatively lower serum testosterone, since the linear regression equations for serum and salivary T differ (Salimetrics, 2019). Fiers et al. (2014) raise concerns about salivary T comparability to serum T levels and its difference between the sexes, where women had relatively raised salivary T compared to serum T – a fact which may explain the difference observed in our study.

In general, observed salivary T-behavior connections in men should be treated as reliable, in women with some caution. Although salivary testing of T in women was considered reliable based on multiple results (e.g. Baxendale et al., 1982; Dabbs Jr, 1990), later studies showed that it could reflect serum levels poorly (Fiers et al., 2014; Flyckt et al., 2009), indeed to such an extent that some authors warned before using it to estimate women's T levels (Davison, 2009). High circadian variability with episodic fluctuations exceeding the typical morning maxima was detected by Al-Dujaili and Sharp (2012). Age may have played a role as well, since T levels drop with age in both sexes to a similar relative degree, but the absolute decrease is much greater in men (Keevil et al., 2017). Inter-laboratory differences may also exist, although results from multiple labs tended to be fairly in agreement in Dabbs Jr et al. (1995). Lastly, the method of collection may have influenced the measured T levels, especially in women; Prasad et al. (2017) found that when using salivettes, there “was inflation of testosterone values at the lower end of the distribution (ostensibly in female participants)”. While the salivette cotton swab technique and passive drool collection exhibited close correlation between measured hormonal levels in their study, it cannot be excluded that the collection technique may have inflated the measured T levels in women in our study. Moreover, while cotton swabs and passive drool lead to very similar mean measured cortisol levels, some studies indicate that the within-participant correlation may be low (Kozaki et al., 2009), as some molecules contained in cotton may have cross-linking with assay antibodies, and the cotton might absorb some of the cortisol. However, as Kozaki et al. (2009) conclude, all hormonal collection methods may introduce their own noise and potentially bias.

Salivary testosterone corresponds to bioavailable testosterone more than serum T levels do, which is why we chose this collection method, but it is very dependent on the experimental conditions, as outlined above. It's possible that women adhered more to the pre-experimental instructions to drink more water before coming to the experimental session and to salivate on the swab in their mouth for the whole three minutes' duration of saliva collection. That could also explain why women's testosterone results measured on average higher. In women, there were also fewer samples where the swab contained too little material for both T and C assays, which would support this potential explanation.

We must stress that this result, while puzzling, doesn't constitute a major problem for the aims of the current study, since the Bayesian analyses were done separately for men and women. One important take-away for future studies is that separate sessions for different participant

groups should not be held, because the effect of the session (despite the carefully maintained identical conditions of the session room, time of day, time of year, instructions, researchers on-site and material equipment) may potentially override the effect of treatment.

Given the more debatable comparability of salivary and serum T levels in women and possibly raised salivary T (as compared to serum T) in women relatively to men (Davison, 2009; Fiers et al., 2014; Flyckt et al., 2009), we suspect that more studies arriving to similarly peculiar results may have been conducted, but possibly unpublished due to publication bias and concerns over data reliability. If that is so, their publication is desirable in order to achieve greater reliability and replicability in future studies.

Another potential limitation could be that only a single measurement of C and T levels was done for each participant, making it potentially prone to fluctuations. However, the experiments were conducted always at the same time of day and season of the year, and all the participants received the same instructions beforehand (that they should only come if healthy, go to sleep at their usual time the previous day, avoid overt stress or physical exertion before the experiment, drink enough water and avoid consuming food and flavored drinks at least an hour before the start of the experiment). Any out-of-the-usual hormonal fluctuation in our sample should therefore only introduce noise, but no systematic error.

Finally, while the stakes were only approx. 2 USD, a recent meta-analysis of stake size in DG and UG found no effect of stakes on the offer size in UG and a statistically significant, but practically very small effect in DG (Larney et al., 2019), although it cannot be excluded that some effect might exist within the context of hormonal levels.

To summarize, our findings provide tentative support for the dual-hormone theory, where high-T men (especially if also low-C) may express their dominance through altruism as a status signal. This finding is not that surprising given different mate choice strategies across sexes, especially pertaining to resource signaling (Todd et al., 2007; Waynforth & Dunbar, 1995; Wiederman, 1993).

Further studies and meta-analyses are sorely needed to ascertain the hormonal effects on altruism (including predictions made by different ultimate explanations of altruism) and the role of the dual hormone hypothesis. The current evidence remains mixed, but differences in study design (e.g. measuring salivary, blood or hair cortisol or testosterone; working with endogenous hormone levels or using hormonal administration; putting the participants in either the role of proposer or responder in a given game, or both; using experimental games or other measures of altruism; anonymous or face-to-face setting in the game; hypothetical or real stakes, and their value; mentioning specifically testosterone before the experiment, etc.) may be responsible for some of the observed differences, highlighting the need for carefully constructed meta-analyses and for far more replication studies. For studying relatively easily measurable proxies of androgen exposure, such as the digit ratio, large-scale studies such as Neyse et al. (2020) are preferable. For measuring hormonal levels directly, this is of course not attainable due to the cost of obtaining and analyzing the samples, and future meta-studies and replications can at least partly overcome the resulting small-sample issues.

The findings ascertained here will hopefully contribute to untangling the current enigma of hormonal influences of altruistic behavior, which appears to be complex and often situationally-dependent. Our results show that both high T and C were related to lower generosity in women, while the relationship of T ran in the opposite direction in men, where it was further modulated by the levels of C in a complex pattern. This outcome tentatively suggests that altruism could be interpreted in terms of status-seeking and display, in line with earlier findings of Durkee et al. (2020) that benefit-generation rather than cost-affliction leads to higher status allocation. However, further research is needed to test these conclusions more thoroughly and in a greater variety of contexts and methods, such as mixed-sex groups or higher stakes.

CRedit authorship contribution statement

Julie Novakova: Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Petr Tureček:** Writing – original draft, Visualization, Software, Formal analysis, Data curation. **Kamila Machová:** Writing – original draft, Project administration, Methodology. **Katerina Šýkorová:** Writing – original draft, Project administration, Methodology. **Vojtěch Zíka:** Writing – original draft, Project administration, Methodology. **Jaroslav Flegr:** Writing – original draft, Validation, Supervision, Resources, Methodology, Investigation, Conceptualization.

Financial support

This work was funded by the project GA UK 1502018, supported by the Charles University's grant agency.

Ethics Approval

The study was approved by the IRB of the Faculty of Science, Charles University.

Research Transparency & Reproducibility

All data and computer code used in the analysis are publicly available at <https://osf.io/pwztm/>.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.evolhumbehav.2024.106615>.

References

- Al-Dujaili, E. A. S., & Sharp, M. A. (2012). Female salivary testosterone: Measurement, challenges and applications. *Steroids-From Physiology to Clinical Medicine*, 129–167. <https://doi.org/10.5772/53648>
- Bailey, P. E., Ruffman, T., & Rendell, P. G. (2013). Age-related differences in social economic decision making: The ultimatum game. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 68(3), 356–363. <https://doi.org/10.1093/geronb/gbs073>
- Baxendale, P. M., Jacobs, H. S., & James, V. H. T. (1982). Salivary testosterone: Relationship to unbound plasma testosterone in normal and hyperandrogenic women. *Clinical Endocrinology*, 16(6), 595–603. <https://doi.org/10.1111/j.1365-2265.1982.tb03176.x>
- Bermejo, J. L., Valldecabres, R., Villarrasa-Sapiña, I., Monfort-Torres, G., Marco-Ahulló, A., & Do Couto, B. R. (2022). Increased cortisol levels caused by acute resistance physical exercise impair memory and learning ability. *PeerJ*, 10, Article e13000. <https://doi.org/10.7717/peerj.13000>
- Brañas-Garza, P., Espín, A. M., García-Muñoz, T., & Kovářík, J. (2019). Digit ratio (2D:4D) and prosocial behaviour in economic games: No direct correlation with generosity, bargaining or trust-related behaviours. *Biology Letters*, 15(8), 20190185. <https://doi.org/10.1098/rsbl.2019.0185>
- Brañas-Garza, P., & Kovářík, J. (2013). Digit ratios and social preferences: A comment on Buser (2012). ESI Working Paper 13–31. Retrieved from http://digitalcommons.chapman.edu/esi_working_papers/29.
- Brunner, R., Schaefer, D., Hess, K., Parzer, P., Resch, F., & Schwab, S. (2006). Effect of high-dose cortisol on memory functions. *Annals of the New York Academy of Sciences*, 1071(1), 434–437. <https://doi.org/10.1196/annals.1364.037>
- Burnham, T. C. (2007). High-testosterone men reject low ultimatum game offers. *Proceedings of the Royal Society B: Biological Sciences*, 274(1623), 2327–2330. <https://doi.org/10.1098/rspb.2007.0546>
- Buser, T. (2012). Digit ratios, the menstrual cycle and social preferences. *Games and Economic Behavior*, 76(2), 457–470. <https://doi.org/10.1016/j.geb.2012.07.006>

- Ceccato, S., Kettner, S. E., Kudielka, B. M., Schwieren, C., & Voss, A. (2018). Social preferences under chronic stress. *PLoS one*, 13(7), Article e0199528. <https://doi.org/10.1371/journal.pone.0199528>
- Cueva, C., Roberts, R. E., Spencer, T. J., Rani, N., Tempest, M., Tobler, P. N., ... Rustichini, A. (2017). Testosterone administration does not affect men's rejections of low ultimatum game offers or aggressive mood. *Hormones and Behavior*, 87, 1–7. <https://doi.org/10.1016/j.yhbeh.2016.09.012>
- Dabbs, J. M., Jr. (1990). Salivary testosterone measurements: Reliability across hours, days, and weeks. *Physiology & Behavior*, 48(1), 83–86. [https://doi.org/10.1016/0031-9384\(90\)90265-6](https://doi.org/10.1016/0031-9384(90)90265-6)
- Dabbs, J. M., Jr., Campbell, B. C., Gladue, B. A., Midgley, A. R., Navarro, M. A., Read, G. F., ... Worthman, C. M. (1995). Reliability of salivary testosterone measurements: A multicenter evaluation. *Clinical Chemistry*, 41(11), 1581–1584. <https://doi.org/10.1093/clinchem/41.11.1581>
- Davison, S. (2009). Salivary testing opens a Pandora's box of issues surrounding accurate measurement of testosterone in women. *Menopause*, 16(4), 630–631. <https://doi.org/10.1210/js.2016-1029>
- Dekkers, T. J., van Rentergem, J. A. A., Meijer, B., Popma, A., Wagemaker, E., & Huizenga, H. M. (2019). A meta-analytical evaluation of the dual-hormone hypothesis: Does cortisol moderate the relationship between testosterone and status, dominance, risk taking, aggression, and psychopathy? *Neuroscience & Biobehavioral Reviews*, 96, 250–271. <https://doi.org/10.1016/j.neubiorev.2018.12.004>
- Del Giudice, M., & Gangestad, S. W. (2022). Hormone ratios suffer from striking lack of robustness to measurement error. *Psychoneuroendocrinology*, 142, 105802. <https://doi.org/10.1016/j.psyneuen.2022.105802>
- Denson, T. F., Mehta, P. H., & Tan, D. H. (2013). Endogenous testosterone and cortisol jointly influence reactive aggression in women. *Psychoneuroendocrinology*, 38(3), 416–424. <https://doi.org/10.1016/j.psyneuen.2012.07.003>
- Diekhof, E. K., Wittmer, S., & Reimers, L. (2014). Does competition really bring out the worst? Testosterone, social distance and inter-male competition shape parochial altruism in human males. *PLoS One*, 9(7), Article e98977. <https://doi.org/10.1371/journal.pone.0098977>
- Durkee, P. K., Lukaszewski, A. W., & Buss, D. M. (2020). Psychological foundations of human status allocation. *Proceedings of the National Academy of Sciences*, 117(35), 21235–21241. <https://doi.org/10.1073/pnas.2006148117>
- Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, 14, 583–610. <https://doi.org/10.1007/s10683-011-9283-7>
- Fiers, T., Delanghe, J., T'Sjoen, G., Van Caenegem, E., Wierckx, K., & Kaufman, J. M. (2014). A critical evaluation of salivary testosterone as a method for the assessment of serum testosterone. *Steroids*, 86, 5–9. <https://doi.org/10.1016/j.steroids.2014.04.013>
- Flegr, J., Lindová, J., Pivoňková, V., & Havlíček, J. (2008). Brief communication: Latent toxoplasmosis and salivary testosterone concentration—Important confounding factors in second to fourth digit ratio studies. *American Journal of Physical Anthropology*, 137(4), 479–484. <https://doi.org/10.1002/ajpa.20888>
- Flyckt, R. L., Liu, J., Frasure, H., Wekselman, K., Buch, A., & Kingsberg, S. A. (2009). Comparison of salivary versus serum testosterone levels in postmenopausal women receiving transdermal testosterone supplementation versus placebo. *Menopause*, 16(4), 680–688. <https://doi.org/10.1097/gme.0b013e318199d5c4>
- Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6(3), 347–369. <https://doi.org/10.1006/game.1994.1021>
- Galizzi, M. M., & Nieboer, J. (2015). Digit ratio (2D: 4D) and altruism: Evidence from a large, multi-ethnic sample. *Frontiers in Behavioral Neuroscience*, 9, 41. <https://doi.org/10.3389/fnbeh.2015.00041>
- Grebe, N. M., Del Giudice, M., Thompson, M. E., Nickels, N., Ponzi, D., Zilioli, S., ... Gangestad, S. W. (2019). Testosterone, cortisol, and status-striving personality features: A review and empirical evaluation of the dual hormone hypothesis. *Hormones and Behavior*, 109, 25–37. <https://doi.org/10.1016/j.yhbeh.2019.01.006>
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3(4), 367–388. [https://doi.org/10.1016/0167-2681\(82\)90011-7](https://doi.org/10.1016/0167-2681(82)90011-7)
- Hamilton, W. D. (1964a). The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7(1), 1–16. [https://doi.org/10.1016/0022-5193\(64\)90038-4](https://doi.org/10.1016/0022-5193(64)90038-4)
- Hamilton, W. D. (1964b). The genetical evolution of social behaviour. II. *Journal of Theoretical Biology*, 7(1), 17–52. [https://doi.org/10.1016/0022-5193\(64\)90039-6](https://doi.org/10.1016/0022-5193(64)90039-6)
- Het, S., Ramlow, G., & Wolf, O. T. (2005). A meta-analytic review of the effects of acute cortisol administration on human memory. *Psychoneuroendocrinology*, 30(8), 771–784. <https://doi.org/10.1016/j.psyneuen.2005.03.005>
- Hönekopp, J., Bartholdt, L., Beier, L., & Liebert, A. (2007). Second to fourth digit length ratio (2D: 4D) and adult sex hormone levels: New data and a meta-analytic review. *Psychoneuroendocrinology*, 32(4), 313–321. <https://doi.org/10.1016/j.psyneuen.2007.01.007>
- Inoue, Y., Burriss, R. P., Hasegawa, T., & Kiyonari, T. (2023). Testosterone promotes dominance behaviors in the ultimatum game after players' status increases. *Scientific Reports*, 13(1), 18029. <https://doi.org/10.1038/s41598-023-45247-4>
- Inoue, Y., Takahashi, T., Burriss, R. P., Arai, S., Hasegawa, T., Yamagishi, T., & Kiyonari, T. (2017). Testosterone promotes either dominance or submissiveness in the ultimatum game depending on players' social rank. *Scientific Reports*, 7(1), 1–9. <https://doi.org/10.1038/s41598-017-05603-7>
- Keevil, B. G., Clifton, S., Tanton, C., Macdowall, W., Copas, A. J., Lee, D., ... Mercer, C. H. (2017). Distribution of salivary testosterone in men and women in a british general population-based sample: The third national survey of sexual attitudes and lifestyles (Natsal-3). *Journal of the Endocrine Society*, 1(1), 14–25. <https://doi.org/10.1210/js.2016-1029>
- Knickmeyer, R. C., Woolson, S., Hamer, R. M., Konneker, T., & Gilmore, J. H. (2011). 2D: 4D ratios in the first 2 years of life: Stability and relation to testosterone exposure and sensitivity. *Hormones and Behavior*, 60(3), 256–263. <https://doi.org/10.1016/j.yhbeh.2011.05.009>
- Knight, E. L., Morales, P. J., Christian, C. B., Prasad, S., Harbaugh, W. T., Mehta, P. H., & Mayr, U. (2022). The causal effect of testosterone on men's competitive behavior is moderated by basal cortisol and cues to an opponent's status: Evidence for a context-dependent dual-hormone hypothesis. *Journal of Personality and Social Psychology*, 123(4), 693. <https://doi.org/10.1037/pspa0000305>
- Kopsida, E., Berrebi, J., Petrovic, P., & Ingvar, M. (2016). Testosterone administration related differences in brain activation during the ultimatum game. *Frontiers in Neuroscience*, 10, 66. <https://doi.org/10.3389/fnins.2016.00066>
- Kozaki, T., Hashiguchi, N., Kaji, Y., Yasukouchi, A., & Tochihara, Y. (2009). Effects of saliva collection using cotton swab on cortisol enzyme immunoassay. *European Journal of Applied Physiology*, 107(6), 743–746. <https://doi.org/10.1007/s00421-009-1178-3>
- Kratochvíl, L., & Flegr, J. (2009). Differences in the 2nd to 4th digit length ratio in humans reflect shifts along the common allometric line. *Biology Letters*, 5(5), 643–646. <https://doi.org/10.1098/rsbl.2009.0346>
- Larney, A., Rotella, A., & Barclay, P. (2019). Stake size effects in ultimatum game and dictator game offers: A meta-analysis. *Organizational Behavior and Human Decision Processes*, 151, 61–72. <https://doi.org/10.1016/j.obhdp.2019.01.002>
- Manning, J. T., Scutt, D., Wilson, J., & Lewis-Jones, D. I. (1998). The ratio of 2nd to 4th digit length: A predictor of sperm numbers and concentrations of testosterone, luteinizing hormone and oestrogen. *Human Reproduction (Oxford, England)*, 13(11), 3000–3004. <https://doi.org/10.1093/humrep/13.11.3000>
- Marwell, G., & Ames, R. E. (1981). Economists free ride, does anyone else? Experiments on the provision of public goods, IV. *Journal of Public Economics*, 15(3), 295–310. [https://doi.org/10.1016/0047-2727\(81\)90013-X](https://doi.org/10.1016/0047-2727(81)90013-X)
- Matsumoto, Y., Yamagishi, T., Li, Y., & Kiyonari, T. (2016). Prosocial behavior increases with age across five economic games. *PLoS One*, 11(7). <https://doi.org/10.1371/journal.pone.0158671>
- McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Oxon: CRC Press. <https://doi.org/10.1201/9781315372495>
- Mehta, P. H., & Beer, J. (2010). Neural mechanisms of the testosterone-aggression relation: The role of orbitofrontal cortex. *Journal of Cognitive Neuroscience*, 22(10), 2357–2368. <https://doi.org/10.1162/jocn.2009.21389>
- Mehta, P. H., & Josephs, R. A. (2010). Testosterone and cortisol jointly regulate dominance: Evidence for a dual-hormone hypothesis. *Hormones and Behavior*, 58(5), 898–906. <https://doi.org/10.1016/j.yhbeh.2010.08.020>
- Millet, K., & Dewitte, S. (2009). The presence of aggression cues inverts the relation between digit ratio (2D: 4D) and prosocial behaviour in a dictator game. *British Journal of Psychology*, 100(1), 151–162. <https://doi.org/10.1348/000712608X324359>
- Montoya, E. R., Terburg, D., Bos, P. A., & Van Honk, J. (2012). Testosterone, cortisol, and serotonin as key regulators of social aggression: A review and theoretical perspective. *Motivation and Emotion*, 36(1), 65–73. <https://doi.org/10.1007/s11031-011-9264-3>
- Nakavachara, V. (2018). The economics of altruism: The old, the rich, the female. *Journal of Human Behavior in the Social Environment*, 28(1), 28–43. <https://doi.org/10.1080/10911359.2017.1393367>
- Neyse, L., Vieider, F. M., Ring, P., Probst, C., Kaernbach, C., van Eimeren, T., & Schmidt, U. (2020). Risk attitudes and digit ratio (2D: 4D): Evidence from prospect theory. *Journal of Risk and Uncertainty*, 60(1), 29–51. <https://doi.org/10.1007/s11166-020-09321-w>
- Novakova, J., Machová, K., & Flegr, J. (2021). Looking like a million dollars: Does attractiveness priming increase altruistic behavior in experimental games? *Frontiers in Psychology*, 12, Article 658466. <https://doi.org/10.3389/fpsyg.2021.658466>
- Ogawa, K., Kawamura, T., & Matsushita, K. (2018). *Higher the cognitive ability of, and/or the younger the dictator, the more selfish: Experimental evidence*. Available at SSRN 3267146. <https://doi.org/10.2139/ssrn.3267146>
- Parslow, E., Ranehill, E., Zethraeus, N., Blomberg, L., von Schoultz, B., Hirschberg, A. L., ... Dreber, A. (2019). The digit ratio (2D: 4D) and economic preferences: No robust associations in a sample of 330 women. *Journal of the Economic Science Association*, 5(2), 149–169. <https://doi.org/10.1007/s40881-019-00076-y>
- Pfaffheicher, S. (2017). Illuminating the dual-hormone hypothesis: About chronic dominance and the interaction of cortisol and testosterone. *Aggressive Behavior*, 43(1), 85–92. <https://doi.org/10.1002/ab.21665>
- Pfaffheicher, S., & Keller, J. (2014). Towards a biopsychological understanding of costly punishment: The role of basal cortisol. *PLoS One*, 9(1), Article e85691. <https://doi.org/10.1371/journal.pone.0085691>
- Pfaffheicher, S., Landhäußer, A., & Keller, J. (2014). Individual differences in antisocial punishment in public goods situations: The interplay of cortisol with testosterone and dominance. *Journal of Behavioral Decision Making*, 27(4), 340–348. <https://doi.org/10.1002/bdm.1811>
- Pornpattananangkul, N., Chowdhury, A., Feng, L., & Yu, R. (2019). Social discounting in the elderly: Senior citizens are good Samaritans to strangers. *The Journals of Gerontology: Series B*, 74(1), 52–58. <https://doi.org/10.1093/geronb/gbx040>
- Prasad, S., Knight, E. L., & Mehta, P. H. (2019). Basal testosterone's relationship with dictator game decision-making depends on cortisol reactivity to acute stress: A dual-hormone perspective on dominant behavior during resource allocation. *Psychoneuroendocrinology*, 101, 150–159. <https://doi.org/10.1016/j.psyneuen.2018.11.012>
- Prasad, S., Narayanan, J., Lim, V. K., Koh, G. C., Koh, D. S., & Mehta, P. H. (2017). Preliminary evidence that acute stress moderates basal testosterone's association

- with retaliatory behavior. *Hormones and Behavior*, 92, 128–140. <https://doi.org/10.1016/j.yhbeh.2016.10.020>
- Raulo, A., & Dantzer, B. (2018). Associations between glucocorticoids and sociality across a continuum of vertebrate social behavior. *Ecology and Evolution*, 8(15), 7697–7716. <https://doi.org/10.1002/ece3.4059>
- RCoreTeam. (2019). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org>.
- Reimers, L., & Diekhof, E. K. (2015). Testosterone is associated with cooperation during intergroup competition by enhancing parochial altruism. *Frontiers in Neuroscience*, 9, 183. <https://doi.org/10.3389/fnins.2015.00183>
- Reimers, L., Kappo, E., Stadler, L., Yaquib, M., & Diekhof, E. K. (2019). Endogenous testosterone correlates with parochial altruism in relation to costly punishment in different social settings. *PeerJ*, 7, Article e7537. <https://doi.org/10.7717/peerj.7537>
- Ronay, R., & Galinsky, A. D. (2011). Lex talionis: Testosterone and the law of retaliation. *Journal of Experimental Social Psychology*, 47(3), 702–705. <https://doi.org/10.2139/ssrn.1872648>
- Roney, J. R. (2016). Theoretical frameworks for human behavioral endocrinology. *Hormones and Behavior*, 84, 97–110. <https://doi.org/10.1016/j.yhbeh.2016.06.004>
- Salimetrics. (2019). Expanded range salivary testosterone enzyme immunoassay kit. Retrieved from: <https://salimetrics.com/wp-content/uploads/2018/03/testosterone-saliva-elisa-kit.pdf>.
- Schulreich, S., Tusche, A., Kanske, P., & Schwabe, L. (2022). Altruism under stress: Cortisol negatively predicts charitable giving and neural value representations depending on mentalizing capacity. *Journal of Neuroscience*, 42(16), 3445–3460. <https://doi.org/10.1523/jneurosci.1870-21.2022>
- Senci, C. M., Breccia, F., & Freidin, E. (2020). Attention to prescriptive norms increases dictator game generosity in women but not men: Using the 2D: 4D digit ratio to test the role of biology. *Journal of Psychology*, 7, 21. <https://doi.org/10.12974/2313-1047.2020.07.2>
- Shirtcliff, E. A., Granger, D. A., & Likos, A. (2002). Gender differences in the validity of testosterone measured in saliva by immunoassay. *Hormones and Behavior*, 42(1), 62–69. <https://doi.org/10.1006/hbeh.2002.1798>
- Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2020). Specification curve analysis. *Nature Human Behaviour*, 4(11), 1208–1214. <https://doi.org/10.1038/s41562-020-0912-z>
- Singer, N., Sommer, M., Döhnel, K., Zänkert, S., Wüst, S., & Kudielka, B. M. (2017). Acute psychosocial stress and everyday moral decision-making in young healthy men: The impact of cortisol. *Hormones and Behavior*, 93, 72–81. <https://doi.org/10.1016/j.yhbeh.2017.05.002>
- Sparrow, E. P., Armstrong, B. A., Fiocco, A. J., & Spaniol, J. (2019). Acute stress and altruism in younger and older adults. *Psychoneuroendocrinology*, 100, 10–17. <https://doi.org/10.1016/j.psyneuen.2018.09.025>
- StanDevelopmentTeam. (2021). Stan modeling language users guide and reference manual. Retrieved from <https://mc-stan.org>.
- Stanton, S. J. (2017). The role of testosterone and estrogen in consumer behavior and social & economic decision making: A review. *Hormones and Behavior*, 92, 155–163. <https://doi.org/10.1016/j.yhbeh.2016.11.006>
- Starcke, K., Polzer, C., Wolf, O. T., & Brand, M. (2011). Does stress alter everyday moral decision-making? *Psychoneuroendocrinology*, 36(2), 210–219. <https://doi.org/10.1016/j.psyneuen.2010.07.010>
- Staufenbiel, S. M., Penninx, B. W., Spijker, A. T., Elzinga, B. M., & van Rossum, E. F. (2013). Hair cortisol, stress exposure, and mental health in humans: A systematic review. *Psychoneuroendocrinology*, 38(8), 1220–1235. <https://doi.org/10.1016/j.psyneuen.2012.11.015>
- Sýkora, Z. (2018). *Hormones and competitive behavior*. Master's thesis. Faculty of Social Sciences, Charles University. Retrieved from <https://dspace.cuni.cz/bitstream/handle/20.500.11956/99536/120298438.pdf>.
- Talarovičová, A., Kršková, L., & Blažeková, J. (2009). Testosterone enhancement during pregnancy influences the 2D: 4D ratio and open field motor activity of rat siblings in adulthood. *Hormones and Behavior*, 55(1), 235–239. <https://doi.org/10.1016/j.yhbeh.2008.10.010>
- Todd, P. M., Penke, L., Fasolo, B., & Lenton, A. P. (2007). Different cognitive processes underlie human mate choices and mate preferences. *Proceedings of the National Academy of Sciences*, 104(38), 15011–15016. <https://doi.org/10.1073/pnas.0705290104>
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46(1), 35–57. <https://doi.org/10.1086/406755>
- Van Honk, J., Montoya, E. R., Bos, P. A., Van Vugt, M., & Terburg, D. (2012). New evidence on testosterone and cooperation. *Nature*, 485(7399), E4–E5. <https://doi.org/10.1038/nature11136>
- Vinkers, C. H., Zorn, J. V., Cornelisse, S., Koot, S., Houtepen, L. C., Olivier, B., ... Joëls, M. (2013). Time-dependent changes in altruistic punishment following stress. *Psychoneuroendocrinology*, 38(9), 1467–1475. <https://doi.org/10.1016/j.psyneuen.2012.12.012>
- Waynforth, D., & Dunbar, R. I. (1995). Conditional mate choice strategies in humans: Evidence from 'Lonely Hearts' advertisements. *Behaviour*, 755–779. <https://doi.org/10.1163/156853995X00135>
- West, S. A., El Mouden, C., & Gardner, A. (2011). Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and Human Behavior*, 32(4), 231–262. <https://doi.org/10.1016/j.evolhumbehav.2010.08.001>
- Wiederman, M. W. (1993). Evolved gender differences in mate preferences: Evidence from personal advertisements. *Ethology and Sociobiology*, 14(5), 331–351. [https://doi.org/10.1016/0162-3095\(93\)90003-Z](https://doi.org/10.1016/0162-3095(93)90003-Z)
- Zahavi, A. (1995). Altruism as a handicap: The limitations of kin selection and reciprocity. *Journal of Avian Biology*, 26(1), 1–3. <https://doi.org/10.2307/3677205>
- Zak, P. J., Kurzban, R., Ahmadi, S., Swerdloff, R. S., Park, J., Efremidze, L., ... Matzner, W. (2009). Testosterone administration decreases generosity in the ultimatum game. *PLoS One*, 4(12), Article e8330. <https://doi.org/10.1371/journal.pone.0008330>
- Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., Von Schoultz, B. O., Hirschberg, A. L., & Johannesson, M. (2009). A randomized trial of the effect of estrogen and testosterone on economic behavior. *Proceedings of the National Academy of Sciences*, 106(16), 6535–6538. <https://doi.org/10.1073/pnas.0812757106>
- Zhang, Q., Ma, J., & Nater, U. (2019). How cortisol reactivity influences prosocial decision-making: The moderating role of sex and empathic concern. *Frontiers in Human Neuroscience*, 13, 415. <https://doi.org/10.3389/fnhum.2019.00415>
- Zilioli, S., Ponzi, D., Henry, A., & Maestriperieri, D. (2015). Testosterone, cortisol and empathy: Evidence for the dual-hormone hypothesis. *Adaptive Human Behavior and Physiology*, 1(4), 421–433. <https://doi.org/10.1007/S40750-014-0017-X>

Behavioural Processes

Testing Trivers: discerning the role of direct reciprocity in the evolution of altruism --Manuscript Draft--

Manuscript Number:	
Article Type:	Research Paper
Keywords:	altruism; experimental games; health; life history theory; reciprocity
Corresponding Author:	Julie Nekola Novakova, MSc. Charles University Faculty of Science CZECH REPUBLIC
First Author:	Julie Nekola Novakova, MSc.
Order of Authors:	Julie Nekola Novakova, MSc. Kamila Machová, MSc. Vojtěch Zíka, Ph.D. Jaroslav Flegr, Professor, Ph.D.
Abstract:	Altruism remains one of the oft-cited evolutionary puzzles, but its evolution can be explained by multiple phenomena. Among them is Triversian reciprocal altruism, relying, however, on conditions such as good memory and recognition, reduced temporal discounting, viscous populations and generally high probability of the same individuals interacting repeatedly and recognizing the fact. These are unlikely to be met in many species and led to widespread criticism of Trivers' hypothesis as a major explanatory factor of altruistic behavior. In humans, however, these conditions can be met relatively easily. We have investigated the possible impact of memory, health (influencing the chances of future interactions), cognitive reflection and temporal discounting on the degree of altruism exhibited in the experimental Dictator and Ultimatum Game, played for real monetary rewards, in a sample of 158 general-population subjects (N=158, 84 F, 74 M). We expected memory, health condition and cognitive reflection to correlate positively with altruistic behavior, while temporal discounting was expected to have a negative relationship with altruism. These hypotheses are additionally in line with the life history theory, which predicts that a harsher environment (likely leading to worse mental and/or physical health) or generally unfavorable conditions would lead to "faster", less cooperative life strategies and greater temporal discounting. However, we found no significant association between working memory or health and behavior in either of the experimental games. Cognitive reflection and intelligence only correlated negatively with minimum acceptable offers in the UG. Temporal discounting, however, correlated negatively with offers in the DG and UG, and positively with minimum acceptable UG offers. The overall findings suggest that if preconditions for Triversian direct reciprocity still play any active role in humans, it is likely a subtle, small effect not discernible in small samples; more likely, it is overshadowed by generalized reciprocity and other processes that don't place such strict demands on individual agents, and by proximate influences of altruistic behavior.

Testing Trivers: discerning the role of direct reciprocity in the evolution of altruism

Julie Novakova^{1*}, Kamila Machová¹, Vojtěch Zíka^{1,2}, Jaroslav Flegr¹

¹ Laboratory of Evolutionary Biology, Department of Philosophy and History of Science, Faculty of Science, Charles University, Prague, Czech Republic

² Laboratories of Behavioral Studies (LABS), Faculty of Social and Economic Studies, Jan Evangelista Purkyně University, Ústí nad Labem, Czech Republic

* Corresponding author

E-mail: julie.novakova@natur.cuni.cz

Address: Faculty of Science, Viničná 7, 128 00 Prague 2, Czech Republic

Abstract

Altruism remains one of the oft-cited evolutionary puzzles, but its evolution can be explained by multiple phenomena. Among them is Triversian reciprocal altruism, relying, however, on conditions such as good memory and recognition, reduced temporal discounting, viscous populations and generally high probability of the same individuals interacting repeatedly and recognizing the fact. These are unlikely to be met in many species and led to widespread criticism of Trivers' hypothesis as a major explanatory factor of altruistic behavior. In humans, however, these conditions can be met relatively easily. We have investigated the possible impact of memory, health (influencing the chances of future interactions), cognitive reflection and temporal discounting on the degree of altruism exhibited in the experimental Dictator and Ultimatum Game, played for real monetary rewards, in a sample of 158 general-population subjects (N=158, 84 F, 74 M). We expected memory, health condition and cognitive reflection to correlate positively with altruistic behavior, while temporal discounting was expected to have a negative relationship with altruism. These hypotheses are additionally in line with the life history theory, which predicts that a harsher environment (likely leading to worse mental and/or physical health) or generally unfavorable conditions would lead to "faster", less cooperative life strategies and greater temporal discounting. However, we found no significant association between working memory or health and behavior in either of the experimental games. Cognitive reflection and intelligence only correlated negatively with minimum acceptable offers in the UG. Temporal discounting, however, correlated negatively with offers in the DG and UG, and positively with minimum acceptable UG offers. The overall findings suggest that if preconditions for Triversian direct reciprocity still play any active role in humans, it is likely a subtle, small effect not discernible in small samples; more likely, it is overshadowed by generalized reciprocity and other processes that don't place such strict demands on individual agents, and by proximate influences of altruistic behavior.

Keywords: altruism; experimental games; health; life history theory; reciprocity

Introduction

Background

Altruistic behavior, fairness preferences and cooperation are, even after decades of research, still not sufficiently explained phenomena. A number of hypotheses have been proposed to shed light on the origins of these traits, including mutualism; kin selection (Hamilton 1964a, b; Rachlin & Jones 2008); interdemic (group) selection (Rachlin 2019); reciprocal altruism, both direct and indirect, based on individual reputation (Trivers 1971); various forms of “greenbeard” altruism (Gardner & West 2010); and costly signaling in sexual selection (Zahavi 1995). For the most part, the above-mentioned hypotheses do not exclude each other and could have contributed to the evolution of altruistic behavior to some extent, varying in different evolutionary lineages.

Reciprocal altruism (Trivers 1971) represents one of these major hypotheses of the evolution of altruistic behavior. Reciprocity, however, requires several nontrivial conditions. Trivers himself listed 1) many chances of reciprocal altruistic interactions in the length of a lifetime, 2) repeated interactions with the same small set of individuals, and 3) symmetrical exposure of individuals to each side of altruistic situations. In addition, Wilkinson (1988) reviewed known examples of reciprocal cooperation and conditions that must be specified before asserting that what has been observed truly constitutes reciprocation: the act must (at the time) reduce the donor's fitness compared to the selfish alternative; the fitness of the recipient must increase by the act; the act does not depend on an immediate benefit for the donor (separating reciprocity from mutualism); individual recognition and memory of previous interaction exist; there is a large but indefinite number of opportunities of reciprocal cooperation in an individual's lifetime. According to Stevens & Hauser (2004), most animals do not fulfill these conditions. Clutton-Brock (2009) was likewise skeptical about the occurrence of reciprocity, arguing that examples of it may be explained by kin selection or donor harassment by begging. However, many cases of direct (as well as indirect and probably less cognitively demanding generalized) reciprocal

1
2
3 altruism have been observed across species (Wilkinson 1984, 1988, Rutte & Taborsky 2007, 2008,
4
5 Schino & Aureli 2008, 2009, Viana et al. 2010, Wilkinson et al. 2016, Carter et al. 2020, Kettler et al.
6
7 2021, Engelhardt & Taborsky 2022, Dolivo et al. 2023).

8
9
10 In humans, notions of reciprocity seem consistent with the observed behavior in experiments
11
12 where it can be tested. Diekmann (2004) found that in a sequential Dictator Game, most people gave
13
14 the other player the same amount they received from him in the first round, and reciprocity enabled
15
16 them to predict the behavior quite well. Similarly to Church (1993), the subjects also cooperated by
17
18 mailing back a questionnaire most frequently when a pre-paid phone card had been included in the
19
20 package and least frequently when it had been promised for a mailed questionnaire – at odds with
21
22 rational payoff maximizing but in line with reciprocal altruism.
23
24
25

26
27 The aim of this study is to test the role of conditions outlined for reciprocal altruism (Trivers
28
29 1971, Stevens & Hauser 2004). Good memory, cognitive reflection and low temporal discounting are
30
31 necessary for individual recognition and assessment of the interaction, while health influences the
32
33 likelihood of future interactions with the same individuals – if it were low, direct reciprocity would be
34
35 unlikely to arise.
36
37
38

39 Post (2005) summarizes the state of knowledge about health and altruism and finds a positive
40
41 connection, which he interprets as a positive influence of helping others on health (modulated by stress
42
43 alleviation through a positive emotional reaction). This direction of influence can be widespread;
44
45 however, the author disregards the opposite direction. Figueredo et al. (2006), taking a different
46
47 approach, studied the heredity of traits related to life strategies, and their life history K-factor included
48
49 a lot of altruism-related variables and correlated with the covitality factor (subjective and objective
50
51 physical and mental health). In addition, a study by Jonason et al. (2015) found an association between
52
53 two Dark Triad traits (psychopathy and Machiavellianism) and negative health outcomes. Prosociality
54
55 and altruism, in contrast, seem connected to better health status, which is corroborated by a meta-study
56
57
58
59
60
61
62
63
64
65

1
2
3 by Hui et al. (2020) who found a modest connection especially between prosociality and psychological
4
5 functioning (a lower one in case of psychological malfunctioning and physical health condition).
6

7
8 Memory, cognition and temporal discounting are also expected to figure prominently in direct
9
10 reciprocity. Kroneisen & Bell (2022) reviewed the role of memory in cooperation, tentatively
11
12 concluding that memory mechanisms oriented at predicting future behavior seem to be supported by
13
14 the evidence, while highly specialized cheater detection modules do not.
15
16

17
18 The complexity of cooperative behavior does appear to be influenced by memory capabilities:
19
20 Milinski & Wedekind (1998) found that if interrupted by a memory-requiring task, student participants
21
22 performed worse in an iterated prisoner's dilemma, and that participants still using more complex
23
24 strategies in the game performed worse in the task requiring working memory. That leads to the
25
26 possible conclusion that memory is vital for assessing such cooperative interactions, and there is a
27
28 trade-off between memory load and cooperation. Related studies (Duffy & Smith 2014, Soutschek &
29
30 Schubert 2015) corroborated this conclusion. Similarly, Murty et al. (2016) found a connection between
31
32 episodic memory and choosing cooperative partners (the participants' own cooperation or altruistic
33
34 behavior was not studied here). A positive connection between memory and cooperation is also
35
36 supported, for instance, by findings by Winke & Stevens (2017), who studied it from the opposite
37
38 direction and found that more cooperative people and people with larger social networks exhibited
39
40 somewhat better memory skills. Computer models suggested that memory increases cooperation as
41
42 well (Moreira et al. 2013). However, some found that it might apply in some games and not in others,
43
44 depending primarily on the cost/benefit ratio (Wang et al. 2014). Finally, animal studies suggest that
45
46 memory and temporal discounting are important factors in not only humans, but also in bird species
47
48 (Larose & Dubois 2011, Stevens & Hauser 2004, Stephens et al. 2002).
49
50
51
52
53
54
55

56
57 As to temporal discounting, Curry et al. (2008) showed that people manifesting low temporal
58
59 discounting contributed more in the Public Goods Game. And while Osiński et al. (2021) found no
60
61 significant effect of temporal discounting rates on offers in the Dictator and Ultimatum Games
62
63
64
65

specifically, the overall literature suggests that steeper discounting lowers cooperation and altruism (Persson et al. 2024). We did not examine social and probability discounting here, but they have been shown to impact altruistic choices as well (Yi et al. 2011, Buddiga & Locey 2024).

A related trait is cognitive reflection; Frederick (2005) found a positive correlation between lower temporal discounting and higher cognitive reflection. However, there has been a long-standing gap in studying the possible connection between cognitive reflection/deliberation and altruistic behavior, addressed only later by several studies (Ben-Ner et al. 2004, Ponti & Rodriguez-Lara 2015, Corgnet et al. 2015, Czerwonka et al. 2018), which found people who scored better on the Cognitive Reflection Test (CRT) or related tests to generally make more selfish offers in the Dictator Game. Whilst one would hypothesize a positive connection based on Triversian direct reciprocity conditions, the current literature already casts a shadow upon it.

In the present study, we focus exactly on the role of pre-conditions of reciprocal altruism and test whether better memory and cognitive scores, lower temporal discounting and better health correlated positively with altruistic behavior in two experimental games, the Dictator Game (DG; Forsythe et al. 1994) and the Ultimatum Game (UG; Güth et al. 1982), both widely used in behavioral sciences. While one could object that these are anonymous one-off interactions, they are widely used measures of altruism and fairness. If direct reciprocity played a key role in the evolution of altruism in humans, one should expect the same mechanisms to kick in even in these settings.

Hypotheses

1. Working memory will correlate positively with altruism.
2. Cognitive reflection and intelligence will correlate positively with altruism.
3. Temporal discounting will correlate negatively with altruism.
4. Better health condition will correlate positively with altruism.

Methods

Participant pool and experimental procedure

The data was collected at the Faculty of Science, Charles University in Prague, in early summer 2018 and 2019. The experiments were conducted in accordance with the Declaration of Helsinki and approved by the faculty IRB (approval no. 2018/10). Participants were invited via online recruitment, with a maximum of 18 participants invited for each session. In total, 158 people participated (74 men with a mean age of 29 years, a median age of 29 years, and ages ranging from 18 to 53; 84 women with a mean age of 31 years, a median age of 28 years, and ages ranging from 18 to 58). On the site, the participants were greeted at the reception and led to the computer lab, where they first read and signed their informed consent and guidelines for the experiment. For a different study looking at testosterone and cortisol levels (Novakova et al., 2024), their saliva was collected into salivettes. After that, they took part in a survey including both experimental games, the Meili Test of short-term memory, a four-question Cognitive Reflection Test (CRT), Cattell's 12-item (A version) intelligence test, self-reported health (physical and mental condition as perceived relative to peers) and self-perceived memory problems, and a temporal discounting task. Another section of the survey focused on participants' views of attractiveness and their number of romantic and sexual relationships, collected for another study (Novakova et al. 2021), which describes the details of that section.

The DG and UG were each played for 400 tokens translatable into 40 CZK (roughly 2 USD), which constituted the participants' reward. They were first presented with the rules and truthfully told that they would be randomly matched with another person in the room and that their earnings at the end of the experiment would reflect the outcome of both games. Each participant was asked how much they want to give (DG) and offer (UG) the other player, and what is their minimum acceptable offer (MAO) in the UG from their co-player. Before the participants played the games, they answered control

questions to assess whether they understood the rules of the games correctly and a question about any previous knowledge of either of the games (from lectures, literature, and so on).

Finally, several indices were calculated before the analysis: A memory score was computed from the number of successfully remembered Meili test images. The number of correct answers in the CRT test amounted to one overall CRT score. An intelligence score was computed from the number of correct choices in Cattell's intelligence test. And, last but not least, two measures of k (the discount rate) were computed from the participants' preferences in two questions on temporal discounting: the first one assessing the indifference point in case of a one-month delay, the second one in case of a one-year delay. The value (V) was fixed at 10,000 CZK (approx. 410 USD), the delay (D) at one, resp. 12 months, and the amount (A) to be hypothetically received after the delay varied between 10,000 and 30,000 CZK. The following discounting function was used to calculate the discount rate for all participants with consistent preferences (participants with inconsistent preferences, e.g. preferring the immediate reward of 10,000 CZK instead of a delayed reward of 20,000 CZK, but preferring a delayed reward of 11,000 CZK, were omitted from the discounting analysis) for each of the questions:

$$k = (A/V - 1)/D$$

Statistical analysis

All the analyses were conducted in R version 4.4.2 (R Core Team 2021). First, the validity of the responses was checked (completion of the survey and answering control questions about understanding the principles of the DG and UG correctly). After descriptive statistics were done (plus a t-test gender differences in the games and correlation of offers in both games), bivariate analyses were conducted using partial Kendall correlation. This nonparametric test examines the relationship between a single dependent variable and one independent variable while controlling for other variables, here sex and age as covariates. The R package "ppcor" was used (Kim 2015). Tau in the results always designates

Kendall's Tau controlled for sex and age. Since the observed significant effects went against the tested hypotheses, we did not perform a correction for multiple testing; the *a priori* hypotheses had to be refuted in any case.

Data availability

The dataset used in this study has been published on FigShare (10.6084/m9.figshare.28292240); more data from the same pool of participants, mentioned above and collected for a different study (Novakova et al., 2021) was also published on FigShare (10.6084/m9.figshare.12067407); data concerning the saliva collection study is published on Open Science Framework (osf.io/pwztm/).

Results

Descriptive statistics of the DG and UG

Participants gave on average 154.4 tokens (38.6%) in the DG, with 200 tokens (50%) being the median allocation. Women gave on average 41.3%, while men gave 35.6% (difference not statistically significant: $t=1.830$, $df=125.6$, $p=0.070$).

In the UG, the mean offer was 202.3 tokens (50.6%), median 200 tokens (50%). Women offered on average 49.4%, men 52% ($t=-1.284$, $df=155.9$, $p=0.201$). The mean minimum acceptable offer (MAO) in the UG was 140.4 tokens (35.1%), median 150 tokens (37.5%); in women, mean MAO was 37 %, while MAO in men was on average 32.9% ($t=1.509$, $df=154.3$, $p=0.133$).

Offers in both games were weakly correlated, as shown by one-sided Pearson's correlation test (positive correlation was expected): $r=0.152$, $t=1.924$, $df=156$, $p=0.028$. Finally, 29% of participants had pre-existing knowledge of both games or either one beforehand (from academic or popular literature, classes, etc.). Neither contributions in the DG (Tau=0.065, Cohen's $f^2=0.102$, $p=0.231$) or the UG (Tau=-0.023, Cohen's $f^2=0.036$, $p=0.675$) were significantly predicted by previous knowledge of

the games; there was only a trend in the UG MAO, where previous knowledge correlated negatively with the MAOs in a two-sided test (Tau=-0.104, Cohen's f^2 =0.165, p =0.055).

Hypotheses testing

Results from the hypotheses testing are summarized below in Table 1. The results do not support a connection between memory and altruism, nor between health (whether mental or physical) and altruism. We observed a negative relationship of the CRT scores with UG offers and MAO, as well as the negative relationship between Cattell's intelligence test scores and UG MAO. Temporal discounting rates exhibited a mostly non-significant, but consistent pattern of a negative connection with DG and UG offers, and a positive one with UG MAO.

	DG contribution	UG offer	UG MAO
Meili memory score	Tau=0.012 Cohen's f^2 =0.019 p =0.826	Tau=-0.028 Cohen's f^2 =0.043 p =0.609	Tau=-0.036 Cohen's f^2 =0.057 p =0.503
Self-reported memory problems	Tau=0.012 Cohen's f^2 =0.018 p =0.831	Tau=0.081 Cohen's f^2 =0.128 p =0.134	Tau=0.010 Cohen's f^2 =0.016 p =0.847
CRT score	Tau=0.033 Cohen's f^2 =0.051 p =0.545	<i>Tau=-0.132</i> <i>Cohen's f^2=0.211</i> <i>p=0.015</i>	<i>Tau=-0.212</i> <i>Cohen's f^2=0.346</i> <i>p<0.0001</i>
Intelligence score	Tau=-0.039 Cohen's f^2 =0.061 p =0.471	Tau=0.028 Cohen's f^2 =0.044 p =0.605	<i>Tau=-0.128</i> <i>Cohen's f^2=0.204</i> <i>p=0.018</i>

One-month discount rate	<i>Tau=-0.108</i> <i>Cohen's f²=0.172</i> <i>p=0.101</i>	<i>Tau=-0.112</i> <i>Cohen's f²=0.178</i> <i>p=0.089</i>	<i>Tau=0.112</i> <i>Cohen's f²=0.178</i> <i>p=0.091</i>
One-year discount rate	<i>Tau=-0.117</i> <i>Cohen's f²=0.187</i> <i>p=0.071</i>	<i>Tau=-0.096</i> <i>Cohen's f²=0.151</i> <i>p=0.142</i>	<i>Tau=0.189</i> <i>Cohen's f²=0.305</i> <i>p=0.004</i>
Physical health	<i>Tau=0.052</i> <i>Cohen's f²=0.083</i> <i>p=0.880</i>	<i>Tau=0.031</i> <i>Cohen's f²=0.049</i> <i>p=0.561</i>	<i>Tau=0.093</i> <i>Cohen's f²=0.148</i> <i>p=0.084</i>
Mental health	<i>Tau=0.008</i> <i>Cohen's f²=0.013</i> <i>p=0.331</i>	<i>Tau=0.011</i> <i>Cohen's f²=0.018</i> <i>p=0.835</i>	<i>Tau=0.007</i> <i>Cohen's f²=0.012</i> <i>p=0.891</i>

Table 1. Hypotheses testing results. The table shows Kendall's Tau, Cohen's f and p values from the nonparametric Kendall correlation (with sex and age as covariates). Statistically significant results are depicted in bold and italics, non-significant statistical trends in italics.

Since there appeared to be a consistent pattern of both temporal discounting rates' effect, Stouffer's meta-analysis method (Stouffer et al. 1949), using the R package "poolr" (Cinar & Viechtbauer 2022), was applied to elucidate the combined significance of the effect. However, as the direction of the effect ran opposite in case of UG MAO (the amount rose with the increasing discount rate, while DG contributions and UG offers decreased with it), a correction for two-tailed tests had to be used (and values of $1-p$ used in the test instead of p in the case of MAO). Therefore, the combined p value for the one-month discount rate k was $p=0.229$, and for the one-year discount rate k $p=0.529$. The relationship of behavior in the games with discounting rates, albeit non-significant, is visualized in Figure 1.

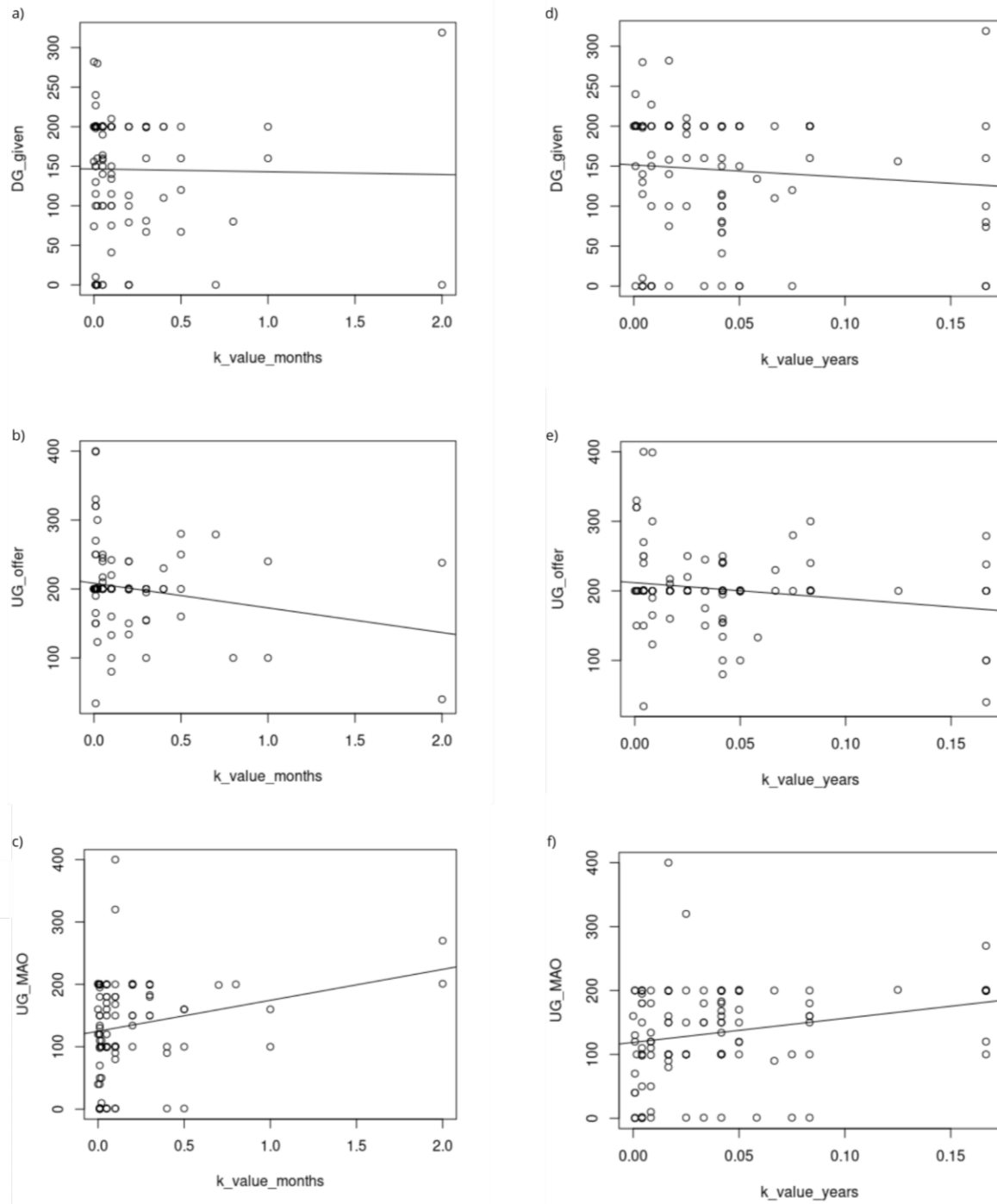


Figure 1. Temporal discounting and behavior in the UG and DG. The scatterplots with linear regression lines above show the offers in the DG and UG and the minimum acceptable offers in the UG by participants' discount rates. The left column shows behavior in the games by the discount rate for one month's time (plot a) DG

contributions, b) UG offers, c) UG MAO), while the right column depicts the same for the one year's discount rate (plot d) DG contributions, e) UG offers, f) UG MAO).

Discussion

Altruism is a complex phenomenon that has likely been shaped by multiple factors (not only) in humans. Our aim was to test whether preconditions of direct reciprocity, as envisaged by Trivers (1971) and others, building up on his concept, play a role in present-day altruistic behavior in humans. We hypothesized that better working memory and cognitive abilities, lower temporal discounting and better health condition would correlate with greater altruism in the experimental Dictator and Ultimatum Games. The overall results, however, do not support these hypotheses.

Working memory, measured through the Meili test (where a better score would have correlated with more altruistic behavior, had the hypothesis been correct) and additionally self-reported memory problems (which would have correlated with altruism negatively if supporting the central hypothesis), did not correlate with offers either in the DG or UG, nor with rejection or acceptance of UG offers. This finding runs contrary to previous studies (Soutschek & Schubert 2015, Duffy & Smith 2014, Milinski & Wedekind 1998). However, we may have observed no connection between memory and altruism in the DG/UG due to the one-shot character of the games; Milinski & Wedekind's (1998) seminal study used an iterated – repeated – Prisoner's Dilemma and found that using more working memory in-between rounds of the games impaired the complexity of the strategies used. Nevertheless, if better memory skills increased altruism in general on a deeper level, not only strategic cooperativeness (as suggested e.g. by studies of one's network size, memory and cooperation – see Winke & Stevens 2017), we should have been able to discern the effect.

At odds with the predictions stemming from direct reciprocity preconditions, performance in cognitive reflection and intelligence tasks had no connection to contributions in the DG, which can be

1
2
3 best interpreted as altruistic, but correlated negatively with MAOs in the UG (and in the case of CRT
4 also UG offers), where strategic thinking comes into play. Rejecting unfair offers has been commonly
5 interpreted in terms of altruistic punishment in the literature (Fehr & Gächter 2002, Fehr & Fischbacher
6 2004, Casal et al. 2019), although some authors warned before this interpretation (Brethel-Haurwitz et
7 al. 2016). If we adhere to it, low MAOs can be interpreted as less altruistic, which runs contrary to our
8 pre-conceived hypothesis of better cognitive abilities predicting greater altruism. In contrast, it would
9 be consistent with an alternative explanation of strategic cooperation rather than altruism (and with
10 high CRT and other tests-scoring people being more selfish in the DG; see Ben-Ner et al. 2004, Ponti &
11 Rodriguez-Lara 2015, Corgnet et al. 2015, Czerwinka et al. 2018).

12
13 Interestingly, MAOs appeared to be most affected by the tested factors in general; if the
14 rejections of low offers can really be deemed to be a marker of altruism (which is a matter of debate –
15 see above), future studies could focus on UG MAO specifically. We encourage further research into the
16 effects of memory, cognitive deliberation, discounting and more traits on the responder behavior in the
17 UG not only in humans, but potentially also in animals where more than remotely feasible (the UG has
18 been adapted in various ways for chimpanzees: see the studies and interpretations of Jensen et al. 2007,
19 Kaiser et al. 2012, Proctor et al. 2013, Milinski 2013, Henrich & Silk 2013, Bueno-Guerra et al. 2019,
20 Sánchez-Amaro et al. 2021, 2024).

21
22 The largely non-significant trends shown by temporal discounting in our study support suggest
23 that people who discount delayed rewards more tend to offer less in both experimental games and
24 accept lower offers. This is in line with our expectations and with the predictions of Triversian direct
25 reciprocity and the life history theory (e.g. Figueredo et al. 2006, 2007). However, we must be careful
26 in the interpretation, since the effect was not statistically significant and the rest of our results did not
27 support the complex preconditions thought necessary for direct reciprocity. On the other hand, the life
28 history theory predicts in a more straightforward manner that people in adverse environments would

1
2
3 resort to “faster” life strategies more suitable in the short term – greater temporal discounting, a
4
5 preference for immediate rewards, less cooperative and altruistic behavior.
6

7
8 Nevertheless, even in this case, we would expect health condition to correlate negatively with
9
10 altruism, and that is not what we observed in the study. We found no connection between altruism and
11
12 either mental or physical health (although the trend in our expected direction in case of physical health
13
14 and UG MAO would merit a more directed follow-up study on a larger sample), unlike a number of
15
16 previous studies that investigated their relationship, hypothesized to run in either direction (Post 2005,
17
18 Figueredo et al. 2006, 2007). One possible explanation is that purely self-reported condition is not a
19
20 good predictor as opposed to more objective measures. However, previous studies such as Figueredo et
21
22 al. (2006, 2007) included subjective measures of health as well. It is more likely that the possible effect
23
24 was too small to discern in our sample.
25
26
27
28

29
30 To follow up on this possibility, our study was potentially limited by being underpowered for
31
32 detecting small, subtle effects. Power analysis (using the “pwr” R package, Champely et al. 2020)
33
34 shows that with our sample of $N=158$, we would have detected effects of a size >0.22 with an 80%
35
36 power (to detect very small effects of ~ 0.1 with the same power, a sample of nearly 800 participants
37
38 would have been needed). Therefore, it would have sufficed to detect medium to strong effects. We
39
40 can thus conclude that if preconditions of Triversian direct reciprocity play any role in human
41
42 interactions, it is likely a small one, coupled with other factors that can influence altruism, such as
43
44 sexual selection. This does not, of course, preclude reciprocity from being an important driving force of
45
46 altruistic behavior; generalized reciprocity doesn’t require complex cognitive processing. Furthermore,
47
48 Schweinfurth & Call (2019) suggest that even in the case of direct reciprocity between two individuals,
49
50 complex cognitive mechanisms need not underlie it if it is based on emotion and attitude rather than
51
52 “calculated” reciprocity. As they note, it is important to better connect the proximate psychological and
53
54 neurological correlates of altruism and cooperation to the ultimate explanations (such as reciprocity
55
56 increasing the survival value of these traits). In this vein, we speculate that said psychological
57
58
59
60
61
62
63
64
65

proximate influences – inequality aversion (Fehr & Schmidt 1999, Bolton & Ockenfels 2000), “warm glow” feelings (Andreoni 1990, Konow 2006) and more – may be too varied to allow for the indirect detection of underlying evolutionary causes such as reciprocity in small samples. How varied would they have been in other species? It might be interesting to adapt this research avenue for animals that showed direct reciprocity, e.g. Norway rats (Kettler et al. 2021, Engelhardt & Taborsky 2022).

To reiterate, the evolution of altruism is a complex issue, and multiple factors likely contributed to it to some extent. Moreover, a “single factor” (itself very complex) such as individual health or memory could be involved in multiple feedbacks going in different directions. Self-reported health condition and memory had no connection to altruistic behavior in our sample, while higher cognitive ability correlated with lower MAOs, interpretable as potential altruistic punishment, but also as a strategic choice. As expected, steeper temporal discounting correlated with lower offers and MAOs, but the effect was not statistically significant. We can therefore conclude that our study did not support a substantial role of Triversian direct reciprocity in humans. Still, we cannot draw any conclusions about other (indirect, attitudinal and especially generalized) forms of reciprocity.

Funding

This work was supported by the grant project GA UK 1502018 of the Charles University’s grant agency, whom we would like to thank.

References

- Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The economic journal*, 100(401), 464-477.
- Ben-Ner, A., Kong, F., & Putterman, L. (2004). Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving. *Journal of Economic Psychology*, 25(5), 581-589.
- Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American economic review*, 91(1), 166-193.
- Buddiga, N. R., & Locey, M. L. (2024). Altruism, reciprocity, and probability: Examining relations through a discounting framework. *Behavioural Processes*, 217, 105024.
- Bueno-Guerra, N., Völter, C. J., de Las Heras, Á., Colell, M., & Call, J. (2019). Bargaining in chimpanzees (*Pan troglodytes*): The effect of cost, amount of gift, reciprocity, and communication. *Journal of Comparative Psychology*, 133(4), 542.
- Carter, G. G., Farine, D. R., Crisp, R. J., Vrtilek, J. K., Ripperger, S. P., & Page, R. A. (2020). Development of new food-sharing relationships in vampire bats. *Current Biology*, 30(7), 1275-1279.
- Casal, S., Fallucchi, F., & Quercia, S. (2019). The role of morals in three-player ultimatum games. *Journal of Economic Psychology*, 70, 67-79.
- Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., ... & De Rosario, M. H. (2018). Package 'pwr'. *R package version*, 1(2).
- Church, A. H. (1993). Estimating the effect of incentives on mail survey response rates: A meta-analysis. *Public Opinion Quarterly*, 57(1), 62-79.
- Cinar, O. & Viechtbauer, W. (2022). The poolr package for combining independent and dependent p values. *Journal of Statistical Software*, 101(1), 1–42. <https://doi.org/10.18637/jss.v101.i01>

- 1
2
3 Corgnet, B., Espín, A. M., & Hernán-González, R. (2015). The cognitive basis of social behavior:
4
5 cognitive reflection overrides antisocial but not always prosocial motives. *Frontiers in*
6
7 *behavioral neuroscience*, 9, 287.
8
9
- 10 Curry, O. S., Price, M. E., Price, J. G. (2008). Patience is a virtue: Cooperative people have lower
11
12 discount rates. *Personality and individual differences*, 44(3), 780-785.
13
14
- 15 Czerwonka, M., Staniszevska, A., & Kompa, K. (2018). Cognitive reflection test in predicting rational
16
17 behavior in the dictator game. In *Problems, methods and tools in experimental and behavioral*
18
19 *economics: Computational methods in experimental economics (CMEE) 2017 Conference* (pp.
20
21 301-312). Springer International Publishing.
22
23
- 24 Diekmann, A. (2004). The Power of Reciprocity Fairness, Reciprocity, and Stakes in Variants of the
25
26 Dictator Game. *Journal of conflict resolution*, 48(4), 487-505.
27
28
- 29 Dolivo, V., Engqvist, L., & Taborsky, M. (2023). When deciding to cooperate by direct reciprocity,
30
31 Norway rats sometimes benefit from olfactory competence and seem not impaired by
32
33 insufficient cognitive abilities. *Animal Cognition*, 26(4), 1119-1130.
34
35
36
- 37 Duffy, S., & Smith, J. (2014). Cognitive load in the multi-player prisoner's dilemma game: Are there
38
39 brains in games?. *Journal of Behavioral and Experimental Economics*, 51, 47-56.
40
41
- 42 Engelhardt, S. C., & Taborsky, M. (2022). Food-exchanging Norway rats apply the direct reciprocity
43
44 decision rule rather than copying by imitation. *Animal Behaviour*, 194, 265-274.
45
46
- 47 Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The quarterly*
48
49 *journal of economics*, 114(3), 817-868.
50
51
- 52 Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, 415(6868), 137-140.
53
- 54 Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and human*
55
56 *behavior*, 25(2), 63-87.
57
58
59
60
61
62
63
64
65

- Figueredo, A. J., Vásquez, G., Brumbach, B. H., Schneider, S. M., Sefcek, J. A., Tal, I. R., Hill, D., Wenner, C., Jacobs, W. J. (2006). Consilience and life history theory: From genes to brain to reproductive strategy. *Developmental Review*, 26(2), 243-275.
- Figueredo, A. J., Vásquez, G., Brumbach, B. H., & Schneider, S. M. (2007). The K-factor, covitality, and personality: A psychometric test of life history theory. *Human Nature*, 18, 47-73.
- Forsythe, R., Horowitz, J., Savin, N.S., Sefton, M. (1994) Fairness in simple bargaining games. *Games and Economic Behavior* 6: 347–369.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4), 25-42.
- Gardner, A., & West, S. A. (2010). Greenbeards. *Evolution*, 64(1), 25-38.
- Güth W, Schmittberger R, Schwarze B (1982) An experimental analysis of ultimatum bargaining. *J Econ Behav Organ* 3: 367–388.
- Henrich, J., & Silk, J. B. (2013). Interpretative problems with chimpanzee ultimatum game. *Proceedings of the National Academy of Sciences*, 110(33), E3049-E3049.
- Hui, B. P., Ng, J. C., Berzaghi, E., Cunningham-Amos, L. A., & Kogan, A. (2020). Rewards of kindness? A meta-analysis of the link between prosociality and well-being. *Psychological Bulletin*, 146(12), 1084.
- Jensen, K., Call, J., & Tomasello, M. (2007). Chimpanzees are rational maximizers in an ultimatum game. *science*, 318(5847), 107-109.
- Jonason, P. K., Baughman, H. M., Carter, G. L., & Parker, P. (2015). Dorian Gray without his portrait: Psychological, social, and physical health costs associated with the Dark Triad. *Personality and Individual Differences*, 78, 5-13.
- Kaiser, I., Jensen, K., Call, J., & Tomasello, M. (2012). Theft in an ultimatum game: Chimpanzees and bonobos are insensitive to unfairness. *Biology Letters*, 8(6), 942-945.

- 1
2
3 Kettler, N., Schweinfurth, M. K., & Taborsky, M. (2021). Rats show direct reciprocity when interacting
4
5 with multiple partners. *Scientific Reports*, 11(1), 3228.
6
7
8 Konow, J. (2006). Mixed feelings: Theories and evidence of warm glow and altruism. *Available at*
9
10 *SSRN 980349*.
11
12 Kroneisen, M., & Bell, R. (2022). Memory as a cognitive requirement for reciprocal
13
14 cooperation. *Current Opinion in Psychology*, 43, 271-277.
15
16
17 Larose, K., Dubois, F. (2011). Constraints on the evolution of reciprocity: An experimental test with
18
19 zebra finches. *Ethology*, 117(2), 115-123.
20
21
22 Milinski, M., Wedekind, C. (1998). Working memory constrains human cooperation in the Prisoner's
23
24 Dilemma. *Proceedings of the National Academy of Sciences*, 95(23), 13755-13758.
25
26
27 Milinski, M. (2013). Chimps play fair in the ultimatum game. *Proceedings of the National Academy of*
28
29 *Sciences*, 110(6), 1978-1979.
30
31
32 Moreira, J., Vukov, J., Sousa, C., Santos, F. C., d'Almeida, A. F., Santos, M. D., & Pacheco, J. M.
33
34 (2013). Individual memory and the emergence of cooperation. *Animal Behaviour*, 85(1), 233-
35
36 239.
37
38
39 Murty, V. P., FeldmanHall, O., Hunter, L. E., Phelps, E. A., & Davachi, L. (2016). Episodic memories
40
41 predict adaptive value-based decision-making. *Journal of Experimental Psychology:*
42
43 *General*, 145(5), 548.
44
45
46 Novakova, J., Machová, K., Sýkorová, K., Zíka, V., & Flegr, J. (2021). Looking Like a Million Dollars:
47
48 Does Attractiveness Priming Increase Altruistic Behavior in Experimental Games?. *Frontiers in*
49
50 *psychology*, 1974.
51
52
53 Novakova, J.; Flegr, J.; Machová, K.; Zíka, V. (2025). Testing-Trivers_Novakova-et-al.csv. figshare.
54
55 Dataset. <https://doi.org/10.6084/m9.figshare.28292240.v1>
56
57
58 Osiński, J., Karbowski, A., Rusek, J., & Reinholz, A. (2021). The effect of delay and social distance in
59
60 the dictator and ultimatum games. *Behavioural Processes*, 193, 104513.
61
62
63
64
65

- Persson, E., Tinghög, G., & Västfjäll, D. (2024). Intertemporal prosocial behavior: a review and research agenda. *Frontiers in Psychology*, 15, 1359447.
- Ponti, G., & Rodriguez-Lara, I. (2015). Social preferences and cognitive reflection: evidence from a dictator game experiment. *Frontiers in behavioral neuroscience*, 9, 146.
- Post, S. G. (2005). Altruism, happiness, and health: It's good to be good. *International journal of behavioral medicine*, 12(2), 66-77.
- Proctor, D., Williamson, R. A., de Waal, F. B., & Brosnan, S. F. (2013). Chimpanzees play the ultimatum game. *Proceedings of the National Academy of Sciences*, 110(6), 2070-2075.
- Rutte, C., & Taborsky, M. (2007). Generalized reciprocity in rats. *PLoS biology*, 5(7), e196.
- Rutte, C., & Taborsky, M. (2008). The influence of social experience on cooperative behaviour of rats (*Rattus norvegicus*): direct vs generalised reciprocity. *Behavioral Ecology and Sociobiology*, 62(4), 499-505.
- Sánchez-Amaro, A., & Rossano, F. (2021). Chimpanzees and bonobos use social leverage in an ultimatum game. *Proceedings of the Royal Society B*, 288(1962), 20211937.
- Sánchez-Amaro, A., Maurits, L., & Haun, D. B. (2024). Chimpanzees engage in competitive altruism in a triadic ultimatum game. *Scientific Reports*, 14(1), 3393.
- Schino, G., & Aureli, F. (2008). Grooming reciprocation among female primates: a meta-analysis. *Biology Letters*, 4(1), 9-11.
- Schino, G., & Aureli, F. (2009). Reciprocal altruism in primates: partner choice, cognition, and emotions. *Advances in the Study of Behavior*, 39, 45-69.
- Soutschek, A., & Schubert, T. (2016). The importance of working memory updating in the Prisoner's dilemma. *Psychological Research*, 80, 172-180.
- Stephens, D. W., McLinn, C. M., & Stevens, J. R. (2002). Discounting and reciprocity in an iterated prisoner's dilemma. *Science*, 298(5601), 2216-2218.

- 1
2
3 Stevens, J. R., Hauser, M. D. (2004). Why be nice? Psychological constraints on the evolution of
4
5 cooperation. *Trends in cognitive sciences*, 8(2), 60-65.
6
7
8 Stouffer, S. A., Suchman, E. A., DeViney, L. C., Star, S. A., & Williams, R. M., Jr. (1949). *The*
9
10 *American Soldier: Adjustment During Army Life (Vol. 1)*. Princeton, NJ: Princeton University
11
12 Press.
13
14
15 Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly review of biology*, 46(1), 35-57.
16
17
18 Wang, T., Chen, Z., Li, K., Deng, X., & Li, D. (2014). Memory does not necessarily promote
19
20 cooperation in dilemma games. *Physica A: Statistical Mechanics and its Applications*, 395, 218-
21
22 227.
23
24
25 Wilkinson, G. S. (1984). Reciprocal food sharing in the vampire bat. *Nature*, 308(5955), 181-184.
26
27
28 Wilkinson, G. S. (1988). Reciprocal altruism in bats and other mammals. *Ethology and Sociobiology*,
29
30 9(2), 85-100.
31
32
33 Wilkinson, G. S., Carter, G. G., Bohn, K. M., & Adams, D. M. (2016). Non-kin cooperation in bats.
34
35 *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1687), 20150095.
36
37
38 Winke, T., & Stevens, J. R. (2017). Is cooperative memory special? The role of costly errors, context,
39
40 and social network size when remembering cooperative actions. *Frontiers in Robotics and AI*, 4,
41
42 52.
43
44
45 Yi, R., Charlton, S., Porter, C., Carter, A. E., & Bickel, W. K. (2011). Future altruism: Social
46
47 discounting of delayed rewards. *Behavioural processes*, 86(1), 160-163.
48
49
50 Zahavi, A. (1995). Altruism as a handicap: the limitations of kin selection and reciprocity. *Journal of*
51
52 *Avian Biology*, 26(1), 1-3.
53
54
55
56
57
58
59
60
61
62
63
64
65

Evolution and Human Behavior

Testing the microbiome-altruism hypothesis: Antibiotic intake doesn't impair altruistic behavior --Manuscript Draft--

Manuscript Number:	
Article Type:	Research Report
Keywords:	altruism; antibiotics; cooperation; manipulation hypothesis; microbiome
Corresponding Author:	Julie Nekola Novakova, MSc. Charles University Faculty of Science Prague, CZECH REPUBLIC
First Author:	Julie Nekola Novakova, MSc.
Order of Authors:	Julie Nekola Novakova, MSc.
	Jaroslav Flegr, Professor, Ph.D.
Abstract:	Microbiomes influence both our physiology and behavior to a great extent. It is therefore not surprising that microbial infections have been hypothesized as an important driving force of evolution of altruism by manipulating their hosts toward altruistic behavior (thus increasing the chances of transmission from one host to another). We have tested this hypothesis, proposed by Lewin-Epstein and collective, first on data from a sample of 604 university student participants concerning both altruistic behavior (with the Dictator Game as a proxy) and antibiotics intake, since antibiotics have been proposed by Lewin-Epstein et al. to hamper altruism if their hypothesis is correct, and then in a follow-up online study containing 8,427 responses from the general public. We found no relationship of antibiotics intake (either recently or in childhood development) and altruistic behavior in experimental games and an opposite-than-expected relationship of self-reported antibiotics intake in childhood with real-life helping, suggesting that while microbiomes remain among vitally important drivers of our behavior, microbial infections likely weren't an important evolutionary influence in the evolution of the complex trait of altruism.
Opposed Reviewers:	

Testing the microbiome-altruism hypothesis: Antibiotic intake doesn't impair altruistic behavior

Julie Novakova^{1*}¶, Jaroslav Flegr^{1¶}

¹ Laboratory of Evolutionary Biology, Department of Philosophy and History of Science, Faculty of Science, Charles University, Prague, Czech Republic

* Corresponding author

E-mail: julie.novakova@natur.cuni.cz

Address: Faculty of Science, Viničná 7, 128 00 Prague 2, Czech Republic

¶ These authors contributed equally to this work.

Acknowledgements

We want to thank Radim Kuba, Robin Kopecký, Lenka Jelena Příplatová, Martin Hůla, Jana Ullmann and Kateřina Šefrová for their help with the data collection in classrooms (Study 1).

Declaration of interest

The authors have no competing interests to declare.

Wordcount: approx. 3600 words (4220 including figure and table captions) excluding references and abstract

Testing the microbiome-altruism hypothesis: Antibiotic intake doesn't impair altruistic behavior

Abstract

Microbiomes influence both our physiology and behavior to a great extent. It is therefore not surprising that microbial infections have been hypothesized as an important driving force of evolution of altruism by manipulating their hosts toward altruistic behavior (thus increasing the chances of transmission from one host to another). We have tested this hypothesis, proposed by Lewin-Epstein and collective, first on data from a sample of 604 university student participants concerning both altruistic behavior (with the Dictator Game as a proxy) and antibiotics intake, since antibiotics have been proposed by Lewin-Epstein et al. to hamper altruism if their hypothesis is correct, and then in a follow-up online study containing 8,427 responses from the general public. We found no relationship of antibiotics intake (either recently or in childhood development) and altruistic behavior in experimental games and an opposite-than-expected relationship of self-reported antibiotics intake in childhood with real-life helping, suggesting that while microbiomes remain among vitally important drivers of our behavior, microbial infections likely weren't an important evolutionary influence in the evolution of the complex trait of altruism.

Keywords: altruism; antibiotics; cooperation; manipulation hypothesis; microbiome

Introduction

From anxiety, unipolar depression and response to stress to schizophrenia or eating disorders, microbiomes seem to have an influence on not just our physical health, but also many aspects of mental health and behavior (see Bailey & Cryan (2017) for a review). It is therefore not surprising that Lewin-Epstein et al. (2017) proposed a hypothesis that microbes could manipulate their hosts toward altruistic behavior, bringing the microbes an evolutionary advantage due to an easier transmission from one host to another. The connection of one's health and altruistic behavior is important to study from the point of view of their work, but also Trivers' (1971) theory of reciprocal altruism (where longer lifetimes, requiring good health, increase the chances of repeated interactions and thus altruistic behavior), or, on the other hand, behavioral medicine and psychology (where the opposite direction of the relationship has been presumed – see e.g. Post (2005)). Other mechanisms such as kin selection, sexual selection and interdemic selection likely contributed to the evolution of altruism as well (see Fig. 1 below), and the extent of their influence is being studied and debated.

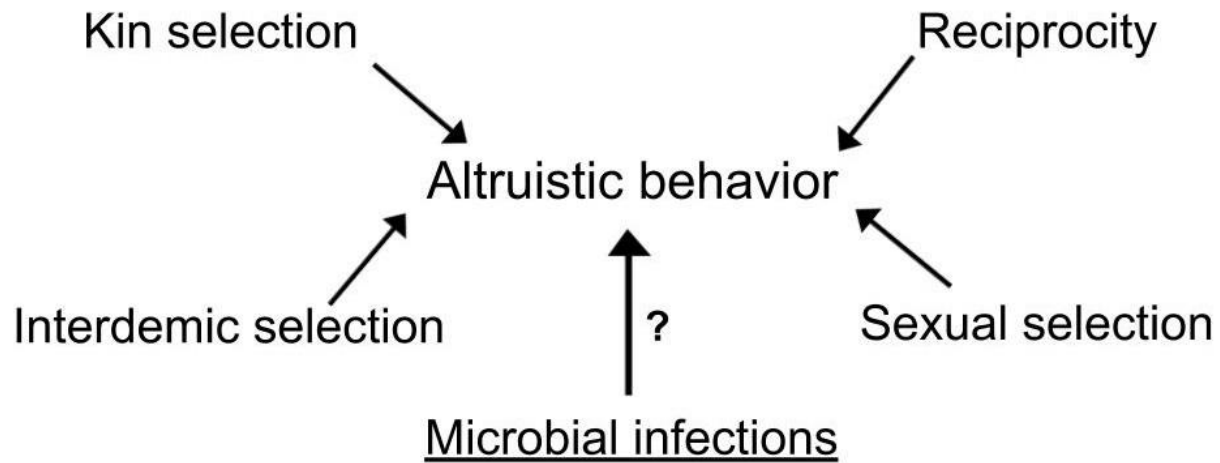


Fig 1. Suggested evolutionary pathways to altruism. This diagram presents an overview of the likely major processes contributing to altruism, including the possibility of microbial infections, tested in the current paper.

The advantage of Lewin-Epstein et al. (2017, further expanded in 2020) lies chiefly in the simplicity of their hypothesis, in which the evolution of altruism doesn't require complex cognitive processes such as memory of previous interactions with other agents, numerical quantification, delayed gratification or decreased temporal discounting. The authors proposed testing their hypothesis by investigating whether taking antibiotics (further abbreviated ATB), leading to the elimination of normal microbiomes, hampers altruism. It would not be unprecedented: broad-spectrum ATB consumption had e.g. a considerable effect on aggression – albeit not on other tested categories of behavior – in hamsters (Sylvia et al. 2017). In animal studies, it would also be possible to distinguish effects of ATB themselves and microbiome disruption by administering ATB to germ-free animals.

Testing the Lewin-Epstein hypothesis in humans is a more complicated process, since we cannot test the effect of ATB alone or usually ethically administer ATB to healthy people.

1
2
3
4 Nevertheless, it is vital not to forgo searching for their (direct or indirect, microbiome-modulated)
5
6 effect on altruism in humans and to attempt to, as the most plausible means of testing the
7
8 hypothesis, compare groups of people who had or hadn't taken ATB for medical reasons, whether
9
10 recently (with potential short-term effects on behavior) or in their childhood development (with
11
12 effects on altruistic behavior, formed then, potentially persisting throughout life).
13
14

15
16 In two related studies (one on university students and one on the general public), we tested
17
18 whether there is any relationship of ATB consumption and altruism in the Dictator Game (DG)
19
20 (Forsythe 1994) in Study 1, and the DG and Trust Game (TG) (Berg et al. 1995) plus a survey of
21
22 altruistic behavior in case of Study 2 (both games are described in Methods). Both games are
23
24 widely used to assess altruistic preferences. As far as we are aware, only a single study so far
25
26 looked into the relationship of ATB (in this case specifically minocycline, a known inhibitor of
27
28 microglial activation) and behavior in the DG and TG: a small-scale controlled ATB
29
30 administration study by Watabe et al. (2011), who found no difference in altruism in the DG
31
32 between the minocycline and placebo groups, but found a possible modulating effect of
33
34 minocycline on trust in others. The rationale for the study was the direct effect of the specific ATB
35
36 on decision-making, not an indirect effect through changes in the microbiome. To our knowledge,
37
38 the present study is the first of its kind.
39
40
41
42
43
44
45
46
47
48
49
50

51 **Methods**

52 **Study 1**

The data in Study 1 was originally collected to explore the connection of health and altruism in light of requirements for Trivers' (1971) reciprocal altruism, but it presented a unique opportunity to test the Lewin-Epstein et al. (2017) hypothesis. The dataset was collected over the years 2011 to 2018 from students of several evolutionary biology and scientific methodology classes (complete N=2,345; 73% F, 27% M; age_{mean}=21.4, age_{median}=21), consenting by filling the questionnaire as stated at its beginning, and the collection was approved by [anonymized]'s IRB.

The students filled an anonymous questionnaire about their health, including the number of times they were taking ATB during the last year, or three years respectively. They also played a hypothetical DG (in which Player 1 is endowed with an amount of money that can be shared with Player 2, whose role in the – typically one-off and anonymous – game is passive; the mean contribution is approximately 28% (Engel 2011)). The hypothetical amount at stake varied by orders of magnitude from 20 CZK to 200,000 CZK. The number of valid responses (= participants who filled both questions about ATB intake and the hypothetical DG, and whose answers were valid – in case of ATB, in integers and with the 1-year intake not exceeding the 3-year intake for the same person, and in case of the DG, corresponding to the amount at stake) was N=604 (73.4% F, 26.6% M; age_{mean}=21.3, age_{median}=21). This dropout rate is predominantly due to many participants leaving one or more questions blank.

We have investigated the connection of ATB use and altruistic giving in the hypothetical Dictator Game (DG) by a nonparametric partial Kendall correlation (R package “ppcor”, Kim 2015), because DG contributions had a significantly non-normal distribution (Kolmogorov-Smirnov normality test: D=0.5, p<0.0001). All analyses were conducted using R version 4.4.2 (R Core Team 2021). The data is available in S1 Dataset_Study1.

Study 2

The data was collected specifically to test the potential relationship between recent ATB intake and altruism, including potential effects from childhood development. In an online Qualtrics survey, respondents were asked to state whether they had taken antibiotics within the last three months. This was followed by a short anamnestic survey including questions about health and ATB intake in childhood, and then a hypothetical DG and TG. In the TG, Player 1 is also endowed with money in the experiment and can share it with Player 2. The shared amount is doubled. Player 2 can then decide to send some part of the doubled amount back to Player 1 (this “transaction” is not doubled). Here, participants were asked to imagine playing round two: how much would they send back to Player 1 if that player sent them 150 out of the total 200 CZK, leaving them with 300 CZK as the amount doubles when reaching Player 2. In addition, the survey included questions assessing the participants’ real-life altruistic behavior, reactions in hypothetical altruistic self-sacrifice situations and close personal contact with others (initiating encounters, proximity avoidance, physical personal boundaries), and further questions about the participants’ health as a control. For the purposes of exploratory analysis, questions about recent intake of antivirals and antiparasitic drugs were also included. The questions are listed in S2 Survey_questions_Study2. An index variable “behavioral altruism” was computed as averaged z-scores of the survey questions regarding real-life and hypothetical altruistic self-sacrifice situations, listed below:

1. *Do you ever donate to charities or otherwise help those in need?* (7-point Likert scale)
2. *Approximately what percentage of your net income did you contribute to charitable causes last year (20 means 20 or more)?* (slider scale 0-20)

3. *Outside of events such as birthday parties or similar events, do you share your possessions with friends (e.g. invite them to meals or to the movies or the theater; give them gifts, etc.) or maybe help them through your effort? (7-point Likert scale)*
4. *Did you share with friends as a child? (7-point Likert scale)*
5. *Have you shared your possessions or efforts with friends, contributed to charity, or otherwise helped those in need (e.g. volunteering) in the last three months? (4-point scale)*
6. *Would you risk your life to save two unknown peers (50% is the risk of dying in the rescue)? (6-point Likert scale)*
7. *Would you risk your life to save 1,000 unknown peers (50% is the risk of dying in the rescue)? (6-point Likert scale)*

The data for this study was being collected from April 2018 to November 2023. Only participants who had stated age of at least 15, took part in both experimental games and completed at least 90% of all questions were included in the analysis. The number of responses, after excluding some by the criteria above, was N=8,427 (65.3% F, 34% M, >1% undisclosed or other; $age_{mean}=37$, $age_{median}=35$). The relationship of ATB intake (recent and in childhood) and behavior in the DG, respectively TG, and the joint variable “behavioral altruism”, was investigated by a nonparametric partial Kendall correlation (R package “ppcor”, Kim 2015), because both DG and TG contributions had a significantly non-normal distribution (Kolmogorov-Smirnov normality test, DG: $D=0.50$, $p<0.0001$; TG: $D=0.51$, $p<0.0001$). Age and sex were included as covariates to control for their effects

Afterwards, an exploratory analysis was conducted in order to look at potential effects of mental and physical health, effects of non-ATB drugs with a potential to affect other parts of the human microbiome, and finally effects of these variables and ATB intake on preferred proximity

1
2
3
4 and initiation of interpersonal contact (as proxies of transmission chances). Since this part of the
5
6 study was purely exploratory, a correction for multiple tests would not have been advisable (see
7
8 Althouse 2016), but it must be firmly noted that any effects observed there should be taken with
9
10 caution and independently tested in later studies.
11
12

13 All analyses were conducted using R version 4.4.2 (R Core Team 2021). The data used in
14
15 the analyses is available in S3 Dataset_Study2. The study was approved by [anonymized]'s IRB.
16
17
18
19
20

21 Results

22 Study 1

23
24
25
26
27 The mean DG allocation was 36.5% of the pie, varying between 20.6% and 42% depending on the
28
29 hypothetical amount at stake. Women gave on average 35.7%, men 38.5%; the difference was not
30
31 statistically significant (Wilcoxon rank sum test: $W=33214$, $p=0.25$, Cohen's $d=0.11$). On average,
32
33 participants took ATB 0.76 times in the past year (56% participants used none, most of the
34
35 remaining 44% used them once), and 1.9 times in the past three years (only 27% not at all, 23%
36
37 once, the rest more frequently).
38
39
40

41 The frequency of ATB use in the past three years had no significant relationship with
42
43 contributions in the DG ($p=0.49$). ATB intake in the preceding one year had a small positive
44
45 correlation with DG contributions (Kendall's Tau=0.061, Cohen's $f^2=0.097$, $p=0.024$). For a visual
46
47 representation of the results, see Fig 2 below.
48
49
50
51
52
53
54
55
56
57

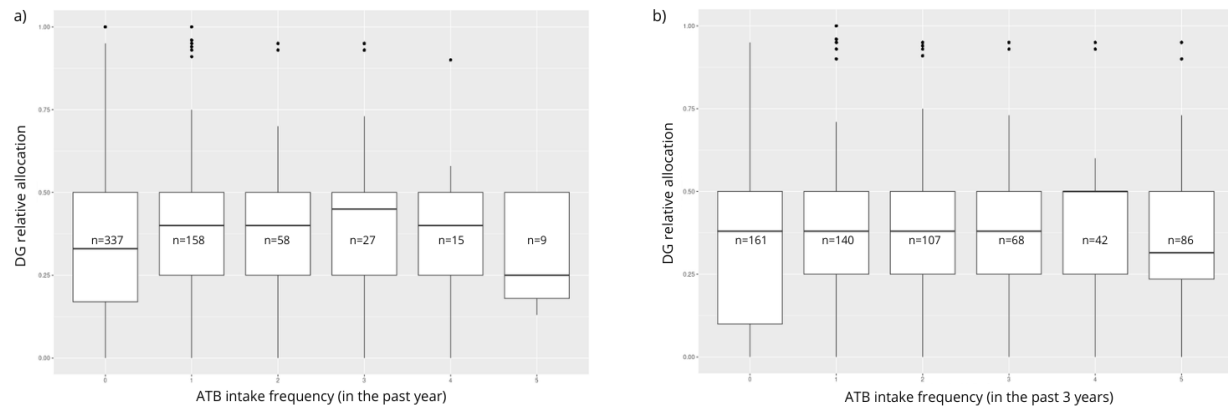


Fig 2. Antibiotic intake and behavior in the DG. The x-axis in both plots shows the frequency of antibiotics use, the y-axis represents the relative allocation in the DG. Fig. 2a (on the left) represents data for antibiotics in the past year, Fig. 2b (on the right) for the past three years. The line within boxplots shows the median value, while the lower and upper hinges correspond to the first and third quartiles, and upper whisker extends from the hinge to the highest value up to 1.5*IQR (inter-quartile range and the lower whisker from the hinge to the lowest value up to 1.5*IQR. Outliers are plotted individually. From the overlap of the boxplots and proximity of their medians in each figure, one can see that there are no meaningful differences or trends in the relationship of the frequency of ATB intake in the past three altruistic behavior in the DG (on the right), and a possible weak positive effect (visually discernible from levels 0 to 3; note the sharper drop for level 5) in case of ATB intake in the preceding year (left).

Study 2

Descriptive statistics

The mean DG allocation was 36.9% of the pie. Women contributed more on average, 38.5%, while men's average contribution was 34%; the difference was significant (Wilcoxon rank sum test: $W=8848847$, $p<0.0001$, Cohen's $d=-0.23$). In the TG, on average 46.8% of the received amount was sent to Player 1, and men sent back more than women (men 50.1% and women 45.1%; Wilcoxon rank sum test: $W=6428266$, $p<0.0001$, Cohen's $d=0.26$).

The proportion of participants who had taken ATB within the three months preceding their participation in the survey was 14.6%. As to ATB intake in childhood, 34.2% reported almost never receiving ATB, 38.5% receiving them once or twice a year, 13.5% three to five times a year, and 5.2% more frequently; 8.6% did not know.

Hypothesis testing

Testing the relationship of ATB intake in the past three months and in childhood, respectively, with the variables to assess altruism (DG contributions, TG contributions, behavioral altruism) yielded one significant result (both pre- and post-Bonferroni correction): childhood ATB consumption correlated positively behavioral altruism (see Figure 4 for its visualisation). The full results are summarized in Table 1 below; in addition, results from the experimental games are visualised in Figure 3.

		ATB intake in childhood	ATB intake in the past 3 months
DG contributions	Kendall's Tau	-0.003	-0.002
	p value	0.736	0.777
	p value (post-Bonferroni corr.)	1	1
	Cohen's f ²	0.004	0.003
TG contributions	Kendall's Tau	0.011	0.003
	p value	0.152	0.631
	p value (post-Bonferroni corr.)	0.910	1
	Cohen's f ²	0.017	0.073
Behavioral altruism	Kendall's Tau	0.047	0.006
	p value	>0.0001	0.441
	p value (post-Bonferroni corr.)	>0.0001	1
	Cohen's f ²	0.073	0.009

Table 1. Testing the hypothesis of ATB intake's effect on altruism. Results from partial Kendall correlation of ATB intake (in childhood and in the past three months) and altruism measures (DG, TG, joint variable behavioral altruism),

with sex and age as covariates. The weak trend between childhood ATB intake and behavioral altruism is highlighted in bold italics.

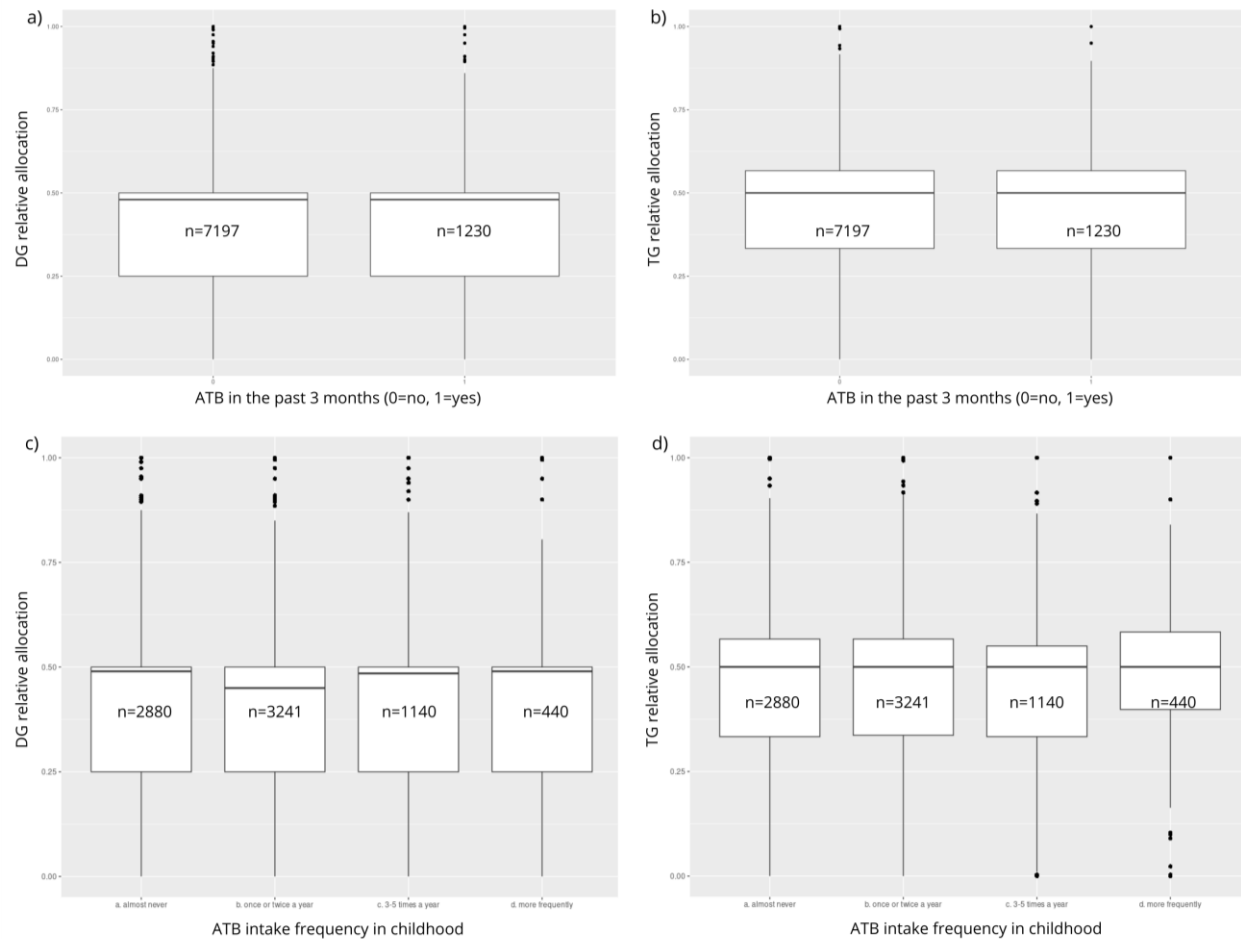


Fig. 3. ATB intake and behavior in the DG and TG. Figure 3a) shows relative DG contributions of participants who did not take any ATB in the past 3 months (left boxplot, denoted 0) and who ingested ATB in that time period (right boxplot, denoted 1). Figure 3b) shows the same relationship for TG contributions. Figures 3c) and 3d) show DG and TG contributions, respectively, for different levels of self-reported ATB intake frequency in childhood (left to right: almost never; once or twice a year; 3-5 times a year; more frequently). See Fig. 2 description for a legend to the boxplots. From the overlap of the boxes and proximity of their medians in each respective figure, one can see that

there are no notable differences or even trends in the relationship of ATB intake (whether recent or in early life) and altruism as measured through the proxy of experimental DG and TG.

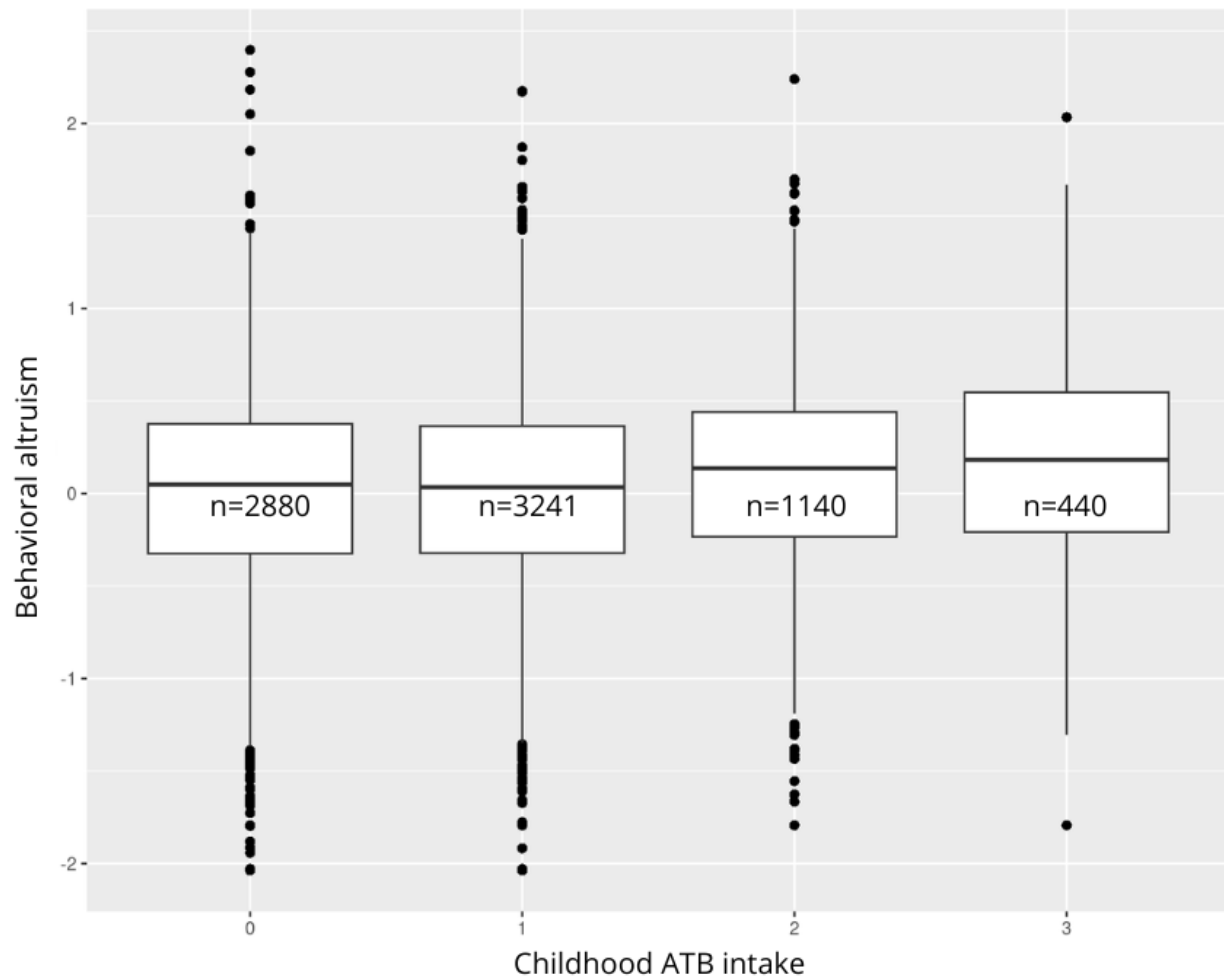


Fig. 4. Childhood ATB intake and behavioral altruism. The boxplot shows the joint variable behavioral altruism divided by categories of the frequency of childhood intake of ATB. One can see that behavioral altruism increases a little with the rising frequency of ATB consumption in childhood. See Fig. 2 description for a legend to the boxplot.

Exploratory analysis

Finally, to gain a better understanding of all the factors that could have shaped altruistic preferences in our sample, ran an exploratory analysis to find potential connections of ATB

consumption, health and behavior beyond the tested hypothesis. Partial Kendall correlation, again with age and sex as covariates, was ran with the three altruism-signifying variables (DG contributions, TG contributions, behavioral altruism), reported tendency to initiate interpersonal encounters, proximity avoidance and preferred physical personal boundary as dependent variables, and sex, age, recent ATB consumption, childhood ATB consumption, recent antivirotics consumption, recent prescribed and unprescribed antiparasitics consumption (divided this way, since antibiotics and antivirotics are prescription-only in the country of study, while some antiparasitics are not), frequency of childhood infections, physical and mental condition and type of birth (vaginal or cesarean; asked because of its role in early microbiome formation) as independent variables. Full results of the exploratory analysis, with Kendall's Tau, Cohen's f^2 , and p values, can be found in supplement S4 Table-exploratory-analysis_Study2. Below, we only report the main findings.

Unsurprisingly, sex, age, mental and physical condition turned out to be significant predictors for most of the dependent variables. Age and sex displayed a similar pattern, where higher age and being male was associated with lower proximity avoidance, behavioral altruism and DG contributions, and larger TG contributions. Better mental condition correlated positively with initiating encounters and behavioral altruism, and negatively with proximity avoidance and width of a preferred personal boundary; physical condition showed the exact same pattern, only to a smaller extent.

Recent antivirotics consumption correlated positively with proximity avoidance, but nothing else; recent unprescribed antiparasitics consumption correlated negatively with DG contributions, and positively with behavioral altruism and initiating encounters; finally, cesarean

birth correlated with lower preferred personal boundaries. It is a question whether these correlations, relatively isolated with no clear cross-variable pattern and very small effect sizes, represent real-life effects, or are merely accidental.

Interestingly, tendency to initiate interpersonal encounters significantly positively correlated with both recent ATB intake (Kendall's $\tau=0.023$, Cohen's $f^2=0.036$, $p=0.002$) and childhood ATB consumption (Kendall's $\tau=0.030$, Cohen's $f^2=0.047$, $p<0.0001$). The relationship is visualized below in Figure 5.

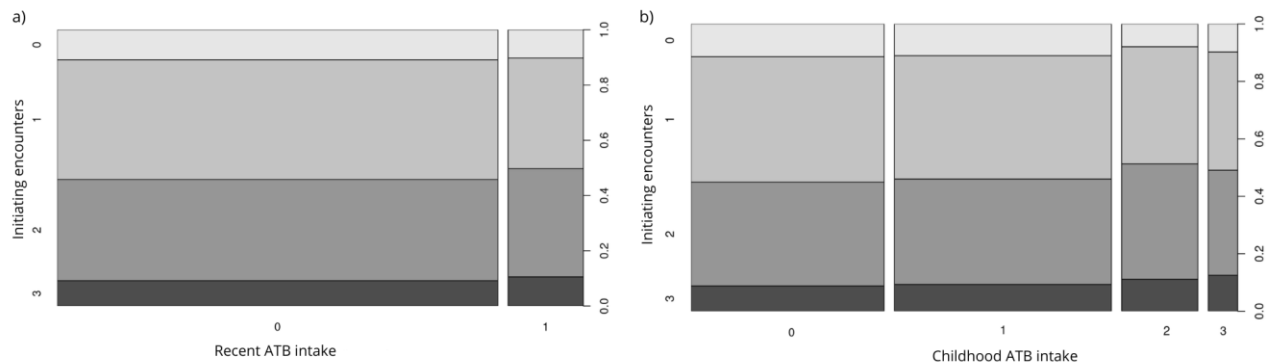


Fig. 5. Childhood ATB intake and initiating interpersonal encounters. The spineplots depict the relative sizes of different combinations of self-reported tendency to initiate interpersonal encounters (from top to bottom: 0, meaning almost never, to 3, meaning often) and recent (Fig. 5a)) or childhood (Fig. 5b)) ATB intake (yes/no in case of recent intake, increasing frequency in case of childhood intake). It can be seen that the relative sizes of more frequent initiation of new encounters typically rise very slightly with the increasing frequency of ATB intake; while statistically significant, the effect does not appear to be practically meaningful.

All effect sizes fell under the “small” category, none having exceeded an effect size of 0.25 and most being notably smaller; any practical impacts of the discerned relationships are therefore likely very limited.

Discussion

Neither of our studies revealed any meaningful impact of recent ATB intake on altruism (with experimental games as proxies, as well as questions on real-life helping and hypothetical situations in case of Study 2), suggesting that the “microbiome-induced altruism” hypothesis probably does not play a significant role in impacting altruistic behavior throughout the course of human life.

However, self-reported childhood ATB intake in Study 2, while having no effect on behavior in the DG and TG, proved to have a non-significant, but notable *positive* relationship with behavioral altruism. In addition, both recent and childhood ATB consumption correlated significantly positively with self-reported tendency to initiate interpersonal encounters. These findings run against the expected direction of influence if the “microbial-induced altruism” hypothesis was correct. We suspect that if it reflects anything, it would be an increased prosociality of people who tended to be frequently sick as children (and some up to the present day) and who benefitted from a larger social network. Nevertheless, that is only one possible interpretation that should be taken with a grain of salt without any further supporting data from an identical sample. As Wilson and Csikszentmihalyi (2007) point out, *context* is key in investigating the relationships of health and altruism, and “*given the right conditions*, altruism can be very good for the altruist” (emphasis by the original authors).

Of course, a pair of studies cannot be used to completely rule out the hypothesis of Lewin-Epstein et al. (2017) even in the specific case of human altruism. Examining this phenomenon in a general human population is made difficult by the interaction of the many factors that likely have

1
2
3
4 influenced and keep influencing human altruism both over the evolutionary timescale and within
5
6 the scope of life of an individual. One way to attempt to overcome this obstacle is to use samples
7
8 whose diversity is limited to minimum. This approach was chosen by Watabe et al. (2011), who
9
10 only recruited young, healthy, college-educated Japanese men for their minocycline administration
11
12 study, and to some extent by us in Study 1 (student population, but of both sexes). In this case, due
13
14 to the possibility to only test a small sample, it made sense to limit its diversity, but one can
15
16 overlook effects more pronounced in other parts of the general population. In a general-population
17
18 study with its many confounders to the potentially small hypothesized effect, striving for large
19
20 samples makes best sense (which was our approach in Study 2).
21
22
23
24

25
26 One possible limitation of both of our studies is the use of “artificial” games to measure
27
28 altruism, and the lack of real incentives for the participants; on the other hand, it was
29
30 complemented by a survey of altruism in real life and imagined lifelike situations (dubbed
31
32 “behavioral altruism”) in our Study 2. Moreover, experimental games have been regularly used in
33
34 a number of studies investigating altruistic behavior and social preferences, and a wide meta-study
35
36 by Engel (2011) found no effects of using real vs. hypothetical stakes in the DG. Nor did our
37
38 earlier experiments using the DG in both real-incentive [anonymized] and hypothetical settings
39
40 [anonymized] substantially deviate from each other and similar experiments in other western
41
42 countries. A TG meta-analysis by Johnson & Mislin (2011) showed that playing against a real
43
44 person (unlike in our hypothetical situation) increases the relative amounts sent in either step of
45
46 the game; here, a simulated game may correspond to real-life behavior less closely. Despite that,
47
48 our TG contributions fall well within the typical behavior summarized in the meta-analysis.
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 A potential limitation of Study 1 (addressed in Study 2) is that the hypothesized effects of
5
6 antibiotics consumption on possible altruism-supporting microbes may not last long, and therefore
7
8 any effects might be measurable on the scale of months or less rather than years. The reported ATB
9
10 intake in the past three months, however, did not prove to have any connection with altruistic
11
12 behavior in the experimental games or with behavioral altruism, and only correlated positively
13
14 with the tendency to initiate interpersonal encounters.
15
16

17
18 In case the influence suggested by Lewin-Epstein et al. (2017) was neurodevelopmental
19
20 rather than acting throughout the whole life, we would not discover an effect, and for this, it was
21
22 useful to look also at effects of antibiotics in childhood. However, we did not find any effect of
23
24 early-life ATB intake on present-day altruism in the DG and TG either, and its connection with
25
26 behavioral altruism ran in the opposite direction than the tested model would suggest.
27
28

29
30 Another potentially useful direction (difficult to implement in humans, though) to test the
31
32 theory would be to compare populations who received antibiotic treatment prenatally or in early
33
34 postnatal development, and who did not. In mice, early-life penicillin intake led to shifts in
35
36 microbiota, behavioral and neurochemical changes (Leclercq et al. 2017).
37
38

39
40 In animals such as mice, such research avenues can be pursued with more rigor, of course.
41
42 There, we can possibly have a genetically very uniform strain with a relatively uniform
43
44 microbiome, grown and kept in precisely known laboratory conditions, with experimental and
45
46 control groups to whom ATB were or weren't administered in precisely known doses at precisely
47
48 known intervals. However, in humans, we do have a complex, real-life population, where
49
50 practically significant effects should be discernible despite all confounding influences with large
51
52 enough samples and careful enough methodology. Our findings tend to run against practical
53
54
55
56
57

1
2
3
4 importance of the “microbial altruism” hypothesis in humans; however, we encourage future
5
6 studies aiming to discern potential effects of different types of ATB or its effects within weeks of
7
8 intake and also looking in greater detail at the effects of ATB consumption in childhood.
9

10
11 While models are always necessarily simplified, they should be able to give us practically
12
13 useful insights into the above-described more complex real-world conditions. In light of the
14
15 current results, we suspect that the model presented by Lewin-Epstein et al. (2017) is probably *too*
16
17 simplified to explain a substantial part of altruistic behavior in humans, which can be influenced
18
19 by a number of factors. In principle, the model most of all demonstrates that if genes can be
20
21 transmitted horizontally, they can by more ‘selfish’ (Dawkins 1982). For example, they can induce
22
23 a type of behavior (including altruistic behavior) that increases the rate of their own future
24
25 transmission, while the same behavior reduces the probability of future transmission of all other
26
27 genes of their carriers by reducing the carriers’ direct fitness. In this respect, its explanatory power
28
29 is similar, for example, to the predictions of cultural evolution models. However, where cultural
30
31 evolution research tries to grasp the complexity of agent-based networks that can represent cultural
32
33 traits or infectious agents (Axelrod 1997, Smith & Conrey 2007, Rorabaugh 2014), the tested
34
35 “infection-modulated altruism” model perhaps remains oversimplified to give unique enough
36
37 predictions.
38
39
40
41
42
43
44
45
46
47
48
49
50

51 **Conclusion**

52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4 It is beyond doubt that our microbiomes have a substantial impact on all aspects of our health and
5
6 behavior (Flegr 2025, Johnson & Foster 2018, Bailey & Cryan 2017, Cryan & Dinan 2012, Archie
7
8 & Tung 2015), however, a simple and straightforward increase in altruism with microbial
9
10 colonization seems unlikely. We consider it possible that this process contributed to the
11
12 phenomenon of altruism, but not as one of the leading forces in humans. It might still play a
13
14 substantial role in cases mentioned by Lewin-Epstein et al. (2017) that are otherwise difficult to
15
16 explain, such as eusociality in species where the queen mates with multiple males. We do hope
17
18 that other research teams will attempt to test the hypothesis of Lewin-Epstein et al. (2017) on more
19
20 diverse sources of data – coming from different groups of human participants, methodologies,
21
22 animal species (including non-model organisms – Cusick et al. 2021), etc. It undoubtedly remains
23
24 an interesting study objective.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

References

- Althouse, A. D. (2016). Adjust for multiple comparisons? It's not that simple. *The Annals of thoracic surgery*, 101(5), 1644-1645.
- Archie, E. A., & Tung, J. (2015). Social behavior and the microbiome. *Current Opinion in Behavioral Sciences*, 6, 28-34.
- Axelrod, R. The dissemination of culture: A model with local convergence and global polarization. *J. Confl. Resolut.* 41, 203-226 (1997).
- Bailey, M. T., & Cryan, J. F. (2017). The microbiome as a key regulator of brain, behavior and immunity: commentary on the 2017 named series. *Brain, behavior, and immunity*, 66, 18-22.
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior*, 10(1), 122-142.
- Cryan, J. F. & Dinan, T. G. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nat. Rev. Neurosci.* 13, 701–712 (2012).
- Cusick, J. A., Wellman, C. L., & Demas, G. E. (2021). The call of the wild: using non-model systems to investigate microbiome–behaviour relationships. *Journal of Experimental Biology*, 224(10), jeb224485.
- Dawkins, R. The Extended Phenotype: The Long Reach of the Gene. *Oxford University Press* (1982).
- Engel, C. Dictator games: a meta study. *Exp. Econ.* 14, 583–610 (2011).
- Fleg, J. (2025). Thirty years of studying latent toxoplasmosis: behavioural, physiological, and health insights. *Folia Parasitologica*, 72, Article 2025.005. <https://doi.org/10.14411/fp.2025.005>
- Forsythe, R. *et al.* Fairness in simple bargaining games. *Game. Econ. Behav.* 6, 347–369 (1994).
- Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of economic psychology*, 32(5), 865-889.

- Johnson, K. V. A., & Foster, K. R. (2018). Why does the microbiome affect behaviour?. *Nature reviews microbiology*, 16(10), 647-655.
- Kim, S. (2015). ppcor: an R package for a fast calculation to semi-partial correlation coefficients. *Communications for statistical applications and methods*, 22(6), 665.
- Leclercq, S., Mian, F. M., Stanisz, A. M., Bindels, L. B., Cambier, E., Ben-Amram, H., ... & Bienenstock, J. (2017). Low-dose penicillin in early life induces long-term changes in murine gut microbiota, brain cytokines and behavior. *Nature Communications*, 8.
- Lewin-Epstein, O. *et al.* Microbes can help explain the evolution of host altruism. *Nature Communications* 8, 14040 (2017).
- Lewin-Epstein, O., & Hadany, L. (2020). Host–microbiome coevolution can promote cooperation in a rock–paper–scissors dynamics. *Proceedings of the Royal Society B*, 287(1920), 20192754.
- Lindová, J. *et al.* Pattern of money allocation in experimental games supports the stress hypothesis of gender differences in *Toxoplasma gondii*-induced behavioural changes. *Folia. Parasit.* **57**, 136–142 (2010).
- Novakova, J. & Flegr, J. How Much Is Our Fairness Worth? The Effect of Raising Stakes on Offers by Proposers and Minimum Acceptable Offers in Dictator and Ultimatum Games. *PLoS ONE* 8, e60966 (2013).
- Post, S. G. Altruism, happiness, and health: It's good to be good. *Int. J. Behav. Med.*, 12, 66-77 (2005).
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Rorabaugh, A. N. Impacts of drift and population bottlenecks on the cultural transmission of a neutral continuous trait: an agent based model. *J. Archaeol. Sci.* 49, 255-264 (2014).
- Smith, E. R., & Conrey, F. R. Agent-based modeling: A new approach for theory building in social psychology. *Pers. Soc. Psychol. Rev.* 11, 87-104 (2007).

- 1
2
3
4 Sylvia, K. E., Jewell, C. P., Rendon, N. M., John, E. A. S., & Demas, G. E. (2017). Sex-specific modulation
5
6 of the gut microbiome and behavior in Siberian hamsters. *Brain, behavior, and immunity*, 60, 51-62.
7
8 Trivers, R. L. The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57 (1971).
9
10 Watabe, M., Kato, T. A., Monji, A., Horikawa, H., & Kanba, S. (2012). Does minocycline, an antibiotic
11
12 with inhibitory effects on microglial activation, sharpen a sense of trust in social interaction?.
13
14 *Psychopharmacology*, 220, 551-557.
15
16
17 Wilson, D. S., & Csikszentmihalyi, M. (2007). Health and the ecology of altruism. *Altruism and health:*
18
19 *Perspectives from empirical research*, 314-331.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65