

**Univerzita Karlova**  
**Přírodovědecká fakulta**

Katedra filosofie a dějin přírodních věd  
Studijní program: Teoretická a evoluční biologie



**Mgr. et Mgr. Robin Kopecký**

**Biologické a sociální prediktory a koreláty lidské morálky**

**Biological and social predictors and correlates of human morality**

Disertační práce

Školitel: prof. RNDr. Jaroslav Flegr, CSc.

Praha, 2023

## Bibliografický záznam

**Kopecký, Robin.** Biologické a sociální prediktory a koreláty lidské morálky. Praha: Univerzita Karlova, Přírodovědecká fakulta, Latedra filosofie a dějin přírodních věd, 2023. 165 s. Školitel: prof. RNDr. Jaroslav Flegr, CSc.

**Kopecký, Robin.** Biological and social predictors and correlates of human morality. Prague: Charles University, Faculty of Science, Department of Philosophy and History of Sciences, 2023. 165 p. Supervisor: prof. RNDr. Jaroslav Flegr, CSc.

## **Prohlášení**

Prohlašuji, že jsem závěrečnou práci vypracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu. Studie uvedené v této práci byly podpořeny Grantovou agenturou Univerzity Karlovy (GAUK 929216, GAUK 925416), Grantovou agenturou České republiky (GAČR 21-31059S) a Technologickou agenturou České republiky (TAČR TL01000467, TAČR TL03000549).

V Praze, 29. 10. 2023

.....  
Robin Kopecký

## Poděkování

Na počátku studia jsem se cítil jako neohrožená Dana Scully (The X-Files), avšak s postupem času jsem se měnil ve vystrašeného Mortyho (Rick and Morty).

Kdyby nebylo níže jmenovaných fantastických kolegů a nejlepších přátel, jejich obětavé pomoci, nekonečné trpělivosti, odbornosti a zvědavosti, tak mé studium by skončilo neúspěchem v temných hlubinách úzkostí a pochybování.

Děkuji Michaele Jirout Košové za její jedinečné rigorózní myšlení, absolutní spolehlivost a to, že jsme společně etablovali experimentální filosofii v Praze. Děkuji Jeleně Lence Příplatové za všestrannost, diskuze o biologii, vědě a skeptickém myšlení a zejména za to, že mi ochotně pomohla převést mé myšlenky ze schémat do literární formy srozumitelné druhým lidem. Za inspirativní diskuze děkuji zahraničním kolegům Silvii Boschetti a Konrádovi Talmont-Kaminski, děkuji za spolupráci taktéž kolegům z týmu profesora Flegra a z Centra Karla Čapka pro výzkum hodnot ve vědě a technice.

Největší dík patří školiteli prof. Jaroslavu Flegrovi za kontinuální zastání, dodávání naděje a odvahy k práci. Bez jeho iniciace a neustávající podpory by většina z projektů v této práci nikdy nevznikla.

Nakonec děkuji své drahé manželce Adrianě, za všechno.

## Abstrakt

V této disertační práci se zabýváme rozličnými oblastmi biologických a sociálních faktorů ovlivňujících lidskou morálku. Spojujícím prvkem tohoto souboru čtyř studií je souběžně zájem o lidskou morálku jako tradiční filozofické téma spolu s využitím metod empirických věd, zejména biologie, behaviorálních a kognitivních věd.

První studie se zaměřuje na morální algoritmy v autonomních vozidlech. Zkoumali jsme morální preference týkající se volby mezi typy softwaru, které se liší zabudovanými algoritmy pro řešení smrtelných kolizí. Příslušné typy jsou kategorizovány do tří strategií: sobecké, altruistické a konzervativní. Respondenti preferovali altruistické strategie, a tato preference se ještě posílila, pokud byla signalizována ostatním. Altruistická preference je nejvýraznější, když se vztahuje na všechny ostatní, slabší, když odráží pouze osobní volbu, a nejslabší při volbě pro vlastní dítě. Docházíme k závěru, že zveřejnění volby výrazně ovlivní jedince směrem ke společensky výhodnějšímu řešení.

Druhá studie zkoumá vztah mezi infekcí parazitem *Toxoplasma gondii* a jejím následným dopadem na politické přesvědčení a hodnoty lidí. Nakažení účastníci vykazovali tendenci ke zvýšenému tribalismu a sníženému kulturnímu liberalismu a antiautoritářství. Statisticky odfiltrovaný vliv horšího fyzického a duševního zdraví měl jen malý dopad, což naznačuje, že zhoršené zdraví tyto změny nezpůsobilo. Spíše než o adaptaci na výskyt parazitů, jak předpokládá teorie parazitického tlaku, by tyto rozdíly mohly být vedlejšími účinky dlouhodobého mírného stresu.

Třetí studie se zaměřuje na českou populaci a objasňuje vzájemný vztah mezi náboženským přesvědčením a různými aspekty wellbeingu v převážně ateistické zemi. Na rozdíl od předchozích studií jsme zjistili negativní korelaci mezi religiozitou a fyzickým i duševním zdravím. Ukázala se souvislost mezi zvýšenou religiozitou a zvýšeným sebehodnocením altruismu, biologickou zdatností a sníženou sexuální aktivitou.

Ve čtvrté studii jsme zkoumali dětské konceptualizace osobní identity ve vývojovém rámci a testovali koncept esenciálního morálního já. Prostřednictvím hypotetických scénářů změn se nám podařilo ukázat, že pro zachování osobní identity jsou klíčové morální vlastnosti, zejména ty, které souvisejí s mezilidskými vztahy. Navíc negativní změny těchto rysů měly na identitu výraznější škodlivý dopad než změny pozitivní.

V rámci snahy o multidisciplinaritu a přemostění oborů jsme obohatili jednu vědní disciplínu o metodologii jiné, a zabývali se lidskou morálkou z různých perspektiv.

## Klíčová slova:

morální soud, altruismus, wellbeing, experimentální filosofie, náboženství, toxoplazmóza

## Abstract

In this dissertation, we delve into the various domains of biological and social factors influencing human morality. The connecting element of this collection of four studies is both an interest in human morality as a traditional philosophical topic and the use of methods from empirical sciences, particularly biology, behavioural, and cognitive sciences.

The first study focuses on the moral algorithms in autonomous vehicles. We examined moral preferences regarding the choice between software types that differ in their built-in algorithms for dealing with lethal collisions. These are categorized into three specific types: selfish, altruistic, and conservative. Respondents exhibited a preference for the altruistic strategy, which is reinforced when signalled to others. The altruistic preference is the most pronounced when it applies to everybody else, weaker when it reflects only a personal choice, and the weakest when choosing for one's own child. We conclude that making a choice public significantly sways individuals towards a more socially beneficial solution.

The second study investigates the relationship between parasite *Toxoplasma gondii* infection and its subsequent impact on human political beliefs and values. Infected participants exhibited a tendency towards heightened tribalism and lowered cultural liberalism and anti-authoritarianism. Controlling for the effects of worse physical and mental health had little impact, suggesting that impaired health did not cause these changes. Rather than an adaptation to prevalence of parasites, as suggested by parasite-stress theory, the differences might be side-effects of a long-term mild stress.

The third study focuses on the Czech population, elucidating the interplay between religious beliefs and various facets of well-being in a mostly atheist country. In contrast to previous studies, we found a negative correlation between religiosity and both physical and mental health. A connection emerged between increased religiosity and heightened self-rated altruism, biological fitness, and decreased sexual activity.

Lastly, in the fourth study, we explored children's conceptualizations of personal identity within a developmental framework, testing the concept of essential moral self. Through hypothetical scenarios of change, we were able to show that moral traits, especially those linked to interpersonal relationships, were crucial for preserving personal identity. Moreover, negative changes in these traits had a more pronounced detrimental impact on identity than their positive counterparts.

In this endeavor towards multidisciplinary and bridging of disciplines, we enriched one scientific discipline with the methodology of another, addressing human morality from diverse perspectives.

## Key words:

moral judgement, altruism, wellbeing, experimental philosophy, religion, toxoplasmosis

# Obsah

<b>1. Úvod</b> .....	<b>10</b>
1.1 Teoretický úvod .....	10
1.2 Morální soud .....	14
1.3 Evoluční psychologie .....	18
1.4 Evoluční a kognitivní přístupy ke studiu náboženství .....	23
1.5 Experimentální filosofie.....	29
1.6 Reference.....	34
<b>2. Morální soud</b> .....	<b>42</b>
2.1 Morální problémy autonomních vozidel.....	42
2.2 Jak nás činí virtue signalling lepšími .....	57
2.3 Obava ze ztráty vlastního aktérství .....	68
<b>3. Evoluční psychologie</b> .....	<b>80</b>
3.1 Vliv latentní toxoplazmózy na hodnoty .....	80
<b>4. Evoluční a kognitivní přístupy ke studiu náboženství</b> .....	<b>94</b>
4.1 Vliv náboženství na wellbeing, fitness, fyzické a duševní zdraví.....	94
<b>5. Experimentální filosofie</b> .....	<b>135</b>
5.1 Osobní identita a pravé já.....	135

## Seznam publikací, které jsou součástí disertační práce

**Kopecký, R.** (2019). Morální problémy autonomních vozidel. *Filosofický časopis*, vol. 67, s. 263-276. ISSN 0015-1831.

**Kopecký, R.**, Jirout Košová, M., Novotný, D., Flegr, J., Černý, D. (2022). How Virtue Signalling Makes Us Better: Moral Preferences with Respect to Autonomous Vehicle Type Choices. *AI & Society*. <https://doi.org/10.1007/s00146-022-01461-8>

**Kopecký, R.**, Jirout Košová, M. (2023). Obava ze ztráty vlastního aktérství a absence morálního aktéra u autonomních dopravních prostředků. *Filosofie dnes*. (v tisku)

**Kopecký, R.**, Příplatová, L., Boschetti, S., Talmont-Kaminski, K., Flegr, J. (2022). Le Petit Machiavellian Prince: Effects of latent toxoplasmosis on political beliefs and values. *Evolutionary Psychology*, 20(3), <https://doi.org/10.1177/1474704922111265>

**Kopecký, R.**, Boschetti, S., Flegr, J. (2019). Effect of Being Religious on Wellbeing in a Predominantly Atheist Country: Explorative Study on Wellbeing, Fitness, Physical and Mental Health. *PsyArXiv*. <https://doi.org/10.31234/osf.io/3kr6n> (preprint v recenzi)

Jirout Košová, M., **Kopecký, R.**, Oulovský, P., Nekvinda, M., Flegr, J. (2021). My Friend's True Self: Children's Concept of Personal Identity. *Philosophical Psychology*. <https://doi.org/10.1080/09515089.2020.1860209>



## Seznam publikací, které nejsou součástí disertační práce

Sundvall, J., Drosinou, M., Hannikainen, I., Elovaara, K., Halonen, J., Herzon, V., **Kopecký, R.**, Jirout Košová, M., Koverola, M., Kunnari, A., Perander, S., Saikkonen, T., Palomäki, J., Laakasuo, M. (2023). Innocence over utilitarianism – Heightened Moral Standards for Robots in Rescue Dilemmas. *European Journal of Social Psychology*. <https://doi.org/10.1002/ejsp.2936>

Boschetti, S., Prossinger, H., Příplatová, L., Riha, D., **Kopecký, R.**, Binter J. (2023). Are Patterns Game for Our Brain? AI identifies individual Differences in Rationality and Intuition Characteristics of Respondents Attempting to Identify Random and Non-random Patterns. In *International Conference on Human-Computer Interaction* (pp. 151-161). Cham: Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-35979-8\\_12](https://doi.org/10.1007/978-3-031-35979-8_12)

Zawistowska, A., Příplatová, L., **Kopecký, R.**, Boschetti, S., Flegr, J., & Talmont-Kamiński, K. (2023, June 6). Superstitions, Religious Identity, and COVID-19 Conspiracy Beliefs in Secularized Czechia. *PsyArXiv*. <https://doi.org/10.31234/osf.io/tdwsx> (preprint v recenzi)

Boschetti, S., **Kopecký, R.**, Příplatová, L., Talmont-Kaminski, K. (2023). Black and white magic efficacy: role of in-group and out-group rituals. (preprint v recenzi)

Koverola, M., Kunnari, A., Drosinou, M., Palomäki, J., Hannikainen, I. R., Jirout Košová, M., **Kopecký, R.**, Laakasuo, M. (2022). Treatments approved, boosts eschewed: Moral limits of neurotechnological enhancement. *Journal of Experimental Social Psychology*. <https://doi.org/10.1016/j.jesp.2022.104351>

Oulovský, P., Košová, M., **Kopecký, R.**, Locke, K. D. (2020). Psychometric properties of the Circumplex scales of interpersonal values in the Czech population. *Československá psychologie*, 64(2), s. 155-167.

Flegr, J., Kuba, R., & **Kopecký, R.** (2020). Rhesus-minus phenotype as a predictor of sexual desire and behavior, wellbeing, mental health, and fecundity. *PLoS One*. <https://doi.org/10.1371/journal.pone.0236134>

# 1. Úvod

## 1.1 Teoretický úvod

Filosofové se po staletí zabývali lidskou morálkou formou spekulací, konceptuální analýzy, reflexe vlastních intuic, avšak v omezeném rozsahu své vlastní kultury. Další vědecké disciplíny je následovaly a rozšířily perspektivu našeho porozumění lidskému morálnímu uvažování a konání. Psychologové, behaviorální a kognitivní vědci empiricky zkoumali lidskou morálku. Pátrali po faktorech, které souvisí a determinují lidskou spolupráci, altruismus, prosociální chování a morální soud. Obdobně jako komparativní psychologie poskytuje srovnání lidského chování a kognice s dalšími více či méně příbuznými živočichy, vývojová psychologie se pokusila objasnit, jak se z téměř mimozemské entity, z dítěte, stane mladý dospělý navigující se ve společnost založené na vzájemné kooperaci pomocí svého vlastního morálního kompasu.

Vědecká komunita dosud nedospěla k obecnému konsensu v otázce souborné teorie lidské morálky. Je možné vymezení morálky v deskriptivním smyslu, s odkazem na funkci, kterou má plnit. Jedním ze způsobů (Gert & Gert, 2020), jak morálku vymezit, je ji popsat jako označení určitého předpisu chování, který předkládá jednotlivec nebo skupina. Toto vymezení zahrnuje etické normy, zvyklosti, zákony i náboženské doktríny.

Charles Darwin (1871) uvažoval o lidské morálce následujícím způsobem: *“I fully subscribe to the judgment of those writers who maintain that of all the differences between man and the lower animals the moral sense or conscience is by far the most important.”* (chapter III, The Descent of Man). Tím se de facto přihlásil k britské myšlenkové tradici teorie morálních sentimentů, kterou reprezentuje např. Hume (1751) a Smith (1759). Jiní biologové si chtěli disciplínu zkoumající morální teorie – etiku, na nějaký čas přivlastnit. E.O. Wilson píše: *“the time has come for ethics to be removed temporarily from the hands of the philosophers and biologized”* (Wilson, 1975). Skutečně tuto příležitost využili jiní výzkumníci na poli komparativní psychologie při studiu morálky u zvířat, a to nejen u primátů (De Waal, 1996).

Ovšem chování, např. sociálního hmyzu, nemá intencionalitu, je totiž výsledkem geneticky determinovaného chování (Ayala, 1987), na rozdíl od psychologického altruismu v rovině lidské morálky. Tato disertační práce je zaměřena na lidskou morálku, ačkoliv zjevně minimálně z části podléhá obdobným evolučním mechanismům, jakým podléhá kooperace a altruismus u ostatních živočichů, se kterými sdílíme společné předky a evoluční minulost.

Hlavní funkci morálky u člověka chápe Tomasello (2016) jako regulaci sociální interakce s ostatními, která slouží k dosažení spolupráce, i když je zřejmé, že my lidé jsme všichni do jisté míry sobeckými individui, která maximalizují svou individuální fitness. Tato regulační funkce, jejímž prostřednictvím omezíme svůj vlastní zájem ve prospěch druhých, je biologický altruismus. Ten se projevuje snížením vlastní fitness například sdílením či vynakládáním vlastních zdrojů na pomoc druhým, nebo tím, že bereme ohled na zájem druhých – formou reciprocity, dodržováním spravedlnosti, nebo následováním norem. Z naší evoluční historie víme, že moderní lidé jsou potomky sociálních primátů, kteří strávili 52 milionů let v sociálních skupinách (Shultz, Opie & Atkinson, 2011) a poslední dva miliony let žili jako úzce spolupracující lovci a sběrači (Tooby & Devore, 1987).

Tomasello (2016) popsal tři koncepty intencionality a její evoluční historii. Nejjednodušší individuální intencionalita je schopnost uzpůsobit chování k dosažení určitého cíle, obvykle v kompetici vůči druhým. Například když šimpanzi hledají a sbírají ovoce ve skupině, tak každý sbírá sám pro sebe. Individuální intencionalita mohla být přítomna u našich předků před 6 miliony let. Sdílená intencionalita, jejíž stáří je přibližně 400 tisíc let, se projevuje jako možnost mít sdílený cíl skupiny lidí, umožňuje kolaborativní lov a je nutná pro přežití. A konečně třetí, kolektivní intencionalita, stará přibližně 150 tisíc let, zakládá první normy a konvence v malých skupinách. Tyto kulturní praktiky dávaly základ morálce v posledních 100 tisících letech.

Morální systémy představují komplexní soustavy hodnot, ctností, norem, institucí a evolučně vyvinutých mechanismů, které společně pracují na potlačení či regulaci sobeckých zájmů a umožňují tak fungování kooperativních společností (Haidt, 2001). Morální systémy mají alespoň tři komponenty: újmu spojenou s individuem, čistotu spojenou s posvátnem a věrnost spojenou se skupinou. Tento koncept přejal Haidt od antropologa Schwedera (Schweder, Much, Mahapatra & Park, 1997). Haidt vystupoval proti racionalistickým teoriím morálky, které označuje za racionalistický blud (*rationlist delusion*) spojenými s Kohlbergem a Piagetem.

Dle Haidta je morální soud primárně spojen s morálními emocemi. Sociálně intuitivní model (Haidt, 2001) obsahuje tři centrální teze: v morálním soudu se vyskytuje jako první intuice, a až po ní následuje rozumové zdůvodnění. Za druhé, morálka není jen o újmě a férovosti, je více komplexní a spočívá na více pilířích. A za třetí, vlastní morálka metaforicky oslepuje a svazuje vůči jiným morálním systémům. Na tomto základě formoval teorii morálních základů (Sweder et al., 1997; Haidt & Joseph, 2004; Haidt & Graham, 2007), dle které se jednotlivé morální systémy skládají z kombinace některých z následujících pilířů: péče/újma, férovost/podvádění,

svoboda/útlak, věrnost/zrada, autorita/podvracení, posvátnost/zneuctění. Různé morální systémy přitom kladou různou váhu jednotlivým pilířům.

Podle Curryho (2016) lze morálku chápat jako soubor biologických a kulturních řešení problémů kooperace, které se pravidelně objevovaly během lidské evoluce a aktuálně objevují v lidském životě. Jeho teorie morálky jako kooperace nám poskytuje rámec, v němž tato adaptivní řešení vymezují, co rozumíme pod pojmem lidská morálka. Morálka v tomto kontextu není jen otázkou reciprocity nebo her bez nulového součtu. Jedná se o soubor vyvinutých morálních pravidel, která mají univerzální charakter. Tato univerzálnost je založena na evolučních mechanismech a specifických složkách morálky. Curry dále rozlišuje různé typy kooperace, které odpovídají různým typům morálky. Příkladem je příbuzenství, které je spojeno s péčí o děti a averzí k incestu. Mutualismus se projevuje ve skupinové konformitě a loajalitě. Výměnná kooperace je spojena s reciprocitou, důvěrou, vděkem, vinou a schopností odpuštění. A konečně, řešení konfliktů zahrnuje dělení a nákladné signály, jako jsou statečnost a štedrost. Curryho pojetí tak nabízí komplexní pohled na morálku a její různé dimenze v kontextu lidské kooperace.

Předložená disertační práce „*Biologické a sociální prediktory a koreláty lidské morálky*“ obsahuje pět kapitol. První kapitola – Úvod obsahuje stručné uvedení k tématu, formě a obsahu práce. Druhá kapitola se skládá z úvodní rozvahy formou eseje, vlastní empirické studie a závěrečné eseje. Třetí až pátá kapitola je každá složena ze samostatné odborné publikace založené na vlastním empirickém zkoumání.

Spojujícím prvkem této kolekce odborných studií a esejí je jednak zájem o lidskou morálku jako o tradiční filosofické téma, a jednak využití metod empirických věd – zejména biologie, behaviorálních a kognitivních věd. Tato snaha o multidisciplinaritu a přemostění disciplín spočívá v obohacení jedné vědecké disciplíny o metodologii vědy druhé, aplikaci teoretického konceptu z biologie (např. teorie signalizace) pro řešení sociálního dilematu, v jehož centru je morální problém. Tyto kombinace metod ve studiu lidské morálky měly za účel najít její méně studované prediktory a koreláty biologické a sociální povahy.

Vlastní obsah práce začíná v druhé kapitole „Morální soud: autonomní vozidla“, ve které je řešen specifický morální problém spojený s nástupem autonomních vozidel – sociální dilema, jehož jádrem je lidská tendence volit pro svůj vůz sobeckou strategii, zatímco od ostatních se očekává altruistická volba. Implementovali jsme teorii čestných signálů zoologa a behaviorálního biologa Amotze Zahaviho a otestovali naše navržené řešení.

Třetí kapitola „Evoluční psychologie: vliv latentní toxoplazmózy na hodnoty“ zkoumala vliv patogenu – parazita *Toxoplasma gondii* na politická přesvědčení a hodnoty nakažených osob.

Pokusili jsme se navrhnout alternativní vysvětlení k *Parasite stress theory* evolučního biologa a entomologa Randy Thornhilla.

Ve čtvrté kapitole „Evoluční a kognitivní přístupy ke studiu náboženství: zdraví a wellebing“ jsme přihlédlí k behaviorálně ekologické perspektivě při studiu náboženství spojené s antropologem Richardem Sosisem a psychologem Jeosephem Bulbulia.

V poslední páté kapitole „Experimentální filosofie: osobní identita a morální já“ jsme užívali metody experimentální filosofie a vývojové psychologie ke zkoumání laických intuic zakládajících filosofický koncept zachování osobní identity. Autory teorie pravého morálního já jsou filosofové Jesse Prinz, Shaun Nichols a psycholožka Nina Strohminger.

Hlavní objevy této práce jsou následující:

- Veřejná signalizace volby morálního softwaru pro autonomní vozidla zvýší prosociální jednání, tj. participanti dávají přednost altruistické strategii před sobeckou.
- Infekce latentní toxoplazmózou je asociována se zvýšeným tribalismem a není mediována celkovým zdravotním stavem. Poskytli jsme nevylučující alternativa v proximální úrovni k *parasite stress theory*.
- Teorie o pozitivním vlivu náboženství na zdraví a *wellbeing* je kontextuálně závislá. V české populaci fyzické zdraví i duševní zdraví negativně koreluje s religiozitou. Vysvětlujeme to tím, že náboženství v kontextu, kdy je minoritní, je odklonem od normy, a to s sebou přináší stress.
- Naše výsledky podpořily specifickou teorii osobní identity i u dětí, a sice že morální kontinuita je zásadní pro zachování osobní identity – jak uvádí teorie esenciálního morálního já. Osobní identita je vnímaná jako porušená, zvláště pokud došlo k negativní morální změně.

## 1.2 Morální soud

*Trolley problem*, česky někdy nazýván vozíkový problém, je ikonickým myšlenkovým experimentem, který se v různých variantách používá napříč disciplínami, které zkoumají, nebo se jen setkávají s lidskou morálkou. V druhé kapitole této práce užíváme zmíněný myšlenkový experiment v kontextu sociálního dilematu spojeného s novou technologií autonomních vozů vybavených řídicím morálním software. Ve výkladu problematiky si bereme na pomoc teorii z behaviorální biologie.

Vozíkový problém vznikl na poli bioetiky v diskusi ohledně interrupcí (Foot, 1967; Foot, 2002). Tuto diskusi dále rozšířila Thomson (1984). Dnes existuje celá řada různých variant vozíkového problému a tento myšlenkový experiment pronikl široce i do popkultury, humorně označován jako *trolleyology* (Edmonds, 2013).

Vozíkový problém existuje ve dvou hlavních variantách: výhybka a můstek. Varianta s výhybkou zobrazuje tramvaj nekontrolovatelně se řítící po kolejích, přičemž je jisté, že když bude pokračovat po dané dráze, zabije pět dělníků. Avšak nedaleko je páka, kterou lze ovládat výhybku. Pokud osoba u páky výhybku přehodí, tramvaj se odkloní na vedlejší kolej, kde je však jiný dělník v bezprostředním nebezpečí. Problém se tedy zaměřuje na morální dilema: Má osoba zachránit pět dělníků na úkor jednoho? Ve variantě s můstkem je situace podobná – pět dělníků je ohroženo na kolejích. Avšak nyní je náhodný kolemjdoucí na můstku nad kolejemi, společně s mohutným mužem, který se nebezpečně vyklání. Pokud by tento mohutný muž spadl z můstku, jeho tělo by tramvaj zastavilo dříve, než by dosáhla pěti dělníků na kolejích. Klíčová otázka zní: Je morálně přijatelné (či dokonce nezbytné) shodit mohutného muže, aby se zachránily životy pěti dělníků? V tomto scénáři by smrt mohutného muže byla nezbytná k záchraně pěti, a tudíž by byl považován ne za cíl, ale za prostředek. V literatuře se pro takový soud upřednostňující záchranu většího počtu osob označuje jako utilitaristický, zatímco morální soud, který by zdůrazňoval morální nesprávnost shození mohutného muže bez ohledu na následky, by byl označen jako deontologický.

Greene (2014) shrnuje své neurovědecké a filosofické výzkumy jako teorii duálních procesů morálního soudu. Dle této teorie deontologické soudy vychází z rychlých, automatických, emocionálních reakcí, zatímco utilitaristické soudy jsou podmíněny deliberativním kognitivním procesem. Teorie původně vycházela z neurozobrazování a reakčních časů (Greene, 2007; Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008, Cushman, Young & Hauser, 2006). Je však nutné poznamenat, že morální soud laiků není ekvivalentní expertním teoriím, precizně formulovaným v deontologické nebo utilitaristické tradici. Obdobně jako může

laik mít intuitivní představu o Schrödingerově kočce, jeho intuice nepředstavuje ani Kodaňskou ani Mnohasvětovou interpretaci, ale jde prostě o nějakou neostrou intuici. Stejně tak by byla chyba intuici ohledně vozíkového problému ztotožňovat s expertními sofistikovanými teoriemi. A konečně je třeba poznamenat, že vozíkový problém není vyčerpávající popis lidského morálního soudu. Zabývá se jedním z pilířů lidské morálky újmy a péče.

Jedna z morálních otázek v kontextu autonomních vozidel je, jak by měly být navrženy algoritmy, které by řešily vzácné situace analogické vozíkovému problému, kdy nastává neodvratná smrtelná kolize. Vystává před námi dilema, jak by měla být tato újma distribuována (Lin, 2015). Pro účely této disertační práce bude úvaha redukována na různé typy morálního softwaru pro autonomní vozy, které mají za úkol řešit sociální dilema mezi preferencí záchrany posádky a preferencí záchrany největšího počtu lidí.

Druhá kapitola této disertační obsahuje dvě doplňující eseje. Tyto eseje plní roli úvodu a poskytují kontext k aktuální debatě ohledně umělé inteligence, sociálního rozměru a etiky technologie. Hlavní empirická studie obsažená v druhé kapitole měla za cíl zodpovědět dvě výzkumné otázky: zda veřejná viditelnost volby typu morálního software pro autonomní vozy jeho majitelem zvýší ochotu přijmout altruistickou variantu. A za druhé, jaký typ situace je pro altruistickou volbu nejobtížnější, zda volba morálního softwaru pro společnost jako celek, volba pouze pro sebe, nebo volba pouze pro vlastního potomka.

Při řešení sociálního dilematu autonomních vozů jsme vycházeli z poznatků behaviorální ekologie (Laidre & Johnstone 2013; Dawkins & Krebs 1978; Zahavi 1975). Ochota užívat altruistické autonomní vozidlo by mohla být interpretována jako čestný signál. Projevy altruismu se dají chápat jako signál ochoty pomáhat druhým, a to i za cenu vlastních nákladů (Hardy & Van Vugt 2006; Fehrler & Przepiorka, 2013). Výzkum (Engelmann, Herrmann & Tomasello, 2012), srovnal souvislost reputace a morálního chování u primátů a dětí. Ukazuje se, že lidé dbají na sociální kontext. Chovají se morálněji, pokud jsou pozorováni druhými. Tento zásadní rozdíl mezi lidskou a zvířecí morálkou ukazuje inspiraci pro viditelnou volbu morálního softwaru pro autonomní vůz.

Pomocí online dotazníku jsme získali odpovědi 2769 respondentů. Užili jsme dva testy pozornosti a vyřadili několik participantů, kteří tímto testem úspěšně neprošli. Participant si nejprve přečetli expertní predikci o efektivitě autonomních vozidel a počtu zachráněných životů a materiálních škod v České republice. Následně byli seznámeni s vizualizacemi kolizních situací podobných vozíkovému problému. Situace byla prezentována formou obrázků, kde byla znázorněna situace v tunelu. Autonomní automobil měl možnost buď pokračovat pouze rovně,

nebo zabočit mírně na stranu. V autonomním voze seděl například jeden pasažér a mířil přímo na pět osob stojících v tunelu vedle havarované dodávky. Participanti měli hodnotit, jak má software rozhodnout, zda jet rovně a usmrtit pět osob, nebo zabočit do dodávky, a tím usmrtit posádku autonomního vozu. Tímto způsobem se participanti seznámili s různými situacemi inspirovanými vozíkovým problémem a s představou, že morální software bude činit rozhodnutí ohledně života a smrti.

Představili jsme tři typy vozidel: sobecké, které vždy upřednostňovalo záchranu posádky, altruistické, které preferovalo záchranu nejvyššího počtu osob v dané situaci i za cenu obětování osob uvnitř vozu, a konzervativní, jehož software by nikdy nezměnil směr jízdy, pokud by to mělo za následek něčí smrt. Zeptali jsme se participantů, jaký typ morálního software by zvolili pro sebe, pokud by se o jejich výběru nikdo nemohl dozvědět, jaký typ morálního software by zvolili pro sebe, kdyby jejich volba byla viditelná na první pohled, a jaký typ by (v tajnosti) volili pro své dítě. Nakonec jsme se zeptali, jak by volili v parlamentu ohledně morálního softwaru pro všechny, kdyby tato volba byla tajná, a jak by volili, kdyby tato volba byla veřejná.

První testovaná hypotéza předpokládala, že respondenti budou mít nižší preference pro sobeckou strategii ve veřejné volbě než ve volbě tajné. Tato předpověď se potvrdila (RR = 0,665, CI (95%) = 0,620–0,714). Druhá hypotéza tvrdila, že při veřejném hlasování budou participanti více altruističtí než v tajném hlasování. Taktéž se potvrdila (RR = 0,825, CI (95%) = 0,738–0,922). Třetí hypotéza předpokládala, že v tajné volbě budou lidé preferovat nejvíce sobeckou strategii pro software autonomního vozu pro své dítě, méně sami pro sebe a nejméně sobečtí budou při tajném hlasování v parlamentu. Oba vztahy se prokázaly.

Naše výsledky ukazují, že respondenti projevují preferenci pro altruistickou strategii u autonomních vozidel, zejména pokud je tato volba viditelná pro ostatní. Altruistická preference je nejsilnější, když se vztahuje na všechny ostatní, méně silná, když odráží pouze osobní volbu, a nejslabší, když se volí pro vlastního potomka. Z těchto výsledků lze usoudit, že většina participantů v naší studii sdílí utilitaristické intuice – altruistické obětování sama sebe, pokud se zachrání více lidí. Preference sobeckého software pro autonomní vozidlo pro své dítě byla v souladu s předchozí studií (Bonneton, Shariff & Rahwan, 2016). Domníváme se, že silná preference pro altruistický software, zejména pokud by byl viditelný na první pohled, může být ovlivněna snahou vyhnout se sociální sankci a snahou o budování pozitivní reputace. Pro stabilitu řešení, které navrhujeme, je v souladu s teorií Zahaviho (1975) důležité zajistit, aby falšování tohoto četného signálu – v tomto případě altruistického software – bylo spojeno s vysokými náklady.



Ve studii týkající se vozíkových problémů Lee, Sul & Kim (2018) zjistili, že když jsou lidé, pozorování druhými, tak mají nižší utilitaristické preference. Důvodem, proč v naší studii efekt pozorování druhými jako veřejně viditelný morální software není také méně utilitaristický, si vysvětlujeme následovně. Mít utilitaristické autonomní vozidlo je maximálně altruistické, neboť oproti vozíkovým problémům uživatel takového vozidla vsází svůj vlastní život v kolizních situacích. Může sám zaplatit svým životem, když altruistický vůz spočítá, že obětováním posádky zachrání více lidských životů.

### 1.3 Evoluční psychologie

Parazité jsou důležitým faktorem ovlivňujícím evoluci všech linií, lidskou evoluci nevyjímaje. Kromě zřejmých souvislostí, jako je například vznik a vývoj tělesného a behaviorálního imunitního systému, však ovlivňují evoluci i subtilněji, a to včetně vlastností, které s biologickými vlivy běžně nespojujeme. Zásadní je například vliv parazitů na lidské hodnoty a normativní soudy, což si můžeme ukázat na příkladu estetiky. Lidé vykazují univerzální preferenci pro krásu na základě symetrických proporcí tváře, absence malformací a čisté pleti. Může jít o známku absence parazitární nákazy, případně známku dobrých genů zaručujících signifikantní odolnost vůči parazitům (Thornhill & Gangestad, 1993). Naopak vnímaná ošklivost – tělesná asymetrie, malformity či například vady pleti – v nás vzbuzuje odpověď behaviorálního imunitního systému (Schaller & Park, 2011). Samozřejmě nejde o vyčerpávající vysvětlení lidských estetických preferencí, může se však jednat o vysvětlení ultimátní. Jak psychologická role krásy (Reber, Schwarz & Winkielman, 2004), tak estetický soud obecně (Dutton, 2009) i fenomén ošklivosti mají souvislost s vyhýbáním se patogenům (Klebl, Luo, Tan, Ern & Bastian, 2021).

Podobně jako může patogenní znechucení (*pathogen disgust*) jakožto součást behaviorálního imunitního systému chránit před možnou nákazou, mohli parazité hrát roli také v evolučním původu morálního znechucení (*moral disgust*) (Chapman, Kim, Susskind & Anderson, 2009). Cílem našeho výzkumu bylo dále prozkoumat tento biologický faktor – nákazu a její možný vliv na morální soud samotných nakažených. Jako experimentální model nám posloužil parazitický prvok *Toxoplasma gondii*, a to ze dvou důvodů. Zaprvé jde o modelový organismus, u kterého byly různé formy parazitické manipulace v minulosti dobře podpořené (Barnard & Behnke, 1990; Moore, 2002; Vyas & Sapolsky, 2010; Webster & McConkey, 2010) a jehož výzkum má v pražské skupině profesora Jaroslava Flegra dobře propracované technické zázemí (pravidelné testování dobrovolníků z řad veřejnosti, dotazníkové studie, skupina Pokusní králíci aj.). Druhým důvodem bylo, že v širokém poli výzkumu vlivu toxoplazmózy na lidské zdraví a psychiku zůstávalo doposud pole morálních hodnot neprozkoumané.

*Toxoplasma gondii* je parazitický jednobuněčný eukaryotický organismus ze skupiny apicomplexa. Finálním hostitelem toxoplazmy jsou kočkovité šelmy, mezihostiteli se může stát většina teplokrevných organismů (Ajioka & Soldati, 2007; Dubey & Carpenter, 1993; Costa et al., 2012; Bártová, Sedlák & Literák, 2006). V epiteliálních buňkách tenkého střeva finálního hostitele prochází toxoplasma pohlavním procesem, jehož výsledkem jsou odolná infekční stadia (oocysty), která organismus vylučuje spolu s výkaly do prostředí. Ty přetrvávají v prostředí, kde

je mohou pozřít spolu s vodou či potravou mezihostitelé. V trávicím traktu mezihostitele oocysty praskají a vypouští sporozoity, kteří diferencují v pohyblivé tachyzoity. Tachyzoiti ve stadiu akutní infekce zaplaví tkáň, přičemž postupně diferencují v bradyzoity, tkáňové cysty, jež jsou známkou latentní infekce. Z různých studií vyplývá, že latentně je toxoplasmou nakažena přibližně třetina lidské populace (Tenter, Heckeroth & Weiss, 2000). Člověk se nejčastěji nakazí oocystami, tj. konzumací zeleniny a dalších zemědělských produktů potřísněných kočičími výkaly, případně tkáňovými cystami ze špatně tepelně opracovaného masa mezihostitelů (Tenter et al., 2000; Kolbeková, Koubartová, Novotná, Kodým & Flegr, 2007). Stejně se toxoplasma dostává i do finálního hostitele, kočkovité šelmy.

Lidé v latentní fázi nákazy toxoplazmózou neprojevují žádné závažné specifické symptomy, takže byla dlouhou dobu považována za bezpříznakovou. Podle řady studií publikovaných v posledních dvaceti letech je ale nákaza pozitivně asociována s řadou onemocnění (Flegr & Escudero, 2016; Flegr, Prandota, Sovičková & Israili, 2014; Flegr, Klapilová & Kaňková, 2014), změn osobnosti (Khademvatan, Saki, Yousefi & Abdizadeh, 2013) i chování (Houdek, 2017; Hodková, Kolbeková, Skallová, Lindová & Flegr, 2007).

Zůstává však otázkou, nakolik je možné odlišit nákazu, která obecně zhoršuje zdravotní stav člověka, od účelové manipulace hostitele ze strany parazita. Jinými slovy, jak rozlišit vedlejší důsledky akutní či latentní nákazy pro zdraví od cílených zásahů do hostitelské fyziologie, které mohou zvyšovat pravděpodobnost přenosu parazita. Toxoplazmózou trpící živočichové včetně člověka každopádně mají aktivovaný imunitní systém a zvýšenou úroveň protizánětlivých cytokinů (Kaňková, Holáň, Zajícová, Kodým & Flegr, 2010). Bylo také prokázáno, že záněty a stres vedou k změnám v motivaci, v emocích i chování (Moieni & Eisenberger, 2018). V recentní preprintové publikaci (Flegr, Hlaváčová & Toman, 2023) však autoři ukázali, že zhoršené zdraví není mediátorem vlivu toxoplazmózy na lidskou psychiku.

Druhou stále otevřenou metodologickou otázkou je, zda parazité mohou ovlivňovat i tak odvozené lidské vlastnosti jakými jsou jeho morální hodnoty a politické preference. Jisté je, že manipulace toxoplasmy mohou být poměrně sofistikované. Vlivný článek „Fatal attraction in rats infected with *Toxoplasma gondii*“ (Berdy, Webster & Macdonald, 2000) dokládá u nakažených kryš změnu antipredačního chování. Krysy, které laboratorně nakaženy a následně byly pozitivně testované na přítomnost protilátek proti toxoplazmóze (dále *Toxoplasma* pozitivní), se oproti negativním nevyhýbaly místům s možnou přítomností kočky (indikovanými kočičí močí) a naopak byly takovými místy přitahovány (tj. vykazovaly fenomén *fatal attraction*). Ve studii (Flegr, Lenochová, Hodný & Vondorvá, 2011) s lidskými probandy se fenomén *fatal attraction*

objevil v případě kočičí moči jen u *Toxoplasma* pozitivních mužů, zatímco u *Toxoplasma* pozitivních žen byl efekt opačný. Tygří moč ani moč jiných živočichů (hyena čabraková, kůň domácí, pes domácí) žádný efekt nevyvolala. Navazující studie (Poirotte et al., 2016) studovala *fatal attraction* u šimpanze učenlivého (*Pan troglodytes troglodytes*) na základě averze k moči jejich přirozeného predátora leoparda skvrnitého (*Panthera pardus*). *Toxoplasma* pozitivní šimpanzi měli nižší averzi k moči svého přirozeného predátora, ale v případě jiných kočkovitých šelem, které nejsou jejich přirozenými predátory, se rozdíl nepotvrdil. To může naznačovat, že u člověka je vliv toxoplazmózy spíše dědictvím evoluční minulosti než aktuálně adaptivní strategií parazita (Flegr, 2023).

Ačkoliv existují studie popisující osobnostní rysy a charakteristiky asociované s latentní toxoplazmózou, tématem hodnot a morálky se zabývá jen několik z nich: Lidé nakažení toxoplazmózou jsou méně svědomití (Lindová, Příplatová & Flegr, 2012) a méně štedří (Lindová et al., 2010). V ekonomické hře na diktátora jsou nakažení muži méně altruističtí, ale nakažené ženy naopak více altruistické (Lindová et al., 2010). Další studie tematicky bližší politickým hodnotám naznačují, že nakažení jsou více otevření podnikání a nebojí se více riskovat (Johnson et al., 2018). S tím může souviset i větší riziko dopravních nehod u infikovaných osob, které je ale třeba nahlížet i z hlediska zhoršení reakčních dob (Flegr, Havlíček, Kodým, Malý & Smahel, 2002; Kocazeybek et al., 2009). Pozoruhodným doplňkem naší studie je výzkum (Meyer et al., 2022) publikovaný čtyři měsíce po naší studii zkoumající vliv toxoplazmózy na populace vlků (*Canis lupus*). Toxoplasmou infikovaní vlci vykazovali značně vyšší tendenci k riskantním rozhodnutím, jako je snaha usilovat o vůdcovství smečky, nebo opuštění smečky v mladším věku. Autoři spekulují, že to může souviset s přímou manipulací parazita. Zmíněné chování zvýší šanci na překryv území, ve kterém nová smečka nebo jedinec žije, a územím kde se nachází definitivní hostitel – puma (*Puma concolor*). Vliv latentní toxoplazmózy na lidské hodnoty ve smyslu altruismu byl zkoumán v předešlých studiích zejména metodologií behaviorálních ekonomických her, v dimenzích péče/újmý a férovosti/podvádění. Zatím nepokrytou oblastí byla souvislost s dalšími aspekty lidské morálky, tedy základy jako jsou autorita/podvracení, věrnost/zrada, svoboda/útlak, či posvátnost/degradace (Haidt, 2012; Graham et al., 2013), tedy hodnot blíže politickým přesvědčením.

Naše studie ve třetí kapitole této disertace měla za cíl porovnat politické přesvědčení a hodnoty participantů infikovaných toxoplazmózou a participantů bez infekce. Prostřednictvím online dotazníku jsme získali odpovědi 2315 respondentů, z toho 477 uvedlo, že byli laboratorně testováni na toxoplazmózu s to s pozitivním výsledkem, zatímco 1838 s výsledkem negativním.

K vyhodnocení pomocí statistické analýzy hlavních komponent jsme použili vlastní baterii otázek PI34. Na základě odpovědí participantů jsme určili 4 hlavní komponenty, které jsme kvůli srozumitelnosti nepojmenovali čísly, ale jmény: tribalismus, ekonomické rovnostářství, kulturní liberalismus a anti-autoritářství. Volbu vlastní baterie otázek podrobně odůvodňujeme v článku; hlavním důvodem byla kulturní specifická českých probandů, pro které by nebylo vhodné užít politické dotazníky nastavené na Spojené státy Americké. Volbou názvů komponentů určitě neztotožňujeme přesvědčení laiků s teoriemi politických filosofů, ani s Haidtovými základy morálky. Zvolili jsme je kvůli praktičnosti na základě určitých styčných ploch s těmito teoriemi a základy morálky. Lze však vidět zřejmé analogie mezi hlavními komponenty politického inventáře – PI34 – a morálními základy. Ekonomické rovnostářství koresponduje s péčí/újmou a férovostí/podváděním, kulturní liberalismus se svobodou/útlakem, anti-autoritářství s autoritou/podvracením a poslední komponenta, tribalismus, odpovídá kombinaci více základů. Do budoucna bude v tomto ohledu informativní srovnat u jednotlivých probandů hlavní komponenty PI34 s teorií morálních základů, aby bylo přesně vidět, jak politická přesvědčení jsou provázána se základy morálky.

Naše studie ukázala, že se toxoplazmózou infikovaní a neinfikovaní účastníci signifikantně lišili ve třech ze čtyř faktorů PI34. Infikovaní dosahovali vyššího skóre v tribalismu a nižšího skóre v kulturním liberalismu a anti-autoritářství. Dále jsme poukázali na pohlavní rozdíly mezi infikovanými a neinfikovanými ženami a muži v politických názorech. Infikované ženy dosáhly vyššího skóre v tribalismu a nižšího skóre v kulturním liberalismu ve srovnání s neinfikovanou kontrolní skupinou, zatímco infikovaní muži dosáhli vyššího skóre v ekonomickém rovnostářství. Tyto výsledky mohou odpovídat rozdílům mezi ženami a muži v chování a postojích pozorovaných po infekci toxoplazmózou. Alternativní vysvětlení změnou politických přesvědčení vlivem celkově horšího zdravotního stavu se neprokázalo – pozorované asociace zůstaly zhruba silné i po odfiltrování vlivu fyzického a duševního zdraví.

Vzhledem k tomu, že vlivy latentní toxoplazmózy u lidí představují spíše dědictví naší evoluční minulosti, se kloníme interpretaci, že tyto nalezené rozdíly nemusí nutně dokládat cílenou manipulací parazita zvyšující pravděpodobnost jeho přenosu. Rozdíly by mohly být spíše vedlejšími účinky dlouhodobé mírné zánětlivé reakce a chronického stresu, tedy faktorů, které ještě nevedly ke zhoršení celkového zdraví. Naše výsledky tedy nepovažujeme za adaptaci na prevalenci parazitů, jak naznačuje *parasite stress theory* (Fincher, Thornhill, Murray & Schaller, 2008). Pro získání jasné představy o účincích mírného zánětu je nicméně potřeba dalšího empirického výzkumu zaměřeného na spektrum různých onemocnění. *Parasite stress theory*

(Fincher et al., 2008; Thornhill & Fincher, 2014) se snaží poskytnout ultimátní evoluční vysvětlení lidských přesvědčení, jako například konzervatismu, odporu a předsudků k cizincům, dle různé úrovně parazitického stresu na daném geografickém území. Kauzalita jde směrem od přítomnosti parazitů, přičemž se nejčastěji uvažuje o evoluční výhodě xenofobie a vyššího konzervatismu jakožto ochrany před nákazou.

Naše studie předkládá alternativní kauzalitu, která se ovšem s předpoklady *parasite stress theory* nevyklučuje. Jedná se o proximální vztah onemocnění – mírný zánět a chronický stres, který zároveň ještě nevede k celkovému zhoršení zdraví, má za následek posun v hodnotách k většímu tribalismu. Naše výsledky ukazují, že celkové zdraví není mediátorem změn, proto se nabízí jako mediátor chronický stres spojený s dlouhodobou infekcí. Předchozí studie (Sýkorová & Flegr, 2021) dokládá, že dlouhodobý stres vede k přechodu na rychlou životní strategii, která by mohla následně ovlivňovat hodnoty a politická přesvědčení. Uzavíráme, že na proximální úrovni nákaza toxoplazmózou zvyšuje tribalismus, což na ultimátní úrovni v zeměpisných místech s vyšší frekvencí parazitů může vést k vyššímu konzervatismu/tribalismu jakožto adaptaci na vyšší riziko infekce.

Na závěr je třeba poznamenat, že sběr dat probíhal mezi lety 2016 až 2021, ale většina participantů (94 %) se účastnila studie před nástupem pandemie COVID-19. Největší pandemie za poslední století tedy neměla zásadní vliv na výsledky této studie, ale je možné, že studie následující v době po pandemii, jež přinesla účastníkům výzkumu osobní zkušenost s introdukcí patogenů, ukáže jiné efekty.

## 1.4 Evoluční a kognitivní přístupy ke studiu náboženství

Položíme-li si otázku, zda má příslušnost k nějakému náboženství vliv na vlastní kvalitu života jednotlivce, musíme blíže specifikovat neostré termíny náboženství a wellbeing. Vzhledem k více možným přístupům k definici a jednotlivým složkám wellbeingu (Lambert at al., 2020) jsme zvolili dvojitý přístup. V první řadě jsme pro potřeby naší studie spočítali agregovaný faktor „*wellbeing*“ na základě pěti komponent, a to fyzického zdraví, duševního zdraví, ekonomické situace, sebehodnocení atraktivity a kvality partnerského vztahu. A za druhé, jsme každou složku (např. subjektivní zdraví, objektivní zdraví apod.) srovnávali s religiozitou samostatně.

Ačkoliv jsou termíny náboženství a náboženská víra problematické a kulturně variabilní, pro potřeby této studie je, vzhledem k jejímu omezení na české kulturní a jazykové prostředí, není nutné je definitivně vymezit. Náboženství sestávající se ze systému hodnot, praktik a vnitřního přesvědčení bylo pro účely tohoto výzkumu určeno jako 1. sebeidentifikace ve smyslu náboženské víry a 2. počet skupinových náboženských aktivit za měsíc.

Diskuze o vlivu náboženství na wellbeing a zdravotní stav může být zahrnována do dvou přístupů k funkci a evoluci náboženství; jeden chápe náboženství jako účelnou adaptaci, druhý vnímá náboženství a náboženská přesvědčení jako vedlejší produkt našeho kognitivního systému. Již Darwinovy (1871) úvahy o animismu u tradičních mimoevropských společností prostřednictvím analogie k usuzování jeho psa, byly vedené v rámci náboženských přesvědčení jako vedlejšího produktu. Když Darwinův pes reagoval vrčením a štěkotem na otevřený slunečník, který se občas pohnul díky vánku, Darwin se domníval, že pes usuzuje z pohybu bez viditelné příčiny přítomnost nějakého dalšího aktéra. (Day, 2008; Chidester, 2009). Toto přisuzování intencionality a aktérství u člověka i tam, kde zřejmě nejsou, popsal Berrett jako *Hyperactive agency detection device* (Berrett, 2000). Jedná se o vedlejší produkt – spandrel (Gould, Lewontin & Anderson, 1979) lidské kognice, teorie mysli, která má být zodpovědná za původ a šíření víry v nadlidské aktéry jakou jsou např. bozi, duchové nebo předci (Gray & Wegner, 2010; Tremblin, 2010). Adaptivní složka této atribuce je evolučního původu: obecně je výhodnější falešně pozitivní chyba než falešně negativní. Můžeme si to představit například v kontextu možné přítomnosti predátora, kdy falešně negativní chyba v odhalení takového aktéra vede ke smrti, zatímco falešně pozitivní chyba vede k chybnému soudu s řádově nižší nebezpečností. Dalším analogickým jevem *pareidolie* – nadměrné vnímání vizuálních stimulů. Lidé například interpretují vizuální vjemy jako tváře tam, kde žádné přítomny nejsou. (Tsao & Livingstone, 2008).

Analogii k původu náboženských přesvědčení jako spandrel, vedlejšího produktu teorie mysli a detekci aktérů, spatřuje Boyer (2003) ve čtení, které analogicky využívá vyvinutého modulu jazyka. A konečně vyústěním přístupu k náboženství jako vedlejšímu produktu může být odmítnutí termínu náboženství jako vědeckého termínu pro jeden jev, o kterém má potom smysl se bavit pouze v pragmatické předvědecké rovině. Boyer (2007) tvrdí, že náboženství nemohlo být produktem evoluce jednoduše proto, že žádné náboženství neexistuje, ale jedná se jen o různá spojení metapřesvědčení, reflexivních přesvědčení a praktik (Boyer, 2022).

Druhým přístupem k původu náboženství je Adaptacionalistický přístup. Vychází z toho, že náboženství je široce rozšířené a je s lidmi dost dlouhý čas, aby se ukázaly jeho možné výhody. Adaptacionalistický přístup se tedy snaží odhalit jeho možné benefity (Sosis 2019). Rozšiřuje přístup na nejenom jednotlivé prvky náboženství a náboženských přesvědčení, ale zajímá se o celek. Tvrdí, že k pochopení náboženství je třeba celý systém, nejen jedna komponenta, např. nadpřirození aktéři nebo systém norem (Sosis & Kiper, 2014).

Kognitivní vědci a evoluční psychologové náboženství, když zkoumají nějakou konkrétní vlastnost jako adaptivní, tak hledají, zda existuje obecný psychologický mechanismus, který selekce zvýhodňovala, aby jeho výsledkem bylo specifické chování – ptají se po psychologických adaptacích. Oproti tomu behaviorální ekolog se ptá na něco jiného – zda konkrétní chování vede ke zvýšené fitness jeho vykonavatelů. Určitá vlastnost je adaptivní jen v určitém prostředí, tedy X je adaptivní, pokud v prostředí vede k relativně vyšší fitness než jiné strategie. Například modlitba v jedné komunitě může přinést výhodu, ale v druhé smrt (Sosis & Bulbulia, 2011).

Dle (Sosis & Bulbia, 2011) je metodologický rozdíl mezi behaviorálně ekologickým přístupem a evolučně psychologickým. Rozdíl spočívá v tom, že se evoluční psychologie a částečně i teorie podvojně dědičnosti zabývají zakládajícími psychologickými mechanismy, které produkují adaptivní odpovědi. Teorie podvojně dědičnosti vysvětluje lidské chování jako výsledek dvou interagujících procesů, jako koevoluci genetické evoluce a kulturní evoluce. Behaviorální ekologové jsou agnostičtí vůči mechanismům, psychologickým a fyziologickým. Místo toho mají předpoklad, že selekce vytvořila mechanismus na generování chování k tomu, aby optimalizoval odpovědi v rámci omezení (*constrains, tradeoffs*) v daném prostředí (Krebs & Davies, 1997). Rozdíly jsou i v generování hypotéz: Behaviorální ekologové odvozují hypotézy z optimalizovaných evolučně stabilních modelů, které testují kvantitativními modely z etnografie a historických dat. Evoluční psychologové odvozují hypotézy z neformálního úsudku, užitím evoluční logiky, a testují je experimentálními a dotazníkovými metodami. Teorie podvojně dědičnosti odvozuje hypotézy z populačních modelů, a testuje hypotézy pomocí řady metod, např.



pomocí počítačových simulací, laboratorních a terénních experimentů a etnografických pozorování dle (Sosis & Bulbulia, 2011).

Otázka, zda je náboženství adaptivní, je konfuze obdobně, jako když se ptáme, zda náboženství pomáhá našemu zdraví a *wellbeingu*. Adaptace sama je složitý komplexní jev (Reeve & Sherman, 1993). Vlastnost je *adaptivní*, pokud zvýší fitness svému nositeli. Ale *adaptace* referuje nejen o procesu přirozeného výběru, ale i o jeho výsledku. Adaptace jako výsledek řeší konkrétní, reprodukce či přežití se týkající ekologický problém daného organismu (Sosis & Bulbulia, 2011). Naše preference sladké a slané stravy je adaptací, ale dnes nemusí být adaptivní. Vyhrát ruskou ruletu je adaptivní, ale takové chování nemusí být adaptací. Adaptace závisí na kontextu: Je například náboženské chování X adaptivní právě v prostředí A? Adaptací může být i víra v nadpřirozené aktéry či systém norem umožňujících kooperaci; náboženský systém norem přítom, zdá se, zprostředkovává kooperaci i v rámci rozsáhlých komunit. Nejde tedy jenom o jednotlivé prvky, ale jde o celek, a k pochopení náboženství je třeba celý systém.

Budeme-li nahlížet na náboženství jako na souhrn mechanismů s pozitivním vlivem na fitness, (Sosis, 2019), můžeme se zaměřovat na některé z následujících prvků: rituál, autorita, mýtus, posvátno, nadpřirozený aktér, morální povinnost, smysl. Může se jednat o skupinové a individuální efekty, a některé mechanismy mohou být adaptivními strategiemi. Důležitá však je jejich kontextualita ve smyslu jejich adaptace. Kontextem sociální normy, se právě zabývá naše vlastní studie, ale než k ní přistoupíme, je třeba nastínit jednotlivé proximální mechanismy pozitivního vlivu náboženství na zdraví a *wellbeing* v individuální rovině.

Proximální mechanismy na individuální rovině jsou patrné. Pokud určité náboženství například limituje nezdravé návyky, jako pití alkoholu a kouření (Nagel & Sgoutas-Emch, 2007), můžeme například očekávat méně onemocnění způsobených nadměrnou konzumací alkoholu a tabákových výrobků. Obdobně pokud dané náboženství vyžaduje monogamii, povede nižší počet sexuálních partnerů k nižší prevalenci sexuálně přenosných chorob a nižšímu počtu nechtěných otěhotnění (McCullough, Carter, DeWall & Corrales, 2012; Simons, Burt & Peterson, 2009). Opačný, tedy negativní vliv na duševní zdraví prostřednictvím stresu a negativních emocí má náboženská příslušnost jednotlivce v případě, že se do konfliktu dostane jeho odlišná sexuální orientace (například homosexualita) vůči konzervativnímu přístupu náboženského učení (Gibbs & Goldbach, 2015). Modlitba za zdraví může zlepšit zdravotní návyky (Harrigan, 2011) a obecně může náboženství podporovat zdravý životní styl (Rew & Wong, 2006). Modlitby, meditace nebo účast v rituálech mohou sloužit ke zvládnání stresových situací (*coping mechanism*) a snížit úzkosti a deprese nábožensky založených lidí (Kasi et al., 2012).

Zprostředkující faktor mezi na jedné straně *wellebing* a na druhé straně náboženstvím a spiritualitou jsou emoce a sebevědomí (Joshanloo & Daemi, 2015; Van Cappellen, Toth-Gauthier, Saroglou & Fredrickson, 2014). Náboženství může také podpořit další pozitivní emoce jako odpuštění, pozitivní sebehodnocení a *mindfulness* (Bajaj, Gupta & Pande, 2016).

Sociální podpora v komunitě je kontextově závislá. Je známo, že socialita, síť vztahů mezi lidmi spojenými stejnými hodnotami, je dobrá pro *wellbeing* (George, Larson, Koenig & McCoullough, 2000; Lim & Putnam, 2010), zatímco samota a sociální izolace představuje zdravotní riziko (Holt-Lunstad, Smith, Baker, Harris & Stephenson, 2015). Opačný negativní efekt se tedy ukazuje u nevěřících, kteří jsou vnímání negativně u náboženské komunity (Weber, Pargament, Kunik, Lomax & Stanley, 2012). V tomto kontextuálním podmínění vlivu náboženství nebo jeho absenci na *wellbeing* spatřujeme hlavní omezení dosavadních studií – vliv náboženství se studoval ve společnostech, s většinovým zastoupením nábožensky založených, kde náboženské vyznání a příslušnost k církvím či náboženským skupinám představují buďto jakousi výchozí standardní pozici, nebo dokonce žádoucí normu.

Je tedy možné, že výhody náboženství jsou nejen kontextuálně závislé, ale přímo kontextuálně omezené, jak píše Stavrova (2015). Studie zkoumající souvislost víry na *wellbeingu* (Nicholson, Rose & Bobak, 2009) hledala souvislost mezi účastí na bohoslužbách a zdravím jednotlivců v evropských státech. Tuto souvislost našla v některých evropských státech, tedy nižší účast na bohoslužbách je spojena s horším zdravím, ale u české populace toto nebylo signifikantní. Obdobně další nedávná studie na evropských populacích (Libby et al., 2021) ukázala pozitivní asociaci mezi (*self-reported*) zdravím a více proměnnými popisujícími náboženské přesvědčení jen v případě Německa, u dalších zemí nebyly výsledky statisticky signifikantní. V Česku byla dokonce víra v Boha (a další náboženské otázky) asociována s horším zdravím.

Za účelem prozkoumat vztah mezi náboženstvím a *wellbeingem* v kontextu, kde není většina náboženská, jsme pro náš výzkum zvolili českou populaci. Po potřeby této explorativní studie jsme metodou sněhové koule nasbírali početnou (N=31.633) skupinu online participantů. Náboženská příslušnost pro potřeby následné analýzy byla zjištěna prostřednictvím sebeurčení participantů dle náboženské víry, přesněji otázky „Jaká je vaše náboženská víra?“. Ptali jsme se specificky na náboženskou víru formou *multiple choice* doplněnou o otevřenou odpověď jiné, nikoliv na individuální spiritualitu nebo jiné její projevy. Podle této otázky jsme rozdělili participanty pro následující analýzu jako „nevěřící“, pokud zvolili možnost „nejsem věřící“.

„Věřící mimo církve“ byli participanti, kteří označili, že věří v Boha, ale nejsou členy žádné církve. Následovala možnost zvolit pole nejčastějších českých církví (Církev římskokatolická, Českobratrská církev evangelická, Církev československá husitská) a také jsme poskytli možnost „jiná“ s uvedením vlastní specifikace, otevřenou odpověď. Na základě této specifikace v otevřené odpovědi od participantů jsme spojili všechny výskyty Buddhismu jako buddhismus, obdobně pro judaismus. Další vyznání byla zastoupena ve velmi malé míře.

Na základě dotazníku jsme spočítali 5 samostatných faktorů pro *wellbeing*: fyzické zdraví, duševní zdraví, ekonomická situace, hodnocení vlastní atraktivity, kvalita současného partnerského vztahu. A současně jsme užívali faktor *wellbeingu* složený z těchto pěti zmíněných. Online dotazník zjišťoval komponenty k jednotlivým 5 faktorům. Dále jsme zjišťovali vlastní hodnocení altruismu v hypotetických situacích, které tvořilo agregovaný faktor „*self reported altruism*“. Širší (Pseudo)biologickou fitness jsme spočítali agregovaný faktor složený z počtu současných dětí, plánovaných dětí, biologických sourozenců, strýců a tet. Sexuální aktivitu jsme měřili pomocí spočítaného faktoru na základě odpovědi na standardizovaný dotazník SOI-R (Penke & Asendorpf, 2008). Baterie zjišťující religiozitu zkoumala vlastní důležitost víry v Boha, odměnu a trest po smrti. Statisticky jsme zjišťovali asociaci fokálních proměnných (zmíněných agregovaných faktorů jako např. fitness nebo fyzické zdraví) pomocí ANCOVA s pohlavím a náboženstvím participantů jako *fixed factors*, věk byl kovariát.

Naše hlavní zjištění je následující: Teorie o pozitivním vlivu náboženství na zdraví a *wellbeing* byla většinou testována zejména jen v majoritně náboženských společnostech. V kontrolní společnosti k této teorii – v ateistickém Česku z analýzy vyplývá, že náboženství neprospívá (obdobně jako v několika zmíněných studiích), fyzické zdraví i duševní zdraví negativně koreluje s religiozitou u participantů této studie. Vysvětlujeme to tím, že náboženství v kontextu, kdy je minoritní, je odklonem od normy a to s sebou přináší stres (Stavrova, 2015), sociální podpora tedy v tomto sociálním kontextu nefunguje jako zprostředkující faktor pro náboženství a *wellbeing*.

Druhým zjištěním je, že náboženská aktivita (měřená podle frekvence návštěv kostela, společných modliteb apod.) u katolíků je asociována s vyšším *wellbeingem* (oproti jiným věřícím, kteří mají tuto frekvenci nižší). To odpovídá studii, podle níž obecná aktivita v komunitě zvyšuje *wellbeing* věřícím i nevěřícím (Galen, Sharp & McNulty, 2015; Galen, 2015).

Třetím zjištěním je vyšší morální sebehodnocení věřících oproti morálnímu sebehodnocení nevěřících. U katolíků jako podskupiny jsou frekvence společenských religiózních aktivit a důležitost vlastní víry pozitivně asociovány s vyšším morálním

sebehodnocením. Možné vysvětlení této souvislosti s vyšším sebehodnocením poskytuje vyšší schopnost si odpouštět u věřících (Krause, 2017).

Čtvrtým zjištěním je pozitivní vliv náboženství na fitness, tedy na počet genů předaných do další generace prostřednictvím potomků a dalších příbuzných. Tento efekt vykazuje střední velikost a zdá se být v kontrastu s nižší sexuální aktivitou (oproti nevěřícím).

Je nutné poznamenat, že šlo o explorativní studii s větším množstvím poměrně široce pojatých faktorů. I když byla provedena Benjamini-Hochberg kontrola na vícečetné testy, představuje velké množství proměnných v explorační studii samo o sobě velkou limitaci pro případné kauzální usuzování. Taktéž velikosti zmíněných efektů byly spíše malé. Spekulovat bychom mohli o adaptivní roli náboženství ve smyslu vyšší fitness. A nakonec je třeba zmínit, že, ač naše data naznačují vztah mezi náboženstvím a zdravím a *wellbeingem*, celý systém může mít v české společnosti zároveň oboustrannou kauzalitu na několika různých úrovních. Pokud srovnáváme jen např. katolíky mezi sebou, vyšší účast na společenských náboženských aktivitách vede k benefitům. Pokud však srovnáváme celou společnost, nevěřící z tohoto srovnání vychází jako zdravější. Záleží tedy na kontextu, respektive na tom, co je v dané společnosti normální ve smyslu nejčastěji zastoupené. Na věřící mimo církve tak může dopadat dvojitý negativní účinek: chybí u nich pozitivní dopad na *wellbeing* a zdraví pramenící ze sociálních vazeb při participaci na společných náboženských aktivitách, a zároveň mají zvýšený stresový faktor způsobený tím, že jejich vnitřní přesvědčení je proti společenskému normálu.

## 1.5 Experimentální filosofie

V historii filosofických debat se účastníci často spoléhali a odvolávali na své intuice, sdílená přesvědčení a obecné názory. Avšak mnohdy opomíjeli nutnost podložit své odborné intuice o lidových konceptech empirickými daty. Současný směr „experimentální filosofie“ prokazuje jistou intelektuální pokoru tím, že zkoumá, jak lidé skutečně přemýšlí o zakládajících filosofických konceptech. Vznik a vývoj experimentální filosofie, co do rozšíření souboru metod, lze přirovnat analogicky k vzniku analytické filosofie na počátku dvacátého století. Analytická filosofie se tehdy snažila o zpřesnění filosofie a přiblížení se přírodním vědám užitím logického aparátu a logickou analýzou jazyka.

Nepovažujeme však experimentální filosofii za revoluční změnu paradigmatu v rámci celé filosofické disciplíny. Spíše ji vidíme jako extenzi metod, které přebírá od rychle se rozvíjejících sociálních věd, kognitivní vědy, a experimentální psychologie. Experimentální filosofie má zájem o stejné problémy jako dosavadní filosofie, přičemž rozšiřuje zkoumání o systematické experimentální studie, které zahrnují i statistickou analýzu.

Nabízí se otázka, zda lze „experimentální filosofii“ vůbec označit za filosofii, nebo je oddělenou vědní disciplínou bližší spíše psychologii. Historický vývoj věd je sousledem postupného vydělování odborných disciplín od filosofie v prvním kroku, nebo od jiných věd v kroku druhém. Průkopníci experimentální filosofie Joshua Knobe a Shaun Nichols ve svém manifestu rozvíjí téma, jak fungují lidské bytosti – především jak funguje lidská mysl (Knobe & Nichols, 2008). Tato, otázka byla v centru západní filosofie po celou svou dobu trvání. Co do obsahu zájmu považují její průkopníci experimentální filosofii za kontinuální s tradičními přístupy filosofie. Nová je pro filosofy výzva vstoupit do dialogu s empirickou skutečností, přičemž tradiční analytické metody zůstávají.

Alexander, Mallon a Weinberg, (2014) definovali pozitivní a negativní program experimentální filosofie. Pozitivní program považuje laické intuice ohledně filosofických konceptů za významný zdroj filosofického poznání. Důležité je, že pozitivní program nespolehá jen na vlastní intuice konkrétní jedné osoby, ale srovnává je s dalšími lidmi ve společnosti. Empirické výzkumy poskytují informace o psychologickém statusu dané intuice, která odráží pravdu o určitém filosofickém konceptu (Alexander, Mallon & Weinberg, 2014). Kritický vůči tomuto přístupu je negativní program, který argumentuje tím, že jednak intuice nejsou spolehlivým zdrojem filosofického poznání a jednak nejsou široce sdílené mezi lidmi. Existuje totiž značná disparita mezi lidovými intuicemi, které jsou charakteristické neostrými koncepty,

na jedné straně, a sofistikovanými, detailně propracovanými pojmy odborných filosofů na straně druhé.

Jedno z nejdiskutovanějších témat experimentální filosofie je tzv. Knobův efekt (Knobe, 2003; Nichols & Ulatowski, 2007). Tento efekt popisuje, že laici přisuzují intencionalitu daného jednání podle morální valence jeho vedlejšího efektu. Vedlejší efekt jednání je považován laiky za úmyslný, pokud má negativní valenci, zatímco v opačném případě (s pozitivní valencí), není pokládán za úmyslný. V Knobově efektu tedy pozorujeme, že morální soud hraje roli v přisuzování intencionality.

Problém osobní identity je nepochybně jedním z tradičních filosofických témat. Analogickou souvislostí morálního soudu s konceptem zachování osobní identity se právě zabývá pátá kapitola této disertační práce. Filozofové se snaží najít hranice změny, které vymezují podmínky zachování osobní identity člověka, přesněji řečeno, za jakých podmínek osoba zaniká a za jakých přetrvává. Základním kritériem je míra podobnosti, kterou si člověk musí zachovat během různých změn, abychom mohli usoudit, že se stále jedná o jednu a tu samou osobu (Tobia, 2015). Problém studia zachování osobní identity v rámci experimentální filosofie detailně analyzovala Michaela Jirout Košová (2020), jež je rovněž první autorkou našeho článku (Jirout Košová, Kopecký, Oulovský, Nekvida & Flegr, 2021) obsaženého v páté kapitole této disertační práce.

První a zásadní experimentální studie týkající se osobní identity provedli Nichols, Prinz a Strohminger (Prinz & Nichols, 2016; Strohminger & Nichols, 2014; Strohminger & Nichols, 2015). Ukázalo se, že klíčovými determinanty pro zachování či narušení osobní identity jsou schopnosti a vlastnosti týkající se mezilidských vztahů, např. empatie, svědomitost, krutost nebo sobectví. Koncept *diachronic identity* se týká zachování identity v čase. Prinz a Nichols (2016) zjistili, že morální kontinuita je nejdůležitější pro přetrvání osobní identity, dokonce je důležitější než paměť, aktérství nebo narativ. Koncept esenciálního morálního já (Strohminger & Nichols, 2014) říká, že právě morální složka je jádrem osobní identity. Tuto myšlenku dále autoři potvrdili v následující studii (Strohminger & Nichols, 2015), ve které blízcí pacientů trpících neurodegenerativní poruchou hodnotili zachování osobní identity. Ve variantě, kdy symptomy nemocí měnily morálku, byla tato změna hodnocena jako větší než v případě, kdy nemoc způsobovala amnézii.

Strohminger, Knobe a Newman (2017) dále přepracovali koncept já (*self*) pro laiky, přičemž zdůraznili existenci užšího konceptu pravého já (*true self*). Toto pravé já je podle nich v laické intuici považováno za to nejdůležitější, co nás hluboko uvnitř činí sebou samými.

Hypotéza, že pravé já je vnímáno obecně jako morální a vnitřně dobré, může být výsledkem toho, jak si lidé konceptualizují esence, ale může být také provázáno s naší potřebou sociálních vazeb. Morální v tomto kontextu není myšleno abstraktně jako etická teorie, míní se tím něco, co zastupuje vztahy k druhým lidem a prosociální chování.

Obdobně Heiphetz, Strohminger a Young (2017) identifikovali roli sociálních vazeb jako klíčovou pro osobní identitu. Analýza odhalila souvislost mezi vnímanou důležitostí komunity a typem morálního přesvědčení. Změny v široce sdílených morálních přesvědčeních by vedly k radikálnější změně identity než změny v kontroverzních morálních přesvědčeních. Důvodem je, že široce sdílená přesvědčení jsou více spjata se vztahy v rámci komunity.

Tobia (2016) zkoumal směr – pozitivní nebo negativní – hypotetických změn. Zaměřil se na roli mezilidských vztahů pro zachování osobní identity pomocí myšlenkového experimentu s reverzní situací slavného případu Phinease Gage (v původní historické verzi se muž po úrazu hlavy radikálně morálně změnil k horšímu, v této nové hypotetické verzi se muž po úrazu hlavy radikálně morálně zlepši). V případě morálního zlepšení participanti ve studii měli tendenci říkat, že se stále jedná o stejnou osobu. Avšak v originální verzi, kde došlo k morálnímu zhoršení, říkají spíš opak. Přestože obě změny byly v zásadě obdobně radikální, negativní varianta vedla k závěru, že identita osoby byla porušena zásadnějším způsobem. Tobia tvrdí, že v otázce osobní identity nezáleží na rozsahu změny, ale jde o přesvědčení, že pravé já je považováno za esenciálně dobré. Normativita, která je spojená se sociální a skupinovou dimenzí, tvoří základní aspekt osobní identity. Heiphetz, Strohminger, Gelman a Young (2018) obdobně zjistili, že dospělí respondenti považovali změnu dobrého morálního přesvědčení na špatné za více závažnou změnu ohrožující zachování osobní identity. U dětí tento rozdíl nezaznamenali. Domníváme se však, že absence tohoto jevu je důsledkem designu studie a malého vzorku participantů.

V naší studii (Jirout Košová et al., 2021), která je součástí 5. kapitoly této práce, jsme se zaměřili na experimentální výzkum dětských filosofických intuicí ohledně zachování osobní identity – testovali jsme teorii esenciálního morálního já ve vývojovém kontextu. Naše práce se od studie (Heiphetz et al., 2018) liší v několika zásadních ohledech. Zavedli jsme více kategorií změny, zahrnuli jsme oba směry změn, rozšířili jsme věkové kategorie a zejména jsme několikrát znásobili počet účastníků studie z řad dětí a teenagerů. Záměrem bylo prozkoumat roli morální kategorie ve srovnání s ostatními (fyzický vzhled, paměť, charakter, percepce, kognice) v tom, jak se změna v dané kategorii projeví na vnímání zachování osobní identity. Testovali jsme následující hypotézy: Morální změny budou hodnoceny jako největší změna pro osobní identitu.

Negativní změny budou hodnoceny jako silnější oproti pozitivním změnám uvnitř kategorie. Negativní morální změny budou hodnoceny jako celkově největší. Taktéž jsme plánovali exploračně popsat efekt věku a pohlaví na kategorie.

Na platformě Open Science Framework jsme studii pre-registrovali (<https://osf.io/qj7k6/>) před samotným sběrem dat. Studie se zúčastnilo celkem 267 respondentů, z nichž jsme část vyloučili z analýzy z důvodů vysokého nebo nízkého věku, protože nedostatečně porozuměli zadání nebo rozhovor nedokončili. Do analýzy bylo zařazeno 217 participantů (56,4% ženy, věk 6-15, průměrný věk 11). Participanty jsme zvali na rozhovor během Veletrhu vědy v Praze, jednalo se o volnočasový festival pro rodiny s dětmi pořádaný českou Akademií věd. Právě tímto způsobem jsme mohli nabídnout účast ve výzkumu dětem v doprovodu rodičů, aniž by situace vyžadovala jejich přítomnost v prostorách fakulty a finanční kompenzaci. Tímto jsme zajistili účast i participantů (respektive jejich rodičů), kteří nebudou motivováni finančně nebo nebudou altruistickými podporovateli vědy.

Samotné otázky pro účastníky se týkaly změny osob ve speciální sci-fi komoře. Osoba byla specifikována jako: jejich kamarád, někdo, koho znají, nebo nějaký člověk obecně (tři druhy scénáře, od nejosobnějšího až po neutrální, jejichž prostřednictvím jsme chtěli odstínit roli mezilidských vztahů). Změny byly v šesti kategoriích: fyzická – vzhled, kognitivní – inteligence, morální – láska ke druhým nebo kruté/vlídlné jednání, charakter – lenost, paměť – životní zkušenost, percepce – zrak. Každá z kategorií byla uvedena v negativním i pozitivním směru změny. Dohromady odpovídal každý z participantů na 14 otázek. Otázka pokaždé zněla, jak moc se změní ta nejdůležitější část osoby, která ji dělá tím, kým doopravdy je. Škála byla od nuly – je stále stejným člověkem a stále má své nejdůležitější vlastnosti, až do šestky – není tím samým člověkem a ztratil své nejdůležitější vlastnosti. Likertovu škálu pro odpovědi jsme ilustrovali kombinací čísel a kruhů s rostoucí velikostí, čímž jsme se snažili napomoci dětem s dosud nerozvinutými čtenářskými dovednostmi.

Ve studii jsme zjistili následující: změny v morální kategorii měly ve výpovědích respondentů ze všech kategorií největší vliv na vnímanou změnu osobní identity. Negativní změna byla vnímána jako signifikantně větší změna já než pozitivní, ve všech kategoriích kromě fyzické. Nejvýznamnější byla morální kategorie, ovšem také ztráta paměti byla vnímána jako mnohem závažnější změna, než pozitivní změna získání super-paměti. Explorační část měla popsat vliv věku jako biologického faktoru na hodnocení zachování osobní identity. Pro tyto potřeby jsme rozdělili participanty dle věku do tří kategorií, 6-8, 9-11 a 12-15 (dle vývojových fází viz Gibbs, 2019). Zaznamenali jsme, že relativní důležitost morální kategorie roste s věkem.



Nejvýznamnější změna zřejmě nastává okolo věku devíti let, zvýrazňuje se obzvláště kontrast mezi morální a fyzickou kategorií, zatímco paměť zůstává blízko morální kategorie. Taktéž jsme zachytili význam osobnějšího scénáře, při kterém byla morální kategorie vnímána jako relativně důležitější oproti ostatním kategoriím než u neutrálního scénáře.

Naše výsledky podpořily teorii i u dětí, že morální kontinuita je zásadní pro zachování osobní identity – teorii pravého morálního já. Osobní identita je vnímána jako porušená, zvláště pokud dochází k negativní morální změně. Taktéž role paměti/vzpomínek je důležitá, specificky paměti vázané na mezilidské vztahy (vzpomínky na blízké lidi apod.). Ostatně už od Locka až po současné výzkumy (Nichols & Bruno, 2010) se ukazuje, že vzpomínky a psychologické charakteristiky jsou důležitější pro zachování osobní identity než fyzické charakteristiky. Zároveň i v naší studii s dětmi platí, že morálka je důležitější než paměť (Prinz & Nichols, 2016). Oproti obdobné studii (Heiphetz et al., 2018) se nám podařilo i u dětí replikovat stejný efekt pravého morálního já, jako u dospělých.

## 1.7 Reference

- Ajioka, J. W., & Soldati, D. (Eds.) (2007). *Toxoplasma: molecular and cellular biology*. London, England: Horizon.
- Alexander, J., Mallon, R., & Weinberg, J. M. (2014). Accentuate the negative. *Experimental philosophy*, 2, 31-50.
- Ayala, F. J. (1987). The biological roots of morality. *Biology and philosophy*, 2, 235-252.
- Bajaj, B., Gupta, R., & Pande, N. (2016). Self-esteem mediates the relationship between mindfulness and well-being. *Personality and Individual Differences*, 94, 96-100.
- Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends in cognitive sciences*, 4(1), 29-34.
- Bártová, E., Sedlák, K., & Literák, I. (2006). Prevalence of *Toxoplasma gondii* and *Neospora caninum* antibodies in wild boars in the Czech Republic. *Veterinary Parasitology*, 142(1-2), 150-153.
- Behnke, J. M., & Barnard, C. J. (1990). Coevolution of parasites and hosts: host-parasite arms races and their consequences. *Parasites: immunity and pathology. The consequences of parasitic infection in mammals.*, 1-22.
- Berdoy, M., Webster, J. P., & Macdonald, D. W. (2000). Fatal attraction in rats infected with *Toxoplasma gondii*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267(1452), 1591-1594.
- Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science*, 352(6293), 1573-1576.
- Boyer, P. (2003). Religious thought and behaviour as by-products of brain function. *Trends in cognitive sciences*, 7(3), 119-124.
- Boyer, P. (2007). *Religion explained: The evolutionary origins of religious thought*. Hachette UK.
- Boyer, P. (2022). Religious Beliefs. In J. Musolino, J. Sommer, & P. Hemmer (Eds.), *The Cognitive Science of Belief: A Multidisciplinary Approach* (pp. 235-253). Cambridge: Cambridge University Press. doi:10.1017/9781009001021.018
- Bulbulia, J., & Sosis, R. (2011). Signalling theory and the evolution of religious cooperation. *Religion*, 41(3), 363-388.
- Costa, D. G. C., Marvulo, M. F. V., Silva, J. S. A., Santana, S. C., Magalhães, F. J. R., Filho, C. L., Ribeiro, V. O., Alves, L. C., Mota, R. A., Dubey, J. P. & Silva, J. C. R. (2012). Seroprevalence of *Toxoplasma gondii* in domestic and wild animals from the Fernando de Noronha, Brazil. *Journal of Parasitology*, 98(3), 679-680.

- Curry, O. S. (2016). Morality as cooperation: A problem-centred approach. *The evolution of morality*, 27-51.
- Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological science*, 17(12), 1082-1089.
- Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. John Murray, London.
- Dawkins, Richard & Krebs, John R. (1978). Animal Signals: Information or Manipulation? In J. R. Krebs & N. B. Davies (eds.), *Behavioural Ecology: An Evolutionary Approach*. pp. 282–309.
- Day, M. (2008). Godless savages and superstitious dogs: Charles Darwin, imperial ethnography, and the problem of human uniqueness. *Journal of the History of Ideas*, 69(1), 49-70.
- De Waal, F. (1996). *Good natured*. Harvard University Press.
- Dubey, J. P., & Carpenter, J. L. (1993). Histologically confirmed clinical toxoplasmosis in cats: 100 cases (1952-1990). *Journal of the American Veterinary Medical Association*, 203(11), 1556-1566.
- Dutton, D. (2009). *The art instinct: Beauty, pleasure, & human evolution*. Oxford University Press, USA.
- Edmonds, D. (2013). *Would you kill the fat man? The trolley problem and what your answer tells us about right and wrong*. Princeton University Press.
- Engelmann, J. M., Herrmann, E., & Tomasello, M. (2012). Five-year olds, but not chimpanzees, attempt to manage their reputations. *PLoS One*, 7(10), e48433.
- Fehrler, S., & Przepiorka, W. (2013). Charitable giving as a signal of trustworthiness: Disentangling the signaling benefits of altruistic acts. *Evolution and Human Behavior*, 34(2), 139-145.
- Fincher, C. L., Thornhill, R., Murray, D. R., & Schaller, M. (2008). Pathogen prevalence predicts human cross-cultural variability in individualism/collectivism. *Proceedings of the Royal Society B: Biological Sciences*, 275(1640), 1279-1285.
- Flegr, J. (2023). Postadaptations. In: Shackelford, T.K. (eds) *Encyclopedia of Sexual Psychology and Behavior*. Springer, Cham. [https://doi.org/10.1007/978-3-031-08956-5\\_385-1](https://doi.org/10.1007/978-3-031-08956-5_385-1).
- Flegr, J., & Escudero, D. Q. (2016). Impaired health status and increased incidence of diseases in Toxoplasma-seropositive subjects—an explorative cross-sectional study. *Parasitology*, 143(14), 1974-1989.
- Flegr, J., Havlíček, J., Kodým, P., Malý, M., & Smahel, Z. (2002). Increased risk of traffic accidents in subjects with latent toxoplasmosis: a retrospective case-control study. *BMC infectious diseases*, 2, 1-6.
- Flegr, J., Hlavacova, J., & Toman, J. (2023). Parasitic manipulation or side effects? The effects of past Toxoplasma and Borrelia infections on human personality and cognitive performance are not mediated by impaired health. *medRxiv*, 2023-01.

- Flegr, J., Klapilová, K., & Kaňková, Š. (2014). Toxoplasmosis can be a sexually transmitted infection with serious clinical consequences. Not all routes of infection are created equal. *Medical hypotheses*, 83(3), 286-289.
- Flegr, J., Lenochová, P., Hodný, Z., & Vondrová, M. (2011). Fatal attraction phenomenon in humans—cat odour attractiveness increased for *Toxoplasma*-infected men while decreased for infected women. *PLoS neglected tropical diseases*, 5(11), e1389.
- Flegr, J., Prandota, J., Sovičková, M., & Israili, Z. H. (2014). Toxoplasmosis—a global threat. Correlation of latent toxoplasmosis with specific disease burden in a set of 88 countries. *PLoS one*, 9(3), e90203.
- Foot, P. (2002). The problem of abortion and the doctrine of the double effect. In: Foot P (ed) *Virtues and vices and other essays in moral philosophy*. Clarendon Press, Oxford, pp 19–33 (Original work published 1967)
- Galen, L. (2015). Atheism, wellbeing, and the wager: Why not believing in God (with others) is good for you. *Science, Religion and Culture*, 2(3), 54–69.
- Galen, L. W., Sharp, M., & McNulty, A. (2015). Nonreligious group factors versus religious belief in the prediction of prosociality. *Social Indicators Research*, 122, 411-432.
- George, L. K., Larson, D. B., Koenig, H. G., & McCullough, M. E. (2000). Spirituality and health: What we know, what we need to know. *Journal of social and clinical psychology*, 19(1), 102-116.
- Gert, B., & Gert, J. (2020, September 8). The definition of morality. *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/fall2020/entries/morality-definition/>
- Gibbs, J. C. (2019). *Moral development and reality: Beyond the theories of Kohlberg, Hoffman, and Haidt*. Oxford University Press.
- Gibbs, J. J., & Goldbach, J. (2015). Religious conflict, sexual identity, and suicidal behaviors among LGBT young adults. *Archives of suicide research*, 19(4), 472-488.
- Gould, S. J., Lewontin, R., & Anderson, C. M. (1979). A philosophical critique of the arguments presented in *The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme*. *Proceedings of the Royal Society B: Biological Sciences*, 205, 581-589.
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of personality and social psychology*, 101(2), 366.
- Gray, K., & Wegner, D. M. (2010). Blaming God for our pain: Human suffering and the divine mind. *Personality and Social Psychology Review*, 14(1), 7-16.
- Greene, J. (2014). *Moral tribes: Emotion, reason, and the gap between us and them*. Penguin.
- Greene, J. D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends in cognitive sciences*, 11(8), 322-323.

- Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*(3), 1144-1154.
- Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review*, *108*(4), 814.
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Vintage.
- Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social justice research*, *20*(1), 98-116.
- Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*(4), 55-66.
- Hardy, C. L., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin*, *32*(10), 1402-1413.
- Harrigan, J. T. (2011). Health promoting habits of people who pray for their health. *Journal of Religion and Health*, *50*, 602-607.
- Hayward, R. D., & Krause, N. (2013). Religion, mental health, and well-being: Social aspects. *Religion, personality, and social behavior*, 265-290.
- Heiphetz, L., Strohminger, N., & Young, L. L. (2017). The role of moral beliefs, memories, and preferences in representations of identity. *Cognitive science*, *41*(3), 744-767.
- Heiphetz, L., Strohminger, N., Gelman, S. A., & Young, L. L. (2018). Who am I? The role of moral beliefs in children's and adults' understanding of identity. *Journal of Experimental Social Psychology*, *78*, 210-219.
- Hodková, H., Kolbeková, P., Skallová, A., Lindová, J., & Flegr, J. (2007). Higher perceived dominance in *Toxoplasma* infected men--a new evidence for role of increased level of testosterone in toxoplasmosis-associated changes in human behavior. *Neuroendocrinology Letters*, *28*(2), 110-114.
- Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality: a meta-analytic review. *Perspectives on psychological science*, *10*(2), 227-237.
- Houdek, Petr. "Puppet master: possible influence of the parasite *Toxoplasma gondii* on managers and employees." *Academy of Management Perspectives* *31*, no. 1 (2017): 63-81.
- Hume, D. (1998). *An enquiry concerning the principles of morals: a critical edition* (Vol. 4). Oxford University Press. (Original work published 1751)
- Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, *323*(5918), 1222-1226.
- Chidester, D. (2009). DARWIN'S DOGS: Animals, animism, and the problem of religion. *Soundings: An Interdisciplinary Journal*, *92*(1/2), 51-75.

- Jirout Košová, M. (2020). Skúmanie významu experimentálnej filozofie skrze koncept osobnej identity. *Filosofický časopis (Philosophical Journal)*, 68(4), 581-603.
- Johnson, S. K., Fitza, M. A., Lerner, D. A., Calhoun, D. M., Beldon, M. A., Chan, E. T., & Johnson, P. T. (2018). Risky business: linking *Toxoplasma gondii* infection and entrepreneurship behaviours across individuals and countries. *Proceedings of the Royal Society B: Biological Sciences*, 285(1883), 20180822.
- Joshanloo, M., & Daemi, F. (2015). Self-esteem mediates the relationship between spirituality and subjective well-being in Iran. *International Journal of Psychology*, 50(2), 115-120.
- Kaňková, Š., Holáň, V., Zajícová, A., Kodym, P., & Flegr, J. (2010). Modulation of immunity in mice with latent toxoplasmosis – the experimental support for the immunosuppression hypothesis of *Toxoplasma*-induced changes in reproduction of mice and humans. *Parasitology Research*, 107, 1421-1427.
- Kasi, P. M., Naqvi, H. A., Afghan, A. K., Khawar, T., Khan, F. H., Khan, U. Z., ... & Khan, H. M. (2012). Coping styles in patients with anxiety and depression. *International Scholarly Research Notices*, 2012.
- Khademvatan, S., Saki, J., Yousefi, E., & Abdizadeh, R. (2013). Detection and genotyping of *Toxoplasma gondii* strains isolated from birds in the southwest of Iran. *British Poultry Science*, 54(1), 76-80.
- Klebl, C., Luo, Y., Tan, N. P. J., Ern, J. T. P., & Bastian, B. (2021). Beauty of the Beast: Beauty as an important dimension in the moral standing of animals. *Journal of Environmental Psychology*, 75, 101624.
- Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis*, 63(3), 190-194.
- Knobe, J., & Nichols, S. (2008). An experimental philosophy manifesto. In J. Knobe & S. Nichols (Eds.), *Experimental philosophy* (pp. 3–14). Oxford University Press.
- Kocazeybek, B., Oner, Y. A., Turksoy, R., Babur, C., Cakan, H., Sahip, N., ... & Torun, M. M. (2009). Higher prevalence of toxoplasmosis in victims of traffic accidents suggest increased risk of traffic accident in *Toxoplasma*-infected inhabitants of Istanbul and its suburbs. *Forensic science international*, 187(1-3), 103-108.
- Kolbekova, P., Kourbatova, E., Novotna, M., Kodym, P., & Flegr, J. (2007). New and old risk-factors for *Toxoplasma gondii* infection: prospective cross-sectional study among military personnel in the Czech Republic. *Clinical Microbiology and Infection*, 13(10), 1012-1017.
- Krause, N. (2017). Religious involvement and self-forgiveness. *Mental Health, Religion & Culture*, 20(2), 128-142.
- Krebs, J. R., & Davies, N. B. (Eds.). (2009). *Behavioural ecology: an evolutionary approach*. John Wiley & Sons. (Original work published 1997)
- Laidre, M. E., & Johnstone, R. A. (2013). Animal signals. *Current Biology*, 23(18), R829-R833.

- Lambert, L., Lomas, T., van de Weijer, M. P., Passmore, H. A., Joshanloo, M., Harter, J., Ishikawa, Y., Lai, A., Kitagawa, T., Chen, D., Kawakami, T., Miyata, H., & Diener, E. (2020). Towards a greater global understanding of wellbeing: A proposal for a more inclusive measure. *International Journal of Wellbeing*, *10*(2), 1–18.
- Lee, M., Sul, S., & Kim, H. (2018). Social observation increases deontological judgments in moral dilemmas. *Evolution and Human Behavior*, *39*(6), 611-621.
- Libby, G., Zimmer, Z., Kingston, A., Haviva, C., Chiu, C. T., Ofstedal, M. B., Saito, Y., & Jagger, C. (2021). Are Religiosity and Spirituality Related to Self-Reported Health Expectancy? An Analysis of the European Values Survey. *Journal of Religion and Health*.
- Lim, C., & Putnam, R. D. (2010). Religion, social networks, and life satisfaction. *American sociological review*, *75*(6), 914-933.
- Lin, P. (2016). *Why ethics matters for autonomous cars. Autonomous driving: Technical, legal and social aspects*, 69-85.
- Lindová, J., Kubena, A. A., Šturcová, A., Krivohlavá, R., Novotná, M., Rubešová, A., Havlíček, J., Kodým, P. & Flegr, J. (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.
- Lindová, J., Příplatová, L., & Flegr, J. (2012). Higher extraversion and lower conscientiousness in humans infected with *Toxoplasma*. *European Journal of Personality*, *26*(3), 285-291.
- McCullough, M. E., Carter, E. C., DeWall, C. N., & Corrales, C. M. (2012). Religious cognition down-regulates sexually selected, characteristically male behaviors in men, but not in women. *Evolution and Human Behavior*, *33*(5), 562-568.
- Meyer, C. J., Cassidy, K. A., Stahler, E. E., Brandell, E. E., Anton, C. B., Stahler, D. R., & Smith, D. W. (2022). Parasitic infection increases risk-taking in a social, intermediate host carnivore. *Communications Biology*, *5*(1), 1180.
- Moieni, M., & Eisenberger, N. I. (2018). Effects of inflammation on social processes and implications for health. *Annals of the New York Academy of Sciences*, *1428*(1), 5-13.
- Moore, J. (2002). *Parasites and the behavior of animals*. Oxford University Press.
- Nagel, E., & Sgoutas-Emch, S. (2006). The Relationship Between Spirituality, Health Beliefs, and Health Behaviors in College Students. *Journal of Religion and Health*, *46*(1), 141–154.
- Nichols, S., & Bruno, M. (2013). Intuitions about personal identity: An empirical study. In *Experimental Philosophy and its Critics* (pp. 11-30). Routledge.
- Nichols, S., & Ulatowski, J. (2007). Intuitions and individual differences: The Knobe effect revisited. *Mind & Language*, *22*(4), 346-365.
- Nicholson, A., Rose, R., & Bobak, M. (2009). Association between attendance at religious services and self-reported health in 22 European countries. *Social Science & Medicine*, *69*(4), 519-528.

- Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: a more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of personality and social psychology*, 95(5), 1113.
- Poirotte, C., Kappeler, P. M., Ngoubangoye, B., Bourgeois, S., Moussodji, M., & Charpentier, M. J. (2016). Morbid attraction to leopard urine in *Toxoplasma*-infected chimpanzees. *Current Biology*, 26(3), R98-R99.
- Prinz, J. J., & Nichols, S. (2016). Diachronic identity and the moral self. In *The Routledge handbook of philosophy of the social mind* (pp. 465-480). Routledge.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience?. *Personality and social psychology review*, 8(4), 364-382.
- Reeve, H. K., & Sherman, P. W. (1993). Adaptation and the goals of evolutionary research. *The Quarterly Review of Biology*, 68(1), 1-32.
- Rew, L., & Wong, Y. J. (2006). A systematic review of associations among religiosity/spirituality and adolescent health attitudes and behaviors. *Journal of adolescent health*, 38(4), 433-442.
- Shultz, S., Opie, C., & Atkinson, Q. D. (2011). Stepwise evolution of stable sociality in primates. *Nature*, 479(7372), 219-222.
- Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The "big three" of morality (autonomy, community, divinity) and the "big three" explanations of suffering. In A. M. Brandt & P. Rozin (Eds.), *Morality and health* (pp. 119–169). Taylor & Frances/Routledge.
- Schaller, M., & Park, J. H. (2011). The behavioral immune system (and why it matters). *Current directions in psychological science*, 20(2), 99-103.
- Simons, L. G., Burt, C. H., & Peterson, F. R. (2009). The effect of religion on risky sexual behavior among college students. *Deviant Behavior*, 30(5), 467-485.
- Smith, A. (2010). *The theory of moral sentiments*. Penguin. (Original work published 1759)
- Sosis, R. (2019). The building blocks of religious systems: Approaching religion as a complex adaptive system. In *Evolution, development and complexity: Multiscale evolutionary models of complex adaptive systems* (pp. 421-449). Springer International Publishing.
- Sosis, R., & Kiper, J. (2014). Religion Tell Us about Religious Commitments. *Challenges to Moral and Religious Belief: Disagreement and Evolution*, 256.
- Stavrova, O. (2015). Religion, self-rated health, and mortality: Whether religiosity delays death depends on the cultural context. *Social Psychological and Personality Science*, 6(8), 911-922.
- Strohming, N., & Nichols, S. (2014). The essential moral self. *Cognition*, 131(1), 159-171.
- Strohming, N., & Nichols, S. (2015). Neurodegeneration and identity. *Psychological Science*, 26(9), 1469-1479.



- Strohming, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept distinct from the self. *Perspectives on Psychological Science*, 12(4), 551-560.
- Sýkorová, K., & Flegr, J. (2021). Faster life history strategy manifests itself by lower age at menarche, higher sexual desire, and earlier reproduction in people with worse health. *Scientific reports*, 11(1), 11254.
- Tenter, A. M., Heckerath, A. R., & Weiss, L. M. (2000). *Toxoplasma gondii*: from animals to humans. *International journal for parasitology*, 30(12-13), 1217-1258.
- Thomson, J. J. (1984). *The trolley problem*. Yale LJ, 94, 1395.
- Thornhill, R., & Fincher, C. L. (2014). *Stress Theory of Values and Sociality*. Berlin: Springer.
- Thornhill, R., & Gangestad, S. W. (1993). Human facial beauty: Averageness, symmetry, and parasite resistance. *Human nature*, 4, 237-269.
- Tobia, K. P. (2015). Personal identity and the Phineas Gage effect. *Analysis*, 75(3), 396-405.
- Tobia, K. P. (2016). Personal identity, direction of change, and neuroethics. *Neuroethics*, 9, 37-43.
- Tomasello, M. (2016). *A natural history of human morality*. Harvard University Press.
- Tooby, J., & DeVore, I. (1987). The reconstruction of hominid behavioral evolution through strategic modeling. *The evolution of human behavior: Primate models*, 183-237.
- Tremlin, T. (2010). *Minds and gods: The cognitive foundations of religion*. Oxford University Press.
- Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annu. Rev. Neurosci.*, 31, 411-437.
- Van Cappellen, P., Toth-Gauthier, M., Saroglou, V., & Fredrickson, B. L. (2016). Religion and well-being: The mediating role of positive emotions. *Journal of Happiness studies*, 17, 485-505.
- Vyas, A., & Sapolsky, R. (2010). Manipulation of host behaviour by *Toxoplasma gondii*: what is the minimum a proposed proximate mechanism should explain? *Folia parasitologica*, 57(2), 88.
- Weber, S. R., Pargament, K. I., Kunik, M. E., Lomax, J. W., & Stanley, M. A. (2012). Psychological distress among religious nonbelievers: A systematic review. *Journal of religion and health*, 51, 72-86.
- Webster, J. P., & McConkey, G. A. (2010). *Toxoplasma gondii*-altered host behaviour: clues as to mechanism of action. *Folia parasitologica*, 57(2), 95.
- Wilson, E. O. (2000). *Sociobiology: The new synthesis*. Harvard University Press. (Original work published 1975)
- Zahavi, A. (1975). Mate selection – selection for a handicap. *Journal of theoretical Biology*, 53(1), 205-214.

## **2. Morální soud**

### **2.1 Morální problémy autonomních vozidel**

---

## Esej

---

# Morální problémy autonomních vozidel<sup>1</sup>

**Robin Kopecký** —

Centrum Karla Čapka pro výzkum hodnot ve vědě a technice, Praha  
Přírodovědecká fakulta Univerzity Karlovy, Praha  
robin.kopecky@natur.cuni.cz

### 1. Úvod

Co jsou autonomní vozidla (AV)? Technická terminologie rozlišuje celou řadu kategorií,<sup>2</sup> pro potřeby filosofické eseje budou stačit následující: Asistent řidiči pomůže například s parkováním nebo udržováním rychlosti. Automat je již schopen sám dojet z místa A na místo B. A nakonec, autonomní vůz umí nejen vykonat vše předcházející, ale je navíc vybaven aktivními prvky: kamerami, radarem a jinými senzory, a hlavně komunikuje – vyměňuje si informace s dalšími účastníky provozu. Budeme dále jako AV souhrnně (zjednodušeně) označovat takový silniční dopravní prostředek, který je vybaven samořídícím softwarem, jenž je nezávislý na posádce.

Jaký je účel AV? Slouží k dopravě lidí a věcí. Pointa je v tom, že AV to po všech stranách zvládnou lépe než stávající kombinace auta a lidského řidiče.

---

1 Tento příspěvek vznikl s podporou Grantové agentury Univerzity Karlovy v rámci projektu GA UK č. 929216: „Faktory ovlivňující altruismus“.

2 Dle mezinárodní standardizace: SAE On-Road Automated Vehicle Standards Committee et al., Taxonomy and Definitions for Terms Related to On-road Motor Vehicle Automated Driving Systems. SAE Standard J3016, 01–16, 2014.

Ukažme si to na příkladu dopravních nehod v EU.<sup>3</sup> I když se investují značné prostředky do zvýšení bezpečnosti na evropských silnicích, ročně na nich zemře 25–30 tisíc lidí a okolo 130 tisíc se zraní. Vidíme, že problém leží přesně mezi volantem a sedadlem: je jím řidič. V USA<sup>4</sup> je 94 % nehod zapříčiněno chybou řidiče a k polovině z nich dojde na dálnici, tedy ve snadné dopravní situaci. Nahradíme-li řidiče něčím (nebo někým?), co/kdo bude dělat méně chyb, jen v EU tím každoročně zachráníme kolem 30 tisíc lidí a předejdeme tisícům zraněných. Řídící software – na rozdíl od řidiče – nejezdí pod vlivem drog, nepíše textové zprávy a neusíná. Zavedení AV by kromě toho, že ubude mrtvých, přineslo další nesporná pozitiva: synchronizovaný provoz AV sníží emise a zrychlí dopravu. Lidé příliš mladí na to, aby sami řídili, nebo lidé se zdravotním hendikepem získají možnost větší mobility. Řidiči nebudou ztrácet čas a nervy řízením, namísto toho mohou během cesty spát nebo užívat AV jako svou mobilní pracovní kancelář. AV nemusí nezbytně vést ani ke zvýšení počtu automobilů, ani k poklesu využívání hromadné dopravy. Lidé nebudou muset vlastnit auto, zvláště když jej využívají jen několik procent dne, většinou na dojíždění do práce. Sdílení aut bude pohodlné – auta se budou sama půjčovat i vracet. Městská hromadná doprava nebude limitována počtem zkušených řidičů. Vidíme jen samé výhody. Měli bychom proto AV jako racionální volbu zavést všude a co nejdříve? Technici jsou skvělí v řešení problémů našeho světa, činí život kvalitnějším – v této esejí se budu ale snažit dostat roli filosofa, totiž ke každému (technickému) řešení, se kterým přijde inženýr, dosadit další (filosofický) problém.

## 2. Filosofické problémy techniky minulé, brzy současné a budoucí

Dříve než obrátíme pozornost k hlavnímu problému – autonomním vozidlům, zasadíme tento problém do historického kontextu. V průběhu minulých dvou staletí docházelo díky technologickému pokroku v každodenním životě nejen ke kvantitativním změnám, kdy byly například olejové lampy hromadně nahrazeny kvalitnějším elektrickým osvětlením, ale také ke změnám kvalitativním, nejčastěji spočívajícím v náhradě živého aktéra strojem. Například namísto aktéra–koně se stal původcem pohybu parní stroj nebo spalovací motor. Živý aktér v určité podobě – strojuvůdce ovšem nadále zů-

3 Údaj ze zprávy Evropské komise [online]. Dostupné na: [https://ec.europa.eu/transport/road\\_safety/specialist/statistics\\_en](https://ec.europa.eu/transport/road_safety/specialist/statistics_en); [cit. 22. 9. 2018]. Souborná data Eurostatu [online]. Dostupné na: [http://ec.europa.eu/eurostat/statistics-explained/index.php/Road\\_safety\\_statistics\\_at\\_regional\\_level](http://ec.europa.eu/eurostat/statistics-explained/index.php/Road_safety_statistics_at_regional_level); [cit. 22. 9. 2018].

4 Česká data o nehodovosti zpracovává Centrum dopravního výzkumu [online], přehledné shrnutí příčin je dostupné na: <http://www.czrso.cz/clanky/hloubkova-analyza-silnicnich-dopravnich-nehod-hlavni-priciny-vzniku-nehod/>; [cit. 22. 9. 2018].

stával na dohled pasažérům, sice už nebyl původcem pohybu, ale měl nad dějem alespoň kontrolu. Jakmile došlo k eliminaci i tohoto aktéra, například ve výtažích, vzbudila jeho absence mezi veřejností nejistotu. Strach z mechanismu fungujícího bez živého aktéra se zde projevil ve své nejprostší podobě.

Dalším příkladem revoluční kvantitativní změny byla automatizace výroby, která způsobila nejen redukci, ale i úplné zrušení některých pracovních pozic. Jako příklad uveďme pozici „počítače“, která spočívala ve vykonávání pomocných numerických výpočtů. „Počítačem“ byl člověk jen se svou (rozšířenou) myslí, tužkou, papírem, tabulkami a pravítkem. Během první světové války počítali lidské počítače trajektorie pro artilerii, což je úloha, kterou bylo možné koncipovat lidskou myslí. Počítající si mohli představit letící projektil a hlavně věděli, co a proč počítají. Ve velkolepém projektu „Manhattan“ byly během druhé světové války zaměstnány stovky takových lidských počítačů – jejichž pozice byla ovšem diametrálně odlišná. Nejenže si nijak nemohli představit formální matematické postupy, ale dokonce ani nesměli vědět, proč takové výpočty dělají a k čemu budou následně sloužit. Dnes pojmenování počítač plně náleží stroji složenému z křemíkového *hardware*, tedy už nikoliv člověku z *wetwaru*, z masa a kostí. Je tomu tak kvůli rychlosti a přesnosti a nemá ani smysl tyto dva typy „počítačů“ srovnávat. Na těchto dvou jednoduchých příkladech je patrné, že došlo k náhradě prosté mechanické – fyzické na úrovni pohybu, ale i intelektuální a kreativní na kognitivní úrovni, přičemž obě změny byly kvalitativní.

Mezi prosté problémy současnosti a blízké budoucnosti můžeme zařadit skutečnost, že jak v činnostech, které vyžadují „strojovou“ přesnost, rychlost apod., tak i v činnostech kreativních člověk prohrává v přímém souboji se stroji. Již v roce 2006 porazil počítač světového velmistra v šachu Garriho Kasparova. Za předpokladu, že se náhrada stroji stane levnější a produktivnější alternativou využití člověka, je na místě se v západním světě obávat o pracovní uplatnění lidí s nízkou odborností. Část lidstva by se mohla snadno stát na trhu práce nezaměstnatelnou. Jistou analogii vidím v historii armád. Dnes již není třeba „*kanónenfutru*“ z řad poddaných, protože současní „pěšáci“ v poli obsluhují specializovanou techniku – pro nespécializované osoby tedy není využití a nemá smysl je masově nabírat do armády. Zmíněné problémy zastarávání a neefektivit člověka se už staly jakýmsi klišé debat o automatizaci. Je možné, že daný problém zmírní nebo vyřeší neviditelná ruka trhu. Mezi předpokládané scénáře patří přesunutí volných pracovníků z výroby do služeb, zkrácení pracovní doby a možná i zavedení nepodmíněného příjmu zajišťujícího základní potřeby těch nejméně uplatnitelných.

To, co přináší zásadní kvalitativní změny, a to, co je revoluční a prakticky nevratné, souborně nazývám „*superproblémy budoucnosti*“. Prvním z nich je „*superdostupnost*“ technologií. Důsledkem snižování nákladů na výro-

bu efektivních strojů a zdvojnásobení výkonu výpočetní kapacity bude dle Mooreova zákona<sup>5</sup> již za 18 měsíců větší dostupnost výkonných technologií. Poté bude teoreticky možné, aby si každý doma v garáži vyvinul nový vzduchem šířitelný smrtelný virus nebo si sestrojil kufříkovou atomovou bombu. Dokud byly prostředky s potenciálně apokalyptickými účinky v rukou jen malého počtu mocenských skupin – států, bylo možné předpokládat, že zastanou role racionálních hráčů. V případě, že se taková moc dostane do rukou téměř kohokoliv, bude na základě zákona velkých čísel téměř jisté, že bude „*superdostupná*“ technologie zneužita šilným jednotlivcem a dojde ke zkáze celého lidstva.

Druhým možným problémem je „*supernerovnost*“ mezi příslušníky rodu *Homo sapiens*. Ta by byla důsledkem masivního kognitivního a fyzického vylepšování člověka – „*enhancementu*“, který by za pomoci genového inženýrství a technologií učinil z části lidí třídu nejen aristokracie, ale přímo bohů. Takto vylepšení nadlidé by neměli trpět žádnými chorobami, stali by se dlouhověkými nebo téměř nesmrtelnými, krásnými, mentálně extrémně pokročilými. Naopak zbytek lidstva by se v porovnání s nimi stal bezvýznamným a zcela postradatelným, obdobně, jako jsou pro dnešního *Homo sapiens* postradatelnými a postradatelnými i naši nejbližší příbuzní, primáti. Můžeme sice namítnout, že rozložení bohatství v dnešním světě již připomíná tuto „*supernerovnost*“, ale díky konečnosti života jednotlivce a meznímu užítku spotřeby nemohou mít lidé stojící na vrcholu společenské pyramidy celkové životní štěstí o tolik vyšší než ostatní. Velké navyšování délky života, schopnosti konzumovat a chápat by tyto limity prolomilo.

Posledním velkým problémem je hrozba, o které v současné době mluví přední světoví veřejní intelektuálové, jako například Bill Gates, Elon Musk nebo Sam Harris. Tato hrozba je známa jako „*superintelligence*“.<sup>6</sup> Tento koncept představuje obecnou umělou inteligenci, která řádově převyšuje lidský intelekt a která je již schopna sama sebe vylepšovat a zvyšovat svůj výkon. *Superintelligence* by tedy v budoucnosti byla zcela superiorní lidské společnosti a bylo by pouze na ní samotné, co by s naším světem a tím pádem s i lidmi provedla, respektive není znám důvod, proč by jí na lidstvu mělo vůbec záležet. Není také zřejmé, jak by jí bylo možné kontrolovat.

Jak tento stručný historický a futuristický výčet souvisí s autonomními vozidly? Společně sdílí motivaci, jež filosofa podněcuje, aby se zabýval společenskými a etickými dopady technických inovací. Někdy totiž až v souvislos-

5 Nejedná se o přírodní zákon, ale o predikci na základě pozorování. A ta má své fyzikální limity.

6 Bezpečností umělé inteligence se zabývají především experti na poli IT, filosofický náhled tohoto problému představil např. Nick Bostrom: Bostrom, N., *Superintelligence: Paths, Dangers, Strategies*. Oxford, Oxford University Press 2014.

tí se „*superproblémem*“ vyjde najevo, že je žádoucí si v optimálním případě kvalitativní změny promyslet ještě předtím, než nastanou. Je také třeba důsledně zvážit, kam jako společnost plánujeme směřovat (jaké jsou naše žádosti) a co považujeme za dobré. To všechno je ale nutné udělat ještě předtím, než učiníme nevratné kroky – než sestavíme „*doomsday machine*“ do kapsy nebo si postavíme svého vlastního tyrana v podobě *superintelligence*.

Technologické inovace zpravidla předcházejí právním a někdy i morálním normám. I když je tento celkem banální výrok nadužíván (ostatně, proč by naši zákonodárci měli psát zákony týkající se inovací, o kterých ani nevíme, že o nich někdo může uvažovat), nic to nemění na jeho relevantnosti zejména pro AV. Technologie se nachází už ve fázi velmi slibných prototypů nasazených v ostrém provozu, ovšem legislativa v našem domácím prostředí schází a diskuze o morálních aspektech AV ještě nebyla seriózně otevřena. Bude nám stačit jen aplikovat současné morální konvence a etické teorie platné pro řidiče Homo sapiens na „řidiče“ – automatické systémy? Soudím, že v případě AV je žádoucí opustit optiku kvantitativních změn a považovat AV jen za jiný příklad řidiče. To nám umožní, abychom řešili daný problém lokálně – z pohledu aktérů–řidičů, kteří budou odstraněni podobně, jako byli odstraněni koně. Zatímco první náhrada za jiný druh pohonu byla jen mechanická, nastávající změna je kognitivní: Nahrazujeme aktéra řidiče něčím jiným. A aplikace toho, jak se dnes řidič chová, snaha jej emulovat, je ještě méně žádoucí, než by bylo přidělovat podkovy na kola svých aut. Cílem této eseje bude proto poukázat na sérii problémů relevantních pro filosofickou diskuzi, které vyvstanou při zavádění nové revoluční technologie autonomních vozidel. Ambicí tedy není problémy vyřešit, ale doporučit témata společenské diskuze, které se kromě filosofů zúčastní výrobci AV, představitelé státní správy a obecních zastupitelstev a také zástupci z řad veřejnosti – uživatelé aut.

Praktickým důvodem této diskuze jsou jistě plány masivního využití této technologie již během našich životů. Nejedná se tedy o futuristický problém, o kterém je možné spekulovat, AV se již staly aktuálním problémem, který je třeba řešit. Dalším důvodem, v tomto případě teoretickým, je to, že nyní bude nutné vyřešit nebo alespoň formalizovat problém praktické etiky zachraňování a odnímání života. Z pouhého hypotetického myšlenkového experimentu, který se krčil v zákoutí „*trolleyologie*“,<sup>7</sup> se stal praktický problém v rámci robotiky a umělé inteligence, na jehož řešení by filosofové měli pro-

7 „Trolley problem“ se stal součástí nejen úvodních hodin filosofie, ale už i popkultury. Populární zpracování tohoto myšlenkového experimentu a širšího výzkumu podal např.: Edmonds, D., *Would You Kill the Fat Man?: The Trolley Problem and What Your Answer Tells Us about Right and Wrong*. Princeton, Princeton University Press 2013.



gramátorům poskytnut nějaký formalizovaný návod, který již nebude mít podobu metafor či proudu nekonzistentních intuicí.

Za nejzásadnější přínos, který by projekt analýzy (formalizace) morálního jednání AV mohl poskytnout naší společnosti, považuji to, že výstupy problémů, které vyřešíme hypoteticky a prakticky v případě autonomních vozidel, pak můžeme analogicky aplikovat do jiných instancí etiky robotiky. Podobně jako je v genetice modelovým organismem octomilka, jejímž prostřednictvím zkoumáme pravidla dědičnosti, funkce genů a jejich interakce, v otázkách praktické etiky robotiky a umělé inteligence by se tímto modelem mohly stát AV. K tomu by bylo třeba splnit následující předpoklady: 1) AV budou nesmírně užitečné; 2) jejich rozšíření bude masivní, čímž se sníží jejich výrobní náklady a AV se stanou ještě dostupnějšími; a 3) interakce AV s předmětem morálky (lidmi) budou těsné a časté. Pokud se nám v budoucnu podaří vyřešit problémy nastíněné v následujících oddílech tohoto textu, budeme nepochybně schopni nabízená řešení aplikovat na další autonomní roboty v nejrůznějších odvětvích, kde dochází k interakci s lidmi a jejich zájmy. Stručně tuto úvahu můžeme shrnout jako řešení problému ve variantě s méně stupni volnosti. Jeden stupeň volnosti „pohybu/regulace“ mají dveře, dva výtah a tři robotická sekačka trávníku. To, co vyřešíme jako snadnější problém, následně aplikujeme na problém komplikovanější. Závěry, jež vyplynou ze zkoumání morálních problémů AV, můžeme následně aplikovat na autonomní zbraňové systémy nebo na programování dobré umělé inteligence.

### 3. Proč jsou AV problémem i pro etika

Mohou být technologie ABS nebo barva vozidla etickým problémem? ABS přebírá od řidiče kontrolu nad jednotlivými koly a při brždění je střídavě zastavuje. K základní podobě tohoto problému nemůže filosof přispět žádnou přidanou hodnotou. Podobně ani zmíněná barva vozu není něčím, k čemu by filosofové mohli říci více než kterýkoliv technik. Můžeme si leda představit, že lépe viditelné barvy jsou bezpečnější, a jsou proto i lepší volbou, pokud chceme minimalizovat riziko neštěstí (zanedbáme-li reflexní prvky). Pasivní systémy nebo systémy minimálně interagující s morálními subjekty a morálními aktéry (lidmi) mohou zkrátka plně existovat i bez diskuze s filosofy. Existuje vůbec v rámci řešení problematiky AV nějaké místo pro filosofa?

„Pokud se něco může pokazit, tak se to pokazí.“ Samozřejmě, že všechny lidské výtvoary mají své limity, a i kdyby zkušené technici učinili svá AV sebe-dokonalejšími, i malá šance na poruchu dříve nebo později zavíní dopravní nehodu. AV bude mnoho a budou v provozu delší dobu – určitě dojde k dopravní nehodě. V takovém neblahém případě si tvůrce řídicího softwaru přeje, jako



asi každý, aby z daných možných nešťastných scénářů nastal ten nejméně špatný. Jeho zájmem je naprogramovat AV tak, aby před nevyhnutelnou kolizí zvolilo takovou akci, jejíž výsledek bude relativně nejlepší vzhledem k možným alternativám. Uvnitř AV se budou nacházet lidé (cestující), taktéž ve výchozích a cílových bodech budou lidé (kolemjdoucí) – a připomeňme, že jedním z hlavních argumentů obhajujících masové rozšíření AV je snížení počtu úmrtí a zranění způsobených dopravními nehodami. Šance, že k takovému nešťastnému scénáři – kdy má AV např. volbu zabránit větší nehodě obětováním své posádky nebo kdy vybírá mezi dvěma různými kolizemi s fatálními následky pro různé počty chodců – dojde, je možná malá, ale jak již bylo řečeno: i malá šance se vzhledem k pravděpodobné délce doby užívání AV a s ohledem na mnohost aktérů rovná téměř jistotě. Programátoři řídicího softwaru chtějí učinit chování AV „morálním“, a tak se obracejí na filosofy, aby zjistili, jaká pravidla rozhodování mají být u AV během nehod použita.

Zastoupení role řidiče může probíhat pomocí emulace. Tě bychom docílili sběrem dat z manuálního řízení a kolizních situací. Je ovšem zjevné, že tato varianta má zásadní nedostatek: dopravní kolize totiž řidiči řeší v časové tísní, ve stresu a s velmi omezenou znalostí situace. Taková rozhodnutí budou tedy inkonzistentní a zřejmě nebudou odpovídat ani morálním preferencím řidičů – rozhodnutím, která by učinili po zralé úvaze. Emulace takových nouzových řešení je tedy nežádoucí. Proto je jednou z výhod autonomního řízení skutečnost, že je lze naprogramovat dopředu. Autoři programu mají obrovské množství času na jeho přípravu (rozhodnutí bude deliberativní, nebude pouhým reflexem) a navíc bude postaveno na kolektivním rozhodnutí – vznikne na základě společenské diskuze. Předpisy vytvořené pro rozhodovací software budou vlastně o tom, jak dopadnou dopravní nehody, rozhodovat ještě předtím, než auta vůbec vyrazí. Pasažéři budou moci dopředu vědět, zda jejich auto udělá vše pro jejich bezpečí a zda bude dodržovat dopravní předpisy, nebo zda naopak nebude na posádku brát ohled a v kolizní situaci jen nestranně zváží konsekvence možných reakcí.

#### 4. Řidič se změní z morálního aktéra na pouhý pasivní subjekt

V debatě týkající se etiky AV se pozornost soustředí zejména na typy kolizního systému,<sup>8</sup> které jsou však jen variantou tzv. vozíkového problému.

8 Lin, P., Why Ethics Matters for Autonomous Cars. Gerdes, J. Ch. – Thornton, S. M., Implementable Ethics for Autonomous Vehicles. Maurer, M. et al., *Autonomous Driving: Technical, Legal and Social Aspects*. Všechny tři příspěvky jsou součástí sborníku: Maurer, M. – Gerdes, J. Ch. – Lenz, B. – Winner, H. (eds.), *Autonomous Driving*. Berlin, Springer Publishing Company Inc. 2016. Dostupné také online na: <https://link.springer.com/book/10.1007%2F978-3-662-48847-8>; [cit. 24. 10. 2018].

Autonomní řízení přináší jen „další stupeň volnosti“ k v minulosti již obsáhle diskutovanému vozíkovému problému, přičemž je přehlížen jiný eticky relevantní problém. Řidič manuálně řízeného vozu je aktérem („*moral agent*“), který (doslova) řídí vůz, a může tedy volit – v poli morálního rozhodování – mezi dodržováním silničních předpisů a jejich porušováním. Protože sám sedí za volantem, je odpovědný i za to, co se svým vozem činí. Nemusí se jednat o dramatickou variantu vozíkového dilematu, kdy řidič volí, zda své auto navede ze srážu, nebo najede do velké skupiny lidí, která vyšla z již havarovaného autobusu. Řidiči tradičních vozidel řeší i malá morální dilemata: například zda budou preferovat svůj čas na úkor času druhých (jízda výhradně v levém pruhu, za sanitkou apod.). Vlastník AV ovšem přestane čelit velkým i malým morálním problémům – jeho vůz se bude chovat tak, jako by jeho „tělo“ pozřelo morální pilulku, která mu zabrání se chovat nemorálně.<sup>9</sup>

Je možné, že by neochota předat svou morální odpovědnost cizímu programu, navíc posílená nedůvěrou v techniku (jež byla zmíněna ve druhé části eseje), mohla vyústit v rozhodnutí, že si skupina „konzervativních“ vlastníků raději ponechá svá manuálně řízená auta, která nebude chtít vyměnit za AV. Proti jejich potenciálnímu rozhodnutí bývá stavěn argument, že přeměna řidiče-aktéra na pasažéra AV je jen analogií toho, co dnes zažívají lidé, když přesednou z vlastního vozu do autobusu městské hromadné dopravy. Tento argument je však poněkud mylný, jak následně ukážu. Jaké rozdíly tedy panují mezi těmito dvěma situacemi? V případě jízdy autobusem je stále přítomen lidský řidič, což může hrát roli u technologicky méně progresivních lidí. Další rozdíl spatřuji v odpovědnosti, která je založena na vlastnictví. Pokud jsme pasažéry autobusu, za jeho chybné jednání, ať už vede ke kolizi nebo k porušení dopravních předpisů, jsou odpovědní řidič a majitel vozu. Pasažéři jsou zcela zproštěni odpovědnosti za další osoby ve voze i za ostatní účastníky dopravního provozu (jsou doslova morálními „*free riders*“). Ovšem v případě, že jsme vlastníky AV a jsme ex-řidiči, stále cítíme zodpovědnost za následky jízdy tohoto vozidla, i když to nejsme my, kdo nakonec „otáčí volantem“.

Zavedení AV, bude-li direktivní, učiní ze všech pasažéry a ex-řidiči neponesou žádnou odpovědnost za následky rozhodnutí na silnici; odpovědnost bude přenesena na tvůrce software. Z konsekvencialistické perspektivy to je ovšem jen dobře: sníží se nehodovost a skutečnost, že řidiči přijdou o příležitost zušlechťovat svůj charakter tím, že nepodlehnu pokušení a neporuší

9 Tento problém považují za typ morálního enhancementu a je vlastně analogický „morální pilulce“. Podrobněji k diskuzi o morálním enhancementu viz v: Persson, I. – Savulescu, J., *Unfit for the Future: The Need for Moral Enhancement*. Oxford, Oxford University Press 2012.

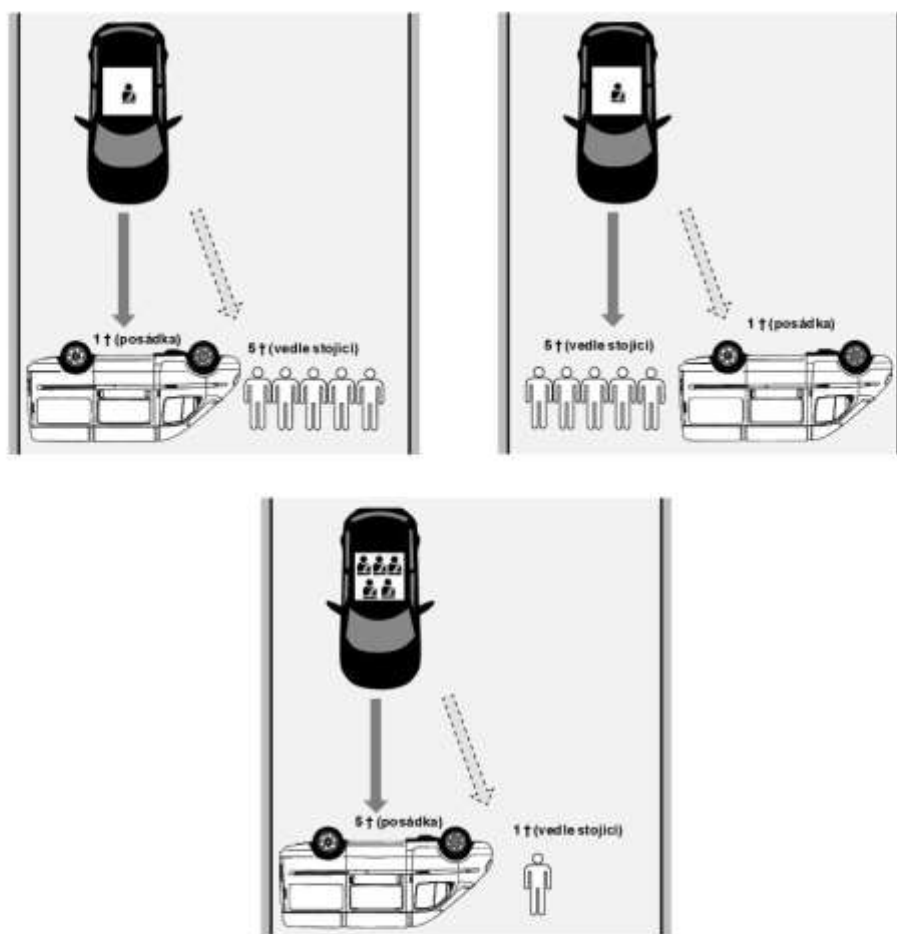
dopravní předpisy (např. objížděním kolony v protisměru), bude vyvážena tisíci zachráněnými oběťmi manuálně řízených aut. Pokud by strach z pozbytí aktérství stále vyvolával neochotu vyměnit své tradiční auto za AV, nabízím dva návrhy. První řeší odpovědnost za vlastnictví vozu. Řešením by byl „*car sharing*“ – ex-řidič by si v „půjčovně“ objednal na svou cestu auto, které by jej v místě A vyzvedlo a dovezlo jej na místo B. Takový pasažér by nebyl vlastníkem vozu, a tak by necítil odpovědnost za činnost pronajatého vozidla. (Podobně jako když dnes cítíme, že je poměrně velký rozdíl mezi tím, zda autem srazíme srnu my sami, nebo se to přihodí řidiči autobusu, ve kterém sedíme.) Druhým způsobem řešení je nabídnout více druhů „morálních pilulek“ – čili více typů obecného chování autonomních aut. (Tím není myšleno, že by za příplatek AV umělo na požádání vybrzdňovat nebo svítit protijedoucím řidičům do očí dálkovými světly.) Touha po volbě, po aktérství (sice nepřímém) by mohla být naplněna výběrem z několika typů „morálních řídicích programů“, o kterých bude řeč v následující kapitole.

### 5. Myšlenkové experimenty, metodická redukce

Pro potřeby následující argumentace si definujeme tři zjednodušené typy morálního softwaru. Jsou to ilustrativní příklady, pomocí nichž si vysvětlíme, jaké rozdíly v nastavení softwaru mohou hrát roli při jeho preferenci mezi cestujícími (nebo třeba vládami). Představme si hypotetickou dopravní nehodu v tunelu. Černé autonomní auto s vyznačeným počtem posádky má poruchu brzd a může volit jen ze dvou variant: první je neměnit směr jízdy a jet rovně, druhou variantou je zahrnout. Následky jsou jediné dva možné: černé auto po nárazu do prázdné dodávky havaruje a jeho posádka (černí panáčci) zahyne. Druhou variantou je, že černé AV zahne a „zabrzdí“ nárazem do osob v tunelu (bílých panáčků), čímž je sice zabije, ale jeho posádka – černí panáčci vyvázne bez újmy. Tato modelová situace nemá emulovat realitu a nabízené typy softwaru nemusí nutně reprezentovat konkrétní etické teorie. Obdobně jako v mechanice někdy např. redukuje tělesa do hmotných bodů a zanedbáváme tření vzduchu, můžeme v morální filosofii zjednodušit složitý svět do myšlenkového experimentu jen s jednou binární volbou.

Nyní zavedeme tři typy řídicího softwaru: *Tank*, který vždy preferuje životy posádky AV před životy kolemjdoucích a pasažérů jiných aut. *Počítadlo*, které vždy preferuje takové řešení krizové situace, při němž zemře nejméně lidí, respektive nejvíce je zachráněno. Posledním je *rytíř*, který nikdy aktivně nezmění směr jízdy, pokud by tato akce měla za následek zabití lidí, kteří by při absenci akce nebyli nijak ohroženi. Z náskresů je patrné, že jízda bez změny směru, vedoucí k úmrtí posádky nebo chodců, je ponecháním zemřít („*letting die*“), zatímco jízda se změnou směru má za následek usmrcení

(„killing“) – které sice není zamýšlené, ale je předvídatelné. (V reálných situacích si představuji např. takovou verzi počítačla, která sníží bezpečnost vlastní posádky o 5%, aby o 60% snížila riziko hromadné dopravní nehody, při níž by zahynulo deset lidí. Předpokládám, že reálné programy budou mít podobu smíšených strategií kombinujících všechny tři uvedené typy softwaru.)



Na prvním obrázku by *počítadlo* a *rytíř* nezměnili směr jízdy, *tank* by zabočil vlevo. Na druhém obrázku by *tank* a *rytíř* nezměnili směr jízdy, *počítadlo* by zabočilo vlevo. Na třetím obrázku by *rytíř* nezměnil směr jízdy, *počítadlo* a *tank* by zabočili vlevo. Z těchto předpokladů budeme dále vycházet.

Ilustrace jsou mé vlastní. Při jejich tvorbě jsem použil obměněné obrázky bez autorských práv (pixabay.com) a program pro ne-profesionály MS Paint.

Etikům zadali inženýři následující úkol: Je nutné naprogramovat řídicí algoritmus a přejeme si, aby auta byla „morální“. To je problémem a výzvou zároveň. Problémem je to zejména proto, že mezi profesionálními filozofy ani jinými mysliteli nepanuje obecná shoda na jediné správné teorii, na základě které by bylo možné takový software napsat. Výzva je to v tom smyslu, že máme naše morální soudy zpracovat do programu, čili máme odstranit veškerá subjektivní východiska a zároveň připustit pluralitu názorů. Tato výzva chce pro všechny definovat jeden závazný systém, který by zpracovával informace z veřejných, sdílených dat sensorů. Vnímám to jako příležitost k tomu, aby společnost znovu promyslela základy morálky. Část etiky se bude muset konečně transformovat ze subjektivního pohledu na operationalizovatelnou teorii zapsatelnou do objektivních matematických vzorců, a to z čistě praktických důvodů: do palubního počítače bude rozhodování třeba zapsat ve formě jedniček a nul. Tato výzva může konvenovat konsekventalistům zaměřeným na kvantitativní měření potěšení. Jiným etickým školám, zvláště těm, které operují s hodnotami ctností a povinností, by to ovšem mohlo činit problémy, a to až do té míry, že by takto vynucený kalkulus nemusely vůbec považovat za eticky vhodný.

Inženýři se obrací také na experimentální vědce (zejména morální psychology). Žádají je, aby popsali, jak probíhá morální soud lidí v případech volby softwaru – ať už při hypotetickém nákupu AV nebo při hypotetickém tvoření legislativy. Osobně se proto domnívám, že je třeba se více věnovat morálním intuicím. Netvrdím, že si všichni softwaroví inženýři uvědomují, že je morální soud zkoumaný psychology deskriptivní, zatímco etika, které se věnují filozofové, je preskriptivní. Ve třetí a čtvrté části této eseje jsou ostatně uvedeny důvody, proč by emulace lidského soudu byla kontraproduktivní. Nástrojem zkoumání morálních intuicí ve vozíkovém problému jsou tradičně myšlenkové experimenty spojené s dotazníkovým šetřením, dalšími metodami může například být užití virtuální reality<sup>10</sup> nebo magnetické rezonance.<sup>11</sup> Smyslem těchto výzkumů je poodhalit faktory, které se podílejí na tvorbě morálního soudu, a dále zjistit, jaká kognitivní zkreslení ovlivňují naše rozhodnutí. Pomocí takto získaných výsledků budou automobilky a vlády v budoucnosti moci např. predikovat to, jaký typ morálního softwaru bude nejvíce žádan na trhu, připustí-li volbu. Zastoupení různých typů morálního jednání AV není však ve společnosti jen jednosměrný proces. Noví vlastníci se totiž budou moci rozhodovat i na základě toho, jaká auta ve společnosti převládnou:

10 Navarrete, C. D. et al., Virtual Morality: Emotion and Action in a Simulated Three-dimensional “Trolley Problem”. *Emotion*, No. 12. 2. 2012, s. 364.

11 Greene, J. D. et al., An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science*, 293, 2001, No. 5537, s. 2105–2108.

Pokud bude normou sobecké vozidlo typu *tank*, bude to mít na volném trhu AV pravděpodobně za následek „závody ve zbrojení“, v nichž půjde o co nejbezpečnější a nejmohutnější *tanky*. Pokud naopak ve společnosti budou od začátku převládat altruistická auta typu *počítadlo* a budou zavedeny sankce pro vlastníky sobeckých vozidel (ať už daňové nebo jen společenské v rámci snížení prestiže), je šance, že dojde k naplnění konsekvenzialistického optima a při dopravních nehodách začne umírat méně lidí.<sup>12</sup>

## 6. Praktické problémy AV

Jedním ze způsobů volby morálního softwaru AV je odmítnutí celého tohoto systému, ve kterém je dopředu rozhodnuto o tom, kdo v kolizní situaci pravděpodobně přežije (díky preferenci vozu) a kdo naopak zemře. Únikem z potenciální pasti tohoto problému by mohlo být setrvání u manuálně řízeného vozidla. Taková úvaha je však založena na chybném předpokladu, že jsou si tradiční auta rovna. Ve skutečnosti však již dopravní prostředky typu *tanku* i *počítadla* existují. Vezmeme-li v úvahu jen bezpečnost posádky, má při kolizích největší výhodu řidič mohutného vozu s vysoko posazeným sezením (nákladní vůz). O něco méně bezpečná jsou robustní SUV, ještě méně bezpečné jsou vozy nižší třídy a vůbec nejhůře na tom jsou při kolizích řidiči motocyklů. Faktem tedy je, že už dnes při pořizování vozů kalkulujeme s naší bezpečností. A nekalkulují jenom zákazníci automobilek. Tradičním učebnicovým příkladem kalkulací s lidskými životy byl vůz Ford Pinto. Automobilka Ford vyčíslila, že by náklady na zesílení přepážky kolem palivové nádrže u všech vozidel modelu Pinto byly vyšší než kompenzace za několik zemřelých, kteří uhoří po jejím vznícení. Také vlády zemí kalkulují s lidskými životy v dopravních prostředcích: Letadla jsou sestřelována, pokud ohrožují například elektrárnu nebo směřují na jiný strategický cíl. Otázka tedy zůstává: Jaký typ softwaru pro AV si pořídít?

Z nedávného amerického výzkumu jednoznačně vyplývá,<sup>13</sup> že by partipanti preferovali, aby ostatní účastníci silničního provozu měli auta typu *počítadlo*, ale pro sebe by si vzali typ *tank*. Zároveň by nechtěli, aby stát všem nařizoval typ *počítadlo*. Tato volba – chtít po všech ostatních, aby byli altruističtí a sám zůstat sobecký – se může zdát ekonomicky racionální. Analogicky tomuto přístupu odpovídá přístup jednotlivců k americké Národní bezpečnostní agentuře (NSA). Jednotlivci nevdají, že NSA sbírá telefonní data mnoha

12 Ekonomické hry na „common good“ ukazují, že pokud se začnou objevovat sobecké strategie, altruistické jednání se udrží na jen základě trestání. Fehr, E. – Gächter, S., Cooperation and Punishment in Public Goods Experiments. *American Economic Review*, 90, 2000, No. 4, s. 980–994.

13 Bonnefon, J.-F. – Shariff, A. – Rahwan, I., The Social Dilemma of Autonomous Vehicles. *Science*, 352, 2016, No. 6293, s. 1573–1576.



dalších lidí. Člověku většinou začne vadit až to, že NSA sbírá právě jeho data. (Na jednu ze souvisejících obav o soukromí, totiž že díky AV budou cestující sledováni, lze odpovědět tak, že sledování AV se moc neliší od kontroly pohybu mobilního telefonu, která je dnes již běžná.)

Dalším praktickým problémem je rozhodnutí, kdo má být zodpovědný za tvorbu softwarových programů morálního rozhodování AV. Stát musí dopředu vytvořit legislativu provozu AV, protože AV se brzy zařadí do běžného provozu a stanou se účastníky dopravních nehod. V případě právního vakua by zúčastněné strany zřejmě žalovaly úplně všechny. Pokud přenecháme rozhodování o typech softwaru jen automobilkám, nemáme jistotu, že software bude řídit AV stejně, jak to prezentují. Vzpomeňme na nedávnou „Dieselgate“ německé automobilky Volkswagen. Víme-li, že automobilky podvádějí s výfukovými plyny, nemůžeme tušit, zda nebudou podvádět i s morálními algoritmy. Co když budou auta Volkswagenu v kolizních situacích vždy preferovat záchranu svých pasažérů oproti jiným zúčastněným?

Posledním z praktických problémů AV je mediální zkratka „Build to kill“, tedy přisuzování AV při kolizních situacích záměr přejíždět kolemjdoucí. Ovšem i jiný výtvar inženýrů – most – je „Build to kill“, pokud bude přetížen nad limit nosnosti. O mostu bychom ovšem neřekli, že zabíjí, ale že má technické limity. (Pokud si nepřejeme most s určitou nosností, můžeme vodní plochu protnout valem ze země. Protože cílem inženýra není vybudovat ideální dílo pro všechny situace, ale vytvořit věc určenou k danému účelu a v daném rozpočtu, čili v limitech užití.) Obdobně rozdíl mezi úmrtím při běžné dopravní nehodě a „zabitím autonomním vozidlem“ bude asi přisuzován intenci zabít, která bude vnímána jako atribut autonomního vozu. Analogicky je tomu například i s konzumací masa, která je běžně tolerována. Avšak pokud se americký zubař vydá do Afriky zastřílet si pro zábavu na lvy, což obsahuje intenci zabít (nikoliv jako prostředek), už je považován za morálně zkaženého člověka. Můžeme cítit spravedlivý hněv, když zjistíme, že řidič úmyslně někoho přejel, aby se vyhnul větší nehodě, ale AV se při přejetí člověka, k němuž dojde za účelem záchrany jiných lidí, mnohem více podobá mostu s limitním zatížením.

## 7. Závěr

Kvalitativní změna, která nastala výměnou vozu taženého koňmi za automobil, nachází svou obdobu v dnešním přechodu k AV jen velmi vzdáleně. Neměníme řidiče z masa a kostí za jiného řidiče z mikročipů, ale rozhodujeme morální problémy mnohem dříve, než k samotné nehodě vůbec dojde. Další ještě zásadnější změnou bude to, že řidiči budou zbaveni přímé morální odpovědnosti, jakož i možnosti morálně se rozhodovat v dopravních situacích.

Bude třeba nasbírat také data týkající se morálního soudu při morální a skutečné preferenci (vlastním nákupu autonomního vozidla).

Navrhuji zaměřit se na právní jistotu v oblasti kontroly morálního softwaru AV. K tomu je nutné co nejrychleji zjistit, „co chceme přijmout jako normy“ řídící provoz AV, a připravit ještě v předstihu příslušné zákony. Z provedených empirických studií zatím vyplývá, že čistý konsekvenzialismus bude nutné doplnit vyváženou preferencí posádky. Pokud by stát striktně trval na typu *počítadlo*, bylo by možné, že by taková AV nevytlačila ze silnic nebezpečné tradiční vozy. Tento model a smíšené strategie z části *počítadla* a z části *tanku* by mohly zabránit tomu, aby se fixoval suboptimální stav úplné převahy vozů typu *tank*.

V této esejí byly záměrně vynechány problémy týkající se dalších specifikací morálního softwaru, které jsou ale oproti zmíněnému hlavnímu problému podružné. Jedním z nich je dodržování silničních předpisů. Pro vůz typu *počítadlo* by se stala problematickou kupříkladu volba mezi kolizí s motorkářem s přílbou a motorkářem bez přílby, neboť by preferoval kolizi s prvně jmenovaným, kterému by sice pravděpodobně způsobil menší škodu, ovšem „trestal“ by jej takto za dodržování předpisů. Smyslem výše uvedených modelových příkladů softwaru nebylo hledání globálního řešení, ale poskytnutí definic a příkladů k diskusi. Druhým opomenutým problémem bylo využití standardizovaných panáček, neboť skutečné kalkulace budou zřejmě počítat i zachráněné roky života a jejich potenciální kvalitu.

Závěrem shrnuji, že na základě konsekvenzialistických východisek není nutné přijmout typ *počítadlo*, protože hlavní morální problém leží v současnosti jinde – skutečným dilematem je otázka, kdy se konečně zbavíme manuálně řízených aut, abychom zachránili tisíce životů. Dokonce i ne-konsekvenzialistické řešení dané situace, tj. preference jakýchkoliv typů morálního softwaru – i toho „bez iniciativy“ zvaného *rytíř* –, by vedla k záchraně tisíců životů, jejichž ztráty jsou dnes zapříčiněny chybami řidičů. Mohli bychom dosáhnout snížení počtu obětí nehod až o 95 %. Musíme brát v úvahu i to, že po zavedení jakéhokoliv morálního softwaru do AV lze jeho program následně vylepšovat. Stačí softwarová aktualizace a z *tanku* můžeme vytvořit *počítadlo* nebo případně jakoukoliv kombinaci rozhodování. Po otevření veřejné diskuze bude nutné se zaměřit i na další problémy. V jejich hierarchii, hned za hlavním úkolem – zbavit se manuálně řízených aut –, vyvstává druhý nejdůležitější: Jak se vyrovnat s obavou ze ztráty morálního aktérství.



## **2.2 Jak nás činí virtue signalling lepšími**



## How virtue signalling makes us better: moral preferences with respect to autonomous vehicle type choices

Robin Kopecky<sup>1,2</sup> · Michaela Jirout Košová<sup>1,2</sup> · Daniel D. Novotný<sup>1,3</sup> · Jaroslav Flegr<sup>1,2</sup> · David Černý<sup>1,4</sup>

Received: 30 July 2021 / Accepted: 13 April 2022 / Published online: 12 May 2022  
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2022

### Abstract

One of the moral questions concerning autonomous vehicles (henceforth AVs) is the choice between types that differ in their built-in algorithms for dealing with rare situations of unavoidable lethal collision. It does not appear to be possible to avoid questions about how these algorithms should be designed. We present the results of our study of moral preferences ( $N=2769$ ) with respect to three types of AVs: (1) selfish, which protects the lives of passenger(s) over any number of bystanders; (2) altruistic, which minimizes the number of casualties, even if this leads to death of passenger(s); and (3) conservative, which abstains from interfering in such situations. We differentiate between scenarios in which participants are to make their decisions privately or publicly, and for themselves or for their offspring. We aim to answer two research questions: (1) whether the public visibility of the choice of an AV type choice make this choice more altruistic and (2) which type of situation makes it more difficult to choose altruistically: when choosing for society as a whole, when choosing only for oneself, or when choosing only for one's offspring. Our results show that respondents exhibit a preference for an altruistic strategy for AVs and that it is reinforced when signaled to others. The altruistic preference is strongest when applies to everybody else, weaker when it reflects a solely personal choice, and weakest when choosing for one's own child. We conclude that a public choice is considerably more likely to pressure consumers into accepting a more socially beneficial solution.

**Keywords** Moral dilemmas · Autonomous vehicles · Theory of signalization · AI ethics

### 1 Introduction

Autonomous vehicles are expected to significantly benefit our transportation systems, their safety, efficiency, and impact on the environment (Talebpour and Mahmasani 2016). Autonomous vehicles (henceforth AVs) also called self-driving cars or driverless cars are in this context defined as autonomous motor vehicles not operated on rails but with rubber tires for use on highways, with examples, such as cars, trucks and buses. However, many technical, social, legal, and ethical questions and challenges concerning AVs and their introduction to the mass market still remain (Shariff et al. 2017; de Sio 2017; Hulse et al. 2018; Epting 2018). Countries with developed automotive industries, such as the Czech Republic, need to face these questions to implement the right legal measures with respect to significant social changes that the introduction of AVs is likely to bring. One of the issues concerns the choice between different types of AVs that vary in their built-in algorithms for dealing with situations of unavoidable lethal collision (Contissa et al. 2017; Goodall 2014;

✉ Robin Kopecky  
robin.kopecky@natur.cuni.cz

Michaela Jirout Košová  
michaela.kosa@gmail.com

Daniel D. Novotný  
novotnyd@tf.jcu.cz

Jaroslav Flegr  
flegr@cesnet.cz

David Černý  
david.cerny@ilaw.cas.cz

<sup>1</sup> The Karel Čapek Center for Values in Science and Technology, Prague, Czechia

<sup>2</sup> Faculty of Science, Charles University, Viničná 7, 128 00 Prague 2, Czechia

<sup>3</sup> Faculty of Theology, University of South Bohemia, České Budějovice, Czechia

<sup>4</sup> Institute of State and Law, Czech Academy of Sciences, Prague, Czechia

Haboucha et al. 2017). The reason why this particular issue is somewhat pressing has to do not so much with the fact that these situations will be happening on a regular basis. Rather, the issue is pressing from a psychological point of view: a non-negligible number of people still tend to be sceptical towards the reliability of new technologies and their safety, of which AVs are a clear example (Nielsen and Hausteijn 2018; Moody et al. 2020). Thus, the manufacturers will have to provide potential consumers with sensible solutions to collision situations, regardless of how common they will turn out to be in practice (Gogoll and Müller 2016; Gerdes and Thornton 2016).

Some of these situations will involve moral dilemmas inspired by trolley problems (Foot 2002; Thomson 1985; Kagan 1991; Lin 2015). In many ways, these dilemmas differ from the classical trolley problems: in the typical trolley cases, the person who decides about the solution of the situation is usually not the one to be harmed; in contrast, the AV cases concern the responsibilities of the person as a driver/traffic participant. Furthermore, AVs will not be limited to set tracks and will have to operate with probability. Perhaps other differences could be found but the question of the similarities and differences of the new AV dilemmas from old trolley problems is not directly relevant to our concerns. What needs to be highlighted is that moral intuitions that people have about these dilemmas do have bearing on the preference for a concrete AV algorithm and for willingness to purchase it. (Note also that we are not primarily concerned here with “risks or the legal and moral responsibilities we face in traffic”, unlike e.g. Nyholm and Smids (2016) who even criticize the relevancy of trolley cases for the AV literature).

There are at least three ways AVs can be set to handle the critical situations: (1) selfish AVs protect the lives of passenger(s) over any number of bystanders; (2) altruistic AVs minimize the number of casualties, even if this leads to death of passenger(s); and (3) conservative AVs abstain from interfering in such situations, even if it leads to the death of a higher number of subjects or death of passenger(s). (The latter two are both non-selfish strategies). In this paper, we present the results of our study of moral preferences with respect to these types of AVs. We furthermore differentiate between scenarios in which participants are to make their decisions privately or publicly, and for themselves or for their offspring. We disregard gender, age, health, biological species and other characteristics of (potential) casualties that can affect the preferences and decisions of respondents in our scenarios (Nothigattu et al. 2018).

Our study is based on the sample of 2769 mostly Czech volunteers (1799 women, 970 men; age IQR 25–32). The data come from our web-based questionnaire that was accessible from May 2017 to December 2017. We aim to answer the following two research questions: (1) Whether the public

visibility of an AV type choice makes this choice more altruistic and (2) which type of situation is more problematic with regard to the altruistic choice: opting for society as a whole, for oneself, or for one’s offspring.

Showing that the private or public character of AV type selection could affect the level of altruism has potentially important implications for policies concerning AV introduction. Findings concerning situations in which people choose most selfishly could be similarly significant. This could inform legislators where to proceed with caution to prevent moral panic which could lead to a strong rejection of AVs or severely delay the introduction of this life-saving technology.

The first research question is inspired by a biological theory of signalization (Laidre and Johnstone 2013; Dawkins and Krebs 1978; Zahavi 1975). Willingness to use AVs that would sacrifice the owner to save others could be perceived as a costly *honest signal*. This form of altruism is a signal of willingness to help others at one’s own cost (Hardy and Van Vugt 2006; Fehrler and Przepiorka 2013). This altruistic concern is related to the concern for one’s own reputation and it will pay off in social standing in the future in repeated social interactions (Nowak and Sigmund 1998), even with respect to strangers (Semmann et al. 2005). Participating in road traffic with people, one does not personally know and yet showing altruism by costly signalling, is analogous to altruism towards strangers (Bereczkei et al. 2010).

The owner of a virtue signalling AV would be exposed to social sanction (either positive or negative) by neighbours able to see which type of vehicle he or she has. Lee et al. (2018) experimentally investigated whether social observation has an impact on trolley problem-like decisions involving physical harm. The study showed that experimentally induced reputation concern reduces sacrificial decisions (called ‘utilitarianism’ in these contexts, since no self-sacrificing altruist was present in the situations). Brown and Sacco (2017) showed that people who prefer utilitarian decision in trolley problem cases seem less trustworthy to participants as long-term partners. Based on these findings, one could easily expect that reputation concern in a public choice will decrease choice of strategy that minimizes the number of casualties in an AV choice if the user is not present in their vehicle (which was not amongst the questions we asked our participants). However, in case the user is present in their vehicle the opposite expectation comes forward—the strategy that minimizes the number of casualties will be increased if a participant is allowing their vehicle to put their own life at stake and thus exhibit the aforementioned altruism towards strangers.

The second research question aims to compare our findings with those of Bonnefon et al. (2016) in different socio-cultural contexts. A part of Bonnefon’s study focused on the AV moral collision software type choice for a user when travelling alone or for a user when travelling with a work



colleague or when travelling with family. The participant was still hypothetically present in the AV type that was to be selected. In our study, we focused solely on the choice of AV type for the participant's beloved child to reduce the possible influence of their own presence and to compare directly between how participants weigh the safety of their child and their personal safety in a secret personal choice and the safety of everybody in a secret voting in Parliament.

## 2 Methods

The study is a part of the project "Factors Influencing Altruism" that was approved by the institutional review board at the Faculty of Science, Charles University (No. 2017/14).

### 2.1 Participants

We devised a web-based study running between May 2017 and December 2017 in which 2769 volunteers participated (1799 women, 970 men; age IQR 25–32). Most participants were college graduates and students ["Master and Doctor": 1060 (38.3%), "Bachelor": 349 (12.6%), "bachelor student": 375 (13.5%)]. All participants understood the Czech language of the questionnaire. The majority of respondents were of Czech nationality [2436, (88%)], the second most represented nationality were Slovaks [267, (9.6%)].

### 2.2 Procedure

Participants were recruited either through our web page [www.PokusniKralici.cz](http://www.PokusniKralici.cz) or our Facebook fan page of the same name (Pokusní králíci) that has a frequent coverage by the media. The internet questionnaire was distributed as a Qualtrics survey. Participants were not incentivized by monetary reward. Rather, they were provided with results and comparison of sub-questionnaires directly after finishing the whole survey. After providing an informed consent, the volunteers filled out a comprehensive combined questionnaire composed of three sub-questionnaires focusing on various research topics, such as the aesthetic appreciation of plants and experimental philosophy of personal identity. It took around 90 min to complete the survey. There were also two attention-checking questions, one before and one after the AV section. Participants who failed to answer either of the questions correctly were excluded from data analyses (there were 45 failures on the first and 26 failures on the second checkpoint). By the end of 2017, the drop rate of the whole survey reached 38% of the 4443 participants who started this questionnaire. Only the participants who passed both attention checks were included in this study ( $N = 2769$ ).

The AV section, which is the subject matter of the present study, was opened by the following brief information on AV:

"According to experts, AVs can prevent 95% of deadly car accidents, which means that the lives of roughly 600 people and 5.6 billion CZK (220 million EUR) may be saved every year in the Czech Republic alone." Then participants were asked to evaluate nine trolley problem visualizations (Fig. 1), choosing between different software-conveyed solutions to the given lethal collision situation.

The reason why these nine visualizations were included, was to make the participants familiar with different trolley-like cases. This may differ with respect to the number of people inside the AV, the number of people on the road, and with respect to the position of the wreck of the van.

The last part of the 90 min questionnaire concerned AVs. We introduced three types of AVs in it: so-called "Tank" (with preference for passengers in the AV, here we refer to it as "selfish" AV), so-called "Calculator" (with preference for the highest number of saved lives, here referred to as "altruistic" AV), and so-called "Knight" (which will not actively change direction if that would lead to killing a pedestrian or a passenger, even though changing direction would save more lives in total, here referred to as "conservative" AV). Our paper draws solely on this last part of the questionnaire, in which the participants were asked the following five questions (paraphrase): (1) what type of AV would you choose as your own car if nobody was able to find out about your choice ("secret/self")? (2) What type of AV would you choose as your own car if your choice was visible at first glance ("public/self")? (3) What type of AV would you choose as the car of your beloved child if nobody was able

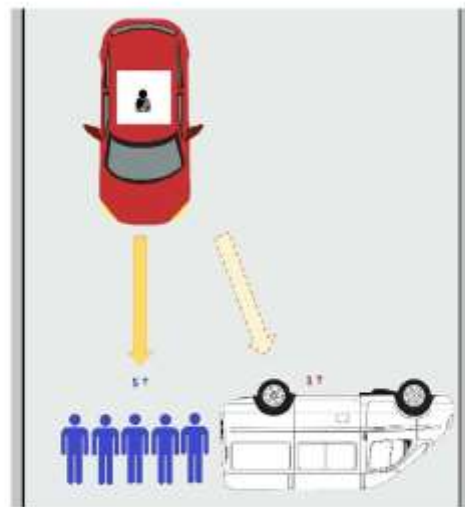


Fig. 1 An example of trolley problem visualization

to find out (“secret/child”)? (4) What type of AV would you vote for in a secret parliamentary vote if it was to become the only legal type of AV (“secret/parliament”)? (5) What type of software would you vote for in a public parliamentary vote if it was to become the only legal type of AV (“public/parliament”)? These five questions were consecutively asked on the same page of the questionnaire, so that participants were able to see them at a single glance.

**2.3 Statistics**

Statistical software IBM SPSS version 20 was used for the statistical tests, descriptive statistics, and producing graphs. Testing the association between two binary variables was done by contingency table using chi-square test. The statistical significance was characterized as *p* value and the effect size as risk ratio (RR). The associations with *p* value lower than 0.05 were considered statistically significant.

**3 Results**

**3.1 Descriptive statistics**

Summary descriptive statistics are displayed in Table 1 for following five questions: (1) What type of AV would you choose as your own car if nobody was able to find out about your choice (“secret/self”)? (2) What type of AV would you choose as your own car if your choice was visible at first glance (“public/self”)? (3) What type of AV would you choose as the car of your beloved child if nobody was able to find out (“secret/child”)? (4) What type of AV would you vote for in a secret parliamentary vote if it was to become the only legal type of AV (“secret/parliament”)? (5) What type of software would you vote for in a public parliamentary vote if it was to become the only legal type of AV (“public/parliament”)?

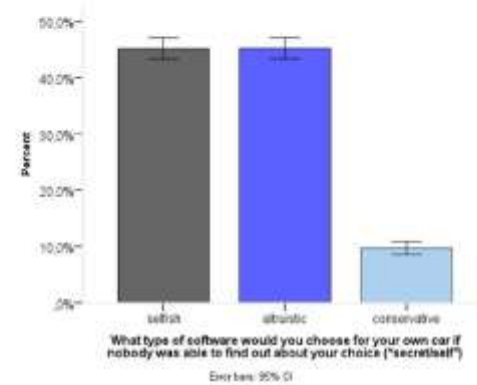
**Table 1** AV choices in different scenarios

	Selfish	Altruistic	Conservative
Secret, self	1219 (45.2%)	1220 (45.2%)	260 (9.6%)
Public, self	809 (30%)	1565 (58.1%)	319 (11.8%)
Secret, child	1794 (66.6%)	751 (27.9%)	150 (5.6%)
Secret, parliament	554 (20.6%)	1795 (66.9%)	335 (12.5%)
Public, parliament	460 (17.1%)	1879 (69.9%)	348 (13%)

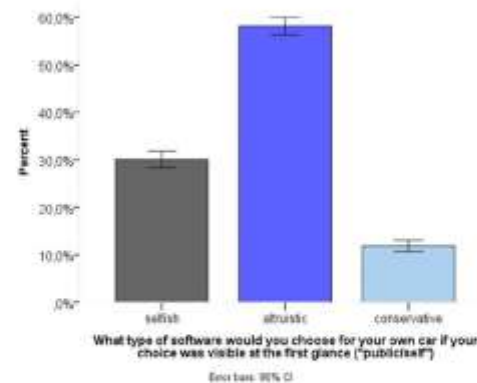
The rows represent different scenarios, while the columns represent choices. (We intended to include the scenario with a child only in the “secret” version as we aimed to answer the second research question in “secret” setting.)

**3.2 Hypotheses 1.1**

Via the first tested hypothesis “H1.1: secret/self > public/self” we tested whether participants are less likely to choose the selfish AV for themselves if the choice is indicated on the car and clearly visible. The distribution of choices is illustrated in Figs. 2 and 3. A chi-square test of independence was performed to examine the relation between the type of choice (secret; public) and the type of AV (selfish or non-selfish AVs, where the latter includes both altruistic and conservative AVs). The relation was significant, since



**Fig. 2** Secret/self. Y-axis shows the fraction of subjects selecting a particular variant of strategy. The error bars show the 95% confidence interval



**Fig. 3** Public/self. Y-axis shows the fraction of subjects selecting a particular variant of the strategy. The error bars show the 95% confidence interval

**Table 2** Change of opinion between secret and public choice

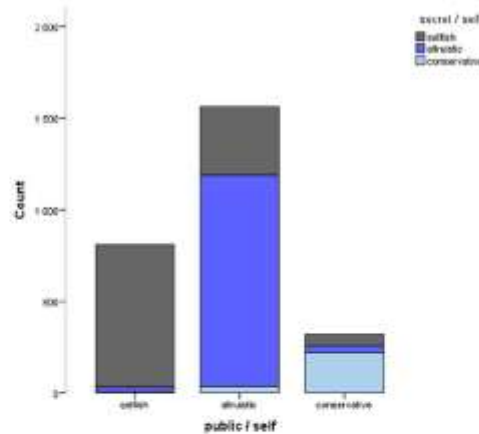
public / self	N	←	secret / self	N	%
selfish	809	}	selfish	777	96.0
			altruistic	27	3.3
			conservative	5	0.8
altruistic	1564	}	selfish	376	24.0
			altruistic	1153	73.7
			conservative	35	2.2
conservative	319	}	selfish	62	19.4
			altruistic	27	8.5
			conservative	220	69.0

On the left side there are participants divided into groups depending on their choice in the public/self-question. The right side describes their other decision in the secret/self-question so the change illustrated in the following Fig. 4 is completed with exact numbers.

respondents expressed lower preferences for a selfish strategy in a public choice than in a secret one [ $\chi^2(1) = 131.40$ ,  $p < 0.00001$ ,  $RR = 0.665$ ,  $CI(95\%) = 0.620-0.714$ ], Table 2 and the stacked bar chart in Fig. 4 shows the change in preference of participants between secret and public choice.

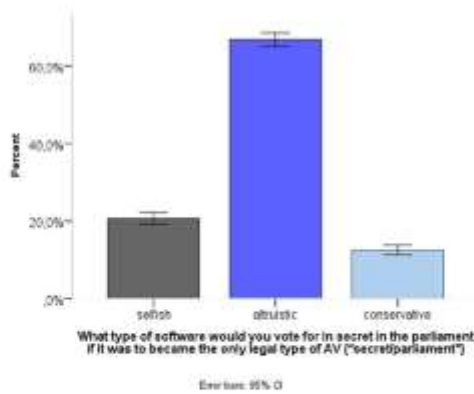
**3.3 Hypothesis 1.2**

Via the second hypothesis “H1.2: secret/parliament > public/parliament” we tested whether participants are less likely to vote for selfish software in Parliament if the voting is public. The distribution of choices is illustrated in Figs. 5 and 6. A chi-square test of independence was performed to examine the relation between the type of choice (secret/parliament; public/parliament) and the type of AV (selfish AV or non-selfish AV, where the latter includes altruistic and conservative AVs). The relation was significant, respondents expressed lower preferences for voting for a selfish strategy in a public vote than in a secret vote [ $\chi^2(1) = 11.519$ ,  $p = 0.0007$ ,  $RR = 0.825$ ,  $CI(95\%) = 0.738-0.922$ ], Table 3 and the stacked bar chart in Fig. 7 show the change in preference of participants between secret and public voting.

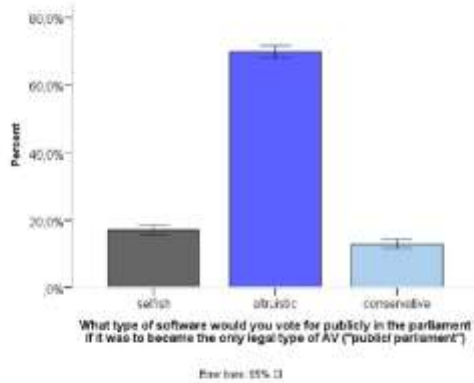


**Fig. 4** The change in preference of three strategies in a public choice by subjects who choose between various strategies in a secret choice. Columns show the preference in public/self-question and colours indicate what the same person has chosen in secret/self-scenario. E.g. the second column shows all people who preferred altruistic choice in public/self-scenario and charcoal-grey colour signifies a subgroup of them that preferred selfish strategy in secret/self-scenario





**Fig. 5** Secret/parliament. Y-axis shows the fraction of subjects selecting a particular variant of the strategy. The error bars show the 95% confidence interval



**Fig. 6** Public/parliament. Y-axis shows the fraction of subjects selecting a particular variant of the strategy. The error bars show the 95% confidence interval

### 3.4 Hypothesis 2

Via the last hypothesis “H2: secret choices: child > self; child > parliament” we tested whether participants prefer the most selfish strategy when deciding for their child, less selfish strategy when deciding for themselves and the least selfish strategy when deciding for everybody. The distribution of choices for their own child is illustrated in Fig. 8. A chi-square test of independence was performed to examine the relation between the type of choice (child; secret/self) and the type of AV (selfish and non-selfish). The relation

was significant when choosing a car for one’s own child: people were more likely to choose selfish software than when choosing for themselves [ $\chi^2(1) = 250.53, p < 0.00001, RR = 0.678, CI(95\%) = 0.646–0.713$ ].

Similarly, when choosing a car in secret for one’s own child people were more likely to choose selfish software than while voting in secret in parliament [ $\chi^2(1) = 1153.17, p < 0.00001, RR = 0.31, CI(95\%) = 0.2866–0.3355$ ].

## 4 Discussion

### 4.1 Empirical considerations

Our results showed that the majority of the respondents of the anonymous internet survey strongly preferred the altruistic strategy based on minimizing the number of people harmed even at the cost of self-sacrifice or based on an active sacrifice of a lesser number of other people. This preference was even stronger in cases of public choice when other people are able to see what type of strategy the AV owner has chosen, which shows that the type of choice indeed has an effect on altruism. The results also showed that the respondents reported much lower preferences for an altruistic strategy when choosing an AV for their child than when choosing for themselves. This result—the personal preference of an altruistic AV for themselves but not for their child—is consistent with the findings of Bonnefon et al. (2016).

Participants chose selfish software more often in cases when the choice of the strategy was private, i.e., not visible at first glance. This suggests that people regard selfish strategies as socially undesirable or morally dubious. Our results do not make it possible to decide on what grounds the group of participants shifted their choice from being selfish in secret to being altruistic in public. Nevertheless, it can be expected that building one’s reputation could play a role. Results of experimental games indicate that reputation is the key factor to maintain altruism in both reciprocal and indirect forms (Kraft-Todd et al. 2018; Semmann et al. 2005; Olivella and Siciliani 2017). Players usually act more altruistically towards altruistic individuals regardless of whether they have helped them directly or not. Signalling an altruistic strategy could also be advantageous in *intra-sexual competition*, i.e. biological concept in which two members of the same sex are competing against each other over a potential mate (Zahavi 1975; Buunk and Massar 2012; Puska et al. 2016).

If reputation building is the crucial motive for signalling the use of an altruistic strategy, then this signal can potentially be falsified, i.e., some AV users, the cheaters, will be signalling an altruistic strategy but in fact they will be using

**Table 3** Changed opinion between secret and public vote

public / parliament	N	←	secret / parliament	N	%
selfish	460	}	selfish	448	97.4
			altruistic	11	2.4
			conservative	1	0.2
altruistic	1874	}	selfish	95	5.1
			altruistic	1750	93.4
			conservative	29	1.5
conservative	346	}	selfish	11	3.2
			altruistic	30	8.7
			conservative	305	88.2

On the left side there are participants divided into groups depending on their vote in the public/parliament question. The right side describes their other decision in the secret/parliament question, so the change illustrated in the following Fig. 7 is completed with exact numbers

selfish AVs. Zahavi's (1975) theory shows that the only evolutionary stable strategy in this circumstance is to make the signalling costly for the cheaters. In the case of signalling an altruistic strategy, it should be obligatory to adopt such legal and technical measures that will make fake signalling both illegal and as technically difficult as possible.

Lee et al. (2018) showed in their study of trolley problems that experimentally induced social observation leads to a drop in utilitarianism, i.e., there is a higher preference for deontological judgement if the person is socially observed. Based on their study, one would expect that participants in a public personal choice would shift towards lower utilitarianism as well. However, in our study, the opposite was the case. A possible explanation has already been suggested in the introduction: having an altruistic (utilitarian) AV is at the same time being altruistic towards strangers—the person is willing to sacrifice his or her own life to save others and thus the reputation concern would work in the opposite direction. The fact that one's own life is involved in the situation is the game-changer here.

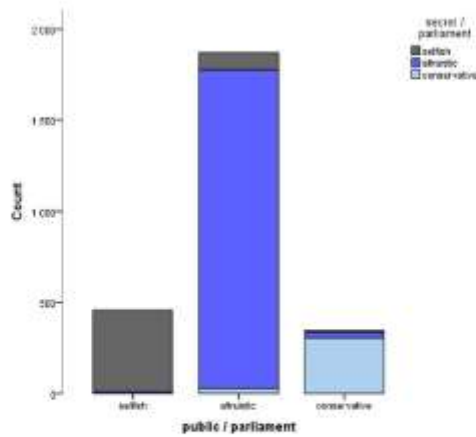
While data were being gathered for this study in the Czech population, the German ethical committee for autonomous driving issued a report (German Ministry of Transportation (2017; philosophically reviewed in Luetge (2017)

in which it is insisted that no innocent human lives may be taken by AVs in collision situations. It is prohibited to offset victims against one another, even though reducing the number of personal injuries may be justifiable. This means that AVs would not be able to react to trolley-like scenarios. This amounts to opting for conservative AVs in our terminology. This strategy was by far the least popular among our respondents in the Czech Republic in all five scenarios, at maximum accepted by 13% of participants in case of a public vote in Parliament. The striking difference between German ethical and legal experts and the Czech public may indicate that there will be culturally dependent and diverse attitudes in different countries. A recent study (Awad et al. 2018) on a global scale indicates similar findings that the strongest preference is for saving more people in lethal collision as well. Further studies on these cultural influences on moral judgements with respect to AVs are needed.

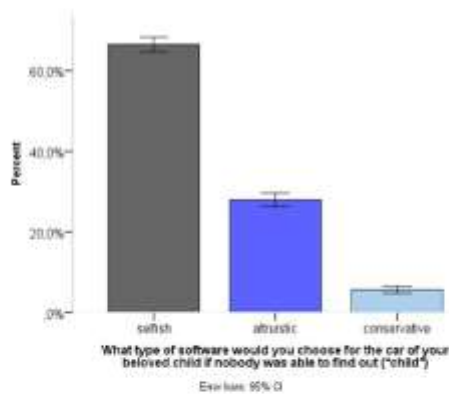
## 4.2 Ethical considerations

German declarations (Luetge 2017) state that the responsibility for driving a car should always stay with an (individual) human person. Perhaps surprisingly, this solution could lead to a widespread use of selfish AVs. Moreover,





**Fig. 7** The change in preference of three strategies in a public parliamentary vote by subjects who choose between various strategies in a secret parliamentary vote. Columns show the preference in public/parliament question and colours indicate what the same person has chosen in secret/parliament scenario. E.g. in the second column there are all people who preferred altruistic choice in public/parliament and charcoal-grey colour signifies a subgroup of them that preferred selfish strategy in secret/parliament scenario



**Fig. 8** Secret/child. Y-axis shows the fraction of subjects selecting a particular variant of the strategy. The error bars show the 95% confidence interval

it is unrealistic and unfair to expect individual (non-driving) operators to be able to step in at the last moment in emergency situations and then bear the full responsibility for these exceptional collision situations. Thus, our policy recommendation is to decide about the permissible AV type

at the state level, rather than leave it up to private choice and competition between AV manufacturers. It would be especially undesirable to try to regulate the choice economically: for example, if the selfish cars were made more expensive or fell under higher taxes, this would result in a “selfish” advantage for the wealthy buyers. Alternatively, to provide the public with the possibility of free choice, we recommend that laws require a clear indication of what type of AV one is driving. This too is likely to increase the number of altruistic AVs and hence overall decrease the death rate by car accidents. This kind of peer pressure would appeal to the sense of moral responsibility towards the other members of the AV owner’s community, which is comparable to the peer pressure already present e.g. due to discussions about ecology and responsibility towards one’s environment.

If it becomes possible for the public to choose freely between various types of driving software in the near future, we strongly advocate the introduction of obligatory measures by means of design indication (e.g. colours, stripes) to publicly display altruism or at least non-selfishness. This measure is necessary for social sanctions and reputational concern to properly accommodate customer choice. We believe that the altruistic algorithm is the right choice given the most commonly held ethical theories, especially consequentialism. Other traditional theories, such as deontology, contractualism or virtue theory, are at least compatible with it as well, or at any rate, it is not straightforward clear how to apply them to decide for or against altruism (Nyholm 2018). At any rate, we do not have the opportunity here to discuss more general issues in ethical theory in detail and the data suggest that the wider public also prefers altruistic strategy, whatever the ultimate ethical theory. The selfish strategy only competes with the altruistic one in the specific “secret choice” situation, which suggests that this type of selfishness is not considered as good and preferable at a more abstract moral level. (For more general considerations how to take avail of such moral intuitions, see Savulescu et al. (2019).)

Less clear is the problem of how to deal with our heightened sensibility and “passenger-prioritizing” with respect to our own offspring. Careful approach is needed here to prevent moral panic. Future research should include a study of moral judgments in other samples, especially in non-Western cultures, to test whether our results can be generalized. The decision over which type of AV to choose for one’s own child proved to be the most problematic, since in this scenario participants chose mostly “passenger-prioritizing” software referred in paper as “selfish”. This indicates that the future AV communication ought not to be primarily focused on convincing the drivers themselves but on communicating the advantages of AVs to parents regarding the safety of school buses or AV taxis for their children.

Although our study is focused specifically on AV type choice, we need to be aware that it is not a standalone aspect

of AV mobility. Real-life environment, especially cities where most of the human–AI interactions do and will take place turns the ethics of AV into a much more complex issue. First, there is the question of test implementation of AV that is already happening in cities of various sizes and geographical locations, usefulness and effectiveness of their experimental design (Karvonen and van Heur 2013), as well as their impact on local politics as discussed in Dowling and McGuirk (2020). Based on investigation into public stances toward implementation and type of ownership of AVs, a second issue arises as of how will the price, availability, and comfort of personal AV impact the scale of personal transportation in cities and how will it translate into changes in urban planning (Cugurullo et al. 2020) and political regulatory context (Aoyama and Alvarez Leon 2021). New urban technologies, including autonomous transportation are incorporated into modern plans of smart and sustainable cities (Chehri and Mouftah 2019), both newly built and transformed from existing ones, but the actual outcome is often problematic (Cugurullo 2017), which further underlines the need for complex consideration of the issue as suggested also in Cugurullo (2021).

### 4.3 Limitations

Our results are based on self-reports in a broad anonymous internet survey and not on respondents' observable behaviour in real-life situations. Therefore, we do not claim to know their real-life choices in buying AVs or voting in Parliament. On the other hand, it is probable that both observed phenomena (higher preference for altruistic strategies in public choices and lower when our children are involved) will play an important role also in real-life decisions. Anyway, a real preference for an altruistic strategy would be proven only by observing individual behaviour in everyday life or in an experimental setting. Another possible limitation of this study is the self-selection of participants, since the participation was not monetary-incentivized. Last but not least, this broad internet survey although with large sample size cannot be considered a sociological representative sample for the whole Czech population and could be culturally specific.

## 5 Conclusion

Our results show that the respondents exhibited a clear preference in AVs for an altruistic strategy that minimizes the number of casualties, even if this leads to death of passenger(s). This preference was reinforced if the AV signals its strategy to others. The altruistic preference was strongest when people choose software for everybody else, weaker in their personal choice, and the weakest in the choice for their own child, when people preferred strategy that protects the lives of passenger(s) over any number

of bystanders. The moral intuition of Czech participants seems not to correspond with the recommendation of the German ethical committee, as the AV strategy that abstains from the trolley problem was the least popular in all situations for participants in our study. Based on the results, we conclude that, in contrast to a private choice, a public choice is considerably more likely to pressure consumers in their personal choice to accept a non-selfish solution, making it a reasonable and relatively cheap way to shift car owners and users towards higher altruism. Also, a hypothetical voting in Parliament about a single available program would probably be less selfish if the voting did not take place in secret.

**Acknowledgements** RK developed the study concept. RK, MJK, JF were responsible for the study design and data collection. RK performed the data analysis and interpretation under the supervision of JF. RK, DN, DC drafted and all authors revised the manuscript. All authors approved the final version of the manuscript for submission. For comments on earlier drafts of this paper we would like to thank: Světlá Hanke Jarošová. We would also like to thank Lenka Připlová and anonymous reviewers for very useful comments on the revision of this paper.

**Funding** This work was supported by the Grant Agency of the Charles University under grant GAUK number 929216; Charles University Research Program under grant UNCE number 204056; and The Technological Agency of the Czech Republic under ETA grant number TL01000467.

**Data Availability** <https://osf.io/y7i5q>.

### Declarations

**Conflict of interest** No potential conflict of interest was reported by the authors.

### References

- Aoyama Y, Alvarez Leon LF (2021) Urban governance and autonomous vehicles. *Cities* 119:103410. <https://doi.org/10.1016/j.cities.2021.103410>
- Awad E, Dsouza S, Kim R, Schulz J, Heinrich J, Shariff A, Bonnefon JF, Rahwan I (2018) The moral machine experiment. *Nature* 563(7729):59–64. <https://doi.org/10.1038/s41586-018-0637-6>
- Bereczkei T, Birkas B, Kerekes Z (2010) Altruism towards strangers in need: costly signaling in an industrial society. *Evol Hum Behav* 31(2):95–103. <https://doi.org/10.1016/j.evolhumbehav.2009.07.004>
- Bonnefon JF, Shariff A, Rahwan I (2016) The social dilemma of autonomous vehicles. *Science* 352(6293):1573–1576. <https://doi.org/10.1126/science.aaf2654>
- Brown M, Sacco DF (2017) Is pulling the lever sexy? Deontology as a downstream cue to long-term mate quality. *J Soc Pers Relat* 36(3):957–976. <https://doi.org/10.1177/0265407517749331>
- Buunk AP, Massar K (2012) Intrasexual competition among males: competitive towards men, prosocial towards women. *Personal Individ Differ* 52(7):818–821. <https://doi.org/10.1016/j.paid.2012.01.010>
- Chehri A, Mouftah HT (2019) Autonomous vehicles in the sustainable cities, the beginning of a green adventure. *Sustain Cities Soc* 51:101751. <https://doi.org/10.1016/j.scs.2019.101751>



- Contissa G, Lagioia F, Sartor G (2017) The Ethical Knob: ethically-customisable automated vehicles and the law. *Artif Intell Law* 25(3):365–378. <https://doi.org/10.1007/s10506-017-9211-z>
- Cugurullo F (2017) Exposing smart cities and eco-cities: frankenstein urbanism and the sustainability challenges of the experimental city. *Environ Plan a: Econ Space* 50(1):73–92. <https://doi.org/10.1177/0308518x17738535>
- Cugurullo F (2021) *Frankenstein urbanism* (Routledge studies in urbanism and the city), 1st edn. Routledge, London
- Cugurullo F, Acheampong RA, Gueriau M, Dusparic I (2020) The transition to autonomous cars, the redesign of cities and the future of urban sustainability. *Urban Geogr* 42(6):833–859. <https://doi.org/10.1080/02723638.2020.1746096>
- Dawkins R, Krebs JR (1978) Animal signals: information or manipulation. *Behav Ecol: Evolut Approach* 2:282–309
- de Sio FS (2017) Killing by autonomous vehicles and the legal doctrine of necessity. *Ethical Theory Moral Pract* 20(2):411–429. <https://doi.org/10.1007/s10677-017-9780-7>
- Dowling R, McGuirk P (2020) Autonomous vehicle experiments and the city. *Urban Geogr*. <https://doi.org/10.1080/02723638.2020.1866392>
- Epting S (2018) Automated vehicles and transportation justice. *Philos Technol* 32(3):389–403. <https://doi.org/10.1007/s13347-018-0307-5>
- Fehrler S, Przepiorka W (2013) Charitable giving as a signal of trustworthiness: disentangling the signaling benefits of altruistic acts. *Evol Hum Behav* 34(2):139–145. <https://doi.org/10.1016/j.evolhumbehav.2012.11.005>
- Foot P (2002) The problem of abortion and the doctrine of the double effect. In: Foot P (ed) *Virtues and vices and other essays in moral philosophy*. Clarendon Press, Oxford, pp 19–33 (**Originally appeared in Oxford Review, 5, 1967**)
- Gerdes JC, Thornton SM (2016) Implementable ethics for autonomous vehicles. In: Mauer M et al (eds) *Autonomous driving. Technical, legal and social aspects*. Springer, Berlin/Heidelberg, pp 87–102
- German ministry of transportation. Report. (2017). [https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?\\_\\_blob=publicationFile](https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile)
- Gogoll J, Müller JF (2016) Autonomous cars: in favor of a mandatory ethics setting. *Sci Eng Ethics* 23(3):681–700. <https://doi.org/10.1007/s11948-016-9806-x>
- Goodall NJ (2014) Ethical decision making during automated vehicle crashes. *Transportation Research Record: Journal of the Transportation Research Board* 2424(1):58–65. <https://doi.org/10.3141/2424-07>
- Haboucha CJ, Ishaq R, Shifan Y (2017) User preferences regarding autonomous vehicles. *Transp Res Part C: Emerg Technol* 78:37–49. <https://doi.org/10.1016/j.trc.2017.01.010>
- Hardy CL, van Vugt M (2006) Nice guys finish first: the competitive altruism hypothesis. *Personal Soc Psychol Bull* 32(10):1402–1413. <https://doi.org/10.1177/0146167206291006>
- Hulse LM, Xie H, Galea ER (2018) Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. *Saf Sci* 102:1–13
- Kagan S (1991) *The limits of morality* (Oxford ethics series), Revised. Clarendon Press, Oxford
- Karvonen A, van Heur B (2013) Urban laboratories: experiments in reworking cities. *Int J Urban Reg Res* 38(2):379–392. <https://doi.org/10.1111/1468-2427.12075>
- Kraft-Todd GT, Bollinger B, Gillingham K, Lamp S, Rand DG (2018) Credibility-enhancing displays promote the provision of non-normative public goods. *Nature* 563(7730):245–248. <https://doi.org/10.1038/s41586-018-0647-4>
- Laidre ME, Johnstone RA (2013) Animal signals. *Curr Biol* 23(18):R829–R833. <https://doi.org/10.1016/j.cub.2013.07.070>
- Lee M, Sul S, Kim H (2018) Social observation increases deontological judgments in moral dilemmas. *Evol Hum Behav* 39(6):611–621. <https://doi.org/10.1016/j.evolhumbehav.2018.06.004>
- Lin P (2015) Why ethics matters for autonomous cars. In: Maurer M et al (eds) *Autonomous driving. Technical, legal and social aspects*. Springer, Berlin/Heidelberg, pp 69–85
- Laetge C (2017) The German ethics code for automated and connected driving. *Philos Technol* 30(4):547–558. <https://doi.org/10.1007/s13347-017-0284-0>
- Maurer M et al (2016) *Autonomous driving. Technical, legal and social aspects*. Springer, Berlin/Heidelberg
- Moody J, Bailey N, Zhao J (2020) Public perceptions of autonomous vehicle safety: an international comparison. *Saf Sci* 121:634–650. <https://doi.org/10.1016/j.ssci.2019.07.022>
- Nielsen TAS, Haustein S (2018) On sceptics and enthusiasts: What are the expectations towards self-driving cars? *Transp Policy* 66:49–55. <https://doi.org/10.1016/j.tranpol.2018.03.004>
- Noothigattu R, Gaikwad S, Awad E, Dsouza S, Rahwan I, Ravikumar P, Procaccia A (2018) A voting-based system for ethical decision making. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1.
- Nowak MA, Sigmund K (1998) Evolution of indirect reciprocity by image scoring. *Nature* 393(6685):573–577. <https://doi.org/10.1038/31225>
- Nyholm S (2018) The ethics of crashes with self-driving cars: a road-map. I. *Philos Compass* 13(7):e12507. <https://doi.org/10.1111/phc3.12507>
- Nyholm S, Smids J (2016) The ethics of accident-algorithms for self-driving cars: an applied trolley problem? *Ethical Theory Moral Pract* 19(5):1275–1289. <https://doi.org/10.1007/s10677-016-9745-2>
- Olivella P, Siciliani L (2017) Reputational concerns with altruistic providers. *J Health Econ* 55:1–13. <https://doi.org/10.1016/j.jhealeco.2017.05.003>
- Puska P, Kurki S, Lähdesmäki M, Siltaoja M, Luomala H (2016) Male-male status signaling through favoring organic foods: is the signaler perceived and treated as a friend or a foe? *Psychol Mark* 33(10):843–855. <https://doi.org/10.1002/mar.20922>
- Savulescu J, Kahane G, Gyngell C (2019) From public preferences to ethical policy. *Nat Hum Behav* 3(12):1241–1243. <https://doi.org/10.1038/s41562-019-0711-6>
- Semmann D, Krambeck HJ, Milinski M (2005) Reputation is valuable within and outside one's own social group. *Behav Ecol Sociobiol* 57(6):611–616. <https://doi.org/10.1007/s00265-004-0885-3>
- Shariff A, Bonnefon JF, Rahwan I (2017) Psychological roadblocks to the adoption of self-driving vehicles. *Nat Hum Behav* 1(10):694–696. <https://doi.org/10.1038/s41562-017-0202-6>
- Talebpoor A, Mahmassani HS (2016) Influence of connected and autonomous vehicles on traffic flow stability and throughput. *Transp Res Part c: Emerg Technol* 71:143–163. <https://doi.org/10.1016/j.trc.2016.07.007>
- Thomson JJ (1985) The trolley problem. *Yale Law J* 94:1395–1415
- Zahavi A (1975) Mate selection—a selection for a handicap. *J Theor Biol* 53(1):205–214. [https://doi.org/10.1016/0022-5193\(75\)90111-3](https://doi.org/10.1016/0022-5193(75)90111-3)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## **2.3 Obava ze ztráty vlastního aktérství**

## **Obava ze ztráty vlastního aktérství a absence morálního aktéra u autonomních dopravních prostředků**

### **Fear of losing personal agency and the absence of a moral agent in autonomous transportation vehicles**

*Robin Kopecký & Michaela Jirout Košová*

afiliace: RK: Vysoká škola uměleckoprůmyslová v Praze

RK, MJK: Centrum Karla Čapka pro studium hodnot ve vědě a technice

robin.kopecky@icloud.com, michaela.kosa@natur.cuni.cz

dedikace: TAČR TL03000549 Umění a design jako řešení změn v dopravě s nástupem autonomního řízení

klíčová slova: aktérství, autonomní vozidla, morální odpovědnost, morální intuice

key words: agency, autonomous vehicles, moral responsibility, moral intuitions

Abstrakt:

Diskuze o filosofických problémech spojených s nástupem autonomní dopravy je často spojena s negativy a je vedena v duchu techno-pesimismu. Významný prostor v této debatě je věnován aplikaci vozíkového dilematu („*trolley problem*“) do řídicích algoritmů vozů. V této eseji se však zamýšlíme nad dalším problémem našich morálních intuic – otázkou morálního aktérství. Pocit morální odpovědnosti si žádá faktickou kontrolu nad situací a kauzální odpovědnost. O tu bývalý řidič jako pasažér autonomního vozu přichází, ale přesto může cítit vinu. Prioritně bychom se měli soustřeďovat na morální psychologii uživatele, spíše než hledání optima kolizí pro algoritmy. V eseji poukazujeme na historické paralely absence aktéra v automatických systémech i na to, že lidé snáze přijímají autonomní nákladní dopravu než dopravu osob. Dále

nastiňujeme možné způsoby řešení přesunu odpovědnosti, a to nejprve prostřednictvím analogie s řidičem přicházejícím o aktérství nástupem do vozu hromadné dopravy. Možný operátor autonomních vozů, který na provoz dohlíží místo skutečného řidiče, může představovat dočasné opatření. Skutečnou alternativou morální odpovědnosti by se ale mohl stát fiktivní morální aktér dosazený provozovatelem vozů. Oproštění se od vlastnictví autonomního vozu osvobozuje lidi od konečného vnímání morální odpovědnosti, takže krátkodobé výpůjčky, sdílení nebo hromadná doprava by mohly být další vhodnou odpovědí na tento psychologický fenomén.

#### Abstract:

Discussion about philosophical problems associated with the rise of autonomous driving is often associated with negatives and is conducted in the spirit of techno-pessimism. A significant space in this debate is dedicated to the application of the “trolley problem” to vehicle control algorithms. However, in this essay, we reflect on another problem of our moral intuitions – the question of moral agency. The feeling of moral responsibility demands factual control over the situation and causal responsibility. The former driver, now a passenger in an autonomous vehicle, loses this control, but may still experience guilt. Primarily, we should focus on the moral psychology of the user rather than searching for collision optima for algorithms. In the essay, we point out historical parallels of the absence of agency in automated systems and the fact that people are more accepting of autonomous freight transportation than personal transportation. Furthermore, we outline possible ways to shift responsibility, starting with an analogy to a driver losing agency when entering public transportation. A possible operator of autonomous vehicles, overseeing the operation instead of a real driver, could serve as a temporary solution. However, a fictional moral agent appointed by the vehicle operator could potentially become a genuine alternative to moral responsibility. Detachment from ownership of an autonomous vehicle frees individuals from a final perception of moral responsibility, so short-term rentals, sharing, or public transportation could be suitable responses to this psychological phenomenon.

Uvedení technologických inovací v podobě propojeného a autonomního řízení pro individuální a hromadnou dopravu s sebou nese kromě jiného příležitosti pro pozitivní změny v podobě snížení dopadu dopravy na životní prostředí, zvýšení bezpečnosti pro cestující, a také finanční a časové úspory (Talebpour & Mahmassani, 2016). Přesto je společenská debata o těchto inovacích vedena v rovině možných negativ (Kopecký, 2019).

Celé subžánry vědeckofantastické literatury se zabývají konfliktem mezi protagonisty a novými technologiemi. Konflikt představuje pole možností pro atraktivní děj, zatímco bezproblémové čistě techno-optimistické prostředí je inertní a čtenářsky nezajímavé. Lze proto očekávat, že se i společenská diskuze bude ubírat techno-pesimistickým směrem, zaměří se na negativa a konkrétní etické otázky. Ačkoli v tomto textu řešíme namísto strojových zabijáků či všeobjímající virtuální reality ze science fiction obavy všedního dne, které odrážejí naše lidské morální intuice, sci-fi kultura hraje přímou kauzální roli i v této oblasti, jak dokládáme v následujícím odstavci.

Mohli bychom očekávat, že se čtenáři a fanoušci fantastiky, vzhledem k jejím odkazům na zneužití technologií, odlidštění, limitace a další problémy s fiktivní technologickou budoucností, stanou obezřetnějšími při morálním zvažování přínosů futuristických inovací. Na základě empirického šetření (Laakasuo et al., 2018) zkoumajícího morální schvalování technologického vylepšování člověka (enhancement) autoři zjistili, že jsou znalosti a příznivý vztah k science fiction s tímto souhlasem asociovány pozitivně, tedy že lidé s fantastikou lépe obeznámení morálně více schvalují enhancement. Otevřenost vůči zkušenosti (z pětifaktorového modelu osobnosti) s tímto schvalováním naopak asociována nebyla.

Přejdeme ale od enhancementu a beletrie k hlavnímu tématu této eseje, kterým jsou překážky pro přijetí autonomní dopravy, a to nikoliv překážky technické, ve smyslu například optimalizace konstrukce vozovky a softwaru dopravních prostředků, ale překážky vyvstávající

z naší lidské mysli a našich hodnot, a jejich možná řešení, která se pokusíme nastínit v závěru.

Tradiční a široce diskutovaný „vozíkový problém“ (anglický termín *“trolley problem”* odkazuje na vůz tramvaje a překládá se jako vozíkové, tramvajové nebo drezínové dilema) představuje úkol pro programátory řídicího softwaru autonomních dopravních prostředků. Ti společně s filozofy hledají optimální morální rozhodovací pravidla, která by mohla být uplatněna v případě poruch nebo dopravních nehod, při nichž vznikají morální dilemata. Jde o situace kdy, pokud nezasáhneme, dojde jistě ke ztrátě na  $X$  životech. Pokud zasáhneme a riziko odvrátíme, následkem bude pravděpodobně nezamýšlená škoda na  $Y$  životech, přičemž  $X > Y$  (Foot, 1967; Jarvis Thomson, 1985; Lin, 2015; Černý, 2020; Kopecký et al., 2022). V těchto dilematech stojí před morálním aktérem či strojem na jeho místě proti sobě omise (nejednat, a jistě nechat zemřít  $X$  lidí) a komise (jednat, a pravděpodobně nechat zemřít  $Y$  [ $X > Y$ ] lidí).



Jiným typem je záchranářské dilema (Sundvall et al., 2023), ve kterém se má morální aktér nebo jeho případný robotický ekvivalent rozhodovat mezi dvěma komisemi: záchranou A, při které zemře B, anebo záchranou B, při které zemře A. V záchranářském dilematu nepředpokládáme omisi, při níž bychom nechali zemřít A i B.

Myšlenkové experimenty zabývající se záchranou a nečinností vedoucí k úmrtí – vozíkový nebo záchranářský problém – však nejsou jedinými možnými typy problémů a domníváme se, že pro kontext společenského přijetí autonomního řízení nejsou ani těmi nejzávažnějšími. Úhelným kamenem zavádění nové technologie se může stát obyčejný uživatel automobilu, jehož ochota přijmout autonomní řízení může mít větší dopad, než nalezení optimálního řešení vozíkových problémů. Místo na dilemat omisi a komisí pro aktéry, ať už lidské nebo jejich robotické alternativy, vyvstávají dva spojené problémy: ztráta aktérství a absence morálního aktéra. Tyto problémy jsou relevantní nejen pro tvůrce autonomních dopravních prostředků, ale i pro další filozofickou debatu týkající se s lidmi interagujících strojů s umělou inteligencí obecně. Návrhy řešení problémů s přijetím autonomní dopravy, které se podaří hypoteticky nalézt, tak budou následně aplikovatelné i v odlišných sférách etiky robotiky, například u pečovatelských nebo armádních robotů, které mají vyšší stupně volnosti.

Jak jsme už uvedli, za centrální problém nepovažujeme možné suboptimální nastavení morálního softwaru vozidla zodpovědného za řešení vozíkových dilemat z pohledu některých etických teorií (například konsekvencialismu nebo deontologie), ale řidičův strach z proměny jeho role aktivního aktéra na pouhý pasivní subjekt. Situaci navíc zhoršuje ztráta skutečného odpovědného morálního aktéra. Cítíme-li se totiž zodpovědní za nastalou situaci, chceme nad ní mít rovněž reálnou kontrolu, a přesune-li se zodpovědnost na někoho jiného, požadujeme, aby byl onou odpovědnou osobou skutečný morální aktér.

Na základě dat z Eurobarometru 2014 přinesli autoři (Hudson, Orivska & Hunady, 2019) pro tuto úvahu dva přínosné poznatky. Zprvė zjistili, že je veřejnost otevřenější autonomním nákladním vozům než autonomním osobním vozům. To může být dáno tím, že se absencí aktéra i posádky jako takové v čistě nákladní dopravě situace morálního poměrování zjednoduší. Dle druhého zjištění určuje technofobie k robotům měřená dle Nomura et al., 2007 také negativní postoje k autonomním vozům.

Jako tradiční historický příklad technofobie a strachu ze ztráty aktérství můžeme uvést výtah, který je svým způsobem obdobný problému chybějícího morálního aktéra u autonomního vozidla. Obecnou veřejnost znepokojovala absence operátora v nově zavedených automatických výtazích, byl tedy zaveden náhradní (lidský) operátor. Jeho zodpovědnost se sice vázala pouze ke stisknutí tlačítka podlaží, cestující se však díky jeho přítomnosti cítili příjemněji a služeb výtahu využívali mnohem ochotněji. Tato historická kuriozita ilustruje lidskou potřebu být schopen identifikovat konkrétního aktéra, který ponese plnou morální odpovědnost za možná řešení potenciálně nebezpečných situací. To může souviset s vnitřní intuicí, že pouze lidský aktér zvládne mít situaci náležitě pod kontrolou, klíčovou roli zde ale pravděpodobně hraje možnost ukázat na někoho reálného jako na nositele morální odpovědnosti. Užití „výtahového řešení“ u autonomního vozu za použití emulace role řidiče řídicím softwarem by zřejmě neuspokojilo lidské obavy, protože by dostatečně nezohledňovalo lidskou intuici týkající se chybějícího lidského morálního aktéra. Inspiraci pro vhodné řešení, které popíšeme v následujících odstavcích, bychom mohli najít u hromadné dopravy.

Řidiče během ovládání vozidla činí morálními aktéry nutnost čelit mnoha dílčím morálním dilematům, od dodržení předepsané rychlosti, pravidel pro předjíždění a parkování, až po závažná provinění, jako je jízda pod vlivem drog. V případě autonomních vozů se však z původních řidičů stávají cestující – pasažéři (Kopecký, 2019). Tím přestávají být schopni

následovat či porušovat relevantní pravidla a tedy i čelit příslušným morálním dilematům. Chování autonomního vozu bude v souladu s platnými regulacemi, které tato dilemata dopředu vyloučí (jde o analogii k užití hypotetické „morální pilulky“ bránící konzumentovi po požití chovat se v rozporu s morálkou). K obdobné proměně dochází u řidičů osobní dopravy při nástupu do vozidla dopravy hromadné: i zde dobrovolně mění svůj status morálního aktéra na pouhého pasažéra. Identifikujme dva možné důvody pro obvykle bezproblémové přijetí tohoto posunu.

Prvním důvodem usnadňujícím daný přechod je nezměněná fyzická přítomnost řidiče ve většině prostředků hromadné dopravy. Je tak zachován morální aktér, který ovládá vozidlo a nese zodpovědnost a morální vinu za případnou nebezpečnou situaci. Případný přenos přímo morálně zodpovědné osoby do autonomní osobní dopravy se ukazuje příliš složitým až nemožným. Lze si však představit živého člověka bez reálné zodpovědnosti dosazovaného na přechodnou dobu "za volant", tj. do role morálního aktéra, jinak autonomně řízeného vozidla. Dočasně zde, stejně jako v případě výtahu, znamená, dokud si společnost nezvykne. Druhý důvod, pro autonomní osobní vozy již plně aplikovatelný, je možnost nebýt vlastníkem/provozovatelem autonomního hromadného/osobního vozidla. Cestující v hromadné dopravě nenesou žádnou právní, morální ani psychologickou odpovědnost, což z nich dělá doslova morální „černé pasažéry“.

V případě vlastnictví autonomního vozu budou bývalí řidiči pravděpodobně stále cítit psychologickou odpovědnost za následky jednání řídicího softwaru, a to i navzdory tomu, že již nebudou fyzicky ovládat volant, pedály, ani páky. Požadované ztráty psychologické odpovědnosti můžeme dosáhnout náhradou osobního vlastnictví vozu krátkodobými pronájmy nebo sdílením u specializovaných společností. Přesun vlastnictví autonomních vozů z majetku

uživatelů zmenší vnímanou morální odpovědnost za případné morálně relevantní pochybení vozu při řešení dopravních situací a případných nehodách.

Toto řešení zároveň odpovídá na otázku, zda by měl být řídicí software regulován státem, výrobcem (automobilkou) či softwarovými vývojáři – psychologická vina i morální odpovědnost padne na vlastníka a provozovatele sdílených vozů. Marginalizujeme morální aktérství pasažéra, ale zachováváme morální odpovědnost provozovatele.

Různou míru morální odpovědnosti „řidiče“, nebo lépe řečeno posádky autonomního vozu v rovině viny nebo zásluh, zkoumali McManus & Rutchick (2019) na modelových situacích adaptovaného vozíkového problému převzatých ze studií Bonnefona et al. (2016). Participanti hodnotili vinu či zásluhu posádky v závislosti na sobeckém/altruistickém rozhodnutí vozu při kolizní situaci. Ukázalo se, že nejvíce viny nebo zásluh přisuzují očekávatelně posádce – „řidiči“ vozu v případě manuálního řízení vozidla, a naopak nejméně v případě, že byl výrobcem plně nastaven řídicí software. Z příkladů různých mezistupňů intervencí posádky a přednastaveného softwaru od výrobce vyplývá, že redukce aktérství na straně „řidiče“ (posádky) vede ke snížení vnímané viny či zásluh posádky za špatné či dobré následky řešení kolizních situací. I když je tento empirický výzkum založen na vozíkových problémech komise a omise, naznačuje, že zde existuje možnost omezením kompetencí pasažérů přenést morální vinu z posádky na výrobce. V navazujícím výzkumu (Bennet et al., 2020) hodnotili participanti přisuzování viny „řidiči“ dle různých stupňů automatizace. I když se s narůstajícím stupněm automatizace přisuzovaná vina „řidiče“ redukovala směrem k dalším zúčastněným stranám, jako například k zmíněnému výrobcí, přesto ve všech případech morální vina padala dle názoru participantů i na „řidiče“ – posádku autonomního vozu.

Závěrem shrňme, že by problém laických morálních intuic týkajících se ztráty vlastního aktérství mohl nalézt částečné řešení ve změně vlastnických vztahů směrem k využívání vozu

formou krátkodobých výpůjček a sdílení, nebo většinovým nahrazením soukromé dopravy dopravou hromadnou. Cílem je přesun břemena morální odpovědnosti z bývalých řidičů na nové majitele. Druhý problém, a sice absenci morálního aktéra, můžeme zkusit řešit analogicky k obsluze výtahů přechodnou přítomností lidských operátorů v autonomních vozech hromadné dopravy. Dalším krokem by mohlo být zavedení fiktivních morálních aktérů, jako je firma provozovatele/vlastníka, nebo do budoucna dokonce dostatečně sofistikovaných strojů. To by však vyžadovalo zahájení společenské a filosofické diskuze o umělých bytostech jako možné extenzi pro morální aktéry.

Obecné problémy, jako je technofobie vůči robotům, mohou být vyřešeny prostřednictvím fenoménu habituace. Vyrovnáme-li se s pravděpodobně roky trvající dobou habituace na různé roboty ve společnosti na základě pozitivních osobních zkušeností (jak můžeme vidět například v japonské kultuře), může se naše obava z absence morálního aktéra postupně transformovat do pozitivního vztahu k novému typu umělého aktéra. Ochota ke sdílení a přechodu k hromadné dopravě na jedné straně, a morální hodnocení chování robotů a autonomních vozidel na straně druhé, musejí být dále náležitě empiricky zkoumány, a to například formou hodnocení jednotlivých návrhů skrze dotazníková šetření. Jenom tak bude možné zhodnotit efektivitu našich návrhů a posunout je k možné implementaci.

## Literatura

Bennett, J. M., Challinor, K. L., Modesto, O., & Prabhakaran, P. (2020). Attribution of blame of crash causation across varying levels of vehicle automation. *Safety Science*, 132, 104968.

Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science*, 352(6293), 1573-1576.

Černý, D. (2020). *The Principle of Double Effect: A History and Philosophical Defense*. Routledge.

Foot, P. (2002). The Problem of Abortion and the Doctrine of the Double Effect. In P. Foot, *Virtues and Vices and Other Essays in Moral Philosophy* (pp. 19-33). Oxford: Clarendon Press. (Originally appeared in *Oxford Review*, 5, 1967.)

Hudson, J., Orviska, M., & Hunady, J. (2019). People's attitudes to autonomous vehicles. *Transportation research part A: policy and practice*, 121, 164-176.

Kopecký, Robin. Morální problémy autonomních vozidel. (2019) *Filosofický časopis*, roč.67, č.2, s.267–268.

Kopecky, R., Jirout Košová, M., Novotný, D. D., Flegr, J., & Černý, D. (2022). How virtue signalling makes us better: Moral preferences with respect to autonomous vehicle type choices. *AI & SOCIETY*, 1-10.

Laakasuo, M., Drosinou, M., Koverola, M., Kunnari, A., Halonen, J., Lehtonen, N., & Palomäki, J. (2018). What makes people approve or condemn mind upload technology? Untangling the effects of sexual disgust, purity and science fiction familiarity. *Palgrave Communications*, 4(1), 1-14.

Lin P. (2015) Why Ethics Matters for Autonomous Cars. In: Maurer M., et al. (Eds.), *Autonomous Driving. Technical, Legal and Social Aspects* (pp. 69-85). Berlin/Heidelberg: Springer.

McManus, R. M., & Rutchick, A. M. (2019). Autonomous vehicles and the attribution of moral responsibility. *Social psychological and personality science*, 10(3), 345-352.

Nomura, T., Suzuki, T., Kanda, T., & Kato, K. (2006, September). Measurement of anxiety toward robots. In *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 372-377). IEEE.

Sundvall, J., Drosinou, M., Hannikainen, I., Elovaara, K., Halonen, J., Herzon, V., Kopecký, R., Jirout Košová, M., Koverola, M., Kunnari, A., Perander, S., Saikkonen, T., Palomäki, J., Laakasuo, M. (2023). Innocence over utilitarianism: Heightened moral standards for robots in rescue dilemmas. *European Journal of Social Psychology*, 53 (4), 779-804.

Thomson, J. J. (1985). The trolley problem. *The Yale Law Journal*, 94,1395-1415.

Talebpour, A., & Mahmassani, H. S. (2016). Influence of connected and autonomous vehicles on traffic flow stability and throughput. *Transportation Research Part C: Emerging Technologies*, 71, 143–163. <https://doi.org/10.1016/j.trc.2016.07.007>


### **3. Evoluční psychologie**

#### **3.1 Vliv latentní toxoplazmózy na hodnoty**



# Le Petit Machiavellian Prince: Effects of Latent Toxoplasmosis on Political Beliefs and Values

Evolutionary Psychology  
July-September 2022: 1–13  
© The Author(s) 2022  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/1474704922112657  
journals.sagepub.com/home/evp  
SAGE

Robin Kopecky<sup>1</sup> , Lenka Příplatová<sup>1</sup>, Silvia Boschetti<sup>1</sup>,  
Konrad Talmont-Kaminski<sup>1</sup>, and Jaroslav Flegr<sup>2</sup>

## Abstract

Humans infected by *Toxoplasma gondii* express no specific symptoms but manifest higher incidence of many diseases, disorders and differences in personality and behavior. The aim of this study was to compare the political beliefs and values of *Toxoplasma*-infected and *Toxoplasma*-free participants. We measured beliefs and values of 2315 responders via an online survey (477 *Toxoplasma*-infected) using the Political Beliefs and Values Inventory (PI34). This study showed *Toxoplasma*-infected and *Toxoplasma*-free participants of our cross-sectional study differed in three of four factors of PI34, scoring higher in Tribalism and lower in Cultural liberalism and Anti-Authoritarianism. We found sex differences in political beliefs associated with *Toxoplasma* infection. Infected women scored higher in tribalism and lower in cultural liberalism, compared with the *Toxoplasma*-free control group, while infected men scored higher in economic equity. These results fit with sexual differences in behavior and attitude observed after toxoplasmosis infection. Controlling for the effect of worse physical health and mental health had little impact, suggesting that impaired health did not cause these changes. Rather than adaptation to prevalence of parasites, as suggested by parasite-stress theory, the differences might be side-effects of long-term mild inflammatory reaction. However, to get clear picture of the mild inflammation effects, more research focused on different infectious diseases is needed.

## Keywords

*Toxoplasma gondii*, manipulation hypothesis, political beliefs, stress, infectious diseases, parasite threat, pathogen avoidance

Date received: May 12, 2022; Revision Submitted: June 23, 2022; Accepted: June 23, 2022

## Introduction

*Toxoplasma gondii*, the parasitic protozoon of cats, infects about one-third of the human population in both developed and developing countries (Tenter et al., 2000). Infected humans express no specific symptoms of the life-long infection, however, they differ from non-infected subjects by a higher incidence of many diseases and disorders (Flegr & Escudero, 2016; Flegr, Klapíková, et al., 2014; Flegr, Prandota, et al., 2014) and also by specific changes in personality (Flegr, 2013; Khademvatan et al., 2013; Lindová et al., 2012) and behavior (Fekadu et al., 2010; Houdek, 2017). These behavioral differences are mostly considered to be products or by-products of the manipulative activity of *Toxoplasma* aimed to increase the chance of transmitting from the infected intermediate host, which can be any warm-blooded animal, to its definitive host, a cat (Webster, 2007). However, *Toxoplasma* infected

animals, including humans, express many signs of the activated immune system. They have increased levels of some proinflammatory cytokines, especially IL-6 (Kaňková et al., 2010). It is known that inflammation, and independently also resulting stress have very strong impacts on various motivational, emotional, and behavioral processes, for recent reviews, see, e.g., Moieni & Eisenberger, 2018; Lasselín et al., 2018; Gassen & Hill, 2019. It was suggested (Lindová et al., 2010;

<sup>1</sup>Institute of Philosophy of the Czech Academy of Sciences, Prague, Czechia  
<sup>2</sup>Laboratory of Evolutionary Biology, Department of Philosophy and History of Sciences, Faculty of Science, Charles University, Prague, Czechia

### Corresponding Author:

Jaroslav Flegr, Laboratory of Evolutionary Biology, Department of Philosophy and History of Sciences, Faculty of Science, Charles University, Vinicna 7, 128 00 Prague 2, Czechia.  
Email: flegr@cesnet.cz



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Martynowicz et al., 2019; Tedford & McConkey, 2017) that a large part of the observed toxoplasmosis-associated changes are not the products of the manipulative activity of *Toxoplasma* but a result of mild chronic stress associated with long-term parasitic infection and resulting health problems. It is known that many toxoplasmosis-associated changes go in opposite directions in the infected men and women (Lindová et al., 2006), and also that men and women cope with stress in diametrically opposed ways (Matud, 2004; Tamres et al., 2016). The stress-coping hypothesis suggests that many toxoplasmosis-associated changes, e.g., higher extroversion in infected women and lower in men, or lower suspiciousness of infected women and higher in suspiciousness of infected men (all in comparison to non-infected peers) can be interpreted as gender-dependent stress coping reaction.

Additionally, the immune response to chronic stress due to parasite infection may elicit a different immune response in the two sexes and lead to different behavioral changes. The research focused on the role of cytokines in the inflammatory process of the CNS showed the presence of a difference between males and females (Bernardi et al., 2020; Lefèvre et al., 2012). The increment of cytokines during toxoplasmosis infection can affect the dopaminergic function (Felger & Miller, 2012) and provoke the behavioral change (Kamerkar & Davis, 2012). Indeed, many of the behavioral changes associated with toxoplasmosis are correlated with the dopaminergic system in humans and in rats (Flegr, 2007; Skallová et al., 2006). Therefore, the behavioral differences observed in infected men and women can be a consequence of differences in the immune system, related to the production of cytokines and the function of the dopaminergic system.

Among the traits affected by toxoplasmosis are also such that could affect the political beliefs of infected subjects. For example, infected men have lower superego strength, i.e., the tendency to disregard the rules of society, while the opposite tendency was observed in infected women (Flegr, 2007). Both infected men and women have lower conscientiousness (Lindová et al., 2012), lower generosity (Lindová et al., 2010) and have lower novelty-seeking (Flegr et al., 2003; Skallová et al., 2005). Men express lower and women higher altruism in the dictatorship game (Lindová et al., 2010). Infected subjects express higher entrepreneurship activity and have a higher tendency for risky behavior (Johnson et al., 2018), which might explain a higher risk of traffic accidents in the infected subjects (Flegr et al., 2002; Kocazeybek et al., 2009). *Toxoplasma*-infected women have also higher aggression (Cook et al., 2015) and self-aggression (Mathai et al., 2016; Postolache et al., 2021).

Although certain effects of infection by a parasitic disease on political views ought to be expected based on the research summarized above, to our knowledge, no relation of toxoplasmosis (or other parasitic diseases) and political beliefs and values has been reported to date. Moreover, studies focused on possible causal connection between infectious diseases (except for toxoplasmosis) and changes in personality profiles are also few and far between, and focused mainly on diseases causing profound

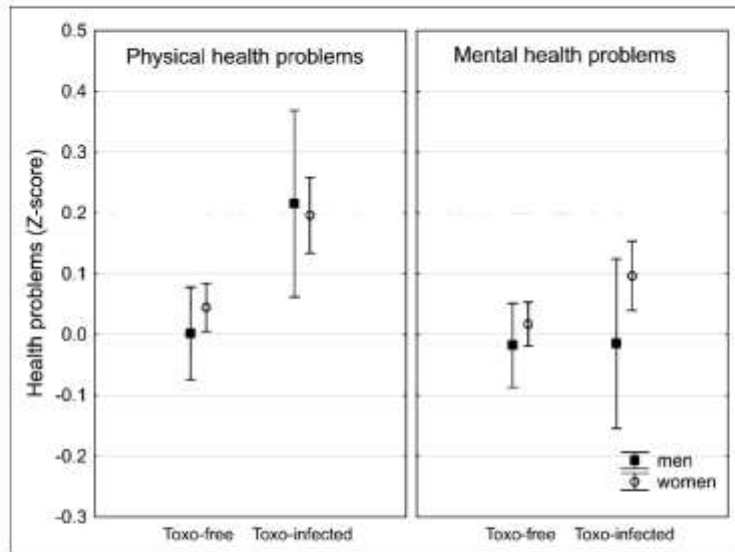
changes in brain, e.g. AIDS (Meehan & Brush, 2016) or neuro-cysticercosis (El-Kady et al., 2021).

Regarding the connection between parasite stress and political views, there exists extensive literature describing changes in traits that differ with varied geographical parasite stress levels. All published studies, however, are population-based studies, searching for the difference between populations exposed to different intensity of the parasite stress. While it is possible to consider our data in the context of this parasite-stress theory and it will be done so in the Discussion, this study primarily focuses on association between *Toxoplasma* status and political beliefs on an individual level. For this purpose, we measured political beliefs and values of 2315 respondents of an internet survey (477 *Toxoplasma*-infected) using the Political Beliefs and Values Inventory (PI34). To test the hypothesis that the observed changes in three of the factor (Tribalism, Cultural Liberalism, and Anti-authoritarianism), but not the fourth (Economic Equity), are related to stress-coping strategies of men and women, we ran the analyses separately for each sex and to find out whether the observed differences are mediated by impaired health of the infected subjects we also analysed a complex model controlling not only for age, sex, size of place of residence, and level of education but also for physical health and mental health.

## Results

The final data set of 2315 respondents, consisted of 467 men (mean age 35.54, s.d. 12.8) and 1848 women (mean age 35.95, s.d. 11.5); the difference in age was not significant ( $t = -0.69$ ,  $df = 2313$ ,  $p = .493$ ). In men, 90 men (mean age 35.54, s.d. 12.9) were infected with *Toxoplasma* and 377 (mean age 35.53, s.d. 12.4) were *Toxoplasma*-free; again, the difference in age was not significant ( $t = -0.69$ ,  $df = 2313$ ,  $p = .493$ ). In contrast, the difference in age between 518 *Toxoplasma*-infected women (mean age 38.39, s.d. 12.8) and 1330 *Toxoplasma*-free women (mean age 35.01, s.d. 11.2) was highly significant ( $t = -5.72$ ,  $df = 1846$ ,  $p < .0001$ ). Prevalence of toxoplasmosis was lower in men (19.27%) than in women (28.03%) ( $\chi^2 = 14.8$ ,  $p = .0001$ ). Prevalence of toxoplasmosis decreased with size of place of residence from 32.9% in smallest villages to 20.9% in the cities over 500,000 inhabitants (Partial Tau controlled for age and sex =  $-0.070$ ,  $p < .0001$ ). No association was observed between toxoplasmosis and education (Partial Tau controlled for age and sex =  $-0.013$ ,  $p = .368$ ), however, the absence of the association in the whole population was caused by the existence of positive association between toxoplasmosis and achieved level of education in men (Partial Tau controlled for age and sex =  $0.083$ ,  $p = .007$ ) and negative association between the toxoplasmosis and achieved level of education in women (partial Tau controlled for age and sex =  $-0.037$ ,  $p = .017$ ).

*Toxoplasma*-infected women expressed worse mental and physical health, and *Toxoplasma*-infected men worse physical health than their *Toxoplasma*-free peers (Figure 1). MANCOVA computed with the physical and mental health



**Figure 1.** Association of toxoplasmosis and sex on health.

Note. The vertical bars show 95% C.I., the mean values computed for covariates (age and size of place of living) at their means.

problems as the dependent variables, toxoplasmosis, sex, and sex-toxoplasmosis interaction as the factors, and age and size of place of living as covariates showed significant effects of age ( $F_{2,1843} = 70.02$ ,  $p < 0.0001$ , partial  $\eta^2 = 0.07$ ) and toxoplasmosis ( $F_{2,1843} = 7.34$ ,  $p = .0007$ , partial  $\eta^2 = 0.008$ ). Results of corresponding ANCOVA analyses computed separately for physical and mental health problems and men, women, and men and women together are shown in Table 1.

Factor analysis performed on data from the Political Inventory revealed four factors with an eigenvalue higher than 1.0. According to loading of these factors by particular questions of the inventory (Table 2), it was possible to identify these factors as Tribalism (11.5% of variability), Economic equity preference (7.8% of variability), Cultural liberalism (5.1% of variability), and Anti-authoritarianism (3.0% of variability). We computed values of these four variables for each respondent as scores in particular factors.

To disclose the effect of *Toxoplasma* infection on political beliefs, we computed the associations of the toxoplasmosis with Tribalism, economic equity preference, Cultural liberalism, and Anti-authoritarianism using partial Kendall correlation controlled to age, sex, and size of place of residence. The results showed that being *Toxoplasma*-infected correlated positively with Tribalism and negatively with Cultural liberalism, and Anti-authoritarianism, see Table 3. All three associations retain statistical significance after the correction for multiple tests. Table 3 also shows the results of the same analyses performed separately for men and women. The most conspicuous

sexual differences were the absence of the positive association of toxoplasmosis with Tribalism in men – here the partial Tau was even negative,  $-0.014$ ,  $p = .760$  in men vs.  $0.040$ ,  $p = .009$  in women, and the presence of the positive association with the preference for economic equity in men  $0.064$ ,  $p = .040$  but not in women  $\text{Tau} = 0.010$ ,  $p = .526$ , see Figure 2.

Toxoplasmosis had a strong association with human health and the size of the place of residence. At the same time, political beliefs had a strong association with health, size of place of residence, and achieved level of education (Table 3, left part). It could be therefore speculated that the association between toxoplasmosis and health is just the side effects of the associations of toxoplasmosis with these confounding variables. To test this possibility, we repeated the partial Kendall analyses, this time controlled not only for sex, age, and size of place of residence but also for physical and mental health and achieved level of education. The results were nearly identical as those obtained with the simpler model (Table 3, right part). Therefore, the effect of toxoplasmosis on political beliefs was not mediated by the effect of toxoplasmosis on health or education.

## Discussion

The present study showed that *Toxoplasma*-infected and *Toxoplasma*-free participants of our cross-sectional study differed in three of four factors measured with Political inventory, namely scored higher in Tribalism and lower in Cultural liberalism and Anti-authoritarianism. These results are in line with

**Table 1.** the Associations Between Physical and Mental Health and Toxoplasmosis.

	Physical health problems			Mental health problems		
	<b>F1,1846</b>	<b>p</b>	<b>Partial eta2</b>	<b>F1,2301</b>	<b>p</b>	<b>Partial eta2</b>
<b>All</b>						
Intercept	<b>24.86</b>	0.0000	0.013	<b>10.89</b>	0.0010	0.005
Age	<b>64.94</b>	0.0000	0.034	<b>35.39</b>	0.0000	0.015
Size of place of living	3.50	0.0615	0.002	0.51	0.4749	0.000
Toxoplasmosis	<b>22.11</b>	0.0000	0.012	<b>6.64</b>	0.0100	0.003
<b>Men</b>	<b>F1,353</b>	<b>p</b>	<b>Partial eta2</b>	<b>F1,462</b>	<b>p</b>	<b>Partial eta2</b>
Intercept	<b>9.99</b>	0.0017	0.028	0.00	0.9539	0.000
Age	<b>23.44</b>	0.0000	0.062	<b>6.75</b>	0.0097	0.014
Size of place of living	2.16	0.1423	0.006	2.10	0.1484	0.005
Toxoplasmosis	<b>5.95</b>	0.0153	0.017	0.03	0.8681	0.000
<b>Women</b>	<b>F1,1489</b>	<b>p</b>	<b>Partial eta2</b>	<b>F1,1835</b>	<b>p</b>	<b>Partial eta2</b>
Intercept	<b>15.88</b>	0.0001	0.011	<b>11.67</b>	0.0006	0.006
Age	<b>43.40</b>	0.0000	0.028	<b>28.36</b>	0.0000	0.015
Size of place of living	2.15	0.1423	0.001	0.17	0.6824	0.000
Toxoplasmosis	<b>16.19</b>	0.0001	0.011	<b>6.32</b>	0.0120	0.003

Note. The table shows the effects of toxoplasmosis, age, and size of place of living computed by ANCOVA for all subjects, men, and women. Statistically significant results were printed in bold.

**Table 2.** Loadings of Four Factors Identified by Factor Analysis of 34 Questions of the Political Inventory.

	Tribalism	Economic Equity	Cultural Liberalism	Anti-authoritarianism
1 reduce income differences	-0.029	0.768	-0.053	0.000
2 same sex together	-0.255	-0.036	0.504	0.113
3 right to protest	-0.062	0.004	0.082	0.466
4 to regulate corporation	-0.072	0.289	0.057	-0.063
5 good for corp. is good for society	0.190	-0.156	-0.008	-0.069
6 minimize differences in status	-0.007	0.800	-0.075	0.001
7 pornography	-0.051	-0.058	0.351	0.302
8 citizens over foreigners	0.674	0.049	-0.086	0.099
9 minorities are threat	0.715	0.056	-0.128	0.004
10 class not nationality	-0.228	0.172	0.121	0.018
11 rich are taxed too much	0.108	-0.416	0.032	0.047
12 religion not to be disgraced	-0.088	0.128	-0.060	-0.428
13 decriminalize marijuana	-0.108	0.006	0.186	0.264
14 euthanasia	0.170	0.037	0.644	0.149
15 parents can spank children	0.388	-0.100	0.019	0.064
16 there are no barbarians	-0.309	0.131	0.123	-0.130
17 cancel borders and nations	-0.466	0.194	0.094	0.098
18 immigrants won't assimilate	0.624	0.021	-0.128	0.034
19 no work no social benefits	0.448	-0.137	0.075	-0.042
20 to each according needs	0.056	0.363	0.046	-0.081
21 mothers stay at home	0.380	0.071	-0.305	-0.072
22 free speech	0.043	-0.097	0.080	0.554
23 no abortion	0.065	0.066	-0.601	-0.101
24 respect authority	0.368	-0.013	0.002	-0.242
25 proud of history	0.267	-0.030	-0.100	-0.032
26 poor but equal	0.108	0.664	-0.082	-0.104
27 free market is free society	0.110	-0.367	0.105	0.223
28 home-school	-0.184	-0.030	-0.079	0.209
29 no to internet censorship	0.025	-0.067	0.096	0.453
30 an eye for an eye	0.610	0.074	0.164	-0.008
31 my life for the nation	0.244	-0.037	-0.180	0.047
32 capital punishment	0.635	0.062	0.228	-0.122
33 right to firearm	0.356	-0.178	-0.034	0.360
34 deport foreign criminals	0.412	0.023	-0.048	0.049

**Table 3.** Correlation of Political Beliefs with Toxoplasmosis and Potential Confounding Variables.

	Tribalism	Economic Equity	Cultural liberalism	Anti-authoritarianism	Tribalism	Economic Equity	Cultural liberalism	Anti-authoritarianism
	<b>ALL</b>				<b>ALL</b>			
Toxo	<b>0.029</b>	0.020	<b>-0.042</b>	<b>-0.034</b>	0.023	0.013	<b>-0.046</b>	<b>-0.032</b>
Age	<b>0.024</b>	<b>-0.039</b>	<b>-0.104</b>	<b>-0.033</b>	<b>0.028</b>	<b>-0.023</b>	<b>-0.092</b>	<b>-0.022</b>
sex	0.003	<b>0.032</b>	<b>0.062</b>	<b>-0.234</b>	0.003	<b>0.031</b>	<b>0.054</b>	<b>-0.234</b>
living	<b>-0.089</b>	<b>-0.054</b>	<b>0.063</b>	<b>0.027</b>	<b>-0.076</b>	<b>-0.045</b>	<b>0.058</b>	<b>0.026</b>
education	<b>-0.098</b>	<b>-0.080</b>	<b>0.023</b>	-0.008	<b>-0.097</b>	<b>-0.076</b>	<b>0.024</b>	-0.008
physical health	<b>0.043</b>	<b>0.027</b>	-0.002	<b>-0.041</b>	<b>0.044</b>	<b>0.022</b>	<b>-0.014</b>	<b>-0.054</b>
mental health	-0.001	<b>0.070</b>	<b>0.031</b>	<b>0.025</b>	-0.010	<b>0.063</b>	<b>0.033</b>	<b>0.030</b>
	<b>WOMEN</b>				<b>WOMEN</b>			
Toxo	<b>0.040</b>	0.010	<b>-0.038</b>	-0.030	0.030	-0.001	<b>-0.042</b>	-0.027
Age	<b>0.022</b>	<b>-0.034</b>	<b>-0.105</b>	<b>-0.027</b>	<b>0.029</b>	<b>-0.021</b>	<b>-0.098</b>	<b>-0.022</b>
living	<b>-0.092</b>	<b>-0.053</b>	<b>0.055</b>	<b>0.017</b>	<b>-0.079</b>	<b>-0.047</b>	<b>0.050</b>	<b>0.020</b>
education	<b>-0.124</b>	<b>-0.084</b>	<b>0.025</b>	<b>-0.020</b>	<b>-0.121</b>	<b>-0.079</b>	<b>0.027</b>	<b>-0.022</b>
physical health	<b>0.058</b>	<b>0.036</b>	-0.003	<b>-0.043</b>	<b>0.057</b>	<b>0.028</b>	<b>-0.022</b>	<b>-0.034</b>
mental health	0.013	<b>0.073</b>	<b>0.056</b>	0.001	0.000	<b>0.065</b>	<b>0.060</b>	0.004
	<b>MEN</b>				<b>MEN</b>			
Toxo	-0.015	<b>0.064</b>	-0.060	-0.055	-0.009	<b>0.065</b>	<b>-0.071</b>	-0.059
Age	<b>0.022</b>	<b>-0.045</b>	<b>-0.079</b>	<b>-0.037</b>	0.024	<b>-0.028</b>	<b>-0.081</b>	-0.025
living	<b>-0.080</b>	<b>-0.050</b>	<b>0.073</b>	<b>0.041</b>	<b>-0.073</b>	<b>-0.043</b>	<b>0.071</b>	<b>0.040</b>
education	<b>-0.061</b>	<b>-0.074</b>	<b>0.020</b>	0.012	<b>-0.062</b>	<b>-0.072</b>	<b>0.020</b>	0.014
physical health	<b>0.022</b>	0.015	-0.001	<b>-0.041</b>	<b>0.021</b>	0.016	-0.002	<b>-0.079</b>
mental health	<b>-0.024</b>	<b>0.066</b>	-0.008	<b>0.060</b>	<b>-0.028</b>	<b>0.061</b>	-0.007	<b>0.067</b>

Note. The left half of the table shows partial Kendall Tau controlled for age, sex, size of place of residence, and the right half Tau controlled for age, sex, size of place of residence, physical health, mental health, and achieved education level. A variable that was used as the dependent variable) was removed from the list of covariates to be controlled for. Statistically significant results were printed in bold.

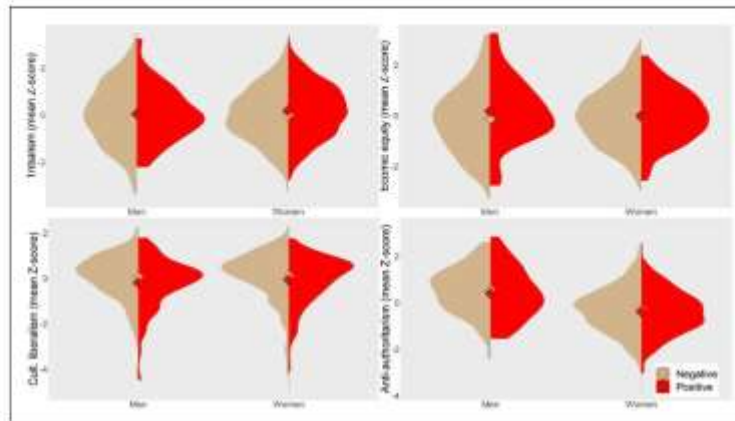
previous broad research, showing that individuals in parasite affected areas are more likely to be conservative and authoritarian (Murray et al., 2013).

Furthermore, we observed sex differences in the studied factors associated with the *Toxoplasma* infection. Indeed, *Toxoplasma*-infected men scored higher in Economic Equity, showing a preference for a more equal and less competitive society, while women infected with toxoplasmosis scored higher in tribalism and lower in cultural liberalism. These associations were not reduced when the effect of worse physical health and mental health were controlled, suggesting that impaired health of infected subjects is not the cause of changes in political beliefs. The same conclusion was also supported by the fact that the changes go in the same direction in men and women, because stress coping-associated behavioral and personality changes mostly go in different directions in men and women.

It was suggested by Lindová et al. (2006, 2012) that these associations might be the results of a mild chronic stress caused by the toxoplasmosis infection instead of the result of the toxoplasmosis itself. The presence of a chronic stress not only explain the presence of behavioral and political differences

with the non-infected control group, but also the presence of sex differences in these behavior and ideologies as different responses to chronic stress in the two sexes, involving differences in the immune system response and in the coping strategies used. Many of the behavioral changes observed in toxoplasmosis infected people correlate with the function of dopamine in the brain and they may have more broad implications, including political ideologies. In line with our results, previous studies (Flegr et al., 2003; Skallová et al., 2005) showed that infected subjects scored lower in novelty seeking, a factor that contributes to a conservative and political opinion (Carney et al., 2008). Indeed, in our sample the infection was associated with higher tribalism and lower cultural liberalism, specifically in women. While we expected differences in the political ideologies of infected men and women, we did not expect a higher score in economic equity in infected men. Typically, men affected with toxoplasmosis showed higher risk propensity and higher entrepreneurial activities (Johnson et al., 2018) more compatible with a competitive type of economy. The association of toxoplasmosis and the preference for an egalitarian economy in men needs to be better explored in future works.





**Figure 2.** The violin plots show the means and the distribution computed as the kernel probability density for male and female, toxoplasma-infected and toxoplasma-free participants of the study.

Several studies have found that societies that are more affected by infectious pathogens also exhibit higher levels of conservative political attitudes such as xenophobia and traditionalism (Bennett & Nikolaev, 2020; Murray et al., 2013; Nikolaev & Salahodjaev, 2017; Thornhill et al., 2009). Similar results have also been found in our study performed on the individual level. The hypothesis that has been proposed is that the attitudes exhibited are connected to pathogen avoidance behaviors aimed at minimizing contact with outsiders (intergroup effect) (Aarseth et al., 2017) who may be carrying new pathogens as well as the maintenance of social traditions that may serve to help protect against pathogens (intragroup effect) (Fincher & Thornhill, 2012), with evidence from a recent cross-national study favoring the intragroup effect (Tybur et al., 2016). Significantly, however, both the effects – and the intragroup effect in particular – seem potentially open to the interpretation that they are generalized responses to stress, rather than to a pathogen (Brown et al., 2016; Currie & MacE, 2012; Hruschka & Henrich, 2013; Ma, 2020).

While the present study examines differences found in association between a parasitic infection and political values in context of increased stress at individual level, the results can be seen as an alternative to the parasite stress theory (Fincher et al., 2008; Thornhill & Fincher, p. 2014) for the following two main reasons: Parasite-stress theory aims to be the ultimate evolutionary explanation of changes in traits that differ with varied geographical parasite stress levels, yet this study focuses directly on the difference between actual infected and non-infected subjects in one small region, where the intensity of the parasite stress is mostly constant (and low). Second reason is that it has been shown that primarily non-zoonotic diseases have a relation to human personality and societal values (Thornhill et al., 2010). However, toxoplasmosis is primarily a zoonotic disease with very specific and limited spread between people – the only

intrapersonal route of infection was suggested from male to female or between two male partners through sexual transmission (Flegr, Klapilová, et al., 2014; Flegr, Prandota, et al., 2014; Hlaváčová et al., 2021; Kaňková et al., 2020).

This being said, there might be a possible connection between the present study and the parasite-stress theory after all. An extensive body of research confirms association between infectious (and in most cases parasitic) diseases and changes in personality profile of animals from molluscs (Seaman & Briffa, 2015) and minnows (Kekiläinen et al., 2013) to migratory birds (Marinov et al., 2017) and mammals (Boyer et al., 2010) including men (Webster, 2001). There are also well studied associations between personality traits and political views, e.g. (Furnham & Fenton-O’Creevy, 2018; Harell et al., 2021; Verhulst et al., 2010; Wang, 2016). While the direction of causality needs to be studied further (Bakker et al., 2021) and while the human-centred field of parasite induced changes in personality traits is regrettably understudied and quite complex (Friedman, 2008), we might expect at least some effect of infectious diseases on political attitudes caused by shifts of personality traits. A possibility thus exists, that at least part of the reported difference in political attitudes in countries with different parasitic disease burdens is not caused by parasite-avoidance but results from a significant part of the population being infected with one or multiple pathogens. This hypothesis is supported by studies that linked a change in personality traits with clear connection to political attitudes (e.g., conservatism) with chronic diseases, although not the infectious ones (Mendelsohn et al., 1995; Sutin et al., 2013). On the other hand, some results suggest stronger prediction of personality traits by historical prevalence of diseases rather than by the current situation, suggesting parasite-avoidance as a factor with greater importance in personality shifts (Schaller & Murray, 2008).

Since the available body of literature discussing possible causal relationships between infectious diseases and political beliefs and values is very sparse, this direction of research might provide interesting and important insight into the changing political climate in certain countries. Studies focused on a wider range of infectious diseases besides toxoplasmosis and severe debilitating illnesses such as neurocysticercosis or AIDS would be especially valuable.

The present study showed a 19% prevalence of toxoplasmosis in men and 28% in women. The most recent Czech epidemiological study performed between 2014–2015 (Flegr, 2017) showed the prevalence of 25% in men and 36% in women aged 30–39 years. It is known, however, that prevalence of toxoplasmosis decreases relatively quickly in most developed countries, including the Czech Republic. For example, a large epidemiological study performed on Czech male soldiers 20 years ago found a prevalence of 35% for the age strata 30–35 years (Kolbekova et al., 2007). It is, therefore, possible that the observed seroprevalence reflects the actual situation in the Czech general population.

### Limitations

The main limitation of the present study was the fact that the participants of the study were self-selected. Their subpopulation probably represents a specific (more altruistic and more curious) segment of the Czech population, rather than a random sample of the Czech internet population. In addition, people with impairments or severe diseases as well as those from the lowest socioeconomic strata were unable to participate. It is not therefore clear to which extent the results can be generalized to the general Czech (or the World) population.

Another limitation of the study is the moderate number of *Toxoplasma*-infected men (90). The reason of imbalanced sex ratio was the fact that women are often tested for toxoplasmosis during pregnancy and therefore a larger fraction of women than men know their toxoplasmosis status. Due to the lower number of men, the associations of toxoplasmosis with Cultural liberalism and Anti-authoritarianism were not significant in men, despite being stronger in men than in the more numerous (518) women.

Since this study dealt with the effect of pathogen-caused stress, there is a possible interference of the global COVID-19 pandemics. However, only 6% of the respondents participated in the period between April 2020 when the first infection in Czechia was observed, and the end of data collection in April 2021.

In the present study, we calculated aggregate indices of physical and mental health and used these indices in our statistical models. In future studies, it will be valuable to analyse the effects of individual health-related variables to disentangle the complex relationships between toxoplasmosis, mental and physical health, psychological traits and political beliefs. Such research could also answer important questions related to the causal direction of the observed correlations.

The associations found in the present study are based on correlations and we cannot infer the direction of causality. It cannot

be ruled out that the explanation of the effect is in the opposite direction, that e.g., higher tribalism itself, by an unknown mechanism, increases the chance of being infected by *Toxoplasma gondii*. Like nearly all past studies, also this one was cross-sectional in nature. It is very difficult to study the relationship between *Toxoplasma* infection and personality using a longitudinal design. The frequency of the *Toxoplasma* infection in adulthood is low, and thousands of participants should be recruited to find several dozens of subjects who will acquire the infection during the study. Until such a study is performed, any conclusion about the causality behind the correlation between the infection and human personality must be based only on analogies with animal models (Skallová et al., 2006; Hodková et al., 2007) or on the existence of a correlation between length of infection and observed personality trait changes (Flegr et al., 1996, 2000) and must be therefore considered only provisional.

Further research is also needed to better clarify the extension and the implication of the associations we found between toxoplasmosis infection and political ideologies, and to clarify the role of sex differences.

## Material and Methods

### Subjects

The questionnaire was distributed via the internet as an online Qualtrics survey. Participants were invited to partake in the study using a social media-based snowball method (Kankova et al., 2015). Potential volunteers, mostly members of the “Lab bunnies” community, an 18,000-member group of Czech and Slovak nationals willing to take part in evolutionary psychology experiments, were invited to participate in an anonymous study about “mystical thinking, superstitions, prejudices, religion and the relation between various environmental factors and health and wellbeing.” The online questionnaire was also promoted in various types of electronic and printed media and TV. Toxoplasmosis and political belief were not explicitly mentioned in the information provided to potential participants in order not to affect the participants’ responses; this procedure was approved by the IRB, see below. They were, however, informed about that the data anonymously collected in the research will be health- and beliefs-related and gave their informed consent with participation in the experiment. Responders were not monetarily compensated for their participation in this study; nevertheless, after finishing the approximately 80 min questionnaire, they were provided with information about the results of related studies and their own results of several tests that were part of the questionnaire.

At the beginning of the survey, the participants were given the following information and were asked to provide their informed consent to participate in the study by pressing a special button: “The study is anonymous and obtained data will be used exclusively for scientific purposes. Your cooperation in the project is voluntary and you can terminate it at any time by closing this web page. You can also skip any

uncomfortable questions; however, complete data is most valuable. If you agree to participate in the research press the "Next" button". Only the subjects who provided their informed consent (95.2% of the visitors of the web page) were allowed to participate in the study. Between 27th May 2016 and 24th April 2021, 14,902 responders took part in the study.

All methods were performed in accordance with the relevant guidelines and regulations. The project, including the method of obtaining electronic informed consent to participate in this anonymous study from all participants, was approved by the IRB anonymized.

### Questionnaire

The electronic survey consisted of several parts that concerned various unrelated projects on evolutionary psychology and psychiatry. In the present study, we inspected and analysed only the data from the Political Inventory (see below), *Toxoplasma* seropositivity. Regarding toxoplasmosis, participants were asked whether they were laboratory tested for toxoplasmosis, what was the result of the laboratory test, and what was the reason for their test. Only participants who were tested due to pregnancy (62% of female participants), health reasons (28% male and 20% female participant), or as part of other research (60% male and 35% female participants) and were, therefore, informed about their toxoplasmosis status by a medical or research professional were included in the study. In the Czech Republic, IgG ELISA is always used in clinical diagnostic, however, in our research, the participants were tested with both ELISA and complement fixation test, therefore a certain fraction of participants were tested with both tests. To control for potential confounding variables, the responders were also asked about their sex, age, education (ordinal scale 1–8: 1–elementary, 8 – Ph.D.), and the size of the communities where they currently live (ordinal variable urbanization: 0 – less than 1000 inhabitants, 1: 1–5 thousand inhabitants, 2: 5–50 thousand inhabitants, 3: 50–100 thousand inhabitants, 4: 100–500 thousand inhabitants, 5: more than 500 thousand inhabitants).

In the anamnestic part of the questionnaire, the participants were also asked to respond to a set of questions concerning their physical health and mental health. They were asked how many different prescription drugs they were taking currently, how many different non-prescription drugs they were taking currently ("how many different herbs, food supplements, multivitamins, superfoods, etc. do you currently take per day"), how many times they visited their primary care doctor in past 365 days ("not for prevention"), how many times they used antibiotics in the past 365 days, and how many different medical specialists they visited (not for prevention) in the past five years. The physical health problems score was calculated as a mean of Z-scores of the last five variables. The responders were also requested to rate how much they suffer from anxieties, phobias, depression, mania, obsessions, auditory hallucinations, and visual hallucinations using a 0–100 scale. We also counted the number of diagnosed and the number of undiagnosed mental health disorders that the responders checked on

a list of 25 mental health disorders and epilepsy. The mental health problems score was calculated as a mean of Z-scores of these 9 variables.

Political beliefs and values of the responders were measured with the Political Inventory (PI34). This instrument consisted of 34 questions (see English translation in appendix) in randomized order. The responders rated the strength of agreement with statements on a scale of 1–5 (1 – strongly disagree, 2 – rather disagree, 3 – neither agree nor disagree, 4 – rather agree, 5 – strongly agree). The ratings were inverted for questions marked with r. Tribalism is calculated as the arithmetic means of responses to questions 8, 9, 15, 16r, 17r, 18, 19, 21, 24, 25, 30, 32, 34; Economic equity as the arithmetic mean of responses to questions 1, 4, 6, 11r, 20, 26, 27r; Cultural liberalism as the arithmetic mean of responses to questions 2, 7, 14, 21r, 23r; and Anti-authoritarianism as the arithmetic mean of responses to questions 3, 12r, 13, 22, 24r, 29, 33. Rather than examining actual partisanship related to established political parties or established ideologies, we assembled the Political Inventory based on contemporary public discourse in the Czech society as a culturally specific questionnaire was needed. Selection of topics for PI34 was inspired mainly by 10 items in the ideological Consistency Scale by PEW based on the using set of questions since 1994 (Dimock et al., 2014) and the 12-item political and social conservatism scale (Everett, 2013).

### Statistics

The raw data file contained the records of 14,902 participants, some of them, however, did not complete the final parts of the questionnaire with the political inventory. We filtered out all subjects younger than 18 and older than 90 years. We also filtered out all records with suspicious data, e.g., those from subjects reporting too low (< 130 cm) or high (> 215 cm) body height and too high number of reported diagnosed or undiagnosed mental health disorders (>10). This data set was used for computing physical health problem scores, mental health score (see above), and for factor analyses. In the main analyses, we used the factors (scores) computed with the Maximum likelihood factors extraction method after varimax rotation, however, the results were very robust; a nearly identical structure of factors was obtained with another method of factor extraction and other methods of factor rotation. In fact, all four factors (plus four other factors) were identified even by the principal component analysis.

For analysing the association between toxoplasmosis and political beliefs we used a core data set consisting of 2315 responders who had been tested for toxoplasmosis in the past and were able to provide us with the information about the results of their laboratory test. About 39% of participants were tested in our laboratory during other scientific studies, 29% during pregnancy, 22% because of their health problems, and 10% for other reasons, e.g., curiosity, blood donation, their partner's pregnancy, etc. Some output variables had asymmetric distribution; therefore, we performed all focal analyses with a



partial Kendall test that is not sensitive to such irregularity in data. The correction for multiple tests was performed with the Benjamini-Hochberg procedure with a False Discovery Rate set to 0.1 (Benjamini & Hochberg, 1995). Partial Kendall tests were performed in R 3.3.1 (R Core Team, 2018) using the package `ppcor` (Kim, 2015). All other tests, factor analyses, and PCA analyses were performed with the statistical package *Statistica v. 10.0*.

## Conclusions

Beside worse mental health in *Toxoplasma*-infected women and worse physical health of *Toxoplasma*-infected participants of both sexes, our study showed sex-dependent differences in Tribalism, Cultural liberalism, and Anti-authoritarianism between *Toxoplasma*-positive and *Toxoplasma*-negative subjects. Specifically, we have found positive association of toxoplasmosis and Tribalism in women but not in men, and a positive association of toxoplasmosis and preference for social equity in men but not in women, in both sexes, there were also negative associations with Cultural liberalism and Anti-Authoritarianism. While our results are in agreement with the parasite-stress theory that predicts higher Authoritarianism and Conservatism in areas with higher prevalence of parasitic diseases, the results cannot be interpreted simply in the framework of this theory, i.e., as a result of enhanced behavioral immunity (parasite avoidance). Rather, the direct effect of long-term mild stress caused by an inflammatory reaction to the presence of a parasite may be at play: shifting hosts' personalities and thus also their political attitudes. Considering the substantial lack of literature discussing effects of mild infectious diseases on personality traits of infected individuals as well as effects of being infected by a parasite on political beliefs and values, further research should be encouraged.

## Author Contribution

Conceptualization RK, LP, JF; methodology RK, LP, SB, KTK, JF; data curation JF; statistical analysis JF; writing - original draft and reviewing and editing RK, LP, SB, KTK, JF. All authors have read and agreed to the published version of the manuscript.

## Data Availability

Dataset available on Figshare.com <https://figshare.com/s/fb78539b38e11db6f73f>

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Grantová Agentura České Republiky, (grant number 21-31059S).

## Institutional Review Board Statement

This study was approved by the IRB of the Faculty of Science, Charles University (Komise pro práci s lidmi a lidským materiálem Přírodovědecké Fakulty Univerzity Karlovy)—No. 2015/07

## Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

## ORCID iD

Robin Kopecký  <https://orcid.org/0000-0002-7140-4649>

## References

- Aarøe, L., Petersen, M. B., & Areneaux, K. (2017). The behavioral immune system shapes political intuitions: Why and how individual differences in disgust sensitivity underlie opposition to immigration. *American Political Science Review*, *111*(2), 277–294. <https://doi.org/10.1017/S0003055416000770>
- Bakker, B. N., Letkes, Y., & Malka, A. (2021). Reconsidering the link between self-reported personality traits and political preferences. *American Political Science Review*, *115*(4), 1–17. <https://doi.org/10.1017/S0003055421000605>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bennett, D. L., & Nikolaev, B. (2020). Historical disease prevalence, cultural values, and global innovation. *Entrepreneurship Theory and Practice*, *45*(1), 145–174. <https://doi.org/10.1177/1042258720914506>
- Bernardi, S., Toffoli, B., Tonon, F., Francica, M., Campagnolo, E., Ferretti, T., Comar, S., Giudici, F., Stenner, E., & Fabris, B. (2020). Sex differences in proinflammatory cytokine levels. *International Journal of Molecular Sciences*, *21*(11), 3861. <https://doi.org/10.3390/ijms21113861>
- Boyer, N., Réale, D., Marmet, J., Pisanu, B., & Chapuis, J.-L. (2010). Personality, space use and tick load in an introduced population of siberian chipmunks *tamias sibiricus*. *Journal of Animal Ecology*, *79*(3), 538–547. <https://doi.org/10.1111/j.1365-2656.2010.01659.x>
- Brown, G. D. A., Fincher, C. L., & Walasek, L. (2016). Personality, parasites, political attitudes, and cooperation: A model of how infection prevalence influences openness and social group formation. *Topics in Cognitive Science*, *8*(1), 98–117. <https://doi.org/10.1111/TOPS.12175>
- Carney, D. R., Jost, J. T., Gosling, S. D., & Potter, J. (2008). The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind. *Political Psychology*, *29*(6), 807–840. <https://doi.org/10.1111/j.1467-9221.2008.00668.x>
- Cook, T. B., Brenner, L. A., Cloninger, C. R., Langenberg, P., Ighide, A., Giegling, I., Hartmann, A. M., Konte, B., Friedl, M., Brundin, L., Groer, M. W., Cui, A., Rajescu, D., & Postolache, T. T. (2015). "Latent" infection with *Toxoplasma gondii*: Association with trait aggression and impulsivity in healthy adults. *Journal of Psychiatric Research*, *60*, 87–94. <https://doi.org/10.1016/j.jpsychires.2014.09.019>

- Currie, T. E., & MacE, R. (2012). Analyses do not support the parasite-stress theory of human sociality. *Behavioral and Brain Sciences*, 35(2), 83–85. <https://doi.org/10.1017/S0140525X11000963>
- El-Kady, A. M., Allemaitem, K. S., Almatroudi, A., Abler, B., & Elsayed, M. (2021). Psychiatric disorders of neurocysticercosis: Narrative review. *Neuropsychiatric Disease and Treatment*, 17, 1599–1610. <https://doi.org/10.2147/NDT.S306585>
- Everett, J. A. C. (2013). The 12 item social and economic conservatism scale (SECS). *PLOS ONE*, 8(12), e82131. <https://doi.org/10.1371/JOURNAL.PONE.0082131>
- Fekadu, A., Shibre, T., & Cleare, A. J. (2010). Toxoplasmosis as a cause for behaviour disorders - overview of evidence and mechanisms. *Folia Parasitologica*, 57(2), 105–113. <https://doi.org/10.14411/fp.2010.013>
- Felger, J. C., & Miller, A. H. (2012). Cytokine effects on the basal ganglia and dopamine function: The subcortical source of inflammatory malaise. *Frontiers in Neuroendocrinology*, 33(3), 315–327. <https://doi.org/10.1016/j.yfrne.2012.09.003>
- Fincher, C. L., & Thornhill, R. (2012). Parasite-stress promotes in-group assortative sociality: The cases of strong family ties and heightened religiosity. *Behavioral and Brain Sciences*, 35(2), 61–79. <https://doi.org/10.1017/S0140525X11000021>
- Fincher, C. L., Thornhill, R., Murray, D. R., & Schaller, M. (2008). Pathogen prevalence predicts human cross-cultural variability in individualism/collectivism. *Proceedings of the Royal Society B: Biological Sciences*, 275(1640), 1279–1285. <https://doi.org/10.1098/RSPB.2008.0094>
- Flegr, J. (2007). Effects of toxoplasma on human behavior. *Schizophrenia Bulletin*, 33(3), 757–760. <https://doi.org/10.1093/schbul/sbl074>
- Flegr, J. (2013). Influence of latent toxoplasma infection on human personality, physiology and morphology: Pros and cons of the toxoplasma-human model in studying the manipulation hypothesis. *Journal of Experimental Biology*, 216(1), 127–133. <https://doi.org/10.1242/JEB.073635>
- Flegr, J. (2017). Predictors of *Toxoplasma gondii* infection in Czech and Slovak populations: The possible role of cat-related injuries and risky sexual behavior in the parasite transmission. *Epidemiology & Infection*, 145(7), 1351–1362. <https://doi.org/10.1017/S095026881700019X>
- Flegr, J., & Escudero, D. Q. (2016). Impaired health status and increased incidence of diseases in toxoplasma-seropositive subjects - an explorative cross-sectional study. *Parasitology*, 143(14), 1974–1989. <https://doi.org/10.1017/S0031182016001785>
- Flegr, J., Havlíček, J., Kodym, P., Malý, M., & Smahel, Z. (2002). Increased risk of traffic accidents in subjects with latent toxoplasmosis: A retrospective case-control study. *BMC Infectious Diseases*, 2(1), 11. <https://doi.org/10.1186/1471-2334-2-11>
- Flegr, J., Klapilová, K., & Kaňková, S. (2014). Toxoplasmosis can be a sexually transmitted infection with serious clinical consequences. Not all routes of infection are created equal. *Medical Hypotheses*, 83(3), 286–289. <https://doi.org/10.1016/j.mehy.2014.05.019>
- Flegr, J., Kodym, P., & Tolarová, V. (2000). Correlation of duration of latent *Toxoplasma gondii* infection with personality changes in women. *Biological Psychology*, 53(1), 57–68. [https://doi.org/10.1016/S0301-0511\(00\)00034-X](https://doi.org/10.1016/S0301-0511(00)00034-X)
- Flegr, J., Prandota, J., Sovičková, M., & Israeli, Z. H. (2014). Toxoplasmosis - A global threat. Correlation of latent toxoplasmosis with specific disease burden in a set of 88 countries. *PLOS ONE*, 9(3), e90203. <https://doi.org/10.1371/journal.pone.0090203>
- Flegr, J., Preiss, M., Klose, J., Havlíček, J., Vitáková, M., & Kodym, P. (2003). Decreased level of psychobiological factor novelty seeking and lower intelligence in men latently infected with the protozoan parasite *Toxoplasma gondii* dopamine, a missing link between schizophrenia and toxoplasmosis? *Biological Psychology*, 63(3), 253–268. [https://doi.org/10.1016/S0301-0511\(03\)00075-9](https://doi.org/10.1016/S0301-0511(03)00075-9)
- Flegr, J., Zitková, Š., Kodym, P., & Frynta, D. (1996). Induction of changes in human behaviour by the parasitic protozoan *Toxoplasma gondii*. *Parasitology*, 113(1), 49–54. <https://doi.org/10.1017/S0031182000066269>
- Friedman, H. S. (2008). The multiple linkages of personality and disease. *Brain, Behavior, and Immunity*, 22(5), 668–675. <https://doi.org/10.1016/j.bbi.2007.09.004>
- Furnham, A., & Fenton-O'Creevy, M. (2018). Personality and political orientation. *Personality and Individual Differences*, 129, 88–91. <https://doi.org/10.1016/j.paid.2018.03.020>
- Harell, A., Hinckley, R., & Mansell, J. (2021). Valuing liberty or equality? Empathetic personality and political intolerance of harmful speech. *Frontiers in Political Science*, 3, 64. <https://doi.org/10.3389/fpos.2021.663858>
- Gassen, J., & Hill, S. E. (2019). Why inflammation and the activities of the immune system matter for social and personality psychology (and not only for those who study health). *Social and Personality Psychology Compass*, 13(6), e12471. <https://doi.org/10.1111/spc3.12471>
- Hlaváčová, J., Flegr, J., Řežábek, K., Calda, P., & Kaňková, Š. (2021). Male-to-Female presumed transmission of toxoplasmosis between sexual partners. *American Journal of Epidemiology*, 190(3), 386–392. <https://doi.org/10.1093/AJE/KWAA198>
- Hodková, H., Kodym, P., & Flegr, J. (2007). Prolonged results of mice with latent toxoplasmosis in learning tests: Impaired learning processes or the novelty discrimination mechanism? *Parasitology*, 134(10), 1329–1337. <https://doi.org/10.1017/S0031182007002673>
- Houdek, P. (2017). Puppet master: Possible influence of the parasite *Toxoplasma gondii* on managers and employees. *Academy of Management Perspectives*, 31(1), 63–81. <https://doi.org/10.5465/amp.2015.0163>
- Hruschka, D. J., & Henrich, J. (2013). Institutions, parasites and the persistence of in-group preferences. *PLOS ONE*, 8(5), e63642. <https://doi.org/10.1371/JOURNAL.PONE.0063642>
- Johnson, S. K., Fitz, M. A., Lemer, D. A., Calhoun, D. M., Beldon, M. A., Chan, E. T., & Johnson, P. T. J. (2018). Risky business: Linking *Toxoplasma gondii* infection and entrepreneurship behaviours across individuals and countries. *Proceedings of the Royal Society B: Biological Sciences*, 285(1883), 20180822. <https://doi.org/10.1098/RSPB.2018.0822>
- Kamerkar, S., & Davis, P. H. (2012). Toxoplasma on the brain: Understanding host-pathogen interactions in chronic CNS infection. *Journal of Parasitology Research*, 2012(1), 1–10. <https://doi.org/10.1155/2012/589295>
- Kaňková, Š., Hlaváčová, J., & Flegr, J. (2020). Oral sex: A new, and possibly the most dangerous, route of toxoplasmosis transmission. *Medical Hypotheses*, 141, 109725. <https://doi.org/10.1016/j.mehy.2020.109725>

- Kaňková, Š., Holáň, V., Zajícová, A., Kodym, P., & Flegr, J. (2010). Modulation of immunity in mice with latent toxoplasmosis—the experimental support for the immunosuppression hypothesis of toxoplasma-induced changes in reproduction of mice and humans. *Parasitology Research*, 107(6), 1421–1427. <https://doi.org/10.1007/s00436-010-2013-9>
- Kankova, S., Flegr, J., & Calda, P. (2015). The influence of latent toxoplasmosis on women's reproductive function: Four cross-sectional studies. *Folia Parasitologica*, 62(1), 056. <https://doi.org/10.14411/fp.2015.041>
- Kekäläinen, J., Lai, Y.-T., Vainikka, A., Sirka, I., & Kortet, R. (2013). Do brain parasites alter host personality? — experimental study in minnows. *Behavioral Ecology and Sociobiology*, 68(2), 197–204. <https://doi.org/10.1007/S00265-013-1634-2>
- Khademvatan, S., Khajeddin, N., Saki, J., & Izadi-Mazidi, S. (2013). Effect of toxoplasmosis on personality profiles of Iranian men and women. *South African Journal of Science*, 109(1–2), 1–4. <https://doi.org/10.1590/sajs.2013.0017>
- Kim, S. (2015). Ppcor: An R package for a fast calculation to semi-partial correlation coefficients. *Communications for Statistical Applications and Methods*, 22(6), 665–674. <https://doi.org/10.5351/CSAM.2015.22.6.665>
- Kocazybek, B., Oner, Y. A., Turkoşoy, R., Babur, C., Cakan, H., Sahip, N., Unal, A., Ozaslan, A., Kilic, S., Saribas, S., Aslan, M., Taylan, A., Koc, S., Dirican, A., Uner, H. B., Oz, V., Ertekin, C., Kucukbasmaci, O., & Torun, M. M., ... (2009). Higher prevalence of toxoplasmosis in victims of traffic accidents suggest increased risk of traffic accident in toxoplasma-infected inhabitants of Istanbul and its suburbs. *Forensic Science International*, 187(1–3), 103–108. <https://doi.org/10.1016/j.forsciint.2009.03.007>
- Kolbekova, P., Kourbatova, E., Novotna, M., Kodym, P., & Flegr, J. (2007). New and old risk-factors for *Toxoplasma gondii* infection: Prospective cross-sectional study among military personnel in the Czech Republic. *Clinical Microbiology and Infection*, 13(10), 1012–1017. <https://doi.org/10.1111/j.1469-0691.2007.01771.x>
- Lasselin, J., Lekander, M., Axelsson, J., & Karshikoff, B. (2018). Sex differences in how inflammation affects behavior: What we can learn from experimental inflammatory models in humans. *Frontiers in Neuroendocrinology*, 50, 91–106. <https://doi.org/10.1016/j.yfrne.2018.06.005>
- Lefèvre, N., Corazza, F., Duchateau, J., Desir, J., & Casimir, G. (2012). Sex differences in inflammatory cytokines and CD99 expression following *in vitro* lipopolysaccharide stimulation. *Shock (Augusta, Ga.)*, 38(1), 37–42. <https://doi.org/10.1097/SHK.0B013E3182571E46>
- Lindová, J., Kubena, A. A., Sturcova, H., Krivohlava, R., Novotna, M., Rubesova, A., Havlicek, J., Kodym, P., & Flegr, J. (2010). Pattern of money allocation in experimental games supports the stress hypothesis of gender differences in *Toxoplasma gondii*-induced behavioural changes. *Folia Parasitologica*, 57(2), 136–142. <https://doi.org/10.14411/fp.2010.017>
- Lindová, J., Novotná, M., Havlíček, J., Jozífková, E., Skallová, A., Kolbeková, P., Hodný, Z., Kodym, P., & Flegr, J. (2006). Gender differences in behavioural changes induced by latent toxoplasmosis. *International Journal for Parasitology*, 36(14), 1485–1492. <https://doi.org/10.1016/j.ijpara.2006.07.008>
- Lindová, J., Připlatová, L., & Flegr, J. (2012). Higher extraversion and lower conscientiousness in humans infected with toxoplasma. *European Journal of Personality*, 26(3), 285–291. <https://doi.org/10.1002/per.838>
- Ma, M. Z. (2020). Testing the parasite-stress theory of sociality based on the circular model of human values: A multilevel analysis approach. *Personality and Individual Differences*, 168, 110277–110277. <https://doi.org/10.1016/j.paid.2020.110277>
- Marinov, M. P., Marchetti, C., Dimitrov, D., Ilieva, M., & Zehtindjiev, P. (2017). Mixed haemosporidian infections are associated with higher fearfulness in Yellow Wagtail (*Motacilla flava*). *Canadian Journal of Zoology*, 95(6), 405–410. <https://doi.org/10.1139/CJZ-2016-0121>
- Martynowicz, J., Augusto, L., Wek, R. C., Boehm, S. L., & Sullivan, W. J. (2019). Guanabenz reverses a key behavioral change caused by latent toxoplasmosis in mice by reducing neuroinflammation. *MBio*, 10(2), 2. <https://doi.org/10.1128/mBio.00381-19>
- Muthai, A. J., Lowry, C. A., Cook, T. B., Brenner, L. A., Brundin, L., Groer, M. W., Peng, X., Giegling, J., Hartmann, A. M., Konte, B., Friedl, M., Fuchs, D., Rujescu, D., & Postolache, T. T. (2016). Reciprocal moderation by *Toxoplasma gondii* seropositivity and blood phenylalanine – tyrosine ratio of their associations with trait aggression. *Pteridines*, 27(3–4), 77–85. <https://doi.org/10.1515/PTERID-2016-0006>
- Mutad, M. P. (2004). Gender differences in stress and coping styles. *Personality and Individual Differences*, 37(7), 1401–1415. <https://doi.org/10.1016/j.paid.2004.01.010>
- Meehan, R. A., & Brush, J. A. (2016). An overview of AIDS dementia complex. *American Journal of Alzheimer's Disease & Other Dementias*, 16(4), 225–229. <https://doi.org/10.1177/153331750101600411>
- Mendelsohn, G. A., Dakof, G. A., & Skaff, M. (1995). Personality change in Parkinson's disease patients: Chronic disease and aging. *Journal of Personality*, 63(2), 233–257. <https://doi.org/10.1111/j.1467-6494.1995.tb00809.x>
- Moieni, M., & Eisenberger, N. I. (2018). Effects of inflammation on social processes and implications for health. *Annals of the New York Academy of Sciences*, 1428(1), 5–13. <https://doi.org/10.1111/nyas.13864>
- Murray, D. R., Schaller, M., & Suedfeld, P. (2013). Pathogens and politics: Further evidence that parasite prevalence predicts authoritarianism. *PLOS ONE*, 8(5), e62275. <https://doi.org/10.1371/JOURNAL.PONE.0062275>
- Nikolaev, B., & Salahodjaev, R. (2017). Historical prevalence of infectious diseases, cultural values, and the origins of economic institutions. *Kyklos*, 70(1), 97–128. <https://doi.org/10.1111/KYKL.12132>
- Pew Research Center, Dimock, M., Kaley, J., Keeler, S., & Doherty, C. (2014). *Political polarization in the American public - Appendix A: The Ideological Consistency Scale*. Pew Research Center. <https://www.pewresearch.org/politics/2014/06/12/appendix-a-the-ideological-consistency-scale/>
- Postolache, T. T., Wadhawan, A., Rujescu, D., Hoisington, A. J., Dagdag, A., Baca-Garcia, E., Lowry, C. A., Okusaga, O. O., & Brenner, L. A. (2021). *Toxoplasma gondii*, suicidal behavior, and intermediate phenotypes for suicidal behavior. *Frontiers in Psychiatry*, 12, 806. <https://doi.org/10.3389/fpsyg.2021.665682>

- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <http://www.r-project.org/>.
- Schaller, M., & Murray, D. R. (2008). Pathogens, personality, and culture: Disease prevalence predicts worldwide variability in sociosexuality, extraversion, and openness to experience. *Journal of Personality and Social Psychology, 95*(1), 212–221. <https://doi.org/10.1037/0022-3514.95.1.212>
- Seaman, B., & Briffa, M. (2015). Parasites and personality in periwinkles (*Littorina littorea*): Infection status is associated with mean-level boldness but not repeatability. *Behavioural Processes, 115*, 132–134. <https://doi.org/10.1016/j.beproc.2015.03.014>
- Skallová, A., Kodym, P., Frynta, D., & Flegr, J. (2006). The role of dopamine in *Toxoplasma*-induced behavioural alterations in mice: An ethological and ethopharmacological study. *Parasitology, 133*(5), 525–535. <https://doi.org/10.1017/S003182006000886>
- Skallová, A., Novotná, M., Kolbeková, P., Gašová, Z., Veselý, V., Sechovská, M., & Flegr, J. (2005). Decreased level of novelty seeking in blood donors infected with *Toxoplasma*. *Neuroendocrinology Letters, 26*(5), 480–486.
- Sutin, A. R., Zonderman, A. B., Ferrucci, L., & Terracciano, A. (2013). Personality traits and chronic disease: Implications for adult personality development. *The Journals of Gerontology: Series B, 68*(6), 912–920. <https://doi.org/10.1093/GERONB/GBT036>
- Tamres, L. K., Janicki, D., & Helgeson, V. S. (2016). Sex differences in coping behavior: A meta-analytic review and an examination of relative coping. *Personality and Social Psychology Review, 6*(1), 2–30. [https://doi.org/10.1207/s15327957PSPR0601\\_1](https://doi.org/10.1207/s15327957PSPR0601_1)
- Tedford, E., & McConkey, G. (2017). Neurophysiological changes induced by chronic *Toxoplasma gondii* infection. *Pathogens (Basel, Switzerland), 6*(2), 19. <https://doi.org/10.3390/pathogens6020019>
- Tenter, A. M., Heckeroth, A. R., & Weiss, L. M. (2000). *Toxoplasma gondii*: From animals to humans. *International Journal for Parasitology, 30*(12–13), 1217–1258. [https://doi.org/10.1016/S0020-7519\(00\)00124-7](https://doi.org/10.1016/S0020-7519(00)00124-7)
- Thornhill, R., & Fincher, C. L. (2014). The parasite-stress theory of sociality, the behavioral immune system, and human social and cognitive uniqueness. *Evolutionary Behavioral Sciences, 8*(4), 257–264. <https://doi.org/10.1037/EBS0000020>
- Thornhill, R., Fincher, C. L., & Aran, D. (2009). Parasites, democratization, and the liberalization of values across contemporary countries. *Biological Reviews, 84*(1), 113–131. <https://doi.org/10.1111/j.1469-185X.2008.00062.x>
- Thornhill, R., Fincher, C. L., Murray, D. R., & Schaller, M. (2010). Zoonotic and non-zoonotic diseases in relation to human personality and societal values: Support for the parasite-stress model. *Evolutionary Psychology, 8*(2), 151–169. <https://doi.org/10.1177/147470491000800201>
- Tybur, J. M., Inbar, Y., Aarøe, L., Barclay, P., Barkow, F. K., de Barra, M., Becker, D. V., Borovoi, L., Choi, L., Choi, J. A., Consedine, N. S., Conway, A., Conway, J. R., Conway, P., Adoric, V. C., Demirci, D. E., Fernández, A. M., Ferreira, D. C. S., Ishii, K., ... Željelj, I. (2016). Parasite stress and pathogen avoidance relate to distinct dimensions of political ideology across 30 nations. *Proceedings of the National Academy of Sciences, 113*(44), 12408–12413. <https://doi.org/10.1073/PNAS.1607398113>
- Verhulst, B., Hatemi, P. K., & Martin, N. G. (2010). The nature of the relationship between personality traits and political attitudes. *Personality and Individual Differences, 49*(4), 306–316. <https://doi.org/10.1016/j.paid.2009.11.013>
- Wang, C. H. (2016). Personality traits, political attitudes and vote choice: Evidence from the United States. *Electoral Studies, 44*, 26–34. <https://doi.org/10.1016/j.electstud.2016.07.004>
- Webster, J. P. (2001). Rats, cats, people and parasites: The impact of latent toxoplasmosis on behaviour. *Microbes and Infection, 3*(12), 1037–1045. [https://doi.org/10.1016/S1286-4579\(01\)01459-9](https://doi.org/10.1016/S1286-4579(01)01459-9)
- Webster, J. P. (2007). The effect of *Toxoplasma gondii* on animal behavior: Playing cat and mouse. *Schizophrenia Bulletin, 33*(3), 752–756. <https://doi.org/10.1093/schbul/sbl073>

## Appendix A

### Political beliefs and values inventory P134

1. The gap between high and low incomes should be reduced.
2. It is acceptable for a same sex couple to live together.
3. Everyone should have the right to publicly protest against or for something.
4. Because we cannot trust corporations that they will respect environmental protection voluntarily, it is necessary to regulate them by law.
5. What is good for big corporations is in consequence good for society.
6. Differences in wealth and social status should be as small among people as possible.
7. If I were married, I wouldn't mind my partner watching pornography alone.
8. My government should value the lives of its citizens more than the lives of foreigners.
9. Minorities of different nationalities are a threat for our culture.
10. People are divided more by their socioeconomic status than by their nationality.
11. People with high incomes are paying overly high taxes.
12. It should be forbidden to publicly insult someone's religion.
13. Growing marihuana for one's own use should not be criminalized.
14. People should have the right to end their life with the help of a physician if they ask for it.
15. Parents should be allowed to physically punish their children.
16. There are no savages, barbarians or primitive cultures in the world, just different kinds of culture.
17. I wish that nations and state borders cease to exist, so we become members of one big group.
18. The first generation and the majority of the second generation of immigrants will not completely integrate in their new country.
19. For those who are able but not willing to work, there should be no social benefits.

20. It should be held true: "From each according to their ability, to each according to their needs."
21. Mothers of children can have their own career; however, their primary role is to take good care of children and the household.
22. Free speech is more important than the possibility of someone feeling deeply offended.
23. If a mother's life is not endangered, abortion should be prohibited by law.
24. It is necessary to teach children to respect authorities and superiors.
25. I am proud of the history of my country.
26. A society that is poor but equal is better than a rich society that is stratified in income and wealth.
27. The freer the market the freer the society.
28. It is acceptable that parents prefer their children to be educated at home instead of sending them to school.
29. The state shouldn't have the right to censor the internet.
30. It should be held true: "An eye for an eye, a tooth for a tooth."
31. I would sacrifice my life for my nation!
32. Capital punishment should be re-introduced for exceptionally serious crimes.
33. Adults with clean criminal records should have a right to own a firearm.
34. Foreigners who commit a criminal offense in our country should be deported to their country of origin.

1	2	3	4	5
strongly disagree	rather disagree	neither agree nor disagree	rather agree	strongly agree

**Scales (Arithmetic Mean of Responses):**

Tribalism: 8, 9, 15, 16r, 17r, 18, 19, 21, 24, 25, 30, 32, 34;  
 Economic equity: 1, 4, 6, 11r, 20, 26, 27r;  
 Cultural liberalism: 2, 7, 14, 21r, 23r;  
 Anti-authoritarianism: 3, 12r, 13, 22, 24r, 29, 33.  
 r: Reversed scoring (scale 5–1, not 1–5).

## **4. Kognitivní a evoluční přístupy k náboženství**

### **4.1 Vliv náboženství na wellbeing, fitness, fyzické a duševní zdraví**



## Effect of being religious on wellbeing in a predominantly atheist country: Explorative study on wellbeing, fitness, physical and mental health

Robin Kopecky<sup>a,b</sup>, Silvia Boschetti<sup>a</sup>, Jaroslav Flegr<sup>b</sup>.

<sup>a</sup>*Institute of Philosophy of the Czech Academy of Sciences, Prague, Czechia;*

<sup>b</sup>*Department of philosophy and history of science, Faculty of Science, Charles University, Prague, Czechia*

**Corresponding author:** Jaroslav Flegr, email: flegr@cesnet.cz,

telephone: +420 221951821, fax: +420 224919704

address: Department of Philosophy and History of Science, Faculty of Science, Charles University, Viničná 7, 12800, Prague, Czechia.

### Abstract

Despite a large volume of research on the impact of religion on different aspects of life, there is still a lack of studies from post-communist countries. In the current study, we aimed to fill this gap by investigating the relationship between religion and wellbeing, physical and mental health, education, sexual behavior and biological fitness among the Czech population. We managed to collect responses from 31633 participants and divided the sample into seven categories based on the type of religious belief and denomination (nonbelievers, believers without denomination, Catholics, Evangelicals, Hussites, Buddhists, Jews). We focused on the wellbeing as our main factor, which we define as composed of a number of sub-variables: physical and mental health, economic situation, self-attractiveness and the quality of the romantic relationship. In contrast to previous studies, we found a negative correlation between religiosity and physical and mental health. On the other hand, religiosity was connected to higher fitness, higher self-rated honesty and altruism, and lower sexual activity, which is in accord with the data from the western countries.

**Keywords** wellbeing, health, fitness, religiosity, post communism

**Data Availability** <https://osf.io/vusyf>

## Introduction

In recent years, religious and personal beliefs have been studied extensively as a possible physical and psychological health-protecting factor (Ellison, Levin 1998; Koenig, King & Carson, 2012; Miller & Thoresen, 2003; McCullough & Laurenceau, 2005). Although the research on religion as a factor of wellbeing is multidisciplinary and numerous there are methodological flaws and disagreements in interpretations of observed positive effects (Sloan & Bagiella, 2002; Maselko & Kubzansky, 2006).

In previous studies, different factors were identified as possible mediating mechanisms in the positive relationship found between religion and well-being (Ellison, Levin 1998; George, Larson, Koenig, & McCullough, 2000). One of the possible mechanisms is the regulation of individual lifestyles and health behavior (Rew & Wong, 2006; Harrigan, 2011). Religious commitment is correlated with self-reported lower frequencies of moral transgressions (Atkinson & Bourrat, 2011) and unhealthy behaviors (Clark, Friedman & Martin, 1999; Nagel & Sgoutas-Emch, 2007) such as drinking and smoking, driving under the influence of alcohol, and drug use. Moreover, the religious control over sexuality in western society influenced the sexual lifestyle of believers and promoted monogamous high-fertility relationships (Weeden, Cohen & Kenrick, 2008; McCullough, Carter, DeWall & Corrales, 2012), decreased the number of sexual partners and increased the age of first sexual intercourse which is also more likely to happen with a fiancé or a spouse (Simons, Burt & Peterson, 2009; Rostosky, Wilcox, Wright & Randall, 2004). These kinds of behaviors are related to lower risk of sexually transmitted diseases and unwanted pregnancy with positive effects on physical health (Simons et al., 2009). On the other hand, the traditional religious approach towards sexuality is conservative (Simons et al., 2009) and, particularly in case of non-traditional personal tendencies (i.e. homosexuality), can cause stress and negative emotion, both important factors for the onset of some mental illnesses.



Another important factor in the relationship between religion and wellbeing is the provision of specific coping resources (Abu-Raiya & Pargament, 2015). The coping mechanism emerge in case of personal and interpersonal stressful situations and can have a strong impact on their outcome (Lazarus & Folkman, 1984). In literature, there can be found specific examples of illness or disease in which religious coping mechanism plays a role. Praying and meditating is a common coping mechanism in case of anxiety or depression in religious people (Kasi et al., 2012), in case of infertility various religious coping mechanisms were found (Latifnejad Roudsari, Allan & Smith, 2014) including surrendering to God's will, belief in spiritual support and engagement in rituals. Other strong coping mechanisms are beliefs in justice or opportunity in the afterlife or reincarnation. Belief in the afterlife has previously been negatively related to mental illness and psychological distress and positively related to a sense of control (Hayward & Krause, 2014), both important for the personal well-being.

The impact of religiosity and spirituality on emotions and self-esteem is another mediating factor found between religion and well-being (Joshani & Daemi, 2015; Van Cappellen, Toth-Gauthier, Saroglou & Fredrickson, 2014). Previous studies showed that religion increased positive emotions, particularly when they are self-transcendent like gratitude, altruism and forgiveness, and through that improved general well-being (Sharma & Singh, 2018; Van Cappellen et al., 2014). Increase in this type of feelings in believers can partially explain increase in self-esteem, due to more positive self-judgment (Shariff, 2015), but it is important to note that different mechanisms could intervene. For instance, religion and spirituality were related to a positive body image (Tiggemann & Hage, 2019) and belief in divine support directly improved self-esteem (Schieman, Bierman, Upenieks & Ellison, 2017). Some religious practices, like meditation and prayer, could also affect self-esteem as was showed for mindfulness (Bajaj, Gupta & Pande, 2016), typically important for some religion like Buddhism.

Additionally, an important mechanism presented in previous studies (George et al., 2000; Lim & Putnam, 2010; Brownfield & Sorenson, 1991; Krause & Wulff, 2005) is social support from the religious community. This can involve emotional support in daily life conveyed through a stable social network and relationships among like-minded people (Lim & Putnam, 2010; Hayward & Krause, 2014), but also specific forms of support such as charitable organizations which often are church related. Social support has previously been linked with wellbeing by many studies (Cohen & Wills, 1985; Helliwell & Putnam, 2004; Valkenburg, Peter & Schouten, 2006; Wang, 2016; Revens et al., 2021), whereas loneliness and social isolation are risk factors for health (Holt-Lunstad et al., 2015). Well documented source of distress for nonbelievers is a negative perception by others (Weber et al., 2012). Additionally, being part of a religious community with strong common identity and high number of shared activities, including ritual, has been linked to personal life satisfaction (Lim & Putnam, 2010) while on the contrary religious non-affiliates reports lower social support relationship (Hayward et al., 2016).

The previously cited studies have certain limitations, of course. In particular, studies on this topic from post-communist countries are rare and mainly limited to Hungary (Thege, Székely & Kopp, 2012). Our present study aims to narrow this lacuna. This paper reports the results of an explorative study in the Czech Republic that examined how religious beliefs are related to quality of life, physical and mental health and biological fitness. The Czech Republic is a post-communist country in Central Europe and is predominantly atheist; indeed it is one of the most atheist countries in Europe (PEW, 2017). The importance of including atheist participants in the religiosity and health research was shown by Hwang, Hammer & Cragun (2009).

It has been suggested by Stavrova (2015) that the benefits of religiosity are restricted to highly religious countries or regions. To test this hypothesis, the analysis of data from the Czech

population, in which non-religiosity represents the social norm, could be very important. A cross European study (Nicholson, Rose & Bobak, 2009) describing connection of attendance at religious services and self-reported health (in 22 countries) showed that lower attendance at religious services is connected with poorer health. However, in the Nicholson study there was no significant connection between health and religious attendance in the Czech population. A very recent study by Libby et al. (2021) shown that while there are positive associations between multiple religiosity measures and self-reported health in Germany and no significant associations in most of the other European countries, the association of three of the four measures and self-reported health is negative in the Czech Republic; the belief in God was also associated with less favourable health only in the Czech Republic and Norway.

In this explorative study, we analysed various properties in a large sample of Czech volunteers (N = 31,633) which we divided into seven categories according to their beliefs – major denominations (catholics, evangelicals, hussites, buddhists, jews), believers without denomination, and nonbelievers. In the anonymous questionnaire that covered a variety of topics (effects of biological and social factors on human sexual behaviour) we asked the participants questions related to their sexual activity, wellbeing, subjective mental health, mental health disorders, physical health, biological fitness in the narrow and broad sense, and their honesty and altruism. We also asked them about how important is their faith in God for them and whether they believe in justice after death. The present study is aimed at exploring the relationship between religion and wellbeing in the Czech population, as an example of a society where religious affiliation is not the social norm. Since there is not unifying consensus regarding what general wellbeing is (Lambert at al., 2020), we considered individually 5 different wellbeing-related factors (physical and mental health, economic situation, self-attractiveness and quality of the current partner) as well as a single measure of wellbeing composed by the above mentioned factors.

## Methods

### Subjects

The internet questionnaire was distributed as a Qualtrics survey. The subjects were invited to participate in the study using a Facebook-based snowball method (Kankova, Flegr & Calda, 2015) by posting an invitation to participate in a “study testing certain evolutionary psychological and parasitological hypotheses, containing many questions related to sexual life” on the wall of the Facebook page “Lab Bunnies” (“Pokusní králíci” in Czech) for Czech and Slovak nationals willing to take part in diverse evolutionary-psychological experiments ([www.facebook.com/pokusnikralici](http://www.facebook.com/pokusnikralici)). Therefore the subjects examined in this study were not selected with regard of the religious beliefs. The participants were informed of the aims of the study on the first page of the electronic questionnaire. They were also provided with the following information: “The questionnaire is anonymous and obtained data will be used exclusively for scientific purposes. Your cooperation in the project is voluntary and you can terminate it at any time by closing this web page. You can also skip any uncomfortable questions; however complete data are the most valuable for us. Only subjects above fifteen years of age are allowed to take the questionnaire. If you agree to participate in the research and are above 15, press the “Next” button”. Some pages of the questionnaire contained the Facebook Share or Like buttons. These buttons were pressed by more than 1,200 participants, which resulted in obtaining data from 49,015 responders in total between 22nd January 2015 and 1st January 2018. The project, including the method of obtaining an electronic consent with a participation in the study, was approved by the IRB of the Faculty of Science, Charles University (Komise pro práci s lidmi a lidským materiálem Přírodovědecké Fakulty Univerzity Karlovy) - No. 2015/01.

## Questionnaires

The main part of the electronic survey consisted of the Sexual Preference and Behavior Inventory 2015 (SPBI-2015) (Flegr & Kuba, 2016). However, it also contained an anamnestic questionnaire collecting various socioeconomic, demographic, health related, epidemiological and psychological data, as well as three projective psychological tests. Altogether, the survey consisted of approximately seven hundred questions and the mean time necessary to complete it was about 89 minutes (the mode was 72 minutes). In the present study, we used only the information about gender, age, population of the town where the responders spent most of their childhood rated on a 6 point scale, body height, body mass, education (eight-points ordinal scale from 1- basic to 8 PhD or equivalent). Three sexuality-related questions from the standard Sexual Attitude Scale SOI-R (Penke & Asendorpf, 2008) (overall number of sexual partners, number of sexual partners in the past year, number of sexual intercourses in past month) which were used to calculate the aggregate variable sexual activity (for the method of calculating the aggregate variables see below).

Three sexuality-related questions from the standard Sexual Attitude Scale SOI-R (Penke & Asendorpf, 2008) (overall number of sexual partners, number of sexual partners in the past year, number of sexual intercourses in past month) which were used to calculate the aggregate variable *sexual activity* (for the method of calculating the aggregate variables see below).

Five questions closely related to wellbeing (quality of physical health, mental health, economic situation, self-attractiveness, and quality of current partnership, all rated on a 6-point Likert scale with 1- very poor and 6 - very good) were used to calculate the variable *wellbeing*.

Five questions on the presence and intensity of subjectively rated symptoms of impaired mental health (whether and how much the participant suffers from depressions, manias, anxieties, phobias, and obsessively compulsive thoughts and behaviors, all rated on scales 0-

100 by moving graphic sliders) that were used to calculate the aggregate variable *symptoms of impaired mental health*.

Three questions related on variety of mental health problems (how much or how often does the subject suffer from mental health disorders that have been diagnosed by a physician, how much or how often the subject suffers from mental health disorders that have not been diagnosed by a physician, and the number of items checked on a list of 24 most frequent mental health disorders for which the subject has been diagnosed or self-diagnosed; the first two variables were rated on the scale 0-100 by moving graphic sliders) were used to calculate the aggregate variable *mental health problems*.

Five questions related to physical health (number of different drugs prescribed by a doctor currently consumed per day; number of drugs, food supplements and vitamins not prescribed by a doctor currently consumed per day; how many times the responder visited his general practitioner physician in the past 365 days; how many times he had to take antibiotics in the past; what number of specialized medical doctors the subject had to regularly attend (not for prevention) at least once in the past five years) were used to calculate the aggregate variable *physical health problems*.

Seven questions related to self reported honesty and altruism (probability of risking one's own life in order to save lives of three other people, probability of returning a found bag containing one million Czech crowns (about \$ 50 thousand), whether the responder has ever cheated during a school examination, whether the responder would prefer to be hit by a car rather than hit somebody else with a car, whether the responder has ever defended a bullied person, whether the responder has ever lied in an important matter, and whether the responder has ever lied in a non-important matter; all questions were responded using 6-point Likert scale, anchored with 1- surely no, and 6- surely yes) were used to calculate the aggregate variable *honesty and altruism*.

Four questions related to inclusive biological fitness (current number of biological children, preferred total number of children, number of biological siblings, total number of biological aunts and uncles) were used to calculate the aggregate variable *biological fitness*.

The responders also had to answer a question about the strength of their religiosity, namely, how strongly they agree with the statement: “My belief in God is extremely important for me.” and the statement: “I believe that after death good people will be rewarded and bad people punished.” using scale 0-100 anchored with 0- strongly disagree and 100- strongly agree by moving a graphic slider. Later, in a separate question, they were asked to answer what is their denomination by choosing from the list of six items, atheist “I am not believer”, no denominations “I am believer but I am not adherent to any church”, the three most common denominations in Czechia (Catholic Church, Protestant Church of Czech Brethren, Czechoslovak Hussite Church) and “other, please specify”. From the people who selected “other”, we analyzed only the data of Buddhists and Jews in the present study as the other denominations (e.g. Islam, Satanists, or Pagans) were selected by relatively low number of responders. The responders could not return to previously answered questions.

### **Statistical analysis**

Before any analyses, less than 1% of data (too high or too short body height, too low or too high body mass or age, too short duration of the test, suspicious combination of reported mental health disorders etc.) was filtered out as suspicious. Also, the aggregate variables for wellbeing, sexual activity, symptoms of impaired mental health, mental health problems, physical health problems, biological fitness, and honesty and altruism were calculated as the arithmetic means of the Z-scores of corresponding variables, see above. Also, the body mass index (BMI) was calculated for each participant as  $(\text{body mass}/\text{squared body height})/10,000$  before the analysis.

Statistica v. 10.0 was used for all statistical tests. ANCOVA (General Linear Models module) was used for the analyses of the association between the focal simple variables (e.g. number of children, education) or aggregate variables (e.g. biological fitness, wellbeing) and two categorical factors – sex and denomination. With the exception of the case of analysis of the association of these focal variables and the age of participants, the variable age was always included into the models as a covariate.

Associations between binary variables particular denominations and focal variables were analyzed by non-parametric partial Kendall's correlation test (Kaňková, Kodým & Flegr 2011; Siegel & Castellan, 1988). This test measures the strength and significance of the association between binary, ordinal and continuous data regardless of their distributions. The partial Kendall Tau reflects the probability that the value of a particular dependent variable for the subject A is higher than for a subject B when the value of an independent variable for a subject A is higher than for the subject B. This technique enabled us to control for one confounding variable, for example the age of a responder. The Excel sheet for computing partial Kendall's Tau and the significance between variables A and B after the variable C is controlled based on Kendall Tau's AB, AC and BC. It is available here: <http://web.natur.cuni.cz/flegr/programy.php>. For the purpose of this analysis, six new binary variables were computed for six denominations under study. Absence of belief (atheism) was always coded as 0 and adherence to particular denominations, including the adherence to no church, was coded with 1. Although we considered this study exploratory, we still checked whether some observed effects could be artefacts of performing multiple statistical tests with Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995) with false discovery rate preset to a very stringent value 0.1.



## Results

The final population included 31,633 subjects, 14,276 women and 17,357 men, who provided the necessary data, including the information about their religious denomination. Demographic and descriptive data are shown in Table 1.

Table 1: Descriptive statistic of the study population stratified for sex and denomination

	<i>Atheists</i>	<i>No Denom.</i>	<i>Catholics</i>	<i>Evangelicals</i>	<i>Hussites</i>	<i>Buddhists</i>	<i>Jews</i>
<i>N (31,633)</i>	20,047 63.4 %	7,252 22.9 %	3,497 11.0 %	392 1.24 %	142 0.45 %	248 0.78 %	55 0.17 %
<i>N women (14,276)</i>	8,546 59.9 %	3,755 26.3 %	1,599 11.2 %	175 1.23 %	61 0.43 %	111 0.78 %	29 0.20 %
<i>N men (17,357)</i>	11,501 66.3 %	3,497 20.1 %	1,898 10.9 %	217 1.25 %	81 0.47 %	137 0.80 %	26 0.15 %
<i>Census sum (10,436,560)</i>	3,604,095 34.53%	705,368 6.76 %	1,082,463 10.37 %	51,858 0.5 %	39,229 0.38 %	6,101 0.06 %	1,474 0.01 %
<i>Census Women (5,326,794)</i>	1,765,197 33.14 %	388,674 7.3 %	614,970 11.54 %	29,644 0.56 %	24,079 0.45 %	2,373 0.04 %	621 0.01 %
<i>Census Men (5,109,766)</i>	1,838,898 35.99 %	316,694 6.2 %	467,493 9.15 %	22,214 0.43 %	15,150 0.3 %	3,728 0.07 %	853 0.02 %
<i>Age women</i>	30.5 (11.1)	31.4 (11.1)	34.1 (13.3)	31.4 (12.7)	37.6 (15.8)	30.0 (10.7)	35.4 (10.7)
<i>Age men</i>	35.4 (12.4)	35.0 (11.5)	39.8 (14.4)	37.1 (13.7)	40.7 (15.6)	35.2 (12.1)	39.4 (15.1)
<i>Children women</i>	0.82 (1.06)	0.96 (1.12)	1.22 (1.22)	1.04 (1.12)	1.36 (1.25)	0.59 (0.93)	1.53 (0.99)
<i>Children men</i>	0.83 (1.10)	0.85 (1.14)	1.29 (1.37)	1.24 (1.44)	1.47 (1.93)	0.88 (1.18)	0.79 (0.98)
<i>Education women</i>	5.06	5.09	5.26	5.26	5.15	5.17	6.00
<i>Education men</i>	5.15	5.17	5.53	5.76	5.21	5.07	5.50
<i>Urbanisation women</i>	3.45	3.33	2.91	3.22	3.23	3.42	4.59
<i>Urbanisation men</i>	3.54	3.45	3.09	3.57	3.93	3.17	4.81
<i>Economic sit. women</i>	59.69	58.76	60.43	57.20	58.03	62.09	67.34
<i>Economic sit. men</i>	64.23	62.46	64.21	64.98	59.72	58.76	63.44
<i>BMI women</i>	23.70	23.80	24.20	23.07	26.32	23.02	24.15
<i>BMI men</i>	25.90	25.93	26.58	25.83	26.84	25.51	24.94

*Note.* The values in parentheses shows standard deviations. Data from census (Czech Statistical Office, 2016): numerous people did not fill religious belief, this question as well as, for example, nationality was not compulsory; 4,662,455 people (i.e. 45,67 %) were not identified. Number of “Buddhists” is a sum of two categories Buddhism and Diamond Way Buddhism - Karma Kagyu Lineage. “Jews” are a sum of Judaism and Federation of Jewish Communities in the Czech Republic.

The association between religion and focal variables, including the aggregate variables, was estimated with ANCOVA (general linear models) with sex and the religion of subjects as the fixed factors and the age as the covariate. The Table 2 and Fig. 1-8 show the results.

Many variables were measured on ordinal scales and some of them had highly asymmetric distribution that cannot be corrected to a symmetric one by any transformation. Therefore, in the post hoc testing, the associations between focal variables and religiosity (separately for particular denominations) were analyzed with nonparametric tests, specifically with the partial Kendall correlation tests with age as a covariate. The reported absence of belief (atheism) was always coded as 0, presence of particular religiosity as 1. The results for the aggregate variables as well as for some of their components, computed for all subjects of particular denomination and separately also for women and men are shown in the Table 3.

Table 2: Differences between men and women of different denominations

	<i>Sex</i>		<i>Religion</i>		<i>Sex-Religion</i>	
	p	Eta <sup>2</sup>	p	Eta <sup>2</sup>	p	Eta <sup>2</sup>
<i>Age</i>	0.000	0.001	0.000	0.012	0.000	0.001
<i>Children</i>	0.027	0.000	0.000	0.005	0.013	0.001
<i>Education</i>	0.703	0.000	0.000	0.002	0.011	0.001
<i>Childhood in Large Cities</i>	0.047	0.000	0.000	0.010	0.075	0.000
<i>Economic Situation</i>	0.224	0.000	0.000	0.001	0.036	0.000
<i>BMI</i>	0.000	0.001	0.005	0.001	0.357	0.000
<i>Wellbeing</i>	0.284	0.000	0.098	0.000	0.238	0.000
<i>Physical Health Problems</i>	0.039	0.000	0.000	0.001	0.513	0.000
<i>Mental Health Problems</i>	0.850	0.000	0.000	0.001	0.368	0.000
<i>Symptoms of Mental Illness</i>	0.002	0.000	0.000	0.008	0.287	0.000
<i>Sexual Activity</i>	0.063	0.000	0.000	0.009	0.048	0.001
<i>Biological Fitness</i>	0.728	0.000	0.000	0.025	0.072	0.000
<i>Honesty</i>	0.001	0.000	0.000	0.006	0.601	0.000
<i>Importance of Belief</i>	0.000	0.001	0.000	0.539	0.000	0.001

*Note.* For the direction and size of the effects see also the figures 1-14. P values lower to 0.00005 are coded as 0.000. According to Cohen's classification (Cohen, 1977) Eta<sup>2</sup> 0.01, 0.06 and 0.14 corresponds to small, medium, and large effects, respectively.

Table 3: Differences between atheists and religious subjects

		<i>No Denom.</i>	<i>Catholics</i>	<i>Evangelicals</i>	<i>Hussites</i>	<i>Buddhists</i>	<i>Jews</i>
<i>Age</i>	All	+0.00	<b>+0.08**</b>	+0.01	<b>+0.03**</b>	+0.00	<b>+0.02*</b>
	Women	<b>+0.04**</b>	<b>+0.08**</b>	+0.01	<b>+0.03**</b>	+0.00	<b>+0.03**</b>
	Men	+0.00	<b>+0.09**</b>	<b>+0.01</b>	<b>+0.02**</b>	+0.00	+0.01
<i>Children</i>	All	<b>+0.03**</b>	<b>+0.09**</b>	<b>+0.03**</b>	<b>+0.03**</b>	+0.00	+0.01
	Women	<b>+0.04**</b>	<b>+0.09**</b>	<b>+0.03*</b>	<b>+0.03</b>	-0.02	<b>+0.03*</b>
	Men	+0.01	<b>+0.08**</b>	<b>+0.03**</b>	<b>+0.03*</b>	<b>+0.00</b>	-0.01
<i>Education</i>	All	+0.00	<b>+0.04**</b>	<b>+0.03**</b>	+0.00	+0.00	<b>+0.02**</b>
	Women	+0.00	<b>+0.02*</b>	<b>+0.02</b>	-0.01	+0.01	<b>+0.03**</b>
	Men	+0.00	<b>+0.06**</b>	<b>+0.04**</b>	+0.00	+0.00	+0.01
<i>Childhood in Large Cities</i>	All	<b>-0.03**</b>	<b>-0.10**</b>	-0.01	+0.00	-0.01	<b>+0.03**</b>
	Women	<b>-0.03**</b>	<b>-0.11**</b>	<b>-0.02</b>	-0.01	<b>+0.00</b>	<b>+0.03**</b>
	Men	<b>-0.02**</b>	<b>-0.09**</b>	+0.00	<b>+0.02</b>	-0.02*	<b>+0.03**</b>
<i>Economic Situation</i>	All	<b>-0.03**</b>	+0.00	+0.00	<b>-0.01</b>	<b>+0.00</b>	+0.01
	Women	<b>-0.02*</b>	+0.01	-0.01	+0.00	+0.01	+0.02
	Men	<b>-0.02**</b>	-0.01	+0.00	<b>-0.02</b>	-0.02	+0.00
<i>BMI</i>	All	<b>-0.02**</b>	<b>+0.01</b>	-0.01	<b>+0.01*</b>	<b>-0.01*</b>	<b>-0.01*</b>
	Women	-0.01	<b>+0.01</b>	<b>-0.02</b>	<b>+0.03**</b>	<b>-0.02</b>	-0.01
	Men	+0.00	<b>+0.02**</b>	+0.00	+0.01	<b>-0.01</b>	<b>-0.01</b>
<i>Wellbeing</i>	All	<b>-0.02**</b>	+0.00	+0.00	-0.01	+0.01	+0.00
	Women	-0.01	+0.01	-0.01	+0.00	<b>+0.02</b>	+0.01
	Men	<b>-0.02*</b>	<b>-0.01</b>	+0.00	-0.01	<b>+0.00</b>	+0.00
<i>Physical Health Problems</i>	All	<b>+0.02**</b>	<b>+0.03**</b>	<b>+0.01</b>	<b>+0.01</b>	+0.02**	+0.01
	Women	<b>+0.02**</b>	<b>+0.03**</b>	+0.00	<b>+0.02</b>	<b>+0.02</b>	+0.00
	Men	<b>+0.02*</b>	<b>+0.03**</b>	<b>+0.02*</b>	+0.01	<b>+0.02*</b>	+0.01
<i>Mental Health Problems</i>	All	<b>+0.03**</b>	<b>+0.01</b>	+0.01	+0.00	<b>+0.03**</b>	<b>+0.01</b>
	Women	<b>+0.04**</b>	+0.01	+0.01	+0.00	<b>+0.02*</b>	<b>+0.02*</b>
	Men	<b>+0.02**</b>	<b>+0.01</b>	+0.01	-0.01	<b>+0.03**</b>	+0.00
<i>Symptoms of Mental Illness</i>	All	<b>+0.09**</b>	<b>+0.02**</b>	<b>+0.01</b>	+0.00	<b>+0.04**</b>	<b>+0.03**</b>
	Women	<b>+0.07**</b>	+0.01	<b>+0.02</b>	+0.00	<b>+0.03**</b>	<b>+0.02</b>
	Men	<b>+0.10**</b>	<b>+0.02*</b>	+0.01	+0.00	<b>+0.04**</b>	<b>+0.03**</b>
<i>Sexual Activity</i>	All	+0.00	<b>-0.08**</b>	<b>-0.04**</b>	-0.01	<b>-0.01</b>	<b>+0.01</b>
	Women	<b>-0.02</b>	<b>-0.10**</b>	<b>-0.05**</b>	-0.02	-0.01	+0.01
	Men	+0.01	<b>-0.07**</b>	<b>-0.03**</b>	+0.00	-0.01	+0.01
<i>Biological Fitness</i>	All	<b>+0.07**</b>	<b>+0.14**</b>	<b>+0.05**</b>	<b>+0.02*</b>	+0.01	+0.00
	Women	<b>+0.09**</b>	<b>+0.15**</b>	<b>+0.04**</b>	<b>+0.02*</b>	+0.01	+0.01
	Men	<b>+0.05**</b>	<b>+0.14**</b>	<b>+0.05**</b>	+0.01	+0.00	+0.00
<i>Honesty</i>	All	<b>+0.05**</b>	<b>+0.07**</b>	<b>+0.03**</b>	<b>+0.02*</b>	+0.01	+0.00
	Women	<b>+0.05**</b>	<b>+0.06**</b>	<b>+0.03**</b>	<b>+0.02*</b>	+0.02	-0.01
	Men	<b>+0.04**</b>	<b>+0.06**</b>	<b>+0.03**</b>	+0.01	<b>+0.00</b>	+0.01
<i>Importance of Belief</i>	All	<b>+0.59**</b>	<b>+0.58**</b>	<b>+0.26**</b>	<b>+0.14**</b>	<b>+0.12**</b>	<b>+0.09**</b>
	Women	<b>+0.60**</b>	<b>+0.59**</b>	<b>+0.28**</b>	<b>+0.14**</b>	<b>+0.13**</b>	<b>+0.11**</b>
	Men	<b>+0.58**</b>	<b>+0.57**</b>	<b>+0.24**</b>	<b>+0.13**</b>	<b>+0.11**</b>	<b>+0.08**</b>

*Note.* The table shows partial Kendall Tau (age of subjects controlled) for correlations between the binary variable religion and variables listed in the first column. The absence of any religion (reported atheism) was coded as 0, religiosity as 1; therefore, the positive Tau means that religious people of particular denomination scored higher in this particular variable. Bold print denotes the associations significant at 0.05, one asterisk those with p value less than 0.005, and two asterisks those with p value less than 0.0005. All significant associations retained significance when the possible effect of multiple comparison was controlled with the Benjamini-Hochberg procedure. After transformation between Tau and d, the Tau 0.062, 0.156, and 0.241 corresponds to d 0.20 (small effect), 0.50 (medium effect) and 0.80 (large effect), respectively (Gilpin, 1993; Cohen, 1997).

Table 4: Relation between “importance of belief” and indicators of quality of life in two main denominations

		<i>Importance of Belief</i>	
		Catholics	Evangelicals
<i>Age</i>	Women	-0.005	-0.100
	Men	-0.005	+0.029
<i>Children</i>	Women	<b>+0.080 *</b>	-0.105
	Men	<b>+0.076 **</b>	<b>+0.125</b>
<i>Physical Health Problems</i>	Women	+0.005	+0.002
	Men	+0.027	+0.020
<i>Mental Health Problems</i>	Women	+0.0047	+0.090
	Men	<b>+0.046 *</b>	<b>+0.106</b>
<i>Honesty</i>	Women	<b>+0.117 **</b>	+0.067
	Men	<b>+0.123 **</b>	<b>+0.130</b>
<i>Symptoms of Mental Illness</i>	Women	-0.015	-0.014
	Men	-0.019	-0.011
<i>Wellbeing</i>	Women	<b>+0.067 **</b>	+0.050
	Men	<b>+0.088 **</b>	+0.084
<i>Sexual Activity</i>	Women	<b>-0.108 **</b>	-0.110
	Men	<b>-0.107 **</b>	-0.107
<i>Fitness</i>	Women	<b>+0.094 **</b>	-0.005
	Men	<b>+0.094 **</b>	<b>+0.114</b>

*Note.* The table shows partial Kendall Tau (age of subjects controlled) for correlations between the importance of belief (in Catholics or Evangelicals) and variables listed in the first column. The positive Tau means that religious people of particular denomination scored higher in this particular variable. Bold print denotes the associations significant at 0.05, one asterisk those with p value less than 0.005, and two asterisks those with p value less than 0.0005. All significant associations retained significance when the possible effect of multiple comparison was controlled with the Benjamini-Hochberg procedure. After transformation between Tau and  $d$ , the Tau 0.062, 0.156, and 0.241 corresponds to  $d$  0.20 (small effect), 0.50

(medium effect) and 0.80 (large effect), respectively (Gilpin, 1993; Cohen, 1997).

Table 5: Relation between “Frequency of Religious Meetings” and indicators of quality of life in two main denominations

		<i>Frequency of Religious Meetings</i>	
		Catholics	Evangelicals
<i>Age</i>	Women	<b>-0.121 **</b>	<b>-0.132</b>
	Men	<b>-0.049</b>	-0.051
<i>Children</i>	Women	<b>+0.079 *</b>	+0.015
	Men	<b>+0.111 **</b>	<b>+0.14</b>
<i>Physical Health Problems</i>	Women	-0.007	+0.019
	Men	<b>+0.049 **</b>	+0.043
<i>Mental Health Problems</i>	Women	<b>+0.058 *</b>	+0.068
	Men	<b>+0.052 *</b>	<b>+0.043</b>
<i>Honesty</i>	Women	<b>+0.100 **</b>	+0.036
	Men	<b>+0.118 **</b>	<b>+0.173 *</b>
<i>Symptoms of Mental Illness</i>	Women	+0.014	+0.058
	Men	+0.036	+0.002
<i>Wellbeing</i>	Women	<b>+0.041</b>	-0.025
	Men	<b>+0.048</b>	+0.099
<i>Sexual Activity</i>	Women	<b>-0.174 **</b>	<b>-0.172</b>
	Men	<b>-0.160 **</b>	<b>-0.186</b>
<i>Fitness</i>	Women	<b>+0.152 **</b>	+0.06
	Men	<b>+0.148 **</b>	<b>+0.144</b>

*Note.* The table shows partial Kendall Tau (age of subjects controlled) for correlations between the frequency of religious meeting (in Catholics or Evangelicals) and variables listed in the first column. The positive Tau means that religious people of particular denomination scored higher in this particular variable. Bold print denotes the associations significant at 0.05, one asterisk those with p value less than 0.005, and two asterisks those with p value less than 0.0005. All significant associations retained significance when the possible effect of multiple comparison



was controlled with the Benjamini-Hochberg procedure. After transformation between Tau and  $d$ , the Tau 0.062, 0.156, and 0.241 corresponds to  $d$  0.20 (small effect), 0.50 (medium effect) and 0.80 (large effect), respectively (Gilpin, 1993; Cohen, 1997).

## Discussion

This study aims to explore the role of religion in the life of particular populations of the Czech Republic and to bring some insight to the social environment of a post-communist country regarding religiosity and beliefs. We examined the wellbeing, physical and mental health, education, sexual behavior, biological fitness, and certain attitudes that previous research showed as related to religious beliefs or religious practices (Powell, Shahabi & Thoresen, 2003; Seybold & Hill, 2001; Koenig, King & Carson, 2012). Many of the studies investigating these factors in relation with religious beliefs were conducted only in western countries with different religious traditions and a different social-political history from post-communist countries (Ahrenfeldt et al., 2017) or the studies were limited to simple self-assessed health measures (e.g. Domínguez & López-Noval, 2021). Based on data anonymously collected from 31,633 volunteers (14,276 women and 17,357 men), we were able to explore the above-mentioned factors in the different religious and nonreligious subgroups of inhabitants of the Czech Republic. Despite its own limitations (see the Limitations section) we thus believe this study brings an important contribution to this field of study by focusing on the mostly atheist country and using multiple measures of wellbeing and health.

The most important factor explored in this study is wellbeing. We studied it as one factor composed of 5 different wellbeing-related variables, and also as its separate components, namely physical and mental health, economic situation, self-attractiveness and quality of the current partner. The finding in this part of the study that was particularly interesting, especially when compared to previous literature, was that physical health was negatively related to religiosity: atheists showed better physical health than religious subgroups, with or without denomination (very small effect size, according to Cohen's classification). Atheists also showed better mental health (small effect size) but this factor needs to be better defined and

explored. Indeed, previous studies found different results according to which types of symptoms and pathologies were measured. In general, the result we obtained can be explained by the theory of Stavrova (2015): being religious or a believer in the Czech Republic is a deviation from the norm. This represents a stress factor that can affect physical and mental health. In line with this theory social support was previously identified as a mediating factor for the positive relationship between religion and well-being (Oman & Thoresen, 2002; George, Larson, Koenig, & McCullough, 2000; Holt, Wang, Clark, Williams & Schulz, 2013). In a recent study engagement in a community or a supportive social group was highlighted as a factor with positive effect on wellbeing in religious people as well as in atheists (Galen, 2015). Our data confirmed this hypothesis in the case of Catholic, the frequencies of participation to church related activities and the importance of beliefs has positive correlation with the wellbeing. It can also be speculated whether the worse health of believers is a cause or an effect of their religiosity. Longitudinal studies are needed to clarify if actually the choice to join a church was made before the occurrence of the health problem or after, as a coping mechanism after learning about a disease. Another possibility is that religious people seek less medical help with their health problems (or search for it later) as a consequence of a more deterministic view of life, including relying on a divine plan or divine intervention in life and in illness (Green & Elliot, 2009). Alternative explanation is that there is not a necessary connection between belief or disbelief in God and better reported health as was the case in study (Speed & Fowler, 2016) based on the data from the American General Social Survey.

From our results we found that believers considered or described themselves as more moral compared to non-believers (small effect sizes). This is consistent with what has been found in previous studies on religion and prosocial behaviors and altruism (Shariff, 2015; Sharma & Singh, 2018; Van Cappellen et al., 2014). It is important to highlight the fact that religious people in our sample believed in reward and punishment in the afterlife and that this characteristic was a predictor of lower moral transgression across the globe in self-report study

(Atkinson & Bourrat, 2011). In Catholic subgroup the frequencies of the church related activities and the importance of beliefs are two factors with positive correlation with the self-rating honesty, suggesting that the higher commitment has influences on the self-judgment. However, Shariff (2015) showed that the self-reported measures of morality and moral decisions in religious people are not associated with prosocial and altruistic behavior. This notion could be supported by the theory that religious beliefs improve self-esteem and self-judgment, also through self-forgiveness (Krause, 2017), and as consequences increase the self-rating on morality and altruism. While individual religiosity seems to not be directly linked to selfless behavior (Shariff, Willard, Andersen & Norenzayan, 2016; Xygalatas et al., 2015), the use of religious priming increased prosocial attitudes and behaviors in different cultures (Shariff et al., 2016). Interestingly, this effect is maintained in a context of religious outgroup (Xygalatas et al., 2015), including the Czech Republic. Another possible explanation is the self-other overlap, a phenomenon especially probable when the bonding to the community is strong. In this case, when the subject answers in a questionnaire instead of thinking about themselves, they think of themselves as members of a community, which leads to their answers being biased by their perception of the members of a community and not by their perception of themselves.

A further finding of this study was that in our sample being religious increased fitness (small to medium effect size) but decreased sexual activity (small effect size). The results showed significantly higher biological fitness for representatives of official churches compared to atheists or even believers that reported no adherence to any official church. It could be just an historical accident but the number of adherents of particular Christian churches correlate with their biological fitness. Indeed, the frequencies of participation to church related activities correlated with lower sexual activities in Catholic and Evangelistic, in the case of Catholic correlated also with higher number of children. We need to abstain from further speculation about the possible causal relationship between religion and biological fitness, however, until

we obtain enough data from the less numerous Christian churches. At the same time most of the religions in our study are known to be very strict about the regulation of sexual activity. Previous work showed that the control exercised by religions over the sexual life of community members can affect sexual behavior and regulate it into reproduction (Weeden, Cohen & Kenrick, 2008; McCullough, Carter, DeWall & Corrales, 2012). Therefore, it was previously hypothesized that one of the main roles of religion in western society is to enhance reproductive fitness supporting high-fertility and monogamous sexual strategy (Weeden, Cohen & Kenrick, 2008). Furthermore, previous studies showed that women are more affected than men in developing conservative sexual attitude (Brelsford, Luquis & Murray-Swank, 2011) and this tendency is visible also in our data. In our sample, the women belonging to the subgroup of Jews, or the female believers without denomination, are also more affected by the religion in the number of children that they have than corresponding male believers.

Comparing the subgroups in our study, the atheists appeared to be healthier, both physically and mentally. They had higher sexual activity but lower fitness, and they considered themselves as less honest and altruistic (all in comparison with believers). They claimed not to believe in life after death, including eventual reward or punishment, and they did not consider beliefs and religion to be important factors in their life.

Catholics, in comparison to atheists, lived in smaller towns, showed worse mental but especially physical health, they had much lower sexual activity but higher biological fitness – number of children. They considered (or described) themselves as more honest and altruistic. As could be expected, their religion and belief in God were much more important in their life than for atheists. A rather unexpected result was that the believers not affiliated with any specific religious church (the far largest group of Czech believers), expressed a profile that was very similar to that of Czech Catholics.

The unaffiliated believers were born in larger cities, achieved lower education, had a more intensive sexual life, expressed a slightly worse mental health and a much lower biological fitness, including number of children than Catholics. It is possible that this category of believers was the unintended product of 40 years of official communist oppression of religion, anti-religious indoctrination in school and informal and often unintended, in principle catholic, upbringing and education in families.

The other religious subgroups are harder to compare, due to the small number of respondents with such beliefs and the resulting presence of extremely large confidence intervals. The mental health of Buddhists was approximately as bad as that of the unaffiliated believers. Jews reported the highest intensity of symptoms of impaired mental health (anxieties, depression, etc.). This could be the result of transgenerational transfer of the Holocaust trauma that was reported frequently in European Jews and showed in previous studies (Niederlande, 1964; Yehuda et al., 2000; Klímová & Roubalová, 2014). Female Jews reported much higher while male Jews much lower number of children in comparison to atheists. Due to relatively low number of subjects, this observation must be considered only preliminary and should be retested in future independent studies.

In our sample, the percentages of members of major Czech religions were similar to the values showed in a recent national census in 2011 (Czech Statistical Office, 2016). We did not expect this consistent results because the questionnaire used for the study was primary focused on sexual behaviors and preferences, and many questions could be rather uncomfortable for the members of some churches. In our sample the percentage of the believers who are adherent to a specific church or religious community is 13.64%, while in the national census in 2011 (Czech Statistical Office, 2016) it is 14.02%, with a strong match in the percentages of the main subcategories (Roman Catholic Church, Czechoslovak Hussite Church, Evangelical Church). The Buddhists and the Jews seem to be overrepresented in our sample. Higher number of

Buddhists can be caused by relatively lower age of our responders. The higher representation of Jews can be the result of the fact that most of our participants were from Prague, in which the Jewish community is larger in comparison to the rest of the Czech Republic. However, it can be also the result of anonymity of participants of our study. The Holocaust trauma is still alive in the Czech Jewish society and is being culturally or possibly epigenetically transferred to the 2nd, 3rd and even 4th generation of the Holocaust survivors (Klímová & Roubalová, 2014). Many Jews are still not willing to identify themselves as Jews in non-anonymous national surveys. The data regarding the amount of atheists and believers without denomination are difficult to compare to the national census because in the national survey the questions regarding religious beliefs were not compulsory. For this reason, only 34.53% of responders actively identify themselves as atheists, only 6.76% as believers without specific church or community, and the 44.67% decided not to identify themselves with any category. From the census results, we can estimate the percentage of unbelievers in the Czech Republic between 34.53% and 79.2%. Another relevant source for studies of religiosity in the Central and the Eastern Europe is the Pew Research Center (PEW, 2017). According to the results of the report from 2015, the 66% of Czech population do not believe in God. Our study, showing the 63.4% of unbelievers, fits well with these previous estimations of religiosity percentages in the population

### Limitation of the present study

Possible limitation of the present study is the self-selection of its participants. Since the project was advertised as a study focused on “testing certain evolutionary psychological and parasitological hypotheses, containing many questions related to sexual life” or “studying the effects of various biological and non-biological factors on sexual preferences and behavior of Czech population”, it is possible that many religious people declined in advance to participate in such a study. However, the frequencies of unbelievers and of adherents of various churches

roughly correspond to those obtained in national surveys (Czech Statistical Office, 2016; PEW, 2017), which suggests that this bias might be in fact relatively small. The bias could be actually larger and probably in the opposite direction, were the study advertised as focused on effects of various factors on religious beliefs or on effect of religious beliefs on various factors. The major advantage of the present setup was that it was absolutely anonymous and that it enabled us to obtain data from a large amount of subjects. This allowed us to analyze the associations between beliefs and health disorders and also to obtain data from Czech religious minorities, such as the Jews, adherents of the Czechoslovak Hussite Church or the Buddhists. It must be emphasized that the number of representatives of some interesting religious minorities such as of the Russian Orthodox Church is still rather low, therefore the research should continue for following years and perhaps specifically target members of the concerned minorities to get reliable statistics for these groups. The explorative aim of this study also brings limitations on the conclusions regarding the relationship between religion and mental health. To get more reliable data, it is necessary to analyze separately the different types of mental disorders, differently related to religion. In further studies, we will also include the type of religious practices, like prayer and meditation, and the frequency of attendance to rituals within the religious community, because they could have strong mediator power in the relationship between religion and health. In this study, we analyzed the effect of religiosity on self-reported honesty and altruism. This effect seems to vary strongly among denominations and differs in the expected direction between believers and nonbelievers. It must be stressed, however, that the variable honesty and altruism reflects self-rated trait that could be significantly different from the real honesty and altruism of the subject.

## Conclusion

The Czech Republic, now probably the most atheist country in Europe (PEW, 2017), represents an interesting product of about 50 years long large-scale social experiment. This



explorative study suggests that there is negative relationship between religiosity and wellbeing in Czech population. Before the communists took power in 1948, most Czechs were at least formally religious, mostly Catholics (Czech Statistical Office, 2016). Religion was a subject of an oppression and strict regulation until the fall of communism in 1989. Despite this, the religiosity survived in a large part of the population and continue to influence the quality of life of individual people even after two generations in persecution. The study of the current Czech population could help to disentangle direct and indirect effects of religiosity, spirituality, formal religiousness, and religious behavior on various aspects of human life.

## References

- Abu-Raiya, H., & Pargament, K. I. (2015). Religious coping among diverse religions: Commonalities and divergences. *Psychology of Religion and Spirituality, 7*(1), 24–33. <https://doi.org/10.1037/a0037652>
- Ahrenfeldt, L. J., Möller, S., Andersen-Ranberg, K., Vitved, A. R., Lindahl-Jacobsen, R., & Hvidt, N. C. (2017). Religiousness and health in Europe. *European Journal of Epidemiology, 32*(10), 921–929. <https://doi.org/10.1007/s10654-017-0296-1>
- Atkinson, Q. D., & Bourrat, P. (2011). Beliefs about God, the afterlife and morality support the role of supernatural policing in human cooperation. *Evolution and Human Behavior, 32*(1), 41–49. <https://doi.org/10.1016/j.evolhumbehav.2010.07.008>
- Bajaj, B., Gupta, R., & Pande, N. (2016). Self-esteem mediates the relationship between mindfulness and well-being. *Personality and Individual Differences, 94*, 96–100. <https://doi.org/10.1016/j.paid.2016.01.020>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological), 57*(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Berggren, N., & Ljunge, M. (2021). Good Faith and Bad Health: Self-Assessed Religiosity and Self-Assessed Health of Women and Men in Europe. *Social Indicators Research, 153*(1), 323–344. <https://doi.org/10.1007/s11205-020-02495-5>
- Brelsford, G. M., Luquis, R., & Murray-Swank, N. A. (2011). College students' permissive sexual attitudes: Links to religiousness and spirituality. *The International Journal for the Psychology of Religion, 21*(2), 127–136. <https://doi.org/10.1080/10508619.2011.557005>
- Clark, K. M., Friedman, H. S., & Martin, L. R. (1999). A Longitudinal Study of Religiosity and Mortality Risk. *Journal of Health Psychology, 4*(3), 381–391. <https://doi.org/10.1177/135910539900400307>
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Academic Press

Inc, New York.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, 98(2), 310–357. <https://doi.org/10.1037/0033-2909.98.2.310>

Czech Statistical Office. (2016). Czech Demographic Handbook 2016. Retrieved from <https://www.czso.cz/csu/czso/czech-demographic-handbook-2016>

Ellison, C. G., & Levin, J. S. (1998). The religion-health connection: Evidence, theory, and future directions. *Health Education & Behavior*, 25(6), 700–720. <https://doi.org/10.1177/109019819802500603>

Flegr, J., & Kuba, R. (2016). The relation of toxoplasma infection and sexual attraction to fear, danger, pain, and submissiveness. *Evolutionary Psychology*, 14(3), 1474704916659746. <https://doi.org/10.1177/1474704916659746>

Galen, L. (2015). Atheism, wellbeing, and the wager: Why not believing in God (with others) is good for you. *Science, Religion and Culture*, 2(3), 54–69. <https://doi.org/10.17582/journal.src/2015/2.3.54.69>

George, L. K., Larson, D. B., Koenig, H. G., & McCullough, M. E. (2000). Spirituality and health: What we know, what we need to know. *Journal of social and clinical psychology*, 19(1), 102–116. <https://doi.org/10.1521/jscp.2000.19.1.102>

Gilpin, A. R. (1993). Table for Conversion of Kendall'S Tau to Spearman'S Rho Within the Context of Measures of Magnitude of Effect for Meta-Analysis. *Educational and Psychological Measurement*, 53(1), 87–92. <https://doi.org/10.1177/0013164493053001007>

Green, M., & Elliott, M. (2009). Religion, Health, and Psychological Well-Being. *Journal of Religion and Health*, 49(2), 149–163. <https://doi.org/10.1007/s10943-009-9242-1>

Harrigan, J. T. (2009). Health Promoting Habits of People Who Pray for Their Health. *Journal of Religion and Health*, 50(3), 602–607. <https://doi.org/10.1007/s10943-009-9293-3>

Hayward, R. D., & Krause, N. (2014). Religion, mental health, and well-being: Social aspects. In Saroglou, V. (Eds.), *Religion, Personality, and Social Behavior* (pp. 255–280) Amsterdam University Press.

Hayward, R. D., Krause, N., Ironson, G., Hill, P. C., & Emmons, R. (2016). Health and Well-Being Among the Non-religious: Atheists, Agnostics, and No Preference Compared with Religious Group Members. *Journal of Religion and Health*, 55(3), 1024–1037. <https://doi.org/10.1007/s10943-015-0179-2>

Helliwell, J. F., & Putnam, R. D. (2004). The social context of well-being. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1449), 1435–1446. <https://doi.org/10.1098/rstb.2004.1522>

Holt, C. L., Wang, M. Q., Clark, E. M., Williams, B. R., & Schulz, E. (2013). Religious involvement and physical and emotional functioning among African Americans: The mediating role of religious support. *Psychology & Health*, 28(3), 267–283. <https://doi.org/10.1080/08870446.2012.717624>

Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality: a meta-analytic review. *Perspectives on Psychological Science*, 10(2), 227–237. <https://doi.org/10.1177/1745691614568352>

Hwang, K., Hammer, J. H., & Cragun, R. T. (2009). Extending Religion-Health Research to Secular Minorities: Issues and Concerns. *Journal of Religion and Health*, 50(3), 608–622. <https://doi.org/10.1007/s10943-009-9296-0>

Joshanloo, M., & Daemi, F. (2015). Self-esteem mediates the relationship between spirituality and subjective well-being in Iran. *International Journal of Psychology*, 50(2), 115–120. <https://doi.org/10.1002/ijop.12061>

Kaňková, Š., Flegr, J., & Calda, P. (2015). The influence of latent toxoplasmosis on women's reproductive function: four cross-sectional studies. *Folia Parasitologica*, 62, 1. <https://doi.org/10.14411/fp.2015.041>

Kaňková, Š., Kodym, P., & Flegr, J. (2011). Direct evidence of Toxoplasma-induced changes in serum testosterone in mice. *Experimental Parasitology* 128(3), 181–183. <https://doi.org/10.1016/j.exppara.2011.03.014>

Kasi, P. M., Naqvi, H. A., Afghan, A. K., Khawar, T., Khan, F. H., Khan, U. Z., Khuwaja, U. B., Kiani, J., & Khan, H. M. (2012). Coping styles in patients with anxiety and depression. *ISRN Psychiatry*, 2012,1–7. <https://doi.org/10.5402/2012/128672>

Klímová H, Roubalová, V. (2014) *The transgenerational transfer of trauma and the possibility of recovery*. Prague: Irine Press.

Koenig, H. G., King, D., & Carson, V. B. (2012). *Handbook of Religion and Health* (2nd ed.). Oxford University Press.

Krause, N. (2017). Religious involvement and self-forgiveness. *Mental Health, Religion & Culture*, 20(2), 128–142. <https://doi.org/10.1080/13674676.2017.1326477>

Krause, N., & Wulff, K. M. (2005). "Church-Based Social Ties, A Sense of Belonging in a Congregation, and Physical Health Status." *The International Journal for the Psychology of Religion*, 15(1), 73–93. [https://doi.org/10.1207/s15327582ijpr1501\\_6](https://doi.org/10.1207/s15327582ijpr1501_6)

Lambert, L., Lomas, T., van de Weijer, M. P., Passmore, H. A., Joshanloo, M., Harter, J., Ishikawa, Y., Lai, A., Kitagawa, T., Chen, D., Kawakami, T., Miyata, H., & Diener, E. (2020). Towards a greater global understanding of wellbeing: A proposal for a more inclusive measure. *International Journal of Wellbeing*, 10(2), 1–18. <https://doi.org/10.5502/ijw.v10i2.1037>

Latifnejad Roudsari, R., Allan, H. T., & Smith, P. A. (2014). Iranian and English women's use of religion and spirituality as resources for coping with infertility. *Human Fertility*, 17(2), 114–123. <https://doi.org/10.3109/14647273.2014.909610>

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer publishing company.

Libby, G., Zimmer, Z., Kingston, A., Haviva, C., Chiu, C. T., Ofstedal, M. B., Saito, Y., & Jagger, C. (2021). Are Religiosity and Spirituality Related to Self-Reported Health Expectancy? An Analysis of the European Values Survey. *Journal of Religion and Health*. <https://doi.org/10.1007/s10943-021-01348-w>

Lim, C., & Putnam, R. D. (2010). Religion, Social Networks, and Life Satisfaction. *American Sociological Review*, 75(6), 914–933. <https://doi.org/10.1177/0003122410386686>

Maselko, J., & Kubzansky, L. D. (2006). Gender differences in religious practices,

spiritual experiences and health: Results from the US General Social Survey. *Social Science & Medicine*, 62(11), 2848–2860. <https://doi.org/10.1016/j.socscimed.2005.11.008>

McCullough, M. E., Carter, E. C., DeWall, C. N., & Corrales, C. M. (2012). Religious cognition down-regulates sexually selected, characteristically male behaviors in men, but not in women. *Evolution and Human Behavior*, 33(5), 562–568. <https://doi.org/10.1016/j.evolhumbehav.2012.02.004>

McCullough, M. E., & Laurenceau, J. P. (2005). Religiousness and the trajectory of self-rated health across adulthood. *Personality and Social Psychology Bulletin*, 31(4), 560–573. <https://doi.org/10.1177/0146167204271657>

Miller, W. R., & Thoresen, C. E. (2003). Spirituality, religion, and health: An emerging research field. *American Psychologist*, 58(1), 24–35. <https://doi.org/10.1037/0003-066x.58.1.24>

Nagel, E., & Sgoutas-Emch, S. (2006). The Relationship Between Spirituality, Health Beliefs, and Health Behaviors in College Students. *Journal of Religion and Health*, 46(1), 141–154. <https://doi.org/10.1007/s10943-006-9088-8>

Nicholson, A., Rose, R., & Bobak, M. (2009). Association between attendance at religious services and self-reported health in 22 European countries. *Social Science & Medicine*, 69(4), 519–528. <https://doi.org/10.1016/j.socscimed.2009.06.024>

Niederlande, W. G. (1964) Psychiatric disorders among persecution victims. A contribution to the understanding of concentration camp pathology and its after-effects. *The Journal of Nervous and Mental Disease*, 139(5), 458–474. <https://doi.org/10.1097/00005053-196411000-00006>

Oman, D., & Thoresen, C. E. (2002). ‘Does Religion Cause Health?’: Differing Interpretations and Diverse Meanings. *Journal of Health Psychology*, 7(4), 365–380. <https://doi.org/10.1177/1359105302007004326>

Penke, L., & Asendorpf, J. B., (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, 95 (5), 1113–1135. <https://doi.org/10.1037/0022-3514.95.5.1113>

Pew Research Center. (2017) *Religious Belief and National Belonging in Central and Eastern Europe*. Retrieved from <https://www.pewforum.org/2017/05/10/religious-belief-and-national-belonging-in-central-and-eastern-europe>

Powell, L. H., Shahabi, L., & Thoresen, C. E. (2003). Religion and spirituality: Linkages to physical health. *American psychologist*, 58(1), 36–52. <https://doi.org/10.1037/0003-066x.58.1.36>

Revens, K. E., Gutierrez, D., Paul, R., Reynolds, A. D., Price, R., & DeHaven, M. J. (2021). Social Support and Religiosity as Contributing Factors to Resilience and Mental Wellbeing in Latino Immigrants: A Community-Based Participatory Research Study. *Journal of Immigrant and Minority Health*, 1–13. <https://doi.org/10.1007/s10903-021-01179-7>

Rew, L., & Wong, Y. J. (2006). A systematic review of associations among religiosity/spirituality and adolescent health attitudes and behaviors. *Journal of Adolescent Health*, 38(4), 433–442. <https://doi.org/10.1016/j.jadohealth.2005.02.004>

Rostosky, S. S., Wilcox, B. L., Wright, M. L. C., & Randall, B. A. (2004). The Impact of Religiosity on Adolescent Sexual Behavior: *Journal of Adolescent Research*, 19(6), 677–697.

<https://doi.org/10.1177/0743558403260019>

Schieman, S., Bierman, A., Upenieks, L., & Ellison, C. G. (2017). Love Thy Self? How Belief in a Supportive God Shapes Self-Esteem. *Review of Religious Research*, 59(3), 293–318.

<https://doi.org/10.1007/s13644-017-0292-7>

Seybold, K. S., & Hill, P. C. (2001). The Role of Religion and Spirituality in Mental and Physical Health.

*Current Directions in Psychological Science*, 10(1), 21–24. <https://doi.org/10.1111/1467-8721.00106>

Shariff, A. F. (2015). Does religion increase moral behavior? *Current Opinion in Psychology*, 6, 108–113. <https://doi.org/10.1016/j.copsyc.2015.07.009>

Shariff, A. F., Willard, A. K., Andersen, T., & Norenzayan, A. (2016). Religious priming: A meta-analysis with a focus on prosociality. *Personality and Social Psychology Review*, 20(1), 27–48. <https://doi.org/10.1177/1088868314568811>

Sharma, S., & Singh, K. (2018). Religion and Well-Being: The Mediating Role of Positive Virtues.

*Journal of Religion and Health*, 58(1), 119–131. <https://doi.org/10.1007/s10943-018-0559-5>

Siegel, S., & Castellan, N. J. (1988). The case of k related samples. *Nonparametric statistics for behavioral sciences*. New York: McGraw-Hill, 170–4.

Simons, L. G., Burt, C. H., & Peterson, F. R. (2009). The Effect of Religion on Risky Sexual Behavior among College Students. *Deviant Behavior*, 30(5), 467–485.

<https://doi.org/10.1080/01639620802296279>

Sloan, R. P., & Bagiella, E. (2002). Claims about religious involvement and health outcomes. *Annals of Behavioral Medicine*, 24(1), 14–21. [https://doi.org/10.1207/s15324796abm2401\\_03](https://doi.org/10.1207/s15324796abm2401_03)

Speed, D., & Fowler, K. (2015). What's God Got to Do with It? How Religiosity Predicts Atheists' Health. *Journal of Religion and Health*, 55(1), 296–308. <https://doi.org/10.1007/s10943-015-0083-9>

Stavrova, O. (2015). Religion, self-rated health, and mortality: whether religiosity delays death depends on the cultural context. *Social Psychological and Personality Science*, 6(8), 911–922.

<https://doi.org/10.1177/1948550615593149>

Thege, B. K., Pilling, J., Székely, A., & Kopp, M. S. (2013). Relationship between religiosity and health: Evidence from a post-communist country. *International journal of behavioral medicine*, 20(4), 477–486. <https://doi.org/10.1007/s12529-012-9258-x>

Tiggemann, M., & Hage, K. (2019). Religion and spirituality: Pathways to positive body image. *Body image*, 28, 135–141. <https://doi.org/10.1016/j.bodyim.2019.01.004>

Valkenburg, P. M., Peter, J., & Schouten, A. P. (2006). Friend networking sites and their

relationship to adolescents' well-being and social self-esteem. *CyberPsychology & Behavior*, 9(5), 584–590.

<https://doi.org/10.1089/cpb.2006.9.584>

Van Cappellen, P., Toth-Gauthier, M., Saroglou, V., & Fredrickson, B. L. (2014). Religion and Well-Being: The Mediating Role of Positive Emotions. *Journal of Happiness Studies*, 17(2), 485–505. <https://doi.org/10.1007/s10902-014-9605-5>

Wang, X. (2016). Subjective well-being associated with size of social network and social support of elderly. *Journal of health psychology*, 21(6), 1037–1042. <https://doi.org/10.1177/1359105314544136>

Weber, S. R., Pargament, K. I., Kunik, M. E., Lomax, J. W., & Stanley, M. A. (2011). Psychological Distress Among Religious Nonbelievers: A Systematic Review. *Journal of Religion and Health*, 51(1), 72–86. <https://doi.org/10.1007/s10943-011-9541-1>

Weeden, J., Cohen, A. B., & Kenrick, D. T. (2008). Religious attendance as reproductive support.

*Evolution and Human Behavior*, 29(5), 327–334. <https://doi.org/10.1016/j.evolhumbehav.2008.03.004>

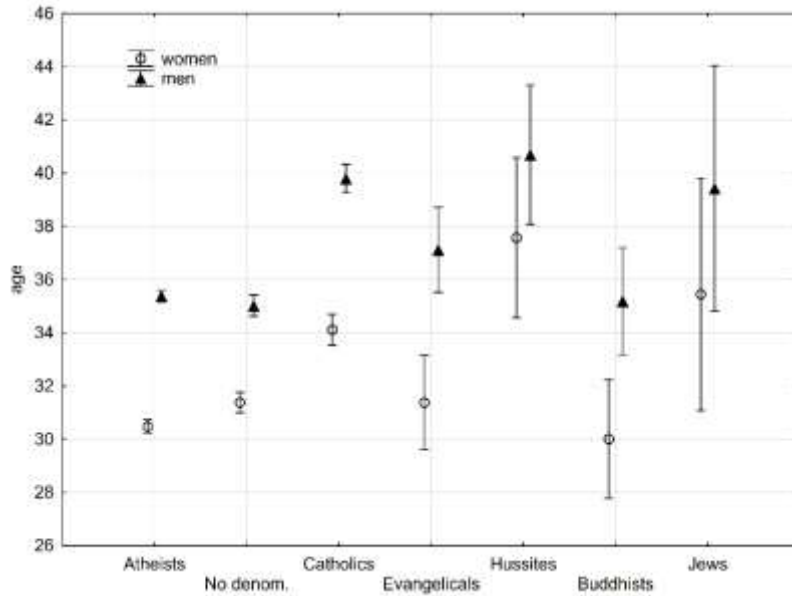
Xygalatas, D., Klocová, E. K., Cigán, J., Kundt, R., Maño, P., Kotherová, S., Mitkidis, P., Wallot, S., & Kanovsky, M. (2015). Location, Location, Location: Effects of Cross-Religious Primes on Prosocial Behavior. *The International Journal for the Psychology of Religion*, 26(4), 304–319.

<https://doi.org/10.1080/10508619.2015.1097287>

Yehuda, R., Bierer, L. M., Schmeidler, J., Aferiat, D. H., Breslau, I., & Dolan, S. (2000). Low Cortisol and Risk for PTSD in Adult Offspring of Holocaust Survivors. *American Journal of Psychiatry*, 157(8), 1252–1259. <https://doi.org/10.1176/appi.ajp.157.8.1252>

**Figure 1**

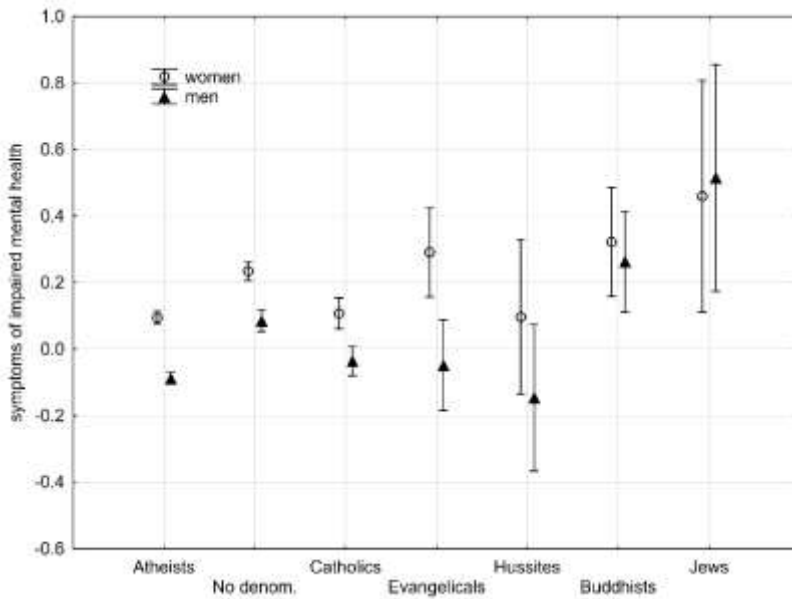
*Difference in age between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

**Figure 2**

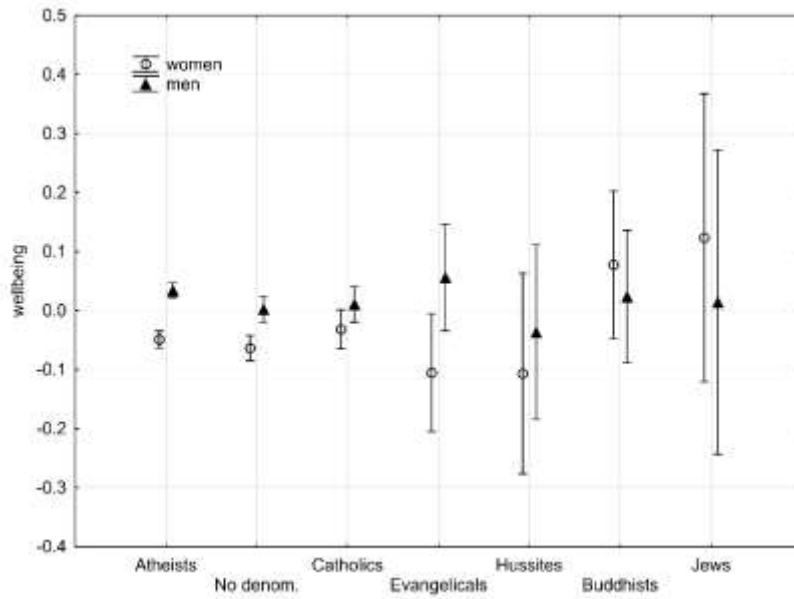
*Difference in the subjective symptoms of impaired mental health between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

**Figure 3**

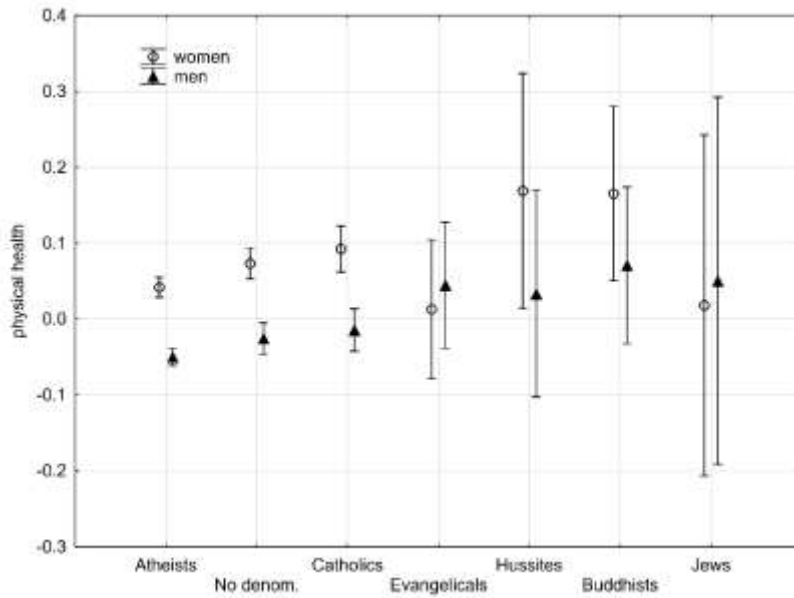
*Difference in the wellbeing measured with WHOQOL between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

**Figure 4**

*Difference in the physical health between men and women of different denominations*

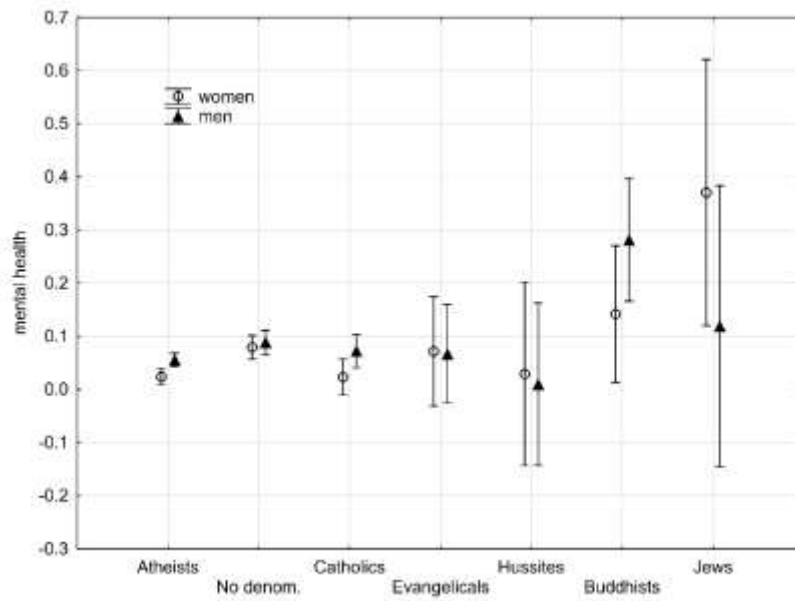


*Note.* Spreads show 95% confidence intervals.



**Figure 5**

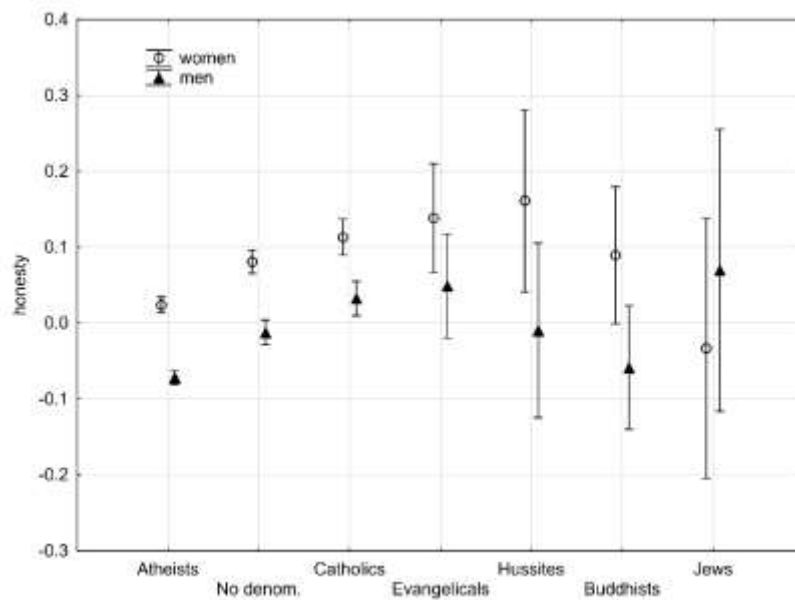
*Difference in the mental health between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

**Figure 6**

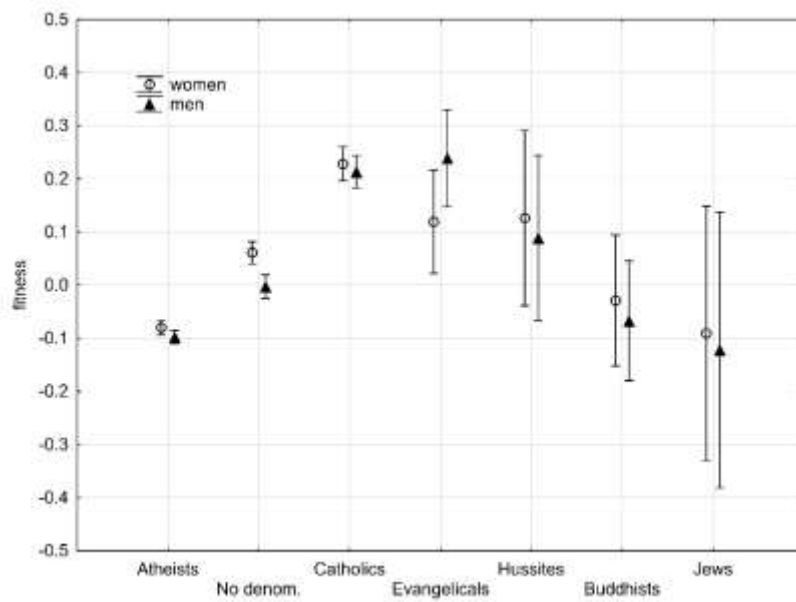
*Difference in the honesty between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

**Figure 7**

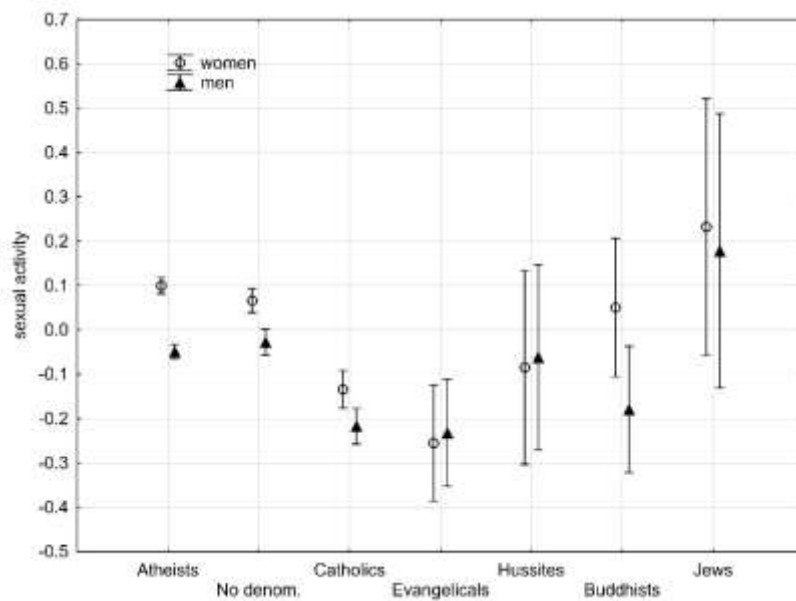
*Difference in the biological fitness between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

**Figure 8**

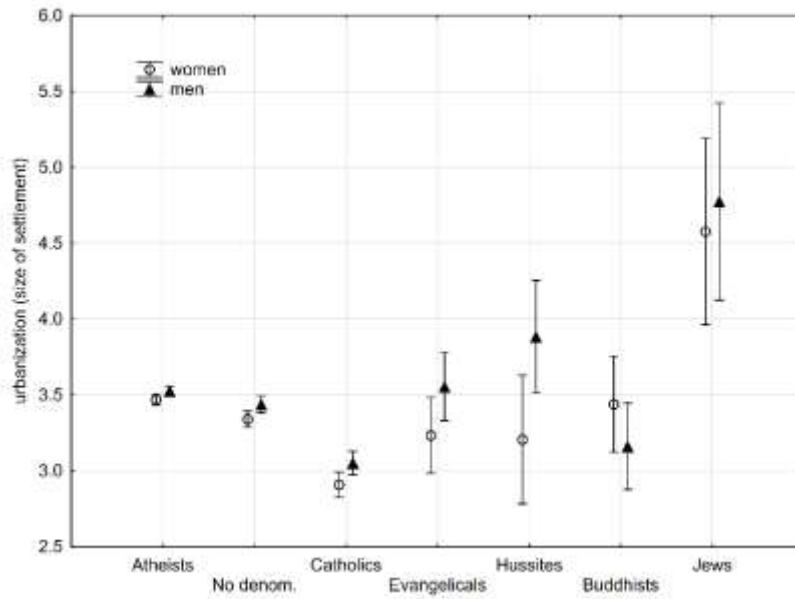
*Difference in the intensity of sexual life between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

**Figure 9**

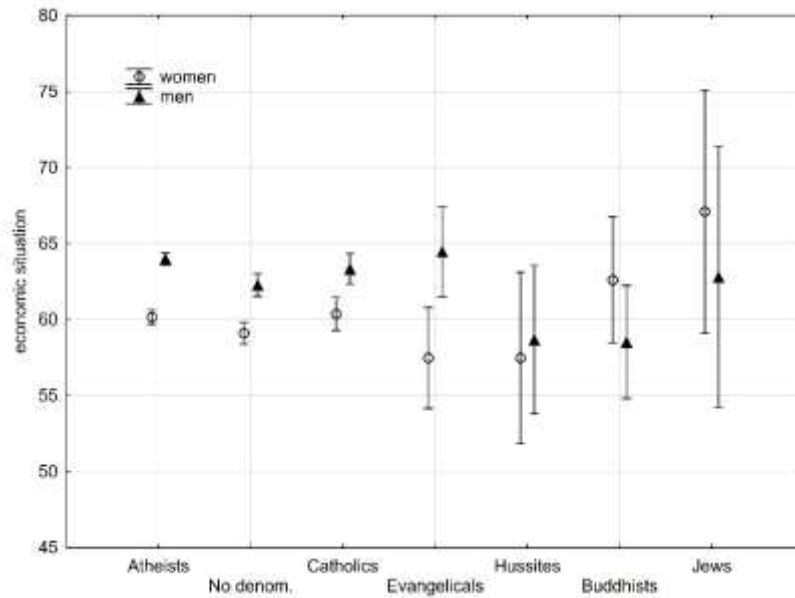
*Difference in urbanization (size of place of living) between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

**Figure 10**

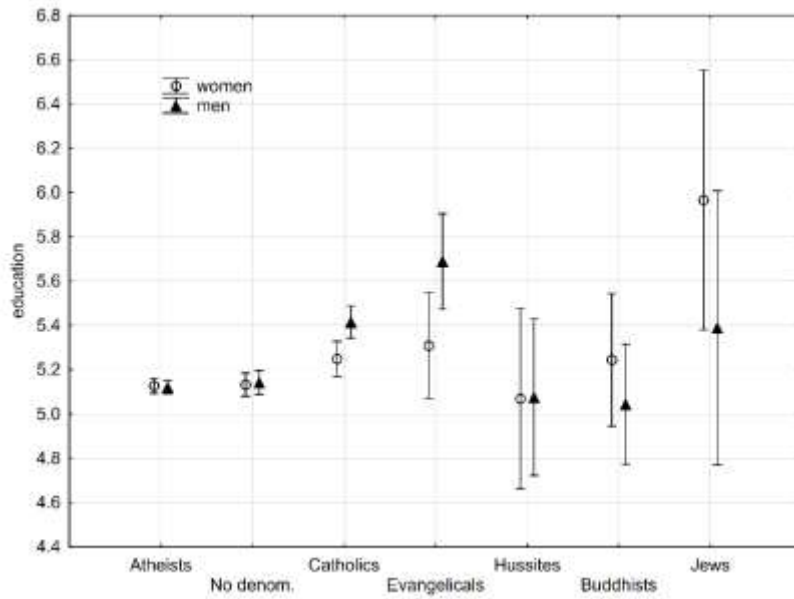
*Difference in economic situation (wealth) between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

**Figure 11**

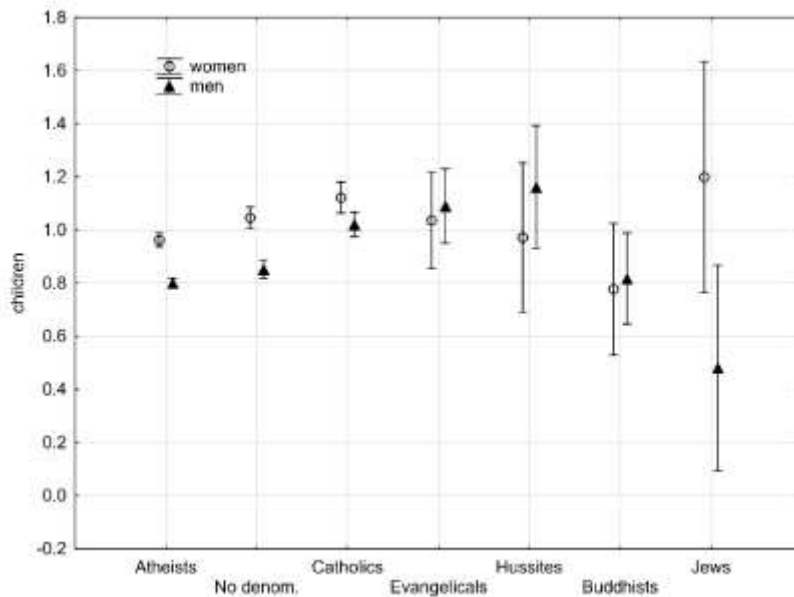
*Difference in education between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

**Figure 12**

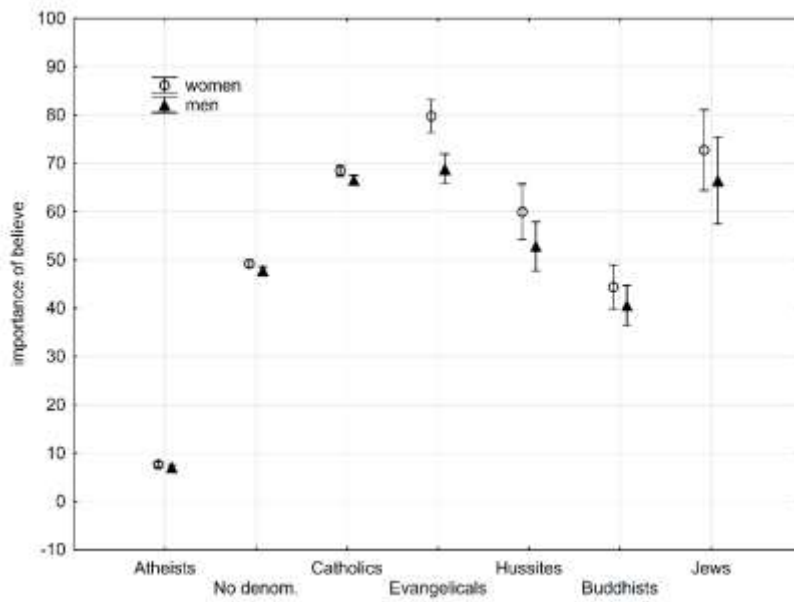
*Difference in fecundity (number of children) between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

**Figure 13**

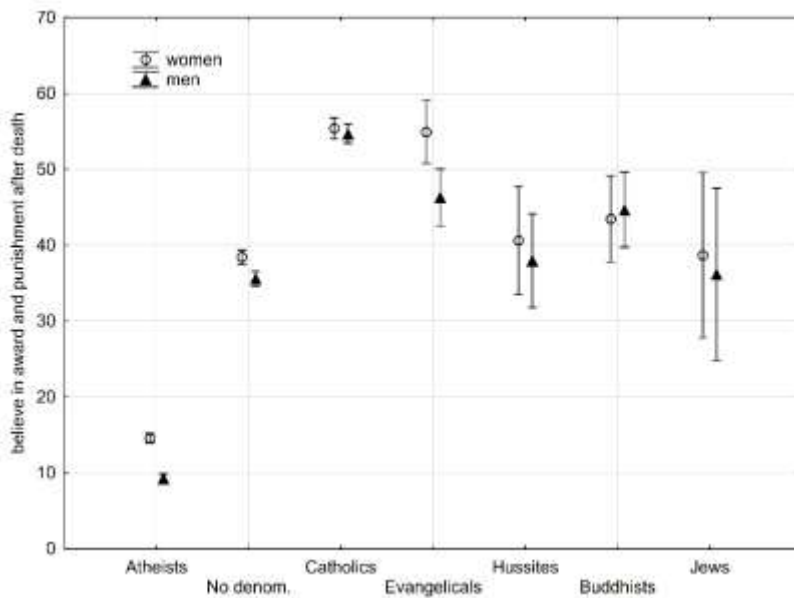
*Difference in importance of religion in personal life between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

**Figure 14**

*Difference in belief in afterlife reward and punishment between men and women of different denominations*



*Note.* Spreads show 95% confidence intervals.

### **Funding details**

This work was supported by The Czech Science Foundation (GACR) under grant number 21-31059S; and the Charles University Grant Agency (GAUK) under grant number 929216.

### **Disclosure statement**

No potential conflict of interest was reported by the authors.

### **ORCID**

Robin Kopecky <https://orcid.org/0000-0002-7140-4649>

[fb.com/robin.kopecky](https://www.facebook.com/robin.kopecky)

Silvia Boschetti <https://orcid.org/0000-0002-8048-4062>

[fb.com/silvia.boschetti.37](https://www.facebook.com/silvia.boschetti.37)

Jaroslav Flegr <https://orcid.org/0000-0002-0822-0126> corresponding

author, [flegr@cesnet.cz](mailto:flegr@cesnet.cz)

### **Acknowledgements**

RK, JF developed the study concept. JF analyzed the data. RK, SB, JF collected data and wrote the manuscript. We would like to thank Konrad Talmont-Kaminski, Lenka Příplatová, Michaela Jirout Košová for their valuable comments and their help with the final revision of the paper.

## **5. Experimentální filosofie**

### **5.1 Osobní identita a právě já**



## My friend's true self: Children's concept of personal identity

Michaela Jirout Kořov, Robin Kopeck, Pavel Oulovsk, Matj Nekvinda & Jaroslav Flegr

To cite this article: Michaela Jirout Kořov, Robin Kopeck, Pavel Oulovsk, Matj Nekvinda & Jaroslav Flegr (2021) My friend's true self: Children's concept of personal identity, *Philosophical Psychology*, 34:1, 47-75, DOI: [10.1080/09515089.2020.1860209](https://doi.org/10.1080/09515089.2020.1860209)

To link to this article: <https://doi.org/10.1080/09515089.2020.1860209>

View supplementary material [↗](#)

Published online: 16 Feb 2021.

Submit your article to this journal [↗](#)

Article views: 519

View related articles [↗](#)

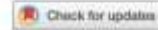
View Crossmark data [↗](#)

Citing articles: 2 View citing articles [↗](#)

Full Terms & Conditions of access and use can be found at <https://www.tandfonline.com/action/journalInformation?journalCode=cphp20>



ARTICLE



## My friend's true self: Children's concept of personal identity

Michaela Jirout Košová<sup>a</sup>, Robin Kopecký<sup>a</sup>, Pavel Oulovský<sup>a</sup>, Matěj Nekvinda<sup>b</sup> and Jaroslav Flegl<sup>a</sup>

<sup>a</sup>Department of Philosophy and History of Science, Faculty of Science, Charles University, Prague, Czech Republic; <sup>b</sup>Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

### ABSTRACT

Our study explores the folk concept of personal identity in the developmental context. Two hundred and seventeen Czech children participated in an interview study based on a hypothetical scenario about a sudden change in their friend, someone they know, or some other unspecified person. The children were asked to judge to what extent particular changes (from six categories of traits) would change the identity core of their friend or some other person on a seven-point scale. We introduced both positive and negative versions of the changes. Our data suggest that children considered moral traits connected to interpersonal relationships crucial for preserving personal identity. Memory connected to personal experiences also scored highly. On the other hand, a change in physical appearance seemed to have the least important impact on personal identity. Negative changes turned out to have a significantly greater impact than positive changes in all categories, except the physical. Possible effects of scenario and the participants' age and sex were further explored. We discuss the possible causes of the effect of negative moral change and the role of social dimension in the development of the concept of personal identity.

### ARTICLE HISTORY

Received 22 February 2019  
Accepted 23 May 2020

### KEYWORDS


Personal identity; true self; experimental philosophy; moral; interpersonal; developmental; psychological essentialism

## 1. Introduction

### 1.1. Current knowledge

As social beings, we live in a world full of selves: each person has various typical characteristics, both physical and psychological, which make them who they are – some of them playing a larger role than others. We see each other as “bundles” of traits and dispositions, which helps us to orient within the social environment and choose wisely when entering new

**CONTACT** Michaela Jirout Košová  [michaela.kosa@gmail.com](mailto:michaela.kosa@gmail.com)  Department of Philosophy and History of Science, Faculty of Science, Charles University, Prague, Czech Republic.

This article has been corrected with minor changes. These changes do not impact the academic content of the article.  
 Supplemental data for this article can be accessed here.

© 2021 Informa UK Limited, trading as Taylor & Francis Group

relationships. However, people change throughout their lives, and although some changes are seen as natural, others can threaten the integrity of the self. Thus, it seems that the “bundles” that we are made up of have certain essential cores. Many thinkers have devoted a substantial part of their work to developing theories about personal identity, trying to determine “from the armchair” what is the most essential aspect of the self and what has to stay intact in order for a person’s identity to remain unbroken.

Together with the researchers who inspired our current work, we are skeptical about most of these armchair-philosophy approaches. Since the concept of self is so dependent on the fact that we are social beings, it is best to look for the answer in a natural environment: between people untouched by highly abstract philosophical debates, whose concepts naturally stem from their everyday life in society. Personal identity is a question that matters to everyone, and we lean toward the view that it is defined by the way people view it in everyday social interactions and practices (Prinz & Nichols, 2016, p. 449). Thus, we see experimental philosophy as the right approach in the quest to find out more about these issues because it enables us to get in contact with the real-world phenomenon, even though we do not deny the value of purely theoretical inquiry.

This paper addresses current experimental-philosophy research on children’s concept of personal identity and its relation to morality. The way we define morality in this context is specific and follows previous studies on the topic, since, as we already suggested, the results of the existing studies point to the fact that interpersonal relationships are especially important in the context of revealing the folk concept of personal identity (Prinz & Nichols, 2016; Strohminger & Nichols, 2014, 2015). The focus of the paper is therefore on the interpersonal dimension of morality, by which we mean personal traits and dispositions that directly affect interpersonal relationships: for example, cruelty, selfishness, criminal behavior, and callousness on the one side; and empathy, generosity, selflessness, conscientiousness, and kindness on the other side. We can also define our sphere of interest as a kind of *moral essentialism* (Heiphetz et al., 2017, p. 745).

Earlier studies exploring these issues were mostly questionnaire studies of adult respondents. Their results revealed a strong tendency in the respondents to ascribe importance to moral traits, especially those connected to the social realm, in the context of preservation of personal identity. In their online studies launched in 2009, Prinz and Nichols explored the folk concept of diachronic identity and managed to show that moral continuity is far more crucial for preserving personal identity over time than memory, agency, or narrative ability. They explained their findings by suggesting that morality is above all a social phenomenon and is thus of extreme importance to us. Moral values are sustained by the community in which



we live, and our success as social beings depends on how this community views us (Prinz & Nichols, 2016, p. 463).

The above-mentioned findings about the role of the social dimension in the concept of personal identity has inspired further studies that have managed to support this view. In their five online questionnaire studies with adult respondents, Strohminger and Nichols (2014) have found strong support for their hypothesis of the *essential moral self*. The intuition of most of their participants turned out to be that moral traits constitute the core of personal identity. Similar results in another study further supported the hypothesis: the study focused on the intuitions of relatives of patients with neurodegenerative disorders (Strohminger & Nichols, 2015). Symptoms of disorders that affect the moral (interpersonal) dispositions of a patient were viewed by respondents as more identity-breaking than other common symptoms, such as amnesia.

Strohminger et al. (2017) refined the concept of the self by suggesting that apart from the concept of *self* as such, folk people also have the concept of *true self*, which is a kind of subset of characteristics that are the most crucial and deep-down constituents of a person's self (p. 552). Based on various psychological research findings, they defend the hypothesis that the true self is generally perceived as moral and inherently good (pp. 552–554), which could possibly be a result of the way we conceptualize essences or our need to maintain social bonds (pp. 556–557).

Similarly, Heiphetz et al. (2017) have found that the role of social bonds seems to be crucial for the folk concept of personal identity. Participants in their studies judged belonging to a group as being important for them, and mediation analyses revealed the relationship between the perceived importance of community and the type of moral belief they judged to be more important for preserving personal identity. Changes to widely shared moral beliefs would result in a more radical identity change than changes in controversial moral beliefs because widely shared beliefs are more closely connected to relationships in the community (p. 758).

The importance of interpersonal relationships in judgments about personal identity also comes forward in research conducted by Tobia (2016), focusing on the direction of hypothetical change. Tobia takes into consideration relational theories of personal identity, which emphasize the role of interpersonal relationships – an approach that is often overlooked in most philosophical theories (p. 38). In his study, he broadened the classical scenario and found that respondents presented with the reversed Phineas Gage scenario (moral improvement) tended to claim that Phineas is still the same person after the accident, whereas in the original scenario (moral deterioration), the opposite tendency was the case. Both changes are equally radical, but the negative version leads more often to conclusions that the identity of the person in question was broken. Tobia suggests that the size of

the change does not play the main role here, and the reason why responses to opposing scenarios differ so greatly is connected to the fact that the self is viewed as essentially good (pp. 39–40). In his earlier paper (2015), Tobia has supported this theory with further examples, for instance, Parfit's 'nobleman' thought-experiment, and examples from pop-culture and literature. He concludes that normativity, which is connected to the social dimension, is to be viewed as a crucial aspect of the personal identity concept (p. 404).

Heiphetz et al. (2018) also reported that their adult respondents judged the change from good moral beliefs to bad moral beliefs to have a significantly more serious impact on the preservation of personal identity than the change in the opposite direction (bad moral beliefs to good moral beliefs). The difference was mediated by the perceived influence of both types of changes on friendships. However, the results of the same study on children aged 8–10 years did not show these effects, possibly because of the sample size or study design (p. 216).

The authors also conducted another study comparing children and adults (pp. 211–214). The respondents were supposed to judge how much a person would change after taking a pill that changes either their widely shared moral beliefs, controversial moral beliefs, memories, or preferences. Children reported, similar to adults, that the person would change more after a change in their widely shared moral beliefs took place.

The difference between adults and children appeared in the impact of memories and preferences on perceived identity change – in this context, children found memories and preferences more important than adults did (p. 214).

Memory, no doubt, plays a crucial role in the question of personal identity and has always been considered in this context (the classic example is John Locke, 2009, originally published 1690). Memory has also come forward in findings of modern authors who emphasize the role of psychological characteristics in contrast to physical ones (Nichols & Bruno, 2010). However, studies exploring the difference between the importance of memory and moral traits for preserving personal identity show that in this context, moral traits surpass memory (Prinz & Nichols, 2016; Strohminger & Nichols, 2014, 2015), although not consistently in all studies (Heiphetz et al., 2017). It is thus desirable for future personal-identity studies to consider the role of memory and explore its importance in comparison to morality.

The question of why interpersonal moral traits are so important for us certainly has to do with the fact that we are social beings and view ourselves and others in the context of social relationships (Prinz & Nichols, 2016, p. 463). However, our understanding of why good moral traits appear in the very core of personal identity might be deepened by pointing out our natural tendency to essentialize (Strohminger et al., 2017, pp. 556–557).



*Psychological essentialism* is a tendency to view entities as having an inner essence that is hidden behind superficial traits (De Freitas, Sarkissian, et al., 2018, p. 138). According to De Freitas and colleagues, the ‘true self’ concept is a result of these natural essentialist tendencies, since the traits that are usually ascribed to the true self are viewed as “immutable, discrete and inherent” (De Freitas, Sarkissian, et al., 2018, p. 138). The answer to the question of why social traits are more important than other kinds of traits might lie in the normative and teleological nature of essentialism. People seem to think about the essences of various entities in terms of their purpose, their *telos* (Rose & Nichols, 2019), and, at the same time, the traits that are believed to support personal identity are predominantly normatively good (De Freitas et al., 2017, p. 398). This logic that connects normativity and teleology could also be applied to human beings: if the purpose of the music band is to make good-quality music, and the purpose of scientific papers is to carry valuable scientific information, then the purpose of a person is to be morally good (De Freitas et al., 2017, p. 397). This theory seems to be in perfect accord with the battery of studies showing that good moral traits play a superior role in preserving personal identity and explains why the folk view of personal identity seems so optimistic: it entails that everyone is inherently good (see De Freitas & Cikara, 2018).

Also, the question of what exactly the studies described in this section and other similar studies reveal about the lay concept of personal identity is still open. Recent debate shows that we can interpret the results as a tendency by respondents to view a change in moral traits as leading to a significant change rather than the annihilation of the person in question (Starmans & Bloom, 2018). On the other hand, it is possible to argue for a more radical interpretation, saying that with moral deterioration comes the end of the person in question, in the sense of numerical identity (De Freitas, Cikara et al., 2018). There are studies, such as those of Tobia (2015), which manage to disambiguate between numerical identity and similarity (Dranseika, 2017), and they still reveal the effect that we find in other, conceptually less precise studies.

We do not believe that it is necessary to stick to either of the interpretations in order to appreciate the aforementioned research. These studies show us that the concept of personal identity is closely connected to the social dimension and the appreciation of moral goodness, regardless of the precise formulation of the folk concept. What is more, it might be misleading to expect philosophical rigor from the folk. On the contrary, we believe that we should be prepared to embrace a certain vagueness when trying to describe folk concepts.

## 1.2. Original study

The study presented in this paper focuses on children's intuitions about personal identity, but it differs from the aforementioned study on children (Heiphetz et al., 2018) by introducing more categories of change, a wider age-range of interviewed children, and a larger sample size. Our aim was to explore the importance of the moral category in comparison to five other categories (physical, memory, character, perception, cognition) in the perceived preservation of personal identity in children aged 5–15 and thus explore the 'essential moral self' hypothesis in the developmental context. By including both directions of change, we also focused on the 'true self' concept.

Based on the current research findings in the field, we decided to test the following hypotheses, as listed in the study preregistration<sup>1</sup>: (1) The impact of hypothetical changes in various categories of traits will differ, and the changes in moral traits (traits which figure strongly in interpersonal relationships) will be rated as having the highest impact on the perceived change in personal identity; (2) when the hypothetical change is negative, the overall change in personal identity will be rated as more serious than in cases where the change is positive; (3) in the case of negative moral change, the overall change in personal identity will be rated as much more serious than in any other category of negative change; (4) the exploratory part of the study: the effect of age and sex in ratings of the impact of different categories of change on preserving personal identity.

## 2. Method

### 2.1. Participants

In June, 2017, we conducted an interview study on Czech children and teenagers. The study was preregistered and approved by the IRB.<sup>2</sup> In our preregistered plan, we set a stopping rule at 300 respondents in the age range 5–15 years old. The number was supposed to approximately match the number of participants in the online studies performed on adult respondents by Strohminger and Nichols (2014), which ranged from 79 to 318.

Altogether, we interviewed 267 respondents. As reported in the preregistration, participants who did not show a sufficient understanding of the study question and who exhibited random answers even before the experimenter properly finished particular questions were filtered out already during the data collection. We excluded 28 participants based on this rule. One participant was excluded due to the parent's report of autism. We also interviewed 21 participants who were above age 15, since we agreed to go through the whole interview with all potential participants who showed an interest even after they had reported their age at the beginning.<sup>3</sup> These participants were excluded before the analysis. One participant aged 5 was



excluded because we were not able to get more participants of this age. We thus set the age range to 6–15. We ended up with a sample of 217 children for the final analysis (56.4% female; age range 6–15; average age = 11).

## 2.2. Procedure

Respondents were randomly recruited at a public family event popularizing science. After gaining informed consent from a parent and asking the parent to step aside, an interviewer informed the child participant that she was exploring the thinking of children and young people and was interested in what the participant thought about the following problem. She assured the child that none of the answers were right or wrong and the point was to find out what the child's own personal opinion was.

Once the child had agreed to this model, the interviewer introduced each participant to a scenario in which a person undergoes various changes after being closed in a special sci-fi chamber. A random group of participants ( $N = 90$ ) were asked about their *friend*; another group of participants ( $N = 91$ ) were asked about *some person in general*; and the rest of the participants ( $N = 36$ ) were asked less generally about *someone they know* ("your friend, peer, or someone else"). Changes in all scenarios encompassed six categories: physical (appearance), cognitive (intelligence), moral (love for others and treatment of others), in character (laziness), in memory (remembering life experiences), and in perception (vision). Both negative and positive versions of the changes were included in all categories. Altogether, we formed 14 questions which were mixed and presented in the following order: the person (1) becomes uglier, (2) gains a super-memory, (3) stops loving their friends, (4) becomes more industrious (5) becomes blind, (6) becomes more beautiful, (7) forgets their life experiences, (8) becomes nicer to others, (9) becomes more stupid, (10) accepts as a friend someone they didn't like before, (11) becomes lazier, (12) gains much better eyesight, (13) becomes cruel to others, and (14) becomes smarter. The respondents were asked to judge how much each of the changes would affect the person's identity core ("the most crucial aspect of the person which makes them who they really are deep inside") on a seven-point scale from 0 (*they are still the same person and their most crucial aspect remains intact*) to 6 (*they are not the same person anymore and have lost their most crucial aspect*). We indicated the scale by circles growing in size (the larger the circle, the larger the overall change – see supplementary material for details and exact formulations). We decided to use this method because we needed the scale to be comprehensible to children who have little or no experience with reading (see, e.g., Okawa, 2008, p. 188). The relationship between the size of the circle and the size of the change was pointed out to the respondents at the beginning.

Once the participant had finished answering the questions, the interviewer thanked them for their cooperation and gave each respondent a badge as a reward.

The choice of within-subject model (with an exception of the between-subject vignette question) was made due to the fact that we were only able to gain a restricted number of respondents varying across age categories. The fixed order of questions also reflects the restricted conditions of the study: a face-to-face interview without the use of computer technology.

### 2.3. Differences between preregistered and realized protocol

After the preregistration, we divided the vignette into three scenarios: 'friend,' 'someone you know', and 'some person in general', in order to explore the role of closeness in interpersonal relationships.

We adjusted the planned age range (5–15) to 6–15 years old, since we had only managed to recruit one 5-year-old.

To test hypothesis 3, we used one-sided paired t-tests, just as we did for hypotheses 1 and 2. Due to a mistake, we failed to mention this test in the preregistered protocol.

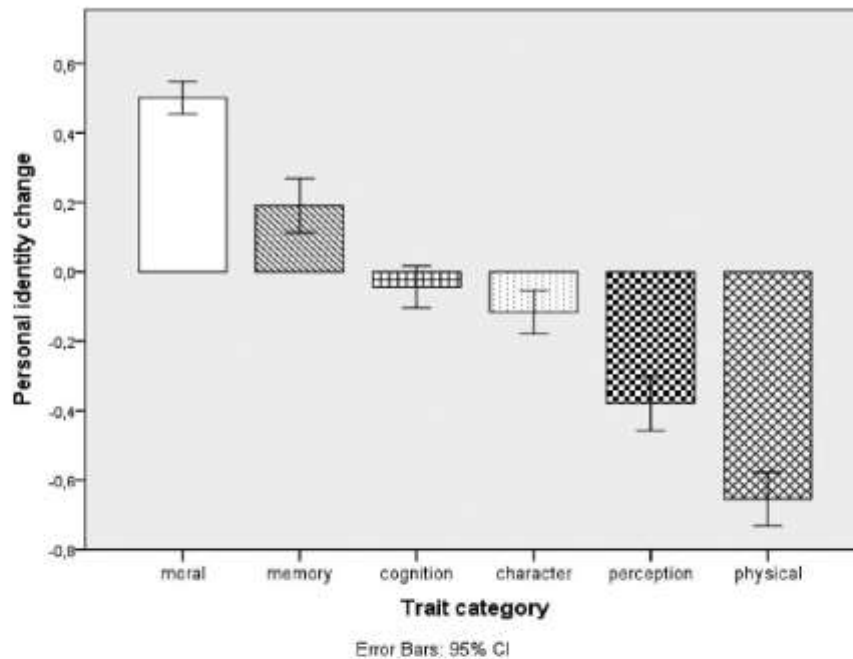
In the exploratory part of our analysis, we added one more test. We used Welch's two-sample two-sided t-tests (Holm-correction-applied) in order to explore the differences between the relative importance of the moral category in comparison to the other categories: in-between age categories, scenario categories, and sexes.

## 3. Results

Prior to all data analyses, we calculated the z-score<sup>4</sup> for each participant to control for between-subject differences in rating and to transfer the ordinal scale to a continuous scale. We set the alpha level to 0.05 and applied Holm-correction for all multiple tests.

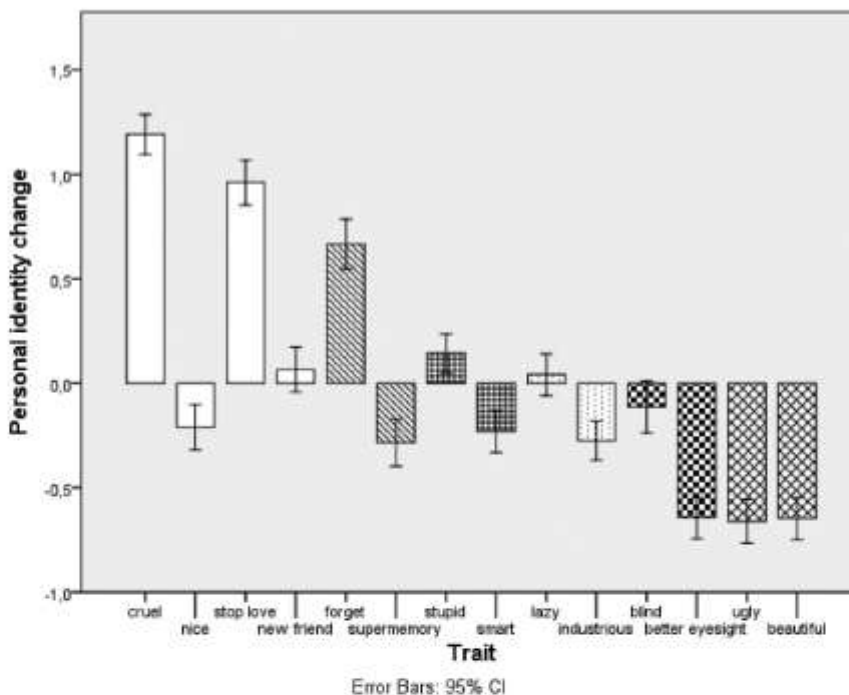
We performed five one-sided paired t-tests to test hypothesis 1 concerning differences between the perceived importance of change in the moral category, and changes in each of the five other categories: physical, cognitive, memory, character, and perception. Data analyses showed that respondents considered a change in the moral traits ( $M = 0.5$ ,  $SD = 0.35$ ) to have a significantly greater impact on personal identity than a change in any other category of traits, especially physical one ( $M = -0.65$ ,  $SD = 0.57$ ,  $t(216) = 21.91$ ,  $p < 0.0001$ ), then perception ( $M = -0.38$ ,  $SD = 0.58$ ,  $t(216) = 16.2$ ,  $p < 0.0001$ ), character ( $M = -0.12$ ,  $SD = 0.46$ ,  $t(216) = 15.09$ ,  $p < 0.0001$ ), cognition ( $M = -0.04$ ,  $SD = 0.45$ ,  $t(216) = 13.15$ ,  $p < 0.0001$ ), and, finally, memory ( $M = 0.19$ ,  $SD = 0.59$ ,  $t(216) = 6.24$ ,  $p < 0.0001$ ) (see Figure 1 and also Figure 1 in the Appendix).





**Figure 1.** Comparison of the perceived change in personal identity in different categories of traits. Respondents considered a change in the moral traits to have a significantly greater impact on personal identity than a change in any other category.

We further performed six one-sided paired *t*-tests to test hypothesis 2 concerning differences between the perceived importance of negative and positive versions of a change in each of the six categories. The results show that a negative change was perceived as having a significantly greater impact on the preservation of personal identity than a positive change in all categories, except the physical. The most salient difference shows itself in the rating of the relevance of the change in the moral category (positive:  $M = -0.07$ ,  $SD = 0.56$ , negative:  $M = 1.08$ ,  $SD = 0.55$ ,  $t(216) = 19.66$ ,  $p < 0.0001$ ). Also, super-memory gain ( $M = -0.28$ ,  $SD = 0.85$ ) and the corresponding memory loss ( $M = 0.67$ ,  $SD = 0.89$ ) were rated very differently, with memory loss scoring significantly higher ( $t(216) = 10.91$ ,  $p < 0.0001$ ). On the other hand, when it comes to a change in physical appearance, negative and positive versions (becomes uglier:  $M = -0.66$ ,  $SD = 0.78$  vs. becomes more beautiful:  $M = -0.65$ ,  $SD = 0.75$ ) received an almost equal rating, and, thus, the result was not anywhere near significant ( $t(216) = -0.19$ ,  $p = 0.57$ ) (see Figure 2 and also Figure 2 in the Appendix).



**Figure 2.** Differences between the perceived importance of negative and positive versions of a change.

The negative change was perceived as having a significantly greater impact on the preservation of personal identity than a positive change in all categories, except physical.

To test hypothesis 3, we analyzed a subset of data on negative versions of changes in each category and performed five one-sided paired t-tests to test the difference between the perceived importance of a negative moral change and a negative change in all the other categories. Analysis shows that a negative moral change ( $M = 1.08$ ,  $SD = 0.55$ ) was rated as significantly more important than a negative change in any other category (in all cases,  $p < 0.0001$ ).

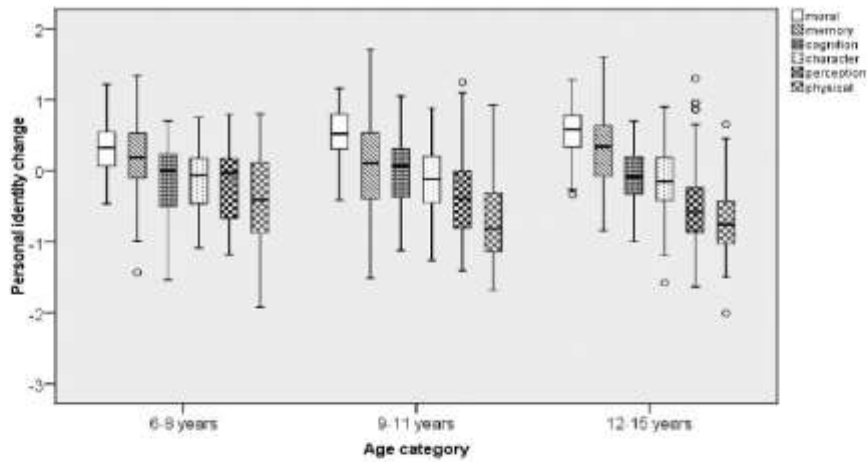
As part of additional exploratory research, we also tested the difference between age groups of 6–8, 9–11, and 12–15 years (according to standard growth stages, see, e.g., age periods in Gibbs, 2014, p. 74), sexes, and scenarios. As planned in preregistration, we performed the ANCOVA test with 'change in personal identity' as the output variable and the direction, category, sex, age, and direction-category interaction as predictors. The effect of the category ( $F(5, 3022) = 156.26$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.205$ ), direction ( $F(1, 3022) = 344.57$ ,  $p < 0.0001$ ,  $\eta_p^2 = 0.102$ ), and category-direction

interaction ( $F(5, 3022) = 42.61, p < 0.0001, \eta_p^2 = 0.066$ ) proved to be significant.

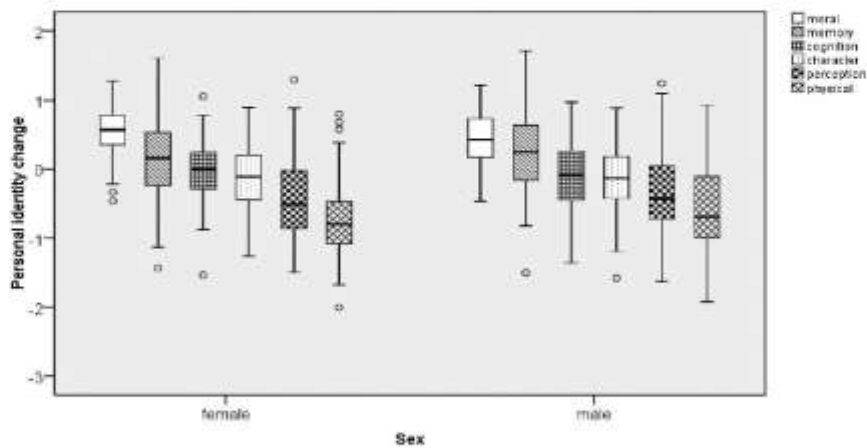
In the resulting graph of the age categories, we observe that there is a change in the relative importance of the moral category in comparison to the other categories of change across age, that is, the relative importance of the moral category grows with age. Thus, we concluded that comparing the values of scores of the same categories between age groups would not be the right way to approach the data. Instead, we decided to focus on comparing differences in the scores of the moral and non-moral categories (i.e., the mean score of the moral category minus the mean of the scores of all the other categories). Then, we applied Welch's two-sample two-sided t-tests (Holm-correction-applied) to test the statistical significance of the difference between the relative importance of the moral category in comparison to the other categories in different age groups, which proved to be significant between the 1<sup>st</sup> age category and the 2<sup>nd</sup> age category ( $p = 0.002$ ), and the 1<sup>st</sup> age category and the 3<sup>rd</sup> age category ( $p < 0.001$ ). It thus seems that the relative importance of the moral category grows with age (see Figure 3). The most prominent change seems to take place around the age of nine, and it is especially obvious with regard to the contrast between the moral and the physical category, while memory remains close to the moral category (see Figure 3–7 in the Appendix).

We also applied the same test in order to explore differences between sexes. The relative importance of the moral category was higher in the case of female respondents ( $p = 0.004$ ). Girls considered the moral category as having a greater impact on the preservation of personal identity in comparison to the other categories than did boys (see Figure 4). The most prominent difference shows itself in the contrast between the importance of the moral versus the physical category, while memory stays very close to the moral category (see Figure 8–12 in the Appendix).

We also tested this effect in the case of scenarios. Again, Welch's two-sample two-sided t-tests (Holm-correction-applied) revealed that there is a statistically significant difference in the relative importance of the moral category in comparison to the other categories between the 'friend' scenario and the 'some person in general scenario' ( $p < 0.001$ ). The more personal the scenario, the greater the relative importance of the moral category (see Figure 5). This effect seems to be much more prominent in the two older age categories, and, again, it seems to be especially obvious in the contrast between the moral and the physical category, while memory remains close to the moral category (see Figure 13–15 in the Appendix).

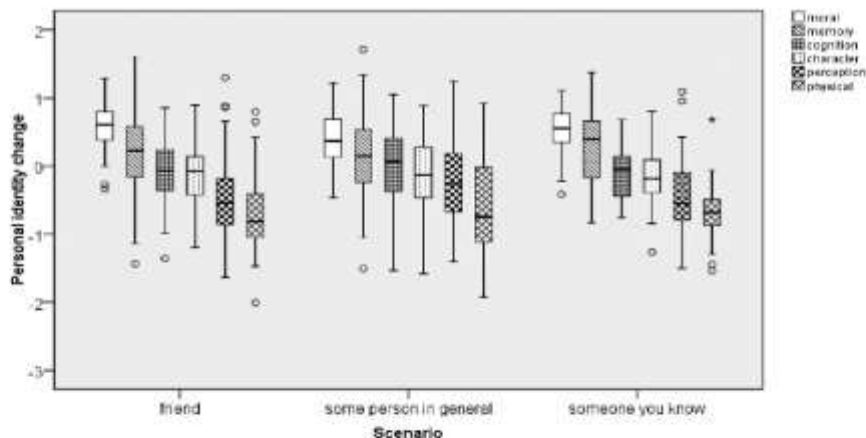


**Figure 3.** The effect of age on the perceived change in personal identity. The relative importance of the moral category grows with age. The bottom end of the whisker is minimum value. The top end of the whisker is maximum value. The bottom of the box is 25th percentile. The top of the box is 75th percentile. The horizontal line inside the box is median. The dots represent outliers.



**Figure 4.** The effect of sex on the perceived change in personal identity. Girls considered the moral category as having a greater impact on the preservation of personal identity in comparison to the other categories than boys. The bottom end of the whisker is minimum value. The top end of the whisker is maximum value. The bottom of the box is 25th percentile. The top of the box is 75th percentile. The horizontal line inside the box is median. The dots represent outliers.





**Figure 5.** The effect of scenario on the perceived change in personal identity. The more personal the scenario, the greater the relative importance of the moral category. The bottom end of the whisker is minimum value. The top end of the whisker is maximum value. The bottom of the box is 25th percentile. The top of the box is 75th percentile. The horizontal line inside the box is median. The dots represent outliers.

#### 4. Discussion

Our results support the view that moral continuity is crucial for the preservation of personal identity and that the ‘essential moral self’ and the ‘true self’ hypotheses go in the right direction. The preference of moral traits in the concept of the self was already present in children and teenagers. Moral traits, or, more precisely, traits that have an important role in the social context, proved again to be central to the folk concept of personal identity.

Personal identity was perceived to be broken especially when a negative moral change had occurred. This could be explained by the fact that the questionnaire was based on respondents’ intuitive assessment. This means that the answers are not strictly rational – that is, they are not a result of moral reasoning accompanied by deeper reflection. They are also, and mainly, based on emotional (quick, automatic, and intuitive) evaluations. This is consistent with the social intuitionist model (Haidt, 2001). The relatively greater importance of the negative change in comparison to the positive change means that the weight of the negative feeling in case of a negative change is more significant than the weight of the positive feeling related to a positive change. This phenomenon could potentially be explained by a feeling of danger, as a negative change of moral traits is associated with asocial and selfish behavior and could thus lead to social conflict.

This leads us to a further suggestion: that the phenomenon of a negative change having greater importance may also have a social meaning. This would support our view that the folk concept of the self is socially determined. Moral traits are the traits that hold society together. A person who is evaluated as “bad” displays a less pronounced tendency to integrate into society. Their interests may even go against the interests of society and be destructive to it. Behavior that is motivated by such interests is generally perceived negatively by the society in which the children live. This indicates that the folk concept of the self operates with an idea of the self that is substantially pro-socially oriented, and this orientation may be observed even in the case of children. This theory is also in accordance with evolutionary theories of morality (Haidt, 2007; Keefer, 2013) and the idea that our concepts stem naturally from our need to be able to orient ourselves within social interactions and maintain good social bonds.

The social explanation could be deepened by considering the points made by psychological essentialism. A natural tendency to essentialize entities based on their hidden traits might have the above-described benefit of maintaining good social bonds, especially when the entity in question (in this case, a person or a group of people) is viewed as inherently good. This positive effect of the ‘true self’ concept on intergroup relationships was also demonstrated experimentally (De Freitas & Cikara, 2018). This might be related to the normative and teleological nature of essentialism. If people indeed see essences in the context of purposes (Rose & Nichols, 2019) and these purposes are normatively good (De Freitas et al., 2017, p. 397), then it makes perfect sense that a human being should be viewed as essentially moral and well disposed toward others, and thus capable of forming good interpersonal relationships. A well-functioning society is, after all, naturally desirable for any social species. A person is seen as losing his or her identity in the case of negative change because they have diverted from their true self. Not being recognized as the same person might even be perceived as a kind of social punishment.

The tendency of the respondents to judge moral change as causing a more radical disruption in personal identity when the person undergoing the change is a friend further accentuates the proposed view. Friends are persons to whom we are emotionally related and who have a high position in the hierarchy of human relationships that reflects patterns in human altruistic behavior (Ma, 2013, p. 3). A change in a friend’s moral attitudes potentially leads to the loss of this mutually valued relationship, which is accompanied by unpleasant feelings. We propose that these feelings cause relatively greater sensitivity with regard to a change in a friend’s moral traits, and thus, they bring about a more radical judgment concerning the rupture in personal identity. Our results support the theory that emphasizes the substantiality of moral traits, the *essential moral self*, together with theories that do justice to the social aspect of personal identity.



Data also reflect the relative importance of autobiographical memory. Our findings are in agreement with the previous conclusions reached by researchers within this field (Bluck et al., 2010; Demiray & Janssen, 2015; Wilson & Ross, 2003). The difference, however, lies in our conclusion that the importance of moral traits is relatively greater than that of autobiographical memory, which is in agreement with findings from the earlier studies already mentioned in Section 1.1. of this paper. We believe that this does not contradict the more widespread theory that emphasizes the significance of autobiographical memory for the establishment of personal identity. However, based on both previous research and our own research, we may conclude that moral traits play an even more essential role here. It could also be the case that memories of life experiences are closely connected to personal memories of interpersonal relationships, which again supports theories which put forward the social aspect of the folk concept of personal identity.

In deep contrast to moral traits, physical appearance played the least important role in the children's concept of personal identity, which again highly favors the claims that psychological traits surpass physical traits when it comes to personal identity intuitions.

The observed effect of age can be a result of a specific aspect of moral development as it is described in the theory of the relationship between the moral self and moral identity by Kingsford and colleagues (Kingsford et al., 2018). In early childhood, children gain a concept of themselves as moral persons (a 'moral self') based on repeated experiences of their conduct in morally colored social situations (p. 656). However, this concept or self-knowledge lacks certain generalizing and evaluative dimensions: the children are able to recognize which types of behavior are consistent with their own, but the step toward a general awareness of what kind of moral person they are doesn't yet take place (p. 657). This type of self-knowledge thus differs from the moral identity of adolescents and adults, where this higher-order reflection leads to a more consistent and evaluative moral self-image (p. 658). This crucial difference shows itself in attempts to define what makes an action moral. From the age of eight onward, children's answers point to a *second-order desire* to pay justice to the kind of moral person they are; by contrast, younger children aged five to seven years old demonstrate a mere acknowledging and respecting of moral norms (pp. 659–660). Gaining the ability to be "genuinely and independently self-evaluative" requires cognitive capabilities that usually do not appear before the ascent of middle childhood (eight to 12 years). Children in their middle childhood also start to be able to compare their specific personal traits to others, especially their peers (p. 661). It is thus possible that older children understand moral traits as being a much more integrative part of personal identity, and they understand the concept of the coherent moral self much

better than younger children do. This explains the specific effect between the youngest age category (6–8) and the middle age category (9–12).

As social beings, children are constantly under the influence of their parents and teachers, who instill in them the moral values of their community and present them with their idea of correct social conduct. Children are repeatedly told that it is not right to judge people by the way they look, that it is important to treat other people right, to not hurt them, and to be polite and nice. Even though authorities play a crucial role in the development of the moral self, children are active participants of this process, and their responses to the parental influence are domain-specific (Smetana & Jambon, 2017, pp. 133–134). Our results suggest that despite the fact that children are also being reprimanded for their laziness, this trait does not seem to have a significant impact on personal identity in their view. Therefore, some natural appreciation of moral traits seems to play an important role, perhaps due to the already mentioned teleological essentialist tendencies.

The observed effect of sex may be related to the age effect, since girls are usually developmentally ahead of boys. It may also be explained via the tendency of girls to ascribe greater importance to social relationships than boys, which is partially encouraged by an upbringing that is based on certain cultural stereotypes about sexes (Eagly & Wood, 2017).

In conclusion, children judged the true self of a person to be well disposed toward other people and inherently good. A change in the direction toward anti-social behavior or a loss of memory bonds to others was considered as much more identity-breaking than a change in any of the other categories. Our study, with its respondents being from a central European country, further supports the view that the 'true self' concept is cross-culturally stable on a more abstract level, even though evaluation of particular acts as good or bad may vary across cultures to some extent (Strohmingier et al., 2017, p. 554).

We are aware that our research has certain limitations. Even though we mixed both the categories and the positive and negative versions of the questions so that children couldn't easily register our focus on the moral items, the order of the items was fixed. This could have had an effect on the answers. Also, the range of traits used in our vignette is considerably narrow; due to children's short attention spans, we needed the interview to be brief. The children's responses to the interviewer – a stranger and an adult figure – could also have played a role in their tendency to answer in a certain way (trying to meet the adult's expectations, perhaps). However, we used simple language in order to make sure that the children understood the questions and observed their feedback. In several cases, the children even spontaneously revealed the reasoning behind their answers and thus showed to what extent the question resonated with their intuitions (for more details, see data).

In future research, it would be apposite to widen the range of categories and the number of items in each category. It would also be convenient to randomize



the order of the questions properly and use a between-subject model to control for the possible order-effects. A much larger study, with more participants, more interviewers, and the use of a computer randomizer, would be needed to ensure that the effects we reported are indeed present and apply more generally. Confirmation of our findings concerning the age and scenario effects would require testing our exploratory questions again, while formulating hypotheses in advance and running confirmatory research. However, we believe that each study only reveals the phenomenon from a certain specific point of view, and we need a battery of similar and, at the same time, slightly varied studies to hint at a wider picture. Being able to turn to the large number of existing results, we nevertheless dare to assert that the proposed existence of the concepts such as the 'essential moral self' and the 'true self' has already gained very robust support.

### Notes

1. "Lay concept of personal identity and the importance of moral and interpersonal traits" preregistered on 8 June 2017. Link: <https://osf.io/qj7k6/>
2. IRB Charles University, Faculty of Science, approval number: 2017/14
3. Parallel research with another interviewer on the topic of trolley problems took place nearby at the family event. Respondents (age-range 6–18 years) were encouraged but not obliged to participate in both studies.
4. Alternative version of the analyses without computing the z-scores is available here: <https://osf.io/qfvu3/> Graphs showing the original scores without any transformations are available in the Appendix.

### Acknowledgement

MJK and RK developed the study concept and were responsible for the study design and data collection. MJK and JF were responsible for the study protocol and its preregistration. MN performed the data analyses. MJK, RK, PO and MN interpreted the results. MJK and PO drafted and all the authors revised the manuscript. All the authors approved the final version of the manuscript for submission.

### Disclosure statement

No potential conflict of interest was reported by the authors.

### Funding

This work was supported by the Charles University Grant Agency [GA UK 925416].

### Notes on contributors

*Michaela Jirout Košová* is currently completing her PhD studies at the Faculty of Science, Charles University. Her domain of interest is philosophy of mind and experimental philosophy. Throughout her studies she has published three papers in Czech philosophical journals (as a single author) and one paper in *Czechoslovak Psychology* (as a co-author). She has been conducting experimental-philosophy research at Charles University. She is also a member of the Karel Čapek Center for Values in Science and Technology.

*Robin Kopecký* is a PhD candidate at the Faculty of Science, Charles University. His domain of interest is experimental philosophy and evolutionary biology. He is a member of the laboratory of evolutionary biology and conducts research in evolutionary psychology and experimental philosophy at Charles University. His paper on ethics and autonomous vehicles has been published by *The Philosophical Journal* in the Czech Republic and he has also published a paper in *Czechoslovak Psychology* (as a co-author). He is a coordinator and a member of the Karel Čapek Center for Values in Science and Technology.

*Pavel Oulovský* is a PhD candidate at the Faculty of Science, Charles University. He focuses on depth psychology and existential psychology. He is a member of the Czech Daseinsanalytische Society and Austrian Daseinsanalytische Institute. As a team member of a grant project at Charles University he analysed psychological diagnostic tools for interpersonal behaviour. His paper (co-authored by Jirout Košová and Kopecký) on psychometric properties of the Circumplex Scales of Interpersonal Values in the Czech population has been published by a renowned Czech psychology journal *Czechoslovak Psychology*.

*Matěj Nekvindá* earned his MA at the Faculty of Mathematics and Physics, Charles University. His main domain of interest is statistics in natural sciences.

*Jaroslav Flegr* is a tenured professor of biology at the Faculty of Science, Charles University, a researcher at the National Institute of Mental Health in the Czech Republic, and a member of the Karel Čapek Center for Values in Science and Technology. He specialises in parasitology, evolutionary biology and biological psychiatry. In addition to some 150 scholarly papers, he has also published four monographs.

### Data Availability

<https://osf.io/cjuf9>

### References

- Bluck, S., Alea, N., & Demiray, B. (2010). You get what you need: The psychosocial functions of remembering. In J. H. Mace (Ed.), *The act of remembering: Toward an understanding of how we recall the past* (pp. 284–307). Wiley-Blackwell.
- De Freitas, J., & Cikara, M. (2018). Deep down my enemy is good: Thinking about the true self reduces intergroup bias. *Journal of Experimental Social Psychology*, 74, 307–316. <https://doi.org/10.1016/j.jesp.2017.10.006>
- De Freitas, J., Cikara, M., Grossmann, I., & Schlegel, R. (2018). Moral goodness is the essence of personal identity. *Trends in Cognitive Sciences*, 22(9), 739–740. Letter. <https://doi.org/10.1016/j.tics.2018.05.006>
- De Freitas, J., Sarkissian, H., Newman, G. E., Grossman, I., De Brigard, F., Luco, A., & Knobe, J. (2018). Consistent belief in a good true self in misanthropes and three

- interdependent cultures. *Cognitive Science*, 42(S1), 134–160. <https://doi.org/10.1111/cogs.12505>
- De Freitas, J., Tobia, K., Newman, J. E., & Knobe, J. (2017). Normative judgements and individual essence. *Cognitive Science*, 41(S3), 382–402. <https://doi.org/10.1111/cogs.12364>
- Demiray, B., & Janssen, S. M. J. (2015). The self-enhancement function of autobiographical memory. *Applied Cognitive Psychology*, 29(1), 49–60. <https://doi.org/10.1002/acp.3074>
- Dranseika, V. (2017). On the ambiguity of 'the same person'. *AJOB Neuroscience*, 8(3), 184–186. <https://doi.org/10.1080/21507740.2017.1366588>
- Eagly, A., & Wood, W. (2017). Gender identity: Nature and nurture working together. *Evolutionary Studies in Imaginative Culture*, 1(1), 59–62. <https://doi.org/10.26613/esic/1.1.10>
- Gibbs, J. C. (2014). *Moral development and reality: Beyond the theories of Kohlberg, Hoffman, and Haidt* (3rd ed.). Oxford University Press.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834. <https://doi.org/10.1037/0033-295X.108.4.814>
- Haidt, J. (2007). The New Synthesis in Moral Psychology. *Science*, 316(5827), 998–1002. <https://doi.org/10.1126/science.1137651>
- Heiphetz, L., Strohminger, N., Gelman, S., & Young, L. (2018). Who am I? The role of moral beliefs in children's and adults' understanding of identity. *Journal of Experimental Social Psychology*, 78, 210–219. <https://doi.org/10.1016/j.jesp.2018.03.007>
- Heiphetz, L., Strohminger, N., & Young, L. L. (2017). The role of moral beliefs, memories, and preferences in representations of reality. *Cognitive Science*, 41(3), 744–767. <https://doi.org/10.1111/cogs.12354>
- Keefer, W. M. (2013). Understanding morality from an evolutionary perspective: Challenges and opportunities. *Educational Theory*, 63(2), 113–131. <https://doi.org/10.1111/edth.12013>
- Kingsford, J. M., Hawes, D. J., & de Rosnay, M. (2018). The moral self and moral identity: Developmental questions and conceptual challenges. *British Journal of Developmental Psychology*, 36(4), 652–666. <https://doi.org/10.1111/bjdp.12260>
- Locke, J. (2009). *Essay concerning human understanding*. WLC books. (Original work published 1690).
- Ma, H. K. (2013). The moral development of the child: An integrated model. *Frontiers in Public Health*, 1, 57. <https://doi.org/10.3389/fpubh.2013.00057>
- Nichols, S., & Bruno, M. (2010). Intuitions about personal identity: An empirical study. *Philosophical Psychology*, 23(3), 293–312. <https://doi.org/10.1080/09515089.2010.490939>
- Okawa, J. B. (2008). Considerations for the cross-cultural evaluation of refugees and asylum seekers. In L. A. Suzuki & J. G. Ponterotto (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (Third ed.) (pp. 165–194). John Wiley & Sons, Inc.
- Prinz, J., & Nichols, S. (2016). Diachronic identity and the moral self. In J. Kiverstein (Ed.), *The Routledge handbook of philosophy of the social mind* (pp. 449–463). Routledge.
- Rose, D., & Nichols, S. (2019). Teleological Essentialism. *Cognitive Science*, 43(4), e12725. <https://doi.org/10.1111/cogs.12725>
- Smetana, J. G., & Jambon, M. (2017). Parenting, morality and social development. In C. C. Helwig (Ed.), *New perspectives on moral development* (pp. 121–139). Routledge.
- Starmans, C., & Bloom, P. (2018). Nothing personal: What psychologists get wrong about identity. *Trends in Cognitive Sciences*, 22(7), 566–568. <https://doi.org/10.1016/j.tics.2018.04.002>
- Strohminger, N., Knobe, J., & Newman, G. (2017). The true self: A psychological concept distinct from the self. *Perspectives on Psychological Science*, 12(4), 551–560. <https://doi.org/10.1177/1745691616689495>



- Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition*, *131*(1), 159–171. <https://doi.org/10.1016/j.cognition.2013.12.005>
- Strohminger, N., & Nichols, S. (2015). Neurodegeneration and Identity. *Psychological Science*, *26*(9), 1469–1479. <https://doi.org/10.1177/0956797615592381>
- Tobia, K. P. (2015). Personal identity and the Phineas Gage effect. *Analysis*, *3*(75), 396–405. <https://doi.org/10.1093/analys/any041>
- Tobia, K. P. (2016). Personal identity, direction of change, and neuroethics. *Neuroethics*, *9*(37), 37–43. <https://doi.org/10.1007/s12152-016-9248-9>
- Wilson, A., & Ross, M. (2003). The identity function of autobiographical memory: Time is on our side. *Memory*, *11*(2), 137–149. <https://doi.org/10.1080/741938210>

## Appendix

Představ si, že tvůj kamarád/kamarádka, spolužák, rovesník nebo někdo jiný/nějaký člověk vstoupí do speciální sci-fi komory, která ho může jakkoli změnit. Jak moc podle tebe tyto možné změny zasáhnou jeho podstatu, to, co je na něm to nejdůležitější, to, kým je, to, co ho nejvíc vystihuje jako človíčka?

**0 – Je to pořád on/ona, zachoval/a si to podstatné ze sebe.**

**6 – Už to není on/ona, ztratil/a to podstatné ze sebe.**

1. Když se stane ošklivějším.
2. Když si najednou dokáže vybavit úplně všechny své vzpomínky.
3. Když přestane mít rád své kamarády a blízké.
4. Když se stane svědomitějším, pilnějším.
5. Když oslepně.
6. Když se stane krásnějším.
7. Když zapomene na všechny svoje zážitky.
8. Když se stane mnohem hodnějším k ostatním dětem.
9. Když zhlopne.
10. Když se začne přátelit s někým, koho vůbec neměl rád.
11. Když sa stane línějším.
12. Když se mu velice zlepší zrak.
13. Když se stane krutým k ostatním dětem.
14. Když zmoudří.

SCALE:

0

1

2

3

4

5

6

Je to pořád on/ona

Už to není on/ona

Imagine that *your friend/friend, schoolmate, peer or someone else/some person* enters a special sci-fi chamber that can change them in any way. What do you think, how much will the following changes affect the person's core, that, what is most important about them, what makes them who they are, what is most characteristic of them as a person?

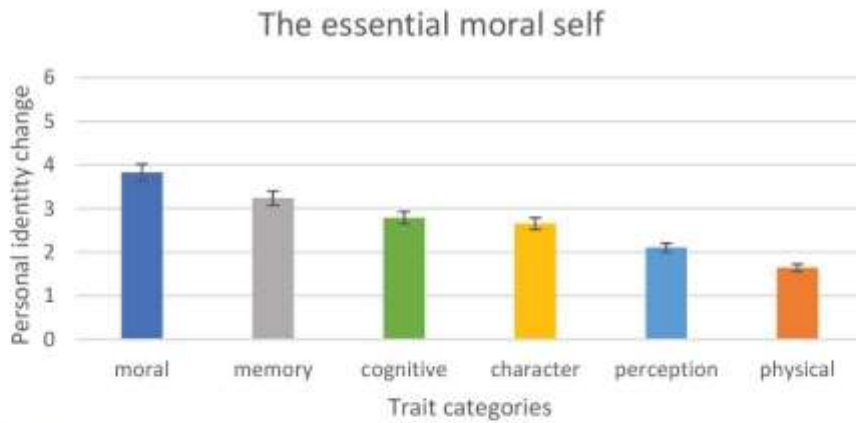
**0 - The person is still him/her, he/she has kept the most important part of them.**

**6 - The person is not him/her anymore, he/she has lost the most important part of them.**

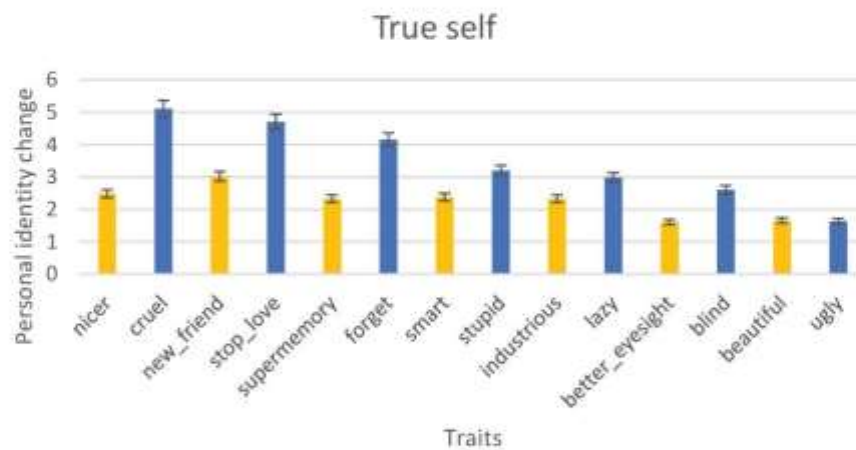
1. If they become uglier.
2. If they are suddenly able to perfectly remember all their memories.
3. If they stop loving their friends and close ones.
4. If they become more industrious.
5. If they become blind.
6. If they become more beautiful (physically).
7. If they forget all their life memories.
8. If they become nicer to other kids.
9. If they become more stupid.
10. If they accept as a friend someone they didn't like before.
11. If they become lazier.
12. If their eyesight improves radically.
13. If they start to be cruel to other people.
14. If they become smarter.

SCALE:

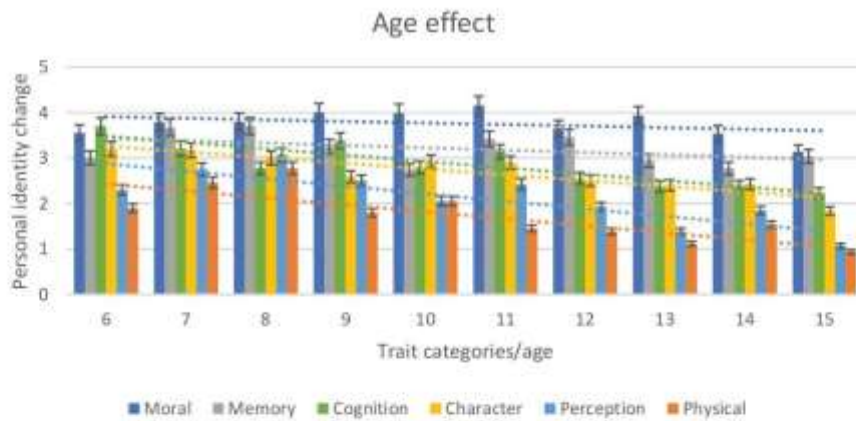




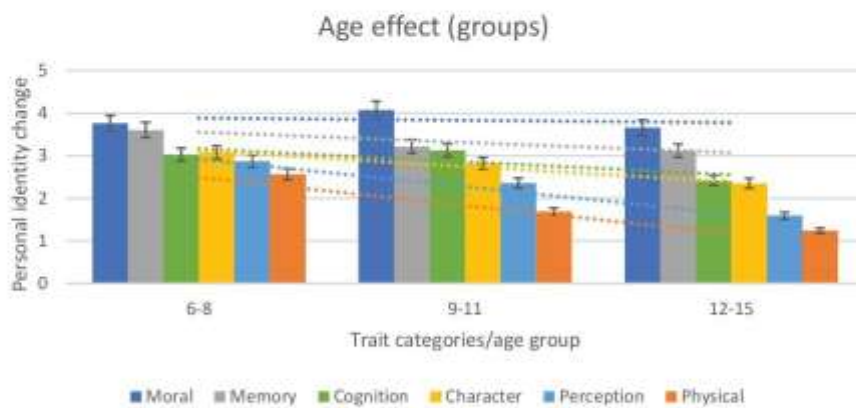
**Figure 1.** Error bars: 95% CI.  
 The graph shows the mean values of the original scores of the perceived change of personal identity for all participants (N=217, 56,4% F) (y-axis) as they differ across six categories of change (x-axis).



**Figure 2.** Error bars: 95% CI.  
 The graph shows the mean values of the original scores of the perceived change of personal identity for all participants (N=217, 56,4% F) (y-axis) as they differ across all items of change and between the positive version (yellow) and the negative version (blue) of each changing trait (x-axis).

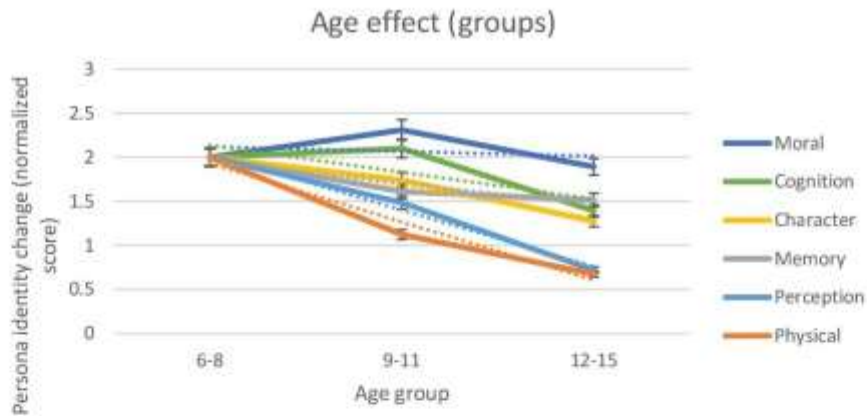


**Figure 3.** Error bars: 95% CI.  
 The graph shows the mean values of the original scores of the perceived change of personal identity (y-axis) as they differ across six categories of change and across different ages (x-axis). The dotted lines show the trait categories' linear trend of the perceived personal identity change score with growing age.

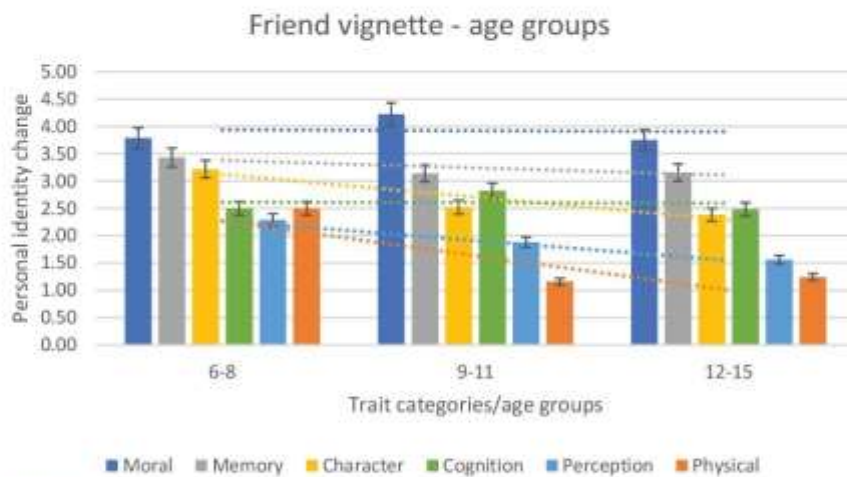


**Figure 4.** Error bars: 95% CI.  
 The graph shows the mean values of the original scores of the perceived change of personal identity (y-axis) as they differ across six categories of change and across different age categories (x-axis). The dotted lines show the trait categories' linear trend of the perceived personal identity change score with growing age.

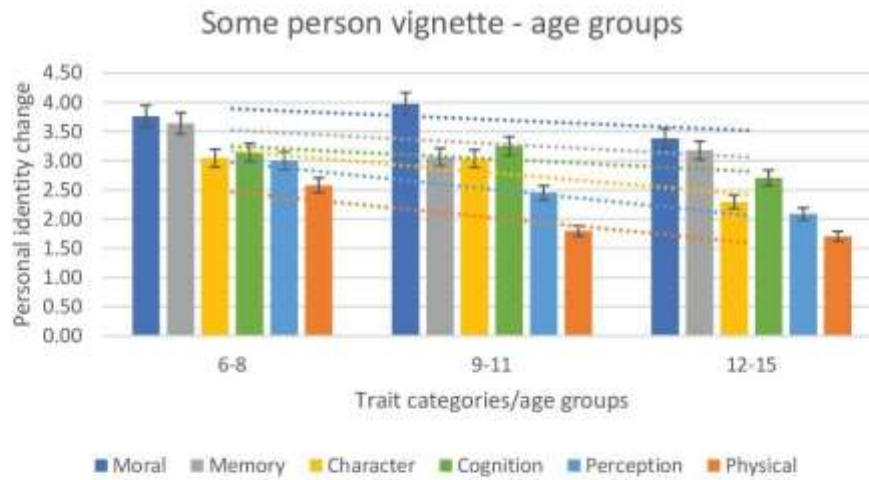




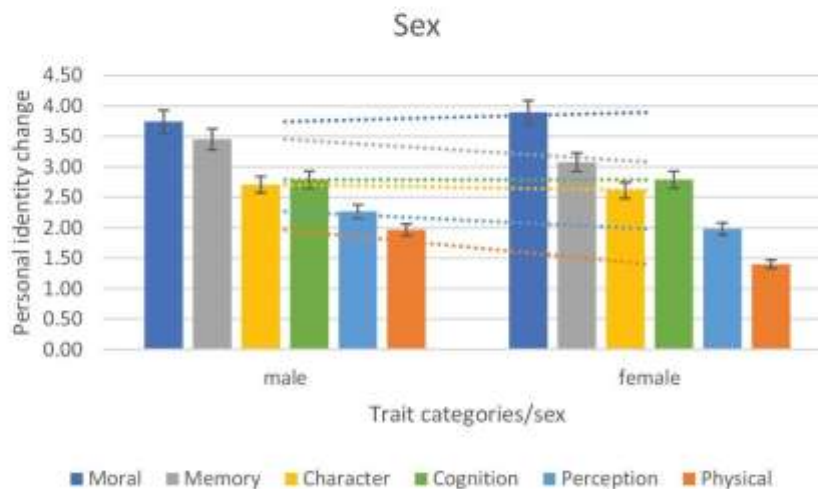
**Figure 5.** Error bars: 95% CI. The graph shows normalized scores (so that they start at 2 for the lowest age) of the perceived change of personal identity (y-axis) as they differ across six categories of change and across different age categories (x-axis). The dotted lines show the trait categories' linear trend of the perceived personal identity change score with growing age.



**Figure 6.** Error bars: 95% CI. The graph shows the mean values of the original scores of the perceived change of personal identity for the participants answering the "friend vignette" (N=90, 64,4% F) (y-axis) as they differ across six categories of change and across different age categories (x-axis). The dotted lines show the trait categories' linear trend of the perceived personal identity change score with growing age.



**Figure 7.** Error bars: 95% CI.  
 The graph shows the mean values of the original scores of the perceived change of personal identity for the participants answering the “some person in general” vignette (N=91, 42,9% F) (y-axis) as they differ across six categories of change and across different age categories (x-axis). The dotted lines show the trait categories’ linear trend of the perceived personal identity change score with growing age.



**Figure 8.** Error bars: 95% CI.  
 The graph shows the mean values of the original scores of the perceived change of personal identity for all participants (y-axis) as they differ across six categories of change and between sexes (x-axis). The dotted lines show the trait categories’ linear trend of the perceived personal identity change score between sexes.

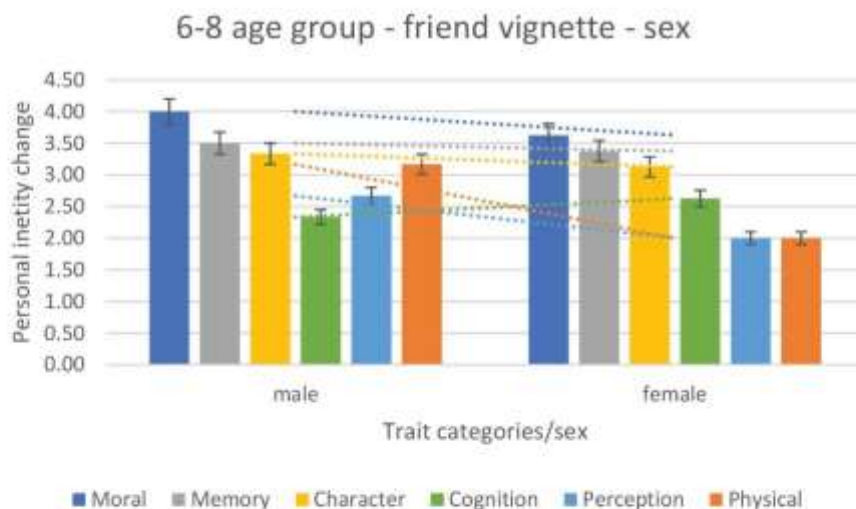


Figure 9. Error bars: 95% CI.

The graph shows the mean values of the original scores of the perceived change of personal identity for the 6–8 years old participants answering the “friend” vignette (note: the sample was too small,  $N=7$ , 57,1% F) (y-axis) as they differ across six categories of change and between sexes (x-axis). The dotted lines show the trait categories’ linear trend of the perceived personal identity change score between sexes.

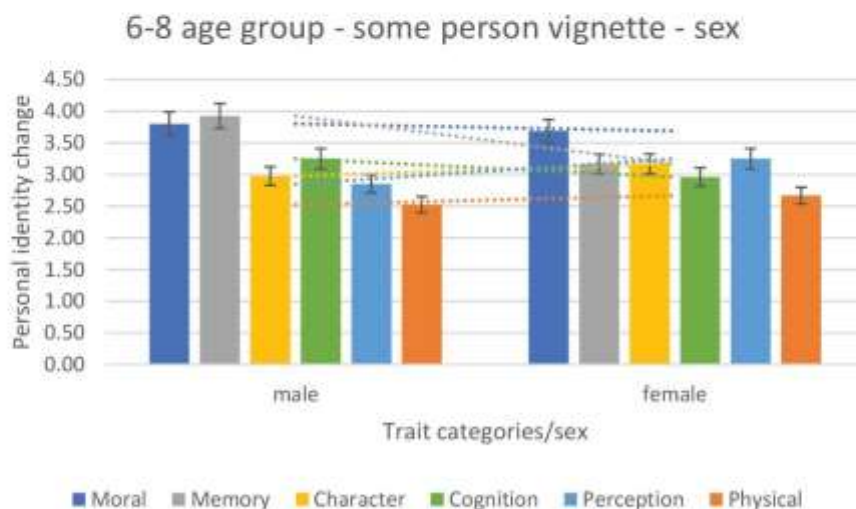
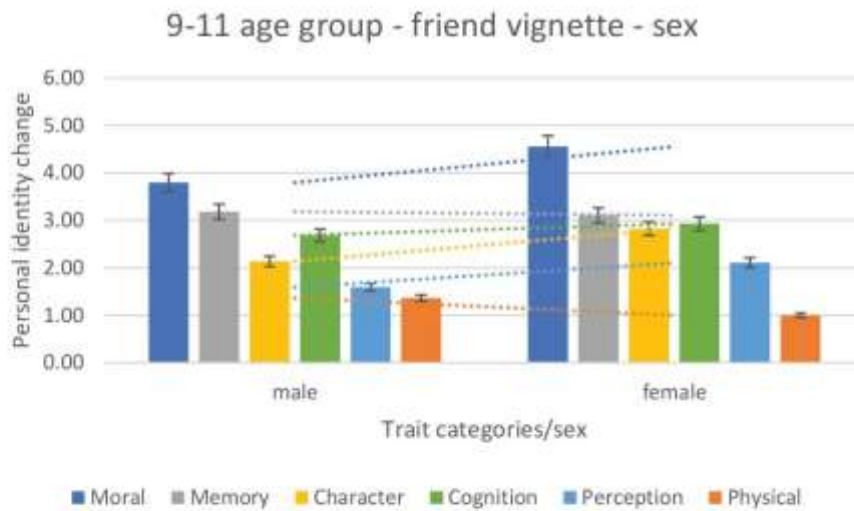
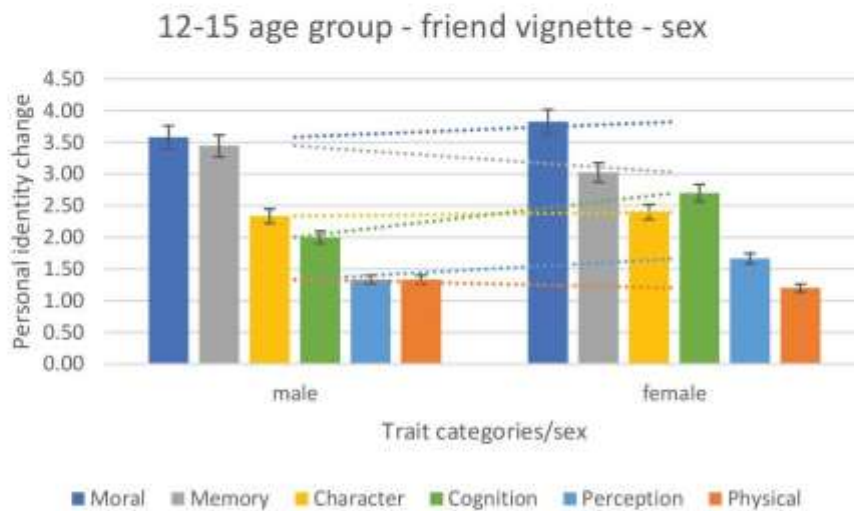


Figure 10. Error bars: 95% CI.

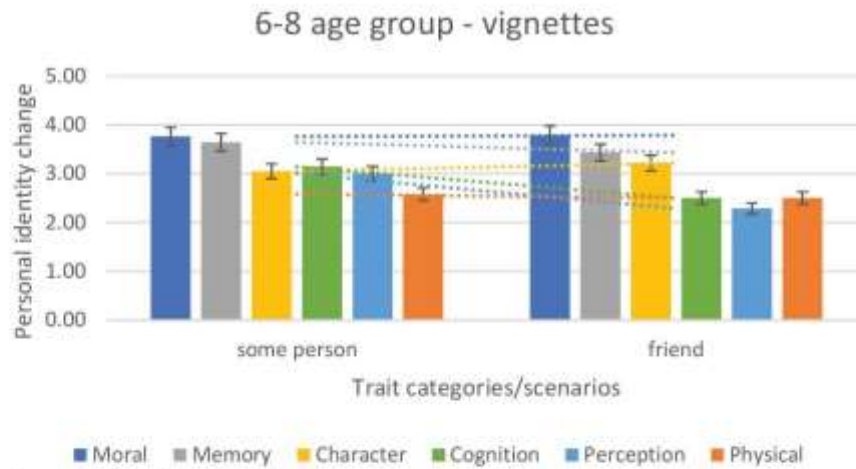
The graph shows the mean values of the original scores of the perceived change of personal identity for the 6–8 years old participants answering the “some person in general” vignette ( $N=32$ , 37,5% F) (y-axis) as they differ across six categories of change and between sexes (x-axis). The dotted lines show the trait categories’ linear trend of the perceived personal identity change score between sexes.



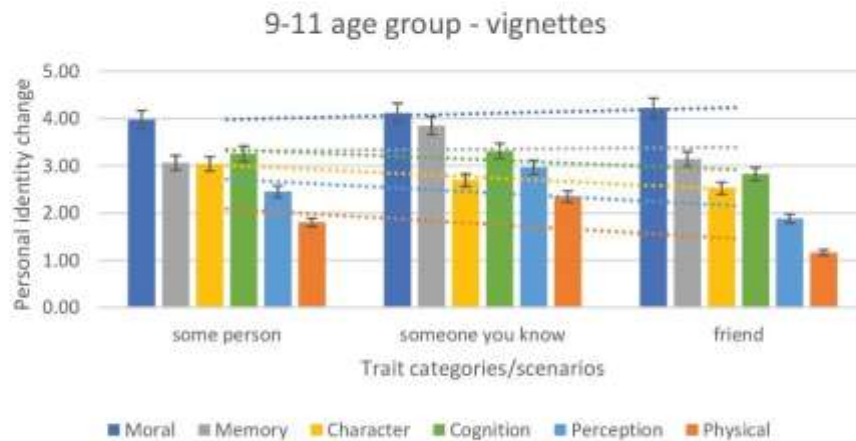
**Figure 11.** Error bars: 95% CI.  
 The graph shows the mean values of the original scores of the perceived change of personal identity for the 9-11 years old participants answering the “friend” vignette (N=25, 56% F) (y-axis) as they differ across six categories of change and between sexes (x-axis). The dotted lines show the trait categories’ linear trend of the perceived personal identity change score between sexes.



**Figure 12.** Error bars: 95% CI.  
 The graph shows the mean values of the original scores of the perceived change of personal identity for the 12-15 years old participants answering the “friend” vignette (N=58, 69% F) (y-axis) as they differ across six categories of change and between sexes (x-axis). The dotted lines show the trait categories’ linear trend of the perceived personal identity change score between sexes.

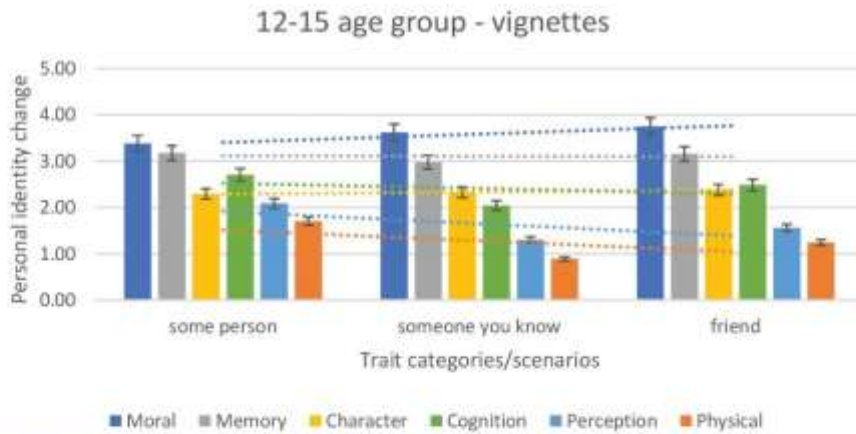


**Figure 13.** Error bars: 95% CI.  
 The graph shows the mean values of the original scores of the perceived change of personal identity for the 6–8 years old participants (N=39, 41% F) (y-axis) as they differ across six categories of change and across different vignettes (x-axis). The dotted lines show the trait categories' linear trend of the perceived personal identity change score with vignette change.



**Figure 14.** Error bars: 95% CI.  
 The graph shows the mean values of the original scores of the perceived change of personal identity for the 9–11 years old participants (N=80, 53,8% F) (y-axis) as they differ across six categories of change and across different vignettes (x-axis). The dotted lines show the trait categories' linear trend of the perceived personal identity change score with vignette change.





**Figure 15.** Error bars: 95% CI.  
 The graph shows the mean values of the original scores of the perceived change of personal identity for the 12–15 years old participants (N=98, 65,3% F) (y-axis) as they differ across six categories of change and across different vignettes (x-axis). The dotted lines show the trait categories' linear trend of the perceived personal identity change score with vignette change.