

## **Experimentální filosofie: hledání návrhu výzkumu v rámci lokálních universit**

Tomáš Zemčík

Vysoká škola baňská – Technická univerzita, CZ

ZEMČÍK, T.: Experimental Philosophy: Looking for A Research Proposal within Local Universities.

*Philosophica Critica*, vol. 2, 2016, no. 2, ISSN 1339-8970, pp. 71-86.

The study dealt with particular potencies of experimental philosophy as a method capable of introducing relevant opinions into protracted disputes not only between different philosophical approaches. The text was focused mainly on the conditions of local universities. The first part of the study introduced the work of the philosopher and psychologist Jonathan Haidt, who through the methods of experimental philosophy implemented a possible solution to the long-term dispute between deontological and sentimentalistic tradition in the area of practical philosophy. For the purpose of this work is experimental philosophy trivially defined as method, whose aim is to verify concrete philosophical standpoints through experiment on the consilience platform of disciplines such as psychology, philosophy or anthropology. The second part introduced three specific bases, areas and topics that could be implemented in local conditions through specific research in the field of experimental philosophy.

*Key words:* Experimental philosophy – Morale – Mind – Konsilience – Psychology – Research

### **Úvod**

Nedávno jsem se seznámil s knihou pojednávající o tématu, které je již delší dobu častým obsahem mých filosofických analýz. Tato kniha nese titul *Digitální demence* (Spitzer, 2014) a jejím autorem je univerzitní profesor, neurovědec, lékař, psycholog a filosof Manfred Spitzer. Spitzer (2014) čtenáři předkládá poměrně komplexní obraz toho, že digitální média a jejich nadužívání zapříčiňují takzvanou digitální demenci. Vedou k dramatickému poklesu schopnosti orientace a učení. Spitzer také ukazuje, že zapříčiňují poruchy čtení, spánku a pozor-

nosti. Mezi další symptomy digitální demence lze podle Spitzera řadit i osamělost, úzkost, otupělost, nadváhu, sklony k násilí a vůbec celkový úpadek společnosti. Témata, o kterých Spitzer pojednává, jsou krajně aktuální a je nezbytné, aby se jimi systematicky zabýval další nejenom základní, ale hlavně aplikovaný (také experimentálně filosofický) výzkum. Na druhou stranu je patrné, že Spitzer argumentuje dosti panicky, přičemž jednostranně vybírá a interpretuje zdroje. Domnívám se však, že chceme-li posoudit celkový status média, nezáleží pouze na jeho formě, ale i na obsahu a způsobu uchopení; s čímž by Spitzer sám dle mé míry v konečném důsledku souhlasil.

Touto tematicky aktuální knihou autor bezesporu znovu otvírá, prohlubuje a aktualizuje společensky zásadní motiv. Je jím Matrixový<sup>1</sup> leitmotiv úpadku lidského světa ve prospěch technologií. Nastolení všepohlcující virtuality vyživované digitálním světem informačních technologií. Počítače, technologické vychytávky, „udělátka a hejbátka“ nás, řečeno slovy klasika ultramoderny, Jeana Baudrillarda, milují. A právě díky tomuto erotickému vztahu se v prostředí systému těchto objektů cítíme tak naživu (Baudrillard 2005, 186). Člověk, který je životně odkázaný na svět digitálních technologií, zvláště pak ztrácí-li své kognitivní schopnosti, se může sice stále považovat za „pána živého tvorstva“, ale stává se také nevyhnutelně existenčně závislým na agendě vlastního výtvoru – pohybu v digitální sféře reality.

Nejedná se však o nic zcela bezprecedentního. Podobné úvahy byly již zprostředkovány a mnohokrát aktualizovány. Za zmínku stojí například Hegelův filosofický mýtus zabývající se dynamikou realizace lidského sebevědomí, který je znám jako příklad „pána a raba“. Připomeňme si, že obsahem této metafory je to, že pán (v tomto případě lidstvo) je odkázán na práci raba (v našem případě sféry digitální, kterou lze též Baudrillardovsky (2005) nazírat jako svébytný subjekt, reprezentovaný systémem objektů, jenž je v neustálém recipročním svůdném kontaktu s lidským subjektem, čímž je ustanoven symbolický řád reality). Pán díky tomu vztahu leniví a postupně se stává zcela nekompetentním. Ráb je zase díky těžkému a invenci vyžadujícímu životnímu poslání nucen se neustále zdokonalovat, až pána jednou zcela přeroste. Stav věcí poté projde jakousi singularitou, kde již není potřeba pána. Rab se zcela emancipuje v novém kvalitativně nesouměřitelném systému. Je však nutné si položit otázku, jestli je tento průměr namístě. Je to tak, že by člověk opravdu bezvýhradně stagnoval, přicházel o své kognitivní schopnosti, zakrňoval a degeneroval? Předává své dominium digitálnímu rabovi? Anebo je tomu spíše tak, že podobné závěry jsou důsledkem jednostranné interpretace faktů? Jednostranné interpretace, která je motivována lidskou tendencí k myšlenkám, jenž vyúsťují buď v revoluční nebo apokalyptické predikce.

---

<sup>1</sup> Poukazuji zde k trilogii filmů *The Matrix* od režisérů Wachowskich.

Sám se kloním spíše k druhé alternativě. Předpokládám, že to, čeho jsme nyní svědky, je zásadní kvalitativní změna ve způsobu myšlení, interpretace vědomostí, jejich zpracování a praktického nasazení. Domnívám se tedy, a nejsem jediný (Lima 2015), že se jedná o komplexní proměnu myšlení založeného na bázi kumulativního vědění v síťový způsob myšlení – otevírá se zde prostor pro filosofii jako platformu síťového myšlení. Podstata síťového způsobu myšlení totiž již neprodlévá v kumulativním memorování a vybavování si obsahu informační databáze či databází oboru zájmu. Podstata síťového myšlení tkví spíše ve snaze obsáhnout širokou logickou strukturu mezi co největším množstvím dynamických a stále se vyvíjejících uzlů jak v informační síti databáze, tak i mezi různými typy databází (viz *konsilience*). Při tomto propojování uzlů je zároveň žádoucí znalost logických struktur databází, které leží mimo dominantní paradigma či spočívají v paradoxu – Feyerabendovský motiv vědeckého anarchismu – a zodpovědnost za důsledky vyplývající z Thomasova teorému (Nutil 2016). V případě potřeby potom síťový myslitel tuto logickou strukturu vztahů plní konkrétním informačním obsahem, který je dostupný digitálně, často online a je šířený mezi mysliteli „rychlostí světla“<sup>2</sup> (Hauer 2013, 39 – 49).

Ovšem tento způsob nakládání s informacemi není rovněž filosoficky bezprecedentní, nový či zcela originální. Myšlenkové zázemí podobných úvah lze hledat například v textu francouzského postmoderního filosofa Jeana-Françoise Lyotarda nazvaném *O postmodernismu*, konkrétně v kapitole *Výuka a její legitimizace performativitou* (1993, 155 – 161). Tento „prorocký“ text poprvé vyšel tiskem již v roce 1979. V principu pojednává o proměně paradigmatu samotného způsobu teoretizování a distribuce vědění napříč společnostmi. Více jsem se touto kvalitativní změnou, prodlévající ve vztahu sedukce a produkce, zabýval ve své studii *Baudrillardova teorie iluzí* (Zemčík 2014, 355 – 368). Ve zkratce pouze naznačím, že se jedná o nakládání s informacemi podle aspektu sedukce, tedy reciproční svůdnosti, který svádí samotný princip produkce (informací, produktů, světónázorů, teorií, sebe sama...) z předem „narýsované“ cesty<sup>3</sup>. Jeff Jarvis, profesor z newyorské City University a jeden ze 100 nejvlivnějších mediálních lídrů planety<sup>4</sup>, ve své knize *Co by udělal Google* (2010) na tento nový princip reciproční svůdnosti často poukazuje. K principu sedukce, která se opět stává dominantní silou v realitě neoddělitelně srostlé s virtualitou, poukazuje v devíti „nových pravidlech nové doby“.

<sup>2</sup> O určujícím vlivu rychlosti na podobu společnost již byla napsána přehledná studie z pera univerzitního profesora a filosofa Tomáše Hauera (Hauer 2013, 39 – 49).

<sup>3</sup> Pro hlubší analýzu vztahu sedukce a produkce mohu odkázat právě k mému článku *Baudrillardova teorie iluzí* (Zemčík 2014).

<sup>4</sup> Podle Světového ekonomického fóra konaného v Davosu v letech 2007 – 2008.

Těmito pravidly jsou:

- „Dnes rozhodují zákazníci. Jejich hlasy jsou slyšet po celém světě a během pár vteřin mají dopad na množství institucí.
- Lidé jsou schopni se vzájemně vyhledat v kterékoliv části světa a vytvořit kolem vás různé koalice – nebo také proti vám.
- Život je veřejnější. Rovněž podnikání.
- Masový trh je mrtvý.
- „Trhy jsou aktivní rozhovory,“ tvrdí *The Cluetrain Manifesto*, klíčová práce internetového věku vydaná v roce 1999. To znamená, že klíčovou dovedností v jakékoliv organizaci už není marketing, ale vedení rozhovoru.
- Dnes se nacházíme v po-nedostatkové ekonomice, v níž kontrola nad distribucí nebo nad výrobky nezaručuje ani odměnu, ani zisk. Musíme se tedy naučit zvládat nadbytek.
- Umožňovat zákazníkům, aby s vámi spolupracovali (při výrobě a distribuci výrobků, při jejich uvádění na trh a podpoře prodeje), je to, co na současných trzích tvoří zisk.
- Nejúspěšnějšími podniky jsou dnes sítě, které odčerpávají co nejméně hodnot, aby mohly co nejvíce růst – a platformy, na kterých jsou tyto sítě vystavěny.
- Vlastnictví produktovodu, lidí a výrobků nebo duševní vlastnictví již není klíčem k úspěchu. Tím je otevřenost“ (Jarvis 2010, 2 – 3).

Tyto pravidla se dají vztáhnout i na nefinanční pojetí ekonomie, jakým je například ekonomie znaku, či libidální ekonomie. Jarvis (2010, 3) ukazuje, že v současném světě mnoho rozhodnutí podle staré produktivní logiky nedává smysl. Ti, co se nepřizpůsobí, jsou odsouzení k zániku ve stínů nových praxí, teorií a myslitelů. Tato nová logika postupně transformuje všechny doposud známé privátní i veřejné instituce a jejich relevance. A tak vyvstává série zásadních otázek k zodpovězení, a to nejenom pro meditační profesory humanitních akademických oborů. Tyto otázky si musí položit celý akademický svět, svět primárního i sekundárního výzkumu, svět teoretické i praktické produkce, distribuce i řízení. Mezi tyto otázky patří například tyto: jak realizovat výzkum v této pro člověka nové situaci? Jak výsledky výzkumu ověřovat? Jak je nasazovat do praxe? O jaké teorii se opřít? Jak výzkum otevřít síle reciprocity? Kde nadále leží hranice mezi tím, co je relevantní a co ne? Tato studie se snaží poskytnout materiál pro jednu takovou konkrétní diskuzi, o možnostech a podobách výzkumu v této nové „post-faktické“ době, o povaze experimentální filosofie a budoucnosti *konsilience* vědních oborů.

V souvislosti s těmito změnami ve společnosti se skloňuje slovo *konsilience* vědění. Tento pojem se používá v širokých kontextech a to i za situace, kdy stále

není vyjasněno, jaké obory by se měly účastnit diskuze jako relevantní a jaké ne. Tak například mají se k diskusi o proliferační civilizaci mimo lékaře a hygieniky přizvat i psychologové, sociologové, historici, filosofové, ekonomové, politologové, síťoví architekti, urbanisté, neurobiologové, etologové či dokonce šamani nebo lidoví vypravěči?

Tato studie představí v základních rysech kulturně-psychologický výzkum morálky lidské mysli (Haidt 2013) amerického filosofa, psychologa a univerzitního profesora působícího na New York University Jonathana Haidta – výzkum, který osobně považuji za vzor toho, co můžeme do budoucna volně též nazývat *experimentální filosofii*. Haidtův výzkum je v mnoha ohledech postaven na principech konsilience vědních oborů, síťového myšlení a svůdnosti. Dále poukáže na potenciál navazujících výzkumů v rámci experimentální filosofie v prostředí středoevropských univerzit.

## **Morálka lidské mysli**

V této kapitole dílo amerického filosofa Jonathana Haidta, který se skrze vlastní doktorské studium začal soustředit na psychologické aspekty morálky. Jeho vědecká činnost gravituje okolo problematiky morálky lidské mysli. V rovině opakovatelných praktických psychologických experimentů přichází s relevantní odpovědí na otázku: která z etických tradic (deontologická, sentimentalistická či instrumentalistická) přesněji koreluje s výsledky vědeckých pokusů? Haidtovy výzkumy považuji za exemplární případ a demonstraci potence experimentální filosofie. Následující kapitola se bude primárně opírat o jeho obsáhlé dílo *Morálka lidské mysli: Proč lidstvo rozděluje politika a náboženství* (Haidt 2013).

## **Výchozí bod**

### **Morální matrice USA – Brazílie**

Jonathan Haidt vychází z výzkumu R. Shwedera, ze kterého se dá jednoduše odvodit tvrzení, že děti nedospívají k uvědomění si morálních povinností samy od sebe skrze racionální vhléd do podstaty toho, že ubližovat sobě a druhým je špatné. Stejně důležitý je i další zřejmý fakt vyplývající ze Shwederova výzkumu, že Američané stejně jako Indové nevnímali rozdíl mezi konvencí a morálkou. Například všichni Američané shodně soudili, že jakkoliv bránit vdovám v jejich

svobodné volbě jídelníčku<sup>5</sup> není otázkou konvence, je to z podstaty špatné. Vždyť vdovy si mohou dělat, co samy chtějí a kdokoliv tvrdí opak, je „tmář“, „zpátečník“ a „zlý člověk“ – takové výroky však patří do oblasti morálky. Jinými slovy, jak v indické Urise, tak i ve Spojených státech amerických je společenský řád roven tomu mravnímu. Schweder tedy tvrdí, že rozdíl mezi společenským a morálním řádem je jedinečným kulturním artefaktem vypovídajícím o vztahu jednotlivce a skupiny. Ze studie vyplývá závěr, že při komparaci výsledků Indie a USA, byly morální soudy nejenom nezachytitelné univerzálním racionálním kódem (který je jakýmsi leitmotivem silně projektivního evropského nároku na pravdivost), ale také se primárně opíraly o emoční intuice respondentů (Haidt 2013, 39 – 42).

Haidt tvrdí, že „Shwederova studie znamenala mohutný útok na celý racionalistický přístup“ (Haidt 2013, 42). V návaznosti na tuto studii vytvořil sérii vlastních příběhů, které se týkaly dvou základních podstat – zhnusení a neúcty. Své historicky z části doplnil o výběr vhodných „dilemat“, které použil například již Schweder. Všechny tyto příběhy se silným morálním podtextem se odehrávaly v soukromí, aby bylo eliminováno to, že samotná observace činnosti aktérů příběhů ublíží jinak nezúčastněnému pozorovateli; pracovně jim říkal „neškodná porušení tabu“ (Haidt 2013, 43 – 44).

Pro demonstraci si zmíníme tři z nich. První se odehrává na pozadí neúcty: „Jedna paní si doma dělá pořádek ve skříni a najde v ní starou americkou vlajku. Už o ní nestojí, a tak ji rozstříhá na kousky a vzniklé hadříčky použije při úklidu toalety“ (Haidt 2013, 44). Druhý příběh evokuje pocit zhnusení: „Jistý muž si jednou týdně zajde do supermarketu a koupí si tam kuře. Ale ještě než ho dá vařit, má s ním pohlavní styk. Potom kuře uvaří a sní“ (Haidt 2013, 25). Poslední evokuje pocit zhnusení a neúcty: „Jedné rodině přejedou před domem psa. Jeho majitelé se doslechli, že psí maso je lahůdka, a tak zvíře naporcují, uvaří a snědí k večeři. Nikdo je při tom nevidí“ (Haidt 2013, 25).

Po zpracování rozsáhlého pole dat se dostal k výsledkům, které bezvýhradně podpořily Shwederovy závěry. Výsledky tohoto brazilsko-amerického experimentu lze shrnout do následujících bodů:

- Američané a Brazilci z vyšších tříd vidí rozdíl mezi přestupky vůči konvencím a morálce. Naopak nižší třídy se projevovaly vetkáním konvencí do morálních matic. Trend byl takový, že čím chudší místo a nižší společenská třída, tím více převládala morálka nad konvencemi.
- Překvapivé však bylo zjištění, že společenské postavení ovlivňovalo názory respondentů výrazně více, než místo odkud pocházeli. Vzdělání a bohatí se

---

<sup>5</sup> „Činy, jež Indové prohlásili za špatné, ale Američané je posuzovali jako přijatelné ... Vdova žijící ve vašem společenství si dvakrát či třikrát týdně dopřeje rybu“ (Haidt 2013, 41). Ryba na jídelníčku je totiž v této sociocentrické společnosti, díky svému falickému tvaru, pro ovdovělou ženu považována za tabu.

tedy shodli se vzdělanými a bohatými a zároveň se shodovali chudí a méně vzdělaní s chudými a méně vzdělanými, a to nezávisle na lokalitě (USA či Brazílie).

- Rozdíly v reakcích se neprojeví ani po položení otázky: „myslíte si, že svým jednáním dotyčný někomu ublížil?“. Morální souzení tedy není podmíněno logikou neublížování. Nemorální je nemorálním, i když svým činem nikdo nikomu neublíží, přičemž i „neškodná porušení tabu“ byla povětšinou souzena jako špatná. Toto zjištění není dle Haidta principiálně vysvětlitelné racionalismem.
- Haidt tedy dochází při interpretaci výsledků svého pozorování k závěru, že východiska deontologické tradice nebyla skrze experiment potvrzena.

Při analýze výsledků, Haidt přichází s teoretickým rámcem obsahujícím tvrzení, že děti mají určitou vrozenou morální kapacitu, která obsahuje intuitivní pocity zhnusení a neúcty. Tato morální kapacita se investuje do konkrétního kulturního prostředí, které potom zapříčiní, že pocit zhnusení a neúcty je evokován rozdílným podnětem. Zastává tedy nativistickou (vrozené morální moduly) a sentimentalistickou pozici (Haidt 2013, 46 – 49).

### **Vykonstruované oběti**

Když Haidt plánoval respondentům položit na závěr každého z příběhů otázku: „myslíte si, že svým jednáním dotyčný někomu ublížil?“, netušil, jak silné reakce tato otázka bude v respondentech probouzet. Ačkoliv většinu příběhů pečlivě koncipoval tak, aby v nich nebylo možno nalézt obět' či aby se na činy aktérů příběhů nemohlo nahlížet jako na poškozování druhých, velké množství respondentů se v příbězích nějakou obět' snažilo najít. „V 38 procentech z 1620 případů, kdy respondent vyslechl příběh o nějakém sice pohoršujícím, ale zároveň neškodném jednání, setkal se s tvrzením, že někomu vznikla újma. Například u zmíněného příběhu s přejetým psem leckdo tvrdil, že rodina by svým počínáním poškodila samu sebe, protože by ji po psím mase bylo špatně“ (Haidt 2013, 49). A tak si Haidt musel položit otázku, jestli respondenti odsoudili chování jako špatné, jelikož předpokládali, že někoho poškodí, nebo tento proces probíhal v opačném pořadí a lidé si předpokládanou újmu vykonstruovali jaksi *ex post*, aby mohli svůj úsudek opřít o nějaký „racionální“ argument (Haidt 2013, 50).

Jelikož se většinou v roli tazatele v USA nacházel sám Haidt, mohl soudit, že si lidé své úsudky vytvářeli velice rychle, ale dodatečná opodstatnění konstruovali poměrně nesnadno, nemotorně a pomalu. Dodatečnou odpověď hledali výrazně delší dobu a formulovali ji váhavě s omluvným tónem v hlasu. „Například tento respondent: „No nevím, ale když ta paní tu vlajku vyhodí, možná se pak bude cítit

provinile?“. Nejedna z těchto potencionálně škodlivých důsledků působil vyloženě groteskně jako například v jednom případě, kdy dítě poté, co odsoudilo ženu, která roztrhala vlajku, prohlásilo, že hadry by při úklidu mohly ucpat záchod a ten by mohl přetéct“ (Haidt 2013, 50). Haidt proto vyškolil speciální tazatele do role jakýchsi „d'áblových advokátů“. Tito tazatelé se stali „experty“ na všechna možná vykonstruovaná odůvodnění a na každé z nich uměli přinést zcela logický a jednoduchý protiargument. „Ale i když respondenti uznali, že tyto údajné oběti jsou ve skutečnosti jen zástupné, stále odmítali označit dané jednání za neškodné ...uchylovali se k tvrzením typu: „Určitě je to špatné, jenom nedokážu říci proč.“ Působilo to jako by pokusné osoby postihlo jakési *morální ochromení* – jako by je neschopnost vyjádřit slovy to, co intuitivně bylo jasné, připravila o řeč“ (Haidt 2013, 50). Je namístě si položit Humeovskou otázku: „Potřebuje vůbec morálka racionální zdůvodnění?“. Není takové zdůvodnění pouze projektivním trikem „lišáctví“ rozumu? Neobelhává rozum sám sebe, považuje-li takovéto zdůvodnění za dostatečné? Haidt má v tomto ohledu jasno.

### **Sociálně-intuicionistický model morálky**

V důsledku pozorování tohoto mentálního mechanismu se Haidt s konečnou platností vzdává konceptu deontologie na poli morálky. Svě závěry dále opírá o sentimentalistickou tradici, která se ve svých východiscích a závěrech shoduje s výsledky pozorování, které mu byly dostupné skrze jeho vlastní experimenty. „Je to týž způsob uvažování, jež v roce 1739 popsal filosof David Hume: „Rozum je a má být jen otrokem vášní a nemá si nikdy nárokovat jinou úlohu než jim sloužit a být jich poslušen“ (Haidt 2013, 50 – 51). Haidt se domnívá, že našel důkaz, že se Hume nemýlil. Jeho zjištění byla v přímém rozporu s dobovým světonázorem, jež v devadesátých letech dominoval na poli psychologie morálky – s paradigmatickým opírajícím se výhradně o racionalismus a deontologickou tradici. Svě důkazy potom v říjnu 1993 uveřejnil v předním časopise zabývajícím se psychologií. Sám očekával vášnivou a bouřlivou reakci, ale stalo se něco, co vůbec neočekával – jeho článek nezaznamenal vůbec žádné reakce. Za prvních pět let od vydání tiskem jeho článek nejenom že vůbec nikdo neocitoval, ale ani na něj nezaútočil.

### **Nová syntéza**

K tomu, abychom pochopili, proč Haidtovy převratné teze po dlouhou dobu nevbuzovaly žádný zájem z pohledu vědecké a akademické obce, si v konturách naznačíme, jaké myšlenkové pozadí vládlo v devadesátých letech v oboru psychologie morálky.



Ve 20. století byly nativistické teorie morálky stigmatizovány přinejmenším v dvou aspektech. Prvním z nich bylo osvojení si nativistické morální pozice představitelem nacistického režimu za dob Německého národního socialismu Adolfem Hitlerem a rozvoj sociálního darwinismu. V podstatě se jedná o stanoviska hlásající, že podpora chudých, nemocných a jinak znevýhodněných skupin je nemorální, jelikož tím v oběhu společenské reprodukce udržujeme jedince a skupiny, kteří by měli být sami eliminováni skrze princip přírodního výběru. Jelikož byl Hitler nacist, používal se dosti naivní argument, že každý přívrženec nativistické morálky je rovněž nacistou<sup>6</sup> (Haidt 2013, 57).

Druhá vlna odporu proti nativismu vychází z radikalismu, který vládl v 60. a 70. letech 20. století na evropských, amerických a latinskoamerických univerzitách. Tyto radikálně reformistické postoje se zakládaly na předpokladu, že lidská přirozenost je nepopsanou tabulí, a proto skrze ní může být realizována jakákoliv utopická vize. Kdyby měla například různá etnika či pohlaví různé vrozené schopnosti, nemohlo by tak snadno dojít k jejich emancipaci. Myšlenky se tak v rámci vědecké a akademické obce soudily nejenom z hlediska „faktičnosti“, ale i ideologické poplatnosti emancipačním hnutím. Tyto ideologické represe postihly například zakladatele sociobiologie Edwarda Osborna Wilsona působícího na Harvardské Univerzitě. Wilson vydal roku 1975 knihu *Sociobiology: The New Synthesis*, kde tvrdí, že přirozený výběr neovlivnil zvířata pouze z hlediska tělesnosti, ale i chování. To, za co byl však odsuzován, bylo přiznání možnosti, že jestliže tak tomu bylo u zvířat, mohlo by tak tomu být i u člověka. Tvrdil, že lidská přirozenost je definována integrálem toho, čeho potencionálně můžeme dosáhnout výchovou svých dětí či vytvářením nových institucí. Jinými slovy poukázal na to, že člověk není ve sféře sociálního konstruktivismu a realizovatelného utopismu principiálně všemocný. Wilsonovi se zdálo evidentní, že racionalisté, které znal jako své kolegy z Harvardu, mezi něž patřili například Lawrence Kohlberg či John Rawls, v principu pouze produkují zdůvodnění morálních intuicí, jenž lze vysvětlit i na základě psychologické evoluce (Haidt 2013, 57 – 58).

„Skutečně lidé věří v existenci lidských práv proto, že tato opravdu existují – obdobně jako matematické pravdy – a sedí v nějaké přihrádce kdesi v kosmu, hned vedle Pythagorovy věty, a jenom čekají, až je nějaký platónští rozumáři objeví? Nebo se lidé při čtení zpráv o mučení otřásají odporem a soucítí s jeho oběťmi, a teprve potom si vycucávají z prstu povídačku o univerzálně platných právech, aby měli své pocity čím ospravedlnit?“ (Haidt 2013, 58 – 59). Wilson, podobně jako Hume, tvrdil, že lidé po přímé emocionální reakci dodatečně konstruují její odůvodnění. Wilson volal po odebrání monopolu na etiku filo-

---

<sup>6</sup> Ačkoliv z logického hlediska tento argument nedává smysl, z pozice sentimentalistického směru v psychologii morálky, tak jak jej zastává Haidt, je motiv tohoto argumentu principiálně snadno pochopitelný.

sofům a předpovídal novou syntézu, která by propojovala studium mnoha oborů, jakými jsou filosofie, biologie, evoluce, atd. – dal jí název *konsilience*; smíšené myšlenky mnoha oborů ve snaze sjednocovat poznání. Tento nový obor má potenci neustále reflexivně explarovat, ověřovat a utvářet normy – v principiálním smyslu Giddensovy reflexivní modernity (Giddens 1990, 42 – 43). Tisk a veřejnost na Wilsona uspořádaly štvanici, nadávaly mu do fašistů, rasistů a chtěly mu odepřít právo veřejného projevu. Odpůrci narušovali jeho přednášky skandováním nenávislivých hesel (Haidt 2013, 59).

V roce 1987, kdy Haidt zahajoval postgraduální studium, byly již sociobiologie, nativismus a konsilience zcela zdiskreditovány. Sociobiologie byla v akademickém světě používána pejorativně, a proto emoce jako produkt evoluce lidského druhu psychologie morálky vůbec neřešila. Jediné „správné“ bádání se odehrávalo na poli vývoje rozumového usuzování a zpracovávání informací. Ovšem mimo psychologii existovala řada excelentních děl, například z pera etologa působícího na atlantské Emory University Franse de Waala, který tvrdil, že šimpanzi mají mnoho z potřebných psychologických základů, které mohou sloužit jako potenciaální materiál k budování morálních systémů a komunit<sup>7</sup>. Nebo například dílo neurovědyce působícího na University of Southern California Antonia Damasia, který v knize nesoucí titul *Descartův omyl* pojednává o případech ztráty emocionality u pacientů s konkrétním poškozením mozku. I když takto postižení lidé mají „zkratované“ emocionální centrum, jejich inteligence zůstává nezměněna. Tito pacienti jednoduše nejsou schopni rozlišit mezi dobrým a špatným i v případech, že excelovali v rozumových úsudcích v oblasti rozlišování pravdivého od nepravdivého<sup>8</sup>. Jediná cesta k učinění rozhodnutí byla pro Damasiovy pacienty možná pouze a čistě skrze racionální kalkul. To však tyto lidi absolutně vyřazovalo ze společnosti, jelikož neustále pouze analyzovali možnosti, ale nebyli se schopní rozhodnout pro žádnou konkrétní. A když se pro nějakou rozhodli, většinou tato možnost byla společensky absolutně nepřijatelná (zvažování výhod a nevýhod vraždy vlastních rodičů a podobně). Damasiovy poznatky zcela

---

<sup>7</sup> Pro rychlé seznámení s tématem doporučuji zhlédnutí krátké přednášky Franse de Waala nazvané *Frans de Waal: Moral behavior in animals*, dostupné přes přednáškovou platformu TED. Viz přednáška s titulkou: <https://www.youtube.com/watch?v=GcJxRqTs5nk>

<sup>8</sup> Toto je principiálně odvozeno od tvrzení, že mozek si vytvořil a neustále udržuje mapy jako vztýčné body společně sdílené reality. A. Damasio ukazuje, jak mozek obsahuje mapu vlastního těla jako stálé jednotky, která je potom používána jako výchozí bod pro zkušenost s tělesností druhých. Podobně je tomu i s mapou emocionality, racionality, ega, morálky či ducha – jak se mapa překrývá s konceptem „objektivní“ reality, se můžeme dočíst například v inspirativním díle Jeana Baudrillarda (1994) *Simulacra and Simulation*. Viz Damasiovu přednášku uveřejněnou na platformě TED nazvanou *The Quest to Understand Consciousness*

[https://www.ted.com/talks/antonio\\_damasio\\_the\\_quest\\_to\\_understand\\_consciousness](https://www.ted.com/talks/antonio_damasio_the_quest_to_understand_consciousness)

jednoznačně popíraly Platónovy představy o rozumu jako řídicím mechanismu emocí (Haidt 2013, 59 – 61).

Intelektuální pole tedy bylo již připraveno, stačilo pouze překonat staré předsudky a realizovat mezioborové spojení poznatků. Haidt přichází s teoretickým modelem založeným na metafoře opačné té, kterou předkládá Platon. S modelem, který koresponduje s jeho vlastním pozorováním.

### **Prastarý slon a mladičský jezdec**

Haidt svou teorii představuje ve článku, který nazval *Emotivní pes a jeho rozumový ocas*, jenž vyšel tiskem roku 2001. S odstupem času si ale uvědomuje, že výstižnější název pro jeho článek by byl *Intuitivní pes a jeho rozumový ocas*. Titul poukazuje k principu toho, že „argumenty, jimiž se pokoušíme podložit morální úsudek, jsou totéž co ocas, jimž vrtí intuitivní pes“ (Haidt 2013, 77). Svou teorii Haidt označuje jako *sociálně-intuicionistický model morálního úsudku*. Staví jej na metafoře obrovského prastarého slona a mladičského jezdce. Platonova metafora o rozumovém vozataji krotícím emotivní koně (Platon 2000, 35) v podstatě vyjadřuje naivní leitmotiv typizovaného evropského zamlčeného filosofického předpokladu, že objektivita vyjádřena skrze racionální konstrukce je objektivně hodnotnější než jakákoliv pozice opírající se o bázi subjektivity. Metafora prastarého intuitivního slona vezoucího mladičského racionálního jezdce pochází původem z asijské buddhistické tradice. Jejím leitmotivem může být naopak předpoklad typu, že subjektivita je subjektivně hodnotnější než objektivita – všimněme si, že tento pohled nechává prostor jak rozvoji objektivit, tak subjektivit – volby výchozí pozice předpokladu (Haidt 2013, 77).

Slon je tedy na metaforické rovině vyjádřením aspektu toho druhu kognice, který prodlévá v realitě automatismů, emocí, intuice a všech forem „vidět, že“. „Automatické procesy pohánějí mysl člověka úplně stejně jako mysl zvířat, již obhospodařují už 500 milionů let“ (Haidt 2013, 74). V posledním miliónu let se u lidí vyvinula schopnost verbální komunikace a rozumového uvažování. Bylo by však naivní se domnívat, že došlo ke zcela nové konfiguraci architektury mozku, jenž by jej uzpůsobila k tomu, aby mohl fungovat na bázi Platonovy představy o rozumu jako primárním principu. Jezdec jako docela nový mentální modul se naopak vyvinul, aby umocňoval síly a dovednosti primordiálního slona (Haidt 2013, 74).

Jezdec je vyjádřením toho aspektu kognice, který zajišťuje chod řízených procesů typu „uvažovat, proč“. Jezdec slouží slonovi tak, že dokáže mentálně dohlédnout za použití abstrakce dále do neurčité budoucnosti. Je schopen lépe stanovit dlouhodobé záměry a cíle, které opírá o bázi vyabstrahovaných informací z imaginace minulých událostí, jejich důsledků a ponaučení z nich plynoucích.

Jezdec slouží skrze svou schopnost osvojovat si nové dovednosti a technologie, díky čemuž může slon snadněji dosahovat svých cílů a vyhýbat se nebezpečným situacím. Jezdec také vystupuje jako tiskový mluvčí slona, aniž by sám často měl sebemenší ponětí o tom, co si slon myslí, nebo co motivuje jeho činy. Tato schopnost je nesmírně důležitá v sociální sféře, která je protkána sítí významů slov. Pokud má totiž slon dostatečně kompetentního jezdce, může strhnout celé stádo dalších slonů i jejich jezdců směrem, kterým si sám „zvolil“. Jestliže totiž chceme změnit něčí názor, musíme například kvůli principu konfirmačního zkreslení<sup>9</sup> působit ne na jeho jezdce, ale slona. Haidt upozorňuje na to, že „totiž ke změně názoru dospíváme pod vlivem našeho okolí. Ostatní nás ustavičně ovlivňují už jen tím, že dávají najevo, koho mají a nemají rádi... ostatní lidé nad námi mají opravdu obrovskou moc... pod vlivem společnosti se i krutost může zdát přijatelná a altruismus působit jako něco trapného“ (Haidt 2013, 76). Podle tohoto modelu, zle dále tvrdit, že konkrétní kulturně singulární „příběh“ evokující například mentální modul tabu je pravděpodobným kulturním artefaktem – produktem časoprostorové souhry mnoha „úspěšných slonů a jejich zručných jezdců“ – který je paradoxně v konkrétní morální matici vždy vnímán jako nezpochybnitelný fakt s nárokem na universalitu.

## **Perspektivy experimentální filosofie**

### **Sociální vrstvy společnosti a morálky**

Pokud se Haidt nemýlí a emoce jsou katalyticky evokovány v důsledku kulturně podmíněných morálních matic, je principiálně možné tyto matrice racionálně popsat a sestavit jakýsi katalog těchto matic. Tento katalog by byl v podstatě souborem oindexovaných meta map morálních matic – byl by to morální atlas založený na logice morálky lidské mysli. Připomeňme si závěr vyplývající z Haidt-ova výzkumu Brazílie – USA poukazující k tomu, že různé sociální vrstvy se ztotožňují s různými morálními maticemi. Proto je třeba vzít v potaz skutečnost, že každá kulturní skupina se nejspíše bude vyznačovat svéráznou morální mapou. Je potom na místě předpokládat, že také každá subkultura, komunita a posléze rodina bude obsahovat nejspíše vlastní morální singularity. Pokud rozvineme tyto domněnky, pak můžeme tvrdit, že v průběhu socializace se morální mapa „divného“<sup>10</sup> (Haidt 2013, 132 – 134) jedince pohybujícího se v prostředí individualistické společnosti stává jedinečnou kombinací mnoha různých map. Tato kombinace je z podstaty symbolického recipročního vztahu jedinec – realita, také

---

<sup>9</sup> Viz podkapitola *Bezmoc morálně ochromených*.

<sup>10</sup> Z ang. WIRED – Western, Educated, Industrialized, Rich, Democratic.

neustále rekombinována a přepisována v žitém kontaktu se svůdnou neurčitostí reality.

Principiálně je však možné pracovat na sestavování orientačních map různých kultur, subkultur, atd. Dále je možné se soustředit na rozklíčování toho, k jakému vzorci a s jakou mírou pravděpodobnosti budou mít konkrétní jedinec či skupina tendence. Ovšem upozorňuji na to, že průkopníci experimentální filosofie budou muset mít neustále na paměti princip sedukce v aspektu toho, že modelem reality není statický strom, ale věčně se uskutečňující síť. A tak se teorie popisující specifickou kombinaci morálního matrixu nebude nacházet ve sféře pravdivosti, ale pravděpodobnosti – nebude staticky pravdivá, ale dynamicky pravděpodobná.

Takováto teorie by podle mého předpokladu mohla dále sloužit jako síťová platforma umožňující přiblížení se pozorovanému. Mohla by také vystihovat aspekt neustálého rekonstruování a přeformování sebe sama v recipročním kontaktu teorie s žitou zkušeností, o které vypovídá. Na bázi této platformy by bylo poté možno vystavovat konkrétní moduly, které by mohly sloužit například k reformaci společenských institucí. Mohla by být také nasazena jako nástroj, testující nakolik jsou legislativní imperativy konzistentní vzhledem k realitě, které se týkají – jako instrument sloužící principům občanské společnosti.

## Bezmoc morálně ochromených

Nyní si v konturách představíme princip *konfirmačního biasu*. Konfirmační zkreslení je dobře prostudovaný a popsáný mentální mechanismus – jedná se o jeden z druhů kognitivního biasu. Označuje se tak automatismus nesymetrie, který zapříčiňuje to, že k potvrzení našeho stanoviska, které emotivně cítíme jako správné, nám stačí pouze jedno dosvědčující tvrzení. Nesymetrie paradoxně prodlévá v tom, že k zamítnutí jakéhokoliv stanoviska, se kterým se naopak intuitivně neztotožňujeme, stačí opět pouze jedna vyvracející teze. Je experimentálně dokázáno, že schopnost najít v jakémkoliv systému buď afirmační nebo negující prvek je přímo úměrná míře inteligence a řečnických vloh testované osoby. Také je potvrzeno to, že jednotlivec je schopen snadno najít mnohé afirmace pro stanovisko, které vnímá jako správné, a zároveň je pro něj nesmírně těžké „vcítit se do kůže protivníka“ a hledat stanoviska potvrzující pozici opačnou. Pokud se však jeden skrze empatii vžije do pohledu druhého, paradoxně často dochází k situaci, kdy přehodnotí vlastní stanovisko. Konfirmační bias tedy není ničím jiným než sofistikovanou a experimentálně potvrzenou formou známého stanoviska *homo mensura* (Haidt 2013, 113 – 115).

Konfirmační bias rovněž souvisí s principem morálního ochromení, jak jsme si jej představili v pododdílu nazvaném *Vykonstruované oběti*. Domnívám se, že studium jejich vzájemného propojení by mohlo být užitečné k porozumění me-

chanismů „morálního“ vlivu určitých subjektů či skupin na jiné subjekty a skupiny (distribuce a redistribuce moci). Můžeme totiž pouze předpokládat, že ochromí-li morálně (například při veřejné debatě) jeden svého oponenta tím, že použije intelektuálně vykonstruovaných prostředků, může s ním poté snadno manipulovat (či s názorem veřejnosti na tohoto oponenta) – konsenzuální pravda či veřejné mínění poté může být nazíráno pouze jako sofistická funkce sumy řečnickovy připravenosti. Tento princip lze v rámci experimentální filosofie extrapolovat například na předpoklady současného sporu východiska a řešení „migrační krize“, Brexitu či všudypřítomné hrozby terorismu. Lze jej také kombinovat například s principy Thomasova teorému (Merton 1948) či „Possibility effectu“ (Kahneman 2012).

### „Kód Lásky“

Ervin László je maďarský filosof, autor 75 knih a 400 studií, zabývající se filosofií vědy, teorií systémů a integrální teorií. László obhájí nikterak originální názor, že se naše globální civilizace nyní pravděpodobně nachází na pokraji možností své existence (vyčerpání surovin, znečištění prostředí, migrační vlny, nebezpečí radikalismu, singularita umělé inteligence atd.). Tvrdí, že první čtvrtina 21. století je ten čas, kdy se rozhoduje o budoucnosti civilizace; přežije-li, transformuje-li se či zanikne. Upozorňuje na to, že se civilizace hroutí také díky tomu, že chybí jazyk či kód dostatečně zřetelný všem – kód, který by umožňoval dostatečně transparentní komunikaci mezi prvky systému. Domnívá se, že pokud si systém neosvojí nějaký univerzální kód, který umožní pracovat s dynamikou jeho prvků, pohltí jej entropie – chaos. László přichází s originálním pohledem. Podle něj je jediným dostatečně univerzálním kódem, který by byl v přijatelné míře srozumitelný všem prvkům v systému globální civilizace, a to na všech úrovních, to co bychom mohli principiálně nazvat láska<sup>11</sup> (László 2013).

Lászlóva myšlenka se mi v kontextu Haidtova výzkumu a perspektivě experimentální filosofie zdá velice inspirativní. Haidt tvrdí, že lidská bytost, ať již žije kdekoliv na světě, je obdařena konkrétními „emočními“ mentálními moduly, které jsou pevně vetkané do architektury lidského mozku. Tyto moduly, principiálně podobné chuťovým pohárkům na lidském jazyku, jsou poté skrze konkrétní morální matrix, který je bezvýhradně kulturním artefaktem, evokovány (Haidt 2013, 164, 214). Situace, ve kterých různé, často nesouměřitelné morální matrice evokují tyto moduly, jsou pro každou tuto matici specifické. Vzpomeňme si na příklad, kdy Indy znepokojovalo to, že vdova jí rybu a Američany zase naopak znepokojovalo to, že by si vdova nemohla svobodně zvolit svůj jídelníček – to, co

---

<sup>11</sup> Pro více informací sleduj [https://www.youtube.com/watch?v=lkA\\_ILHfcfl](https://www.youtube.com/watch?v=lkA_ILHfcfl).

mají tyto dva pohledy společné, je pocit znepokojení v pozadí. A tak se domnívám, že hledali-li bychom onen Lászlóvský univerzální kód podírající transparentní výměnu informací a energie v naší globální civilizaci, měli bychom jej právě hledat ve sféře těchto modulů, které jsou nezávislé na konkrétních morálních maticích.

Zdali je vůbec tento kód možný, je hodno dalšího výzkumu. Stejně jako otázky typu: jakou by mohl mít podobu? Jak jej prakticky aplikovat v rámci komunikace či univerzálních principů? Jaké jsou jeho hranice? Jaká je úloha racionalizace v rámci tohoto kódu? Co má tento projekt společného s hledáním čistého racionálního meta kódu logických pozitivistů působících od 30. let 20. století?

Existují předpovědi, že někdy okolo roku 2035 – 2050 nastane singularita ve vývoji umělé inteligence (UI). Touto singularitou se myslí událost, kdy umělá inteligence ve svých schopnostech přesáhne inteligenci lidskou. Vše vypovídajícím příkladem může být UI jménem TAY vypuštěná společností Microsoft 23. 3. 2016, aby se na sociální síti Twitteru učila od uživatelů. TAY musela být za necelý den vypnuta, jelikož inspirovaná twitty lidí začala šířit rasistické a nábožensky nesnášenlivé názory či obhajovat genocidu (Microsoft 2016). Pokud se lidstvo bude chtít vyhnout katastrofickým scénářům známých zatím pouze z ponurých futuristických vizí, ve kterých se inteligentní stroje z pomocníků stanou tyrany, bude muset být i mysl budoucích super UI dvojvrstvá. Přičemž vrstva logických a racionálních operací (jezdec) bude muset hrát z hlediska rozhodovacích procesů sekundární roli vzhledem k vrstvě morální intuice (slon). Morálka této super UI bude muset být také mnohem pevnější než ta lidská. Bude se také muset vyhnout stranění konkrétním morálním maticím. Jako východisko vidím to, že by sama operovala právě na onom univerzálním „kódu lásky“ – jeho základními moduly by mohly být právě: úcta, soucit, smysl pro svobodu a autoritu, péče, férovost, loajalita, tabu a láska.

## Literatura

- ARENDR, H. (1995): *Eichmann v Jeruzalémě: zpráva o banalitě zla*. Přel. M. Palouš. Praha: Mladá fronta.
- BAUDRILLARD, J. (1994): *Simulacra and Simulation*. Michigan: The University of Michigan.
- DAMASIO, A. (2015): The Quest to Understand Consciousness. In: *TED*. Online stream. Online. 25. May 2015. <[https://www.ted.com/talks/antonio\\_damasio\\_the\\_quest\\_to\\_understand\\_consciousness](https://www.ted.com/talks/antonio_damasio_the_quest_to_understand_consciousness)>.
- DE WAAL, F. (2014): Moral behavior in animals. In: *TED*. Online stream. 23. Apr. 2015. <<https://www.youtube.com/watch?v=GcJxRqTs5nk>>.
- HAIDT, J. (2013): *Morálka lidské mysli: proč lidstvo rozděluje politika a náboženství*. Přel. H. Čížková, Praha: Dybbuk.

- HAUER, T. (2013): Dromology. In: Schneiderová, M. (ed.): *Člověk, ideje, společnost I*. Ostrava: Vysoká škola báňská – Technická univerzita Ostrava.
- JARVIS, J. (2010): *Co by udělal Goole?*. Přel. P. Ivan. Litomyšl: Eastone Books.
- KAHNEMAN, D. (2012): *Myšlení: rychlé a pomalé*. Přel. E. Nevrlá. Brno: Jan Melvil.
- LÁSZLÓ, E. (2013): A New Love Declaration: Ervin Laszlo at TEDxNavigli. In: *TEDx*. Online stream. 28. May 2013. <[https://www.youtube.com/watch?v=lkA\\_ILHfcfl](https://www.youtube.com/watch?v=lkA_ILHfcfl)>.
- LIMA, M. (2015): A visual history of human knowledge. In: *TED*. Online stream. 28. May 2015. <[http://www.ted.com/talks/manuel\\_lima\\_a\\_visual\\_history\\_of\\_human\\_knowledge#t-404599](http://www.ted.com/talks/manuel_lima_a_visual_history_of_human_knowledge#t-404599)>.
- MERTON, R. (1948): Self-Fulfilling Prophecy. In: *The Antioch Review*, 8 (2), 193-210.
- MICROSOFT. (2016): *TAY*. Online. 11. Apr. 2016. <<https://www.tay.ai/>>.
- NUTIL, P. (2016): Thomasův teorém: Proč naše předpovědi mění svět. In: *Manipulátoři.cz*. Online. 11. Apr. 2016. <<http://manipulatori.cz/thomasuv-teorem-proc-nase-predpovedi-meni-svet/>>
- ZEMČÍK, T. (2014): Baudrillardova teorie iluzí. In: *Filosofický časopis*, 62 (3), 355-368.

**Mgr. Tomáš Zemčík, PhD.**

VŠB-TU Ostrava  
Katedra společenských věd  
17. listopadu 15/2172  
708 33 Ostrava – Poruba  
Česká republika  
tomas.zemcik@vsb.cz