

JAK JE MOŽNÉ VĚDĚT, JAKÉ TO JE

Vít Bartoš (*Katedra filosofie, Fakulta přírodovědně-humanitní a pedagogická,
Technická univerzita v Liberci, Česká republika*)

Abstract. This article is criticizing philosopher Frank Jackson's thought experiment (TE) "The knowledge argument". Jackson's goal of the experiment was to prove that qualia are non-physical entities. The TE turned out as having been built on premise which was not realizable from physical reasons. Furthermore, it appears that the experiment implicitly assumes biologically inadequate theory of the hierarchy of cognitive functioning domains of a (human) organism. David Hume's theory was reminded, contrasting to Jackson's related concept of propositional contents and phenomenal mental states. There were also two modern naturalistic qualia interpretations proposed. The first one gives qualia into direct relationship with the neuroscientist Antonio Damasio's theory of feelings. The second one refers to the neuroscientist Gerald Edelman's thoughts of the evolutionary value and structure of qualia. Naturalistic interpretation of qualia concludes that qualia are multidimensional cognitive contents – processes whose meaning lies in the comprehensive evaluation of an experienced situation that is accessible to conscious experience. It is also argued in favor of the proposition that no amount of propositional information can replace qualia – therefore sensory deprivation cannot be substituted by knowledge. This fact is due to limitations of reach of conscious or introspective structures of the human organism. Finally, the fundamental function of physicality in our mental life is reflected.

V tomto článku bych rád provedl kritickou analýzu předpokladů slavného myšlenkového experimentu filosofa Franka Jacksona „The Knowledge Argument“. Ve své podstatě půjde o to ukázat, v jakých momentech nás tento konkrétní myšlenkový experiment může zavádat nebo již přímo zavádí na scesti nekonstruktivními či dokonce spornými předpoklady. Taková kritika ale nemá být pouhým sémantickým cvičením v přesnějším vymezení pojmů. Jejím cílem je zároveň redefinice problému tak, aby se ukázala faktická důležitost a přitažlivost otázek nastíněných v myšlenkovém experimentu.

Obecné problémy myšlenkových experimentů

Rád bych, aby bylo zřejmé, že v žádném případě nejsem odpůrcem myšlenkových experimentů. Naopak, dle mého soudu jsou naprosto zásadní, protože jejich intelektuální potenciál spočívá v jejich provokativnosti a schopnosti excitovat představivost, což jsou faktory, které (osobně) považuji za dostatečný důvod pro jejich provozování.

Nejhlubší problém myšlenkových experimentů spočívá v jejich konstruktivitě. Otázka by měla znít: Za jakých podmínek je myšlenkový experiment konstruktivní? Ještě přesněji

se problém dá formulovat v termínech vztahu mezi logickou a nomickou (fyzikální) nutností. Myšlenkové experimenty (zvláště ve filosofii) jsou založeny totiž na tom, že pokud v našich úvahách nevidíme logický spor, tak taková experimentem simulovaná situace může být chápána jako možná – jako možný svět, kde je principiálně možné to, co popisuje myšlenkový experiment. Z takového, na první pohled myšlenkově možného světa, se pak v druhém – poněkud nejistém – kroku úvahy náhle stává svět „ontologicky alternativní“ k tomu našemu (jiný typ vesmíru s jinými ontologickými a fyzikálními principy). Nakonec z toho celá řada filosofů odvozuje, že standardní fyzikální zákony našeho světa nemusí být brány v úvahu při konstrukci myšlenkového experimentu, protože prostě není rozporné si představovat jinak ontologicko-fyzikálně uspořádané světy¹, které mohou v jistém smyslu paralelně existovat s tím naším.

Domnívám se, že tento přístup je nesprávný a v každém případě nekonstruktivní. Důvody bych uvedl alespoň dva. První je epistemologický a říká, že pokud někde nevidíme prima facie logický spor, tak to ještě neznamená, že tam fakticky není. V případě vysoce komplexních problémů, kam jistě záležitost lidské mysli spadá, prostě zcela principiálně operujeme s nedostatkem informací a naše výchozí premisy, pokud budou širším způsobem analyzovány, se mohou ukázat jako kontradiktorické – domnívám se, že to je případ řady myšlenkových experimentů včetně toho, o kterém tu budu psát.

Druhý důvod je komplikovanější a popravdě řečeno esoteričtější. Jednoduše tvrdím, že mezi logickou a nomickou nutností je ontologický vztah a nikoli ontologická propast, kterou má příkré terminologické (a také pojmové) rozlišení tendenci spíše zdůrazňovat. Pro velké teoretické fyziky 20. století jako byl A. Einstein, P. A. M. Dirac nebo A. S. Eddington, abych zmínil jen některé, je takové příkré rozdělení ne zcela pochopitelné. Fundamentální zákony přírody jsou z jejich úhlu pohledu něčím, co nemůže být jiné, a tak neexistují ani žádné alternativní světy či vesmíry, které by umožňovaly zásadně jiné fyzikální procesy, než se dějí v našem vesmíru. Bůh prostě nehraje v kostky, jak říká Albert Einstein. Myšlenkový experiment, jehož premisy jsou v rozporu s fundamentálními zákony fyziky², je pak z tohoto hlediska dokonce kontradiktorický (fundamentální zákony přírody vyjadřují v tomto pojetí zároveň to, co nazýváme „logickou nutností“). Přijmeme-li tento silně deterministický předpoklad o fungování světa, pak z toho plynou jasná omezení pro konstrukci smysluplných myšlenkových experimentů. Myšlenkové experimenty, které nedbají fyzikálních omezení tak, jak o nich mluví již současná fyzika, a zároveň vytváří zdání logické bezespornosti, jsou ve skutečnosti rozporné v silném (logickém) smyslu slova. Závěr je pak jasný: Dobrý myšlenkový experiment by měl tedy respektovat alespoň ty zákony přírody, které chápeme jako fundamentální – v opačném případě je možné cokoliv a myšlenkový experiment se mění v ničím nespoutané fantazírování a ještě spíše ve „flatus vocis“ – existenci prázdných znaků (spojení slov) bez významu.

¹ Pozice tzv. modálního realismu.

² Je pravda, že zatím žádnou jednotnou fyzikální teorii nemá lidstvo k dispozici, a tak uvažovat o fundamentálních zákonech je možná předčasně. Stoupenci popperovského falibilismu (přesvědčení, že každá teorie bude nakonec vyvrácena) budou tvrdit, že dost možná žádné takové zákony ani neexistují, a že další vývoj vědy ukáže, jak moc jsme se v tomto předpokladu mýlili – jako se to standardně dělo po celou historii moderní vědy. Na druhou stranu obecnost a enormní explanatorní i prediktivní úspěch moderní fyziky i s jejími finalistickými pretencemi je něčím, co nelze přehlédnout. Tento problém tu nelze detailně řešit. Osobně se přikláním k tezi o fundamentálních principech fyziky, k jejichž rozluštění nemá lidstvo daleko – popperovská teorie falibilismu bude pravděpodobně dle vlastního popperovského schématu vývoje vědy falzifikována.

Marie v černobílém pokoji

Po metodologických poznámkách se konečně můžeme věnovat experimentu samotnému. Když dovolíte, ocituji originální formulaci: „*Mary is a brilliant scientist who is, for whatever reason, forced to investigate the world from a black and white room via a black and white television monitor. She specializes in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like ‘red’, ‘blue’, and so on. She discovers, for example, just which wavelength combinations from the sky stimulate the retina, and exactly how this produces via the central nervous system the contraction of the vocal chords and expulsion of air from the lungs that results in the uttering of the sentence ‘The sky is blue’.* ... ***What will happen when Mary is released from her black and white room or is given a color television monitor? Will she learn anything or not? It seems just obvious that she will learn something about the world and our visual experience of it. But then is it inescapable that her previous knowledge was incomplete. But she had all the physical information. Ergo there is more to have than that, and Physicalism is false.***”³

Takže když to shrneme do nějaké srozumitelné formulace, dostaneme asi následující úvahu:

- (1) Marie se naučila všechny fyzikální informace o procesu vnímání (barev například).
→ tj. každý proces vnímání dokáže popsat v termínech fyziky, chemie, neurobiologie atd.
- (2) Poté, co je Marie propuštěna ze svého černobílého pokoje a uvidí nové barvy, se dozví něco **nového**.
→ dozvěděla se informaci typu „*jaké to je vnímat*“ jinou než černou či bílou barvu
- (3) Pokud se tedy Marie dozvěděla něco nového, pak nová informace není fyzikální informací (informací o fyzikálních procesech), protože všechny fyzikální informace měla dopředu k dispozici (dle první premisy).
- (4) Redukcionistický fyzikalismus nemůže platit.

Tento závěr je samozřejmě z hlediska moderní naturalistické vědy skandální. Jednoznačně popírá v podstatě materialistickou tezi o „fyzikální podstatě“⁴ veškeré ontologické reality.

Řekněme si rovnou, že obě výchozí premisy jsou poměrně dost problematické. První premisa je dle mého soudu neuskutečnitelná v žádném možném fyzikálním vesmíru podobného stupně complexity, jako je ten náš. Druhá premisa je z hlediska analýzy zajímavější, protože obsahuje intuitivně přijatelnou pravdu o tom, že Marie se skutečně dozví něco nového, nicméně je založena na neadekvátní (a explicitně nevyslovené) teorii o fungování kognitivních procesů.

K první premise jen krátce; Předpokládat, že Marie má „všechny“ fyzikální informace o konkrétním⁵ procesu vnímání, je prakticky to samé jako tvrdit, že Marie je laplaceovský

³ <<http://plato.stanford.edu/entries/qualia-knowledge/>>

⁴ Pod pojmem „fyzikální podstata“ si představte elementární částice a pole, ze kterých je vše ostatní složeno.

vševědoucí démon. Již to, co dnes víme o procesu vnímání, postačuje k tomu, abychom mohli konstatovat neuvěřitelnou komplexitu tohoto procesu. Samozřejmě je tu možné se dovolávat „jen“ dosavadních empirických dat, ale pokud před námi stojí komplexní adaptivní systém (mozek, organické tělo), ve kterém dle jistých předpokladů mohou hrát zásadní roli i procesy na subatomární úrovni, pak spojení „*mít všechny fyzikální informace (o procesu)*“ je prostě nesmyslné, protože je v rozporu s principy fyziky komplexních systémů a principy kvantové fyziky. Kompletní fyzikální informace je pohled z pozice božího oka. Lze s trochou přisnosti konstatovat, že již tímto předpokladem Frank Jackson svůj myšlenkový experiment diskvalifikoval.

Interpretace experimentu

Je, myslím, v celku jasné, co chtěl Jackson svým myšlenkovým experimentem dokázat; omezenost fyzikalismu! Ve skutečnosti v samotném zadání experimentu je určitý rozpor, který žádá objasnění. Posud'te sami:

V druhé premise Jackson spoléhá na všeobecně sdílenou intuici, která říká, že Marie po propuštění ze svého černobílého světa skutečně získá nové poznání (zažije, jaké to je vnímat barvy celého spektra). Naprosto drtivá většina lidí přistoupí na tuto interpretaci a jen zlomek (z vlastní zkušenosti při výuce je to kolem 10% dotázaných) odpoví, že Marie;

- a) se nemůže dozvědět nic nového, protože dle první premisy již všechno ví;
- b) se sice dozví něco nového, ale pak to tedy znamená, že neměla kompletní fyzikální informace, jak se mylně předpokládalo v první premise.

Vtip je v tom, že pravdu má i v tomto případě většina a Jacksonova intuice je tedy správná. Důležité je ovšem říci, co ve skutečnosti stojí za všeobecně sdíleným přesvědčením. A to po pravdě není vůbec těžké. Každý totiž ví, že je rozdíl mezi abstraktní propoziční znalostí a smyslovým zážitkem, stejně jako je rozdíl mezi „vědět, že“ a uměním praktické realizace tohoto „že“. Prostě je všeobecně známo, že můžete teoreticky velmi dobře vědět, jak se jezdí na kole (vědět, že máte dělat to a to), ale při tom se neumět v praxi na kolo ani posadit (**jak** mám držet rovnováhu a zároveň šlapat?!)⁶. Nakonec úplně analogicky s barvami – můžete mít velmi dobré znalosti o procesu vnímání barev (od dopadu fotonů na sítnici až po procesy v různých oblastech vizuálního kortexu), ale samotný zážitek vnímání barvy je něco zcela jiného. Model není modelované, reprezentace není reprezentované. Propoziční obsahy jsou symboly, modely a to, k čemu se vztahují, jsou buď re-

⁵ Na termín „konkrétní“ je nutno dát důraz. Mentální stavy mají jistě svoji typiku, nicméně vzhledem k všeobecně známé plasticitě mozku je nutné brát v potaz, že identický mentální stav (identický ve smyslu „způsobující identický výstup organismu v jeho interakci se světem“) může být realizován mnoha způsoby (Gerald Edelman tuto vlastnost nazývá kuriózně „degenerací“).

⁶ Jen tak pro zkoušku si můžeme představit, jak by asi vypadal abstraktní propoziční popis chování vašeho těla při technicky komplikovaném sjezdu na kole. Zastánci možnosti funkční ekvivalence abstraktního propozičního popisu a reálného procesu prožívání by pak např. museli tvrdit, že znalost diferenciálních rovnic popisujících faktický proces je postačující a ekvivalentní s procesem „prožívání sjezdu na kole“. Hloubka problému spočívá v poslední instanci v jistém kognitivním atavismu – přesvědčení, že model či symbol může kompletně nahradit reálný modelovaný proces, symbolizovanou věc či událost.

prezentace nižší úrovně, nebo přímo fakta externí reality. V každém případě tyto úrovně reference nelze zaměňovat. Tyto triviální úvahy pak zdůvodňují Jacksonovo: „*just obvious that she will learn something about the world*“.

Jenže teď si musíme povšimnout právě jistého rozporu v zadání. Životnost druhé premisy spočívá právě v obecně sdílené intuici o nesouměřitelnosti komplexního smyslového zážitku organismu a propozičních znalostí. Pak ale není jasné, jak může mít první premisa o kompletních fyzikálních znalostech nějakou relevanci vzhledem k druhé premise?! Jinak vyjádřeno: V první premise se předpokládá souměřitelnost propozičních znalostí (fyzikální popis) a smyslového zážitku či pocitu. V druhé premise se tato souměřitelnost popírá. Závěrem tedy: První premisa musí mít relevantní vztah k druhé (bez první premisy by neexistoval ontologicky fatální závěr). „Mít relevantní vztah“ ovšem znamená mluvit o entitách souměřitelné úrovně. A tak tu máme spor.

Jako ukázkou, že celá záležitost je velmi jednoduchá a má i konceptuálně jednoduché pozadí, si připomeňme (na potvrzení našich intuicí) úvahu Davida Huma ...

Co by si o Marii v černobílém pokoji pomyslel David Hume

David Hume by letos oslavil třísté narozeniny a vzhledem k tématu, jímž se zabýváme, je vhodné připomenout jeho i z dnešního pohledu velmi naturalistickou teorii poznání.

Hume jednoduše tvrdí, že základní surovinou kognice jsou tzv. dojmy – živé vjemy. Z těchto dojmů se (ne zcela jasným způsobem) oddělují ideje, které jsou výrazně méně živé – jsou to svým způsobem kopie dojmů. Důležité je, že ideje – představy (tedy něco jako pojmy) jsou v poslední instanci vždy závislé na dojmech. Hume tvrdí (1972: 47) „*Laponec nebo černocho nemá představu o chuti vína. ... Člověk mírné povahy si nedovede představit nesmiřitelnou pomstu nebo krutost, podobně si sobecké srdce nedovede představit vznešenost přátelství a šlechtnosti. Rádi přiznáváme, že jiné bytosti mohou mít mnohé smysly, o nichž nemáme ponětí, protože jejich ideje nám nikdy nebyly sděleny tím jediným způsobem, kterým idea může vstoupit do ducha, do inteligence, totiž skutečným cítěním a zážitkem.*“ Hume tu zdůrazňuje, moderním slovníkem řečeno, fundamentální funkci základních biologických nastavení organismu. Přirozené morální sklony například nemohou vygenerovat „představu“ (ideu), která by byla v rozporu („sobecké srdce ... vznešenost a přátelství“) s touto přirozeností a zcela obecně; Neexistuje kognitivní, intelektuální obsah, který by neměl svůj původ v „zážitku“. Komplexní zážitek organismu je nutnou podmínkou ideje – tedy něčeho jako „pojmu“ či propoziční znalosti v moderní terminologii.

Z tohoto hlediska by se Humovi jevil problém Marie v černobílém pokoji jako krystalicky jasný. Podle Huma by byla Marie „dojmový“ (v humovském smyslu slova) deprivant. Platilo by pro ni ve vztahu k vnímání barev úplně to samé jako to, co platí pro „Laponce či černocho“ ve vztahu k chuti vína – nemá o ní žádnou **představu**, neví „jaké to je“. Co se tedy stane, když pustíme Marii z jejího černobílého pokoje do našeho barevného světa? V termínech Humovy filosofie platí, že v každém případě získá nové dojmy, není již ovšem vůbec jisté, jestli by dle Huma Marie také automaticky získala novou ideu či propoziční znalost⁷! Že Marie získá nové dojmy, je ovšem triviální fakt a je také triviální fakt, že

⁷ Pravděpodobně nezískala, protože singulární zážitek (vnímání červeně) nevytváří hned žádný asociační vzor. Jednotlivý zážitek prostě není zkušeností.

nový dojem (diference vzhledem k předchozímu prožívání černobílého světa) musí být chápán jako nová informace⁸. Zde se tedy můžeme ještě shodnout s Jacksonem: Ano, Marie skutečně získá novou informaci. Plyne ale z toho, že tato informace je nefyzikálního charakteru? Tvrdil by Jackson, že humovské „dojmy“ jsou nefyzikální entity? Vsadím se, že netvrdil.

Celá problematika příběhu o Marii ale ukazuje řadu zajímavých paradoxů. Všimněme si jen krátce, jak reagují někteří fyzikalisté (např. Daniel Dennett). V podstatě mají dvě možnosti (pokud zároveň neproblematizují vnitřní povahu konceptů experimentu, jako jsem to udělal já výše), které jsme již zmínili. Buď budou tvrdit, že Marie se nic nového nedozví (protože všechno již per definitione ví), anebo pokud se něco nového dozví, tak to znamená, že neměla kompletní fyzikální informace. Není opět těžké odhalit jistou naivitu této „fyzikalistické“ kritiky. V obou případech totiž spočívá na stejném předpokladu o souměřitelnosti (nebo dokonce ekvivalenci – teoretické a snad i praktické zaměnitelnosti) různých kognitivních kapacit organismu – přesně jako to předpokládá Jackson v první premise. Kdyby tomu tak nebylo, tak by první premisa vůbec nemusela být brána v úvahu. Pokud totiž přijmeme tezi o kvalitativní rozličnosti kognitivních domén propozičních obsahů a uvědomovaných smyslových komplexů, pak je zcela jasné, že žádné množství propoziční informace nemůže nahradit prožívaná smyslová data. Ještě jednou: ekvivalence či souměřitelnost znamenají, že model může bezesbytku zastoupit modelované, symbol symbolizované, reprezentující reprezentované, označující označované a nakonec myšlení to, co je myšleno. To je ovšem minimálně problematický předpoklad a připomíná prastarou filosofickou tezi o identitě „myšlení a bytí“, která je daleko spíše spojována s pojmem „idealismu“, než materialismu (fyzikalismu).

Qualia – naturalistická interpretace

Marie se dle Jacksonova předpokladu má dozvědět „jaké to je vnímat barvy“, tj. má získat nová qualia – vědomé fenomenální stavy. Filozofové včetně Jacksona mají tendenci tyto specifické kvalitativní mentální stavy vykládat jako entity speciálního druhu. Důvodem, proč tomu tak je, je velmi pravděpodobně jedinečnost privátních obsahů lidských myslí, jejichž popis je přístupný pouze z pozice první osoby, a systematicky nepřístupný z pozice třetí osoby. Jedinečnost subjektivních privátních mentálních stavů není redukovatelná na fyzikální popis v termínech obecných zákonů a speciálních podmínek. A co nepodléhá standardnímu fyzikalistickému popisu, to nemůže být „tak úplně“ fyzikální. Tak zní pravděpodobný závěr. Nutno dodat, že tento závěr nevyplývá – to, že se něco vzpírá⁹ dosavadním

⁸ Informaci budeme obecně chápat jako vnímanou rozdílnost (percipovanou diferenci) v nejobecnějším smyslu slova. Každý systém přírody na všech hierarchických stupních organizace, který reaguje na změnu, je příjemcem, původcem či mediátorem informace. Informace je fakticky způsobem systematické interakce mezi strukturami přírody.

⁹ Matematik Gregory Chaitin říká, že vysvětlení či pochopení jakéhokoliv procesu v přírodě spočívá v možnosti tento proces symbolizovat – převést na data a u těchto dat ukázat, že je možná jejich informační komprese (algoritmická stlačitelnost). Biologická evoluce a lidské dějiny jsou dost možná právě ten typ procesu, který nelze stlačit buď z toho důvodu, že jsou objektivně nestlačitelné (ireducibilní, v datech není žádná pravidelnost), anebo jsou nestlačitelné díky omezeným kognitivním kapacitám pozorovatele. To se může

možnostem popisu, neimplikuje, že máme co do činění s entitou jiného ontologického statusu, než jsou věci pozorovaného fyzikálního univerza. Pokud existuje hypotéza, která se odvolává na empirická zjištění a vysvětluje qualia, bude konstruktivnější se zaměřit na ni.

Damasiovy pocity jako základ qualií

Pokud se máme přiblížit k pochopení qualií, musíme v každém případě spojit struktury našeho introspektivního vnímání sebe samých s poznatky (neuro)vědy – tedy spojit poznatky z pozice první osoby s poznatky z pozice třetí osoby. Cílem pak je logicky teorie, která umí interpretovat slovník z pozice první osoby slovníkem z pozice třetí osoby.

Neurovědec Antonio Damasio nabízí teorii pocitů, která velmi dobře odpovídá našim intuitivním požadavkům na pochopení qualií. Chci říci, že chceme-li pochopit qualia, Damasiova teorie pocitů by měla být v každém případě součástí této pravděpodobně velmi komplexní teorie – Damasio (2004: 102-104) tvrdí: „*pocit je vnímání určitého stavu těla společně s vnímáním určitého způsobu myšlení i myšlenek s určitou náplní*“ A dále: „*pocity jsou funkčně odlišné od ‚seskupení myšlenek‘ ... podstata pocitů sestává z myšlenek reprezentujících tělo v průběhu procesu ... Odstraní-li tuto podstatu, pojem pocitu zmizí. ... pak by nám nemělo být již nikdy umožněno říkat ‚cítím se‘ šťasten, ale měli bychom říkat ‚myslím se‘ šťasten. ... Co udělá myšlenky ‚šťastné‘? Jestliže neprožívám jistý tělesný stav v jisté kvalitě, kterou nazýváme slasti a kterou v rámci běhu života shledáváme ‚dobrou‘ a ‚pozitivní‘, nemáme vůbec žádný důvod považovat jakoukoli myšlenku za šťastnou – nebo smutnou.*“

Z řečeného plyne, že význam termínů a vět (propozic), které se vztahují k našemu hodnotovému prožívání světa, je vázán na toto tělesné mapování, které je samozřejmě nepropozičního – nepojmového charakteru. Pocity jsou komplexní hierarchicky členěné entity, které mají co do činění s globálním procesem metareprezentování¹⁰, které provádí náš mozek v interakci s tělem. Otázka „co dělá myšlenky šťastné“ je klíčová. Je totiž jasné, že myšlenky samy o sobě to být nemohou. Pouze myšlenky ve spojení s mapami těla mohou být „šťastné“ či „nešťastné“. Opět se ukazuje, že Humova vize spojení dojmů a idejí je víceméně správná.

Pokud jde o qualia: Qualia jsou na úrovni introspekce definována jako „vědění, jaké to je“. Pokud na otázku „jaké to je?“ odpovídáme ve verbálních termínech pro pocitové stavy, je souvislost mezi qualií a Damasiovy pocity zřejmá. Pocity jako spojení tělesného mapování a pojmové reprezentace tohoto mapování (včetně omezené možnosti verbalizace) jsou v podstatě s qualií identické.

A nakonec v termínech bližších fyziologii – Damasio (2004: 104): „*Bezprostředním substrátem pocitů je mapování obrovského počtu tělesných stavů v sensorických oblastech určených k příjmu tělesných signálů.*“ Z toho paradoxně plyne, že pokud si tradičně filosofové představují qualia jako jakési éterické imateriální entity, pak se velmi pravděpodobně hluboce mýlí. Qualia jsou bytostně spojena s tělesnými strukturami.

týkat i procesů lidských myslí a to je také ten důvod, proč se mohou qualia vymykat popisu z pozice vědy (třetí osoby).

¹⁰ Metareprezentace znamená možnost reprezentovat samotné reprezentace.

Edelmanova koncepce qualií

Nositel Nobelovy ceny a jeden z předních odborníků v oblasti neurověd Gerald Edelman nabízí ještě širší koncepci qualií než Damasio – Edelman (2010:120): „*Qualia jsou diskriminace, rozlišení vyššího řádu a na vědomé scény lze nahlížet jako na řadu qualií. Taková řada je prožívána v kontextu obrovského množství událostí odehrávajících se v okolním prostředí, v těle a mozku samotném. Rozmanitost a potenciální souběžnost těchto událostí jsou na vysoké úrovni a qualia, která jsou zahrnuta v integrovaných, a přesto rozličných scénách, s sebou nesou širokou škálu prožitků. Zahrnují tak vnímání, představy, vzpomínky, pocity, emoce, nálady, myšlenky, přání, pohybové scénáře a množství neurčitých tělesných stavů.*“ Domnívám se, že obecně odpovídá naší introspektivní zkušenosti, že qualia jsou mnohoúrovňové obsahy – procesy. V tomto ohledu panuje pozoruhodná shoda mezi Damasiem i Edelmanem. Hume rozlišuje dojmy a ideje, přičemž pod termín „dojem“ se vejde celá řada entit a procesů našeho kognitivního života, pro které již máme jemnější terminologii. I Hume by ovšem byl zastáncem konceptu qualií jako hierarchické struktury různých kognitivních domén.

Qualia se dle Edelmana rozprostírají v celém poli kognitivních akcí organismu – od pojmových struktur (Humových idejí) až k emocím (dojmům) a i níže. Tento „blending“ kognitivních obsahů má dle Edelmana jasný evoluční původ. Qualia mají adaptivní funkci. Organismus, který má tak složitou kognitivní mašinerii, že může vytvářet qualia, získává selektivní výhodu. Qualia jsou v tomto pojetí velice komplexní scénáře obsahující data z mnoha úrovní kognice a jejich funkcí je hierarchizované hodnocení události. Díky tomu, že v qualiích jsou integrovány různé doménové úrovně poznání, je prožívaná událost efektivně kategorizována. Dovolím si malé vysvětlující přirovnání: Qualia můžeme přirovnat k výsledku z procesu rozhodování týmu rozhodčích (fakticky řady hierarchicky uspořádaných kognitivních modulů¹¹), jejichž cílem je prožívané události dát výslednou známku (jako to dělá tým rozhodčích při krasobruslení např.). Quale je právě touto výslednou známkou. Naš vědomý i podvědomý život je neustále provázen hodnocením událostí týkajících se našeho organismu. Pod termínem „hodnocení“ rozumím obecně *rozhodování o biologicky hierarchizovaných sekvencích*¹² celkové interakce organismu s okolním prostředím. Qualia jako komplexní a vědomě přístupné obrazy těchto širokých a mnohadimenzionálních interakcí vytyčují historické momenty prožívání individuálního (lidského) organismu. Jsou tak fundamentálním principem individuální integrity lidského organismu v plynoucím čase.

¹¹ Jako příklad modulární hierarchie můžeme uvést tuto pyramidu (odspoda nahoru): imunitní odpovědi, základní reflexy, řízení látkové výměny, homeostáze, chování odpovídající bolesti a strasti, pudy, pohnutky, motivy, primární emoce, sociální emoce, pocity, qualia, názor-představa, pojmy, propozice – mentalese, přirozený jazyk slov a vět.

¹² Tímto ne zcela jasným spojením myslím jednoduchou věc: biologický organismus musí nejprve vyřešit problémy týkající se bezprostředních požadavků na přežití a sebe-reprodukcí. Až po dosažení homeostáze v těchto základních biologických determinantech se může orientovat na nadstavbové procesy – to je podstata pojmu „hodnocení“.

Co se Marie dozví

Potom, co jsem zde nastínil, bych rád odpověděl na otázku, kterou si klade náš experiment: Co se Marie doopravdy dozví¹³ po svém propuštění z černobílého pokoje?! Na otázku odpovím rovnou: Marie získá nové stavy typu „jaké to je“, ale nebude je umět přiřadit k doméně sebebohatších propozičních znalostí, pokud nejsou splněny určité podmínky. To bych rád vysvětlil:

Představme si např., že Marie vyjde ven ze své černobílé místnosti a uvidí poprvé v životě rudé tulipány (tvar, velikost apod. – primární kvality postačují k identifikaci tulipánu). Co může maximálně usoudit, je to, že pokud jsou to tulipány (primární kvality bezpečně pozná), pak jsou velmi pravděpodobně červené a tomuto komplexnímu zážitku odpovídá určitý neurobiologický proces, který dokáže dokonale popsat. Vtip je v tom, že Marie neví, jaké neurobiologické procesy v ní při procesu vnímání probíhají, protože samotné komplexní neurobiologické procesy nevstupují do domény vědomí a sebevědomí (nejsou přístupné introspekci)¹⁴ – tedy do domény propozičních informací! Marie se na sebe potřebuje nejprve podívat z pozice třetí osoby a naměřit na sobě všechny relevantní charakteristiky procesu vnímání rudého tulipánu – prostě potřebuje pozorovací přístroje nebo druhou osobu, která jí řekne, jak vypadá její neurobiologický stav, o kterém z hlediska sebevědomé činnosti nemůže nic vědět. Pak teprve může říci, že tulipán je opravdu červený (a ne zcela jiný) a získané quale tak efektivně spojit s pojmem „červeně“ a popisem neurobiologického stavu. Do té doby ví v podstatě jen to, že termínu „červená“ odpovídá komplementární popis v termínech fyziky, chemie, neurobiologie apod. V rámci samotného reálného prožitku vnímání barev však neexistuje dostatečný důvod, proč by Marie daný aktuální zážitek neměla spojovat s jakýmkoliv možným popisem – ať již v termínech lidové psychologie nebo v poslední instanci v termínech fyziky. To je zřejmý důsledek faktu, že introspekce (a obecně vědomé procesy) má ve skutečnosti nesmírně omezený dosah a může operovat jen ve své doméně. Vědomá činnost prostě „nevidí“ do biofyzikálních procesů na mikroúrovni, které reálně konstituují proces vnímání (barev) a které by Marii na základě jejích znalostí dovolili identifikovat strukturu jejího aktuálního vjemu. A tak jsou Marii dopředu naučené propoziční informace z pozice třetí osoby zcela k ničemu při reálném aktu vědomé identifikace nových barev. Marie prostě sama o sobě nemůže vědět, co se v ní odehrává na úrovni biologie, chemie a fyziky.

Marie, která vyleze ze svého frustrujícího pokoje, tedy nepozná vůbec žádnou (kromě bílé a černé přirozeně) barvu, i kdyby měla popis všech neuronálních korelátů barevného vnímání. Získá ovšem nové dojmy (stavy „jaké to je“), které musí začlenit do svého vědění – a to nemůže udělat sama nebo bez pomoci přístrojů. Qualia jsou až výsledkem spojení onoho „jaké to je“ a propozičního vědění, které jediné má Marie k dispozici.

¹³ V rámci interpretační vstřícnosti zde připustíme předpoklady experimentu, které, jak jsem ukázal, jsou fakticky nepřipustné...

¹⁴ Kdyby tomu bylo naopak, pak by se introspektivní vhléd musel týkat i biochemických a možná ještě nižších úrovní procesů organismu, což vede ke zjevným absurditám. Konec konců neexistovala by principiálně pak žádná překážka, kde by se introspektivní poznávání mělo zastavit – což je mimochodem v rozporu s triviálním rozlišením na pohled z pozice první a třetí osoby. To, myslím, nepotřebuje dalšího komentáře, i když to otevírá prostor pro zajímavé metafyzické spekulace.

Co dál s qualii?

Realistická (tj. naturalistická) analýza našeho experimentu nás přivedla k několika zjištěním, které je dobré zevšeobecnit.

Především jsme zjistili, že qualia nelze nahradit propozičním popisem prožívání. Je to tím, že naše mysl (mozek a tělo) má hierarchickou a doménovou strukturu. Každá doména v hierarchii má evoluční původ a plní svoji funkci, jejíž stupeň důležitosti (hodnoty) je obecně dán pozicí této domény mezi jinými doménami ve fylogenetické historii biologických organismů. Doména procesu introspektivního vhledu či doména propozičních obsahů jsou z hlediska dějin utváření kognitivních kapacit druhů naprostými inovacemi. Jejich nejvlastnější funkcí je proces reprezentace a metarepresentace, což mohou být abstraktní operace, které vytváří dojem naprosté nezávislosti na ostatních doménách, které jsou striktně spjaté s časoprostorovými či termodynamickými omezeními organismu. Nicméně v našem typu vesmíru platí určitá zcela obecná pravidla. Jedno toto fundamentální pravidlo můžeme nazvat „asymetrií vazby“. Toto pravidlo jednoduše říká, že podkladová doména silněji determinuje doménu svrchní či emergentní – kvarky determinují strukturu atomů, a ne naopak, nukleové kyseliny determinují bílkovinné struktury buňky, a ne naopak, a nakonec v oblasti mentálních entit smyslové obsahy determinují intelektuální obsahy – jako to předpokládal kdysi Hume. To je obecné pravidlo. Interakce mezi doménami různé úrovně existuje, ale je asymetrická – nerovnocenná. Proto také znalosti nemohou nahradit qualia jako entity obsahující i domény níže v hierarchii. Samotná operace „nahrazování“ (substituce) patří striktně do domény vědomých obsahů – propozičních obsahů. Jen v případě této domény může být obsah vyjádřen několika způsoby (jako např. v matematice jedno číslo můžeme vyjádřit nekonečně způsoby za použití základních operací aritmetiky a konceptu reálných čísel). Tato vícená realizace jednoho obsahu ale není možná v doménách nižších a biologicky fundamentálnějších – emoci strachu není možné vyjádřit několika různými způsoby. To je jeden ze základních rozdílů ve fungování domén naší mysli.

Proto je možné obecně konstatovat, že dané quale, obsahující strukturu smyslových obsahů provázených myšlenkovými obsahy, je díky svému založení v nižších kognitivních strukturách možné realizovat jen za předpokladu úplné základní báze smyslových elementů. Pokud vám tedy jako např. Marii chybí elementy barevného vnímání, pak velmi pravděpodobně neexistuje cesta, jak to nahradit jinak než přirozenou zkušeností. V tomto ohledu bude zajímavé zkoumat synestetické procesy, kdy jeden typ smyslového počítka je částečně převáděn a spojován s počítkem jiného smyslového typu. Ale i zde se dá předpokládat, že tam, kde chybí kompletní báze (jako u Marie), nebude synestézie možná.

Na závěr bych zmínil ještě jednu důležitou věc. Na příkladu qualií se ukazuje, jak moc jsou procesy naší mysli „vtělené“. Tělesná zkušenost organismu představuje velmi pravděpodobně nutnou podmínku existence qualií. Na druhou stranu je jasné, že to není podmínka postačující. Jak jsme řekli, qualia jsou komplexní stavy prolínajících se kognitivních domén a tělesné mapování bude jen nutnou podkladovou strukturou. V každém případě mají tato zjištění jistou hodnotu. Vedou nás totiž k otázce, zda může vůbec existovat vědomý obsah, který by nebyl v poslední instanci spjat s tělesným mapováním a tedy s něčím, co na jedné straně tvoří jádro pocitů (dle Damasia) a na druhé straně je předpokladem intuitivní geometrie, skrze kterou interpretujeme svět. Řada výzkumů v oblasti kognitivních věd naznačuje, že to možné není. George Lakoff to demonstruje na tak abstrakt-

ní disciplíně, jako je matematika – viz Lakoff, Núñez (2000:315): „*If mathematics is taken to include the ideas that arithmetic expresses – that is, if calculus is taken to be about something – namely, change – then Weierstrass did not eliminate geometry at all. From the conceptual perspective, he just hid it. From the perspective of mathematical idea analysis, no one could eliminate the geometry metaphorically implicate in the very concept of change in classical mathematics.*“ A stejně jako není pravděpodobně možné eliminovat geometrickou představivost v oblasti matematiky, není možné ani eliminovat pocity jako základ umělecké tvorby; Arnold Gehlen (1972:51) cituje Stephana Mallarmé: „*báseň nevzniká z pocitů, ale ze slov.*“ A dodává: „*A to především znamená, že se básník distancuje od bezprostředního přírodního faktu v sobě i kolem sebe, neboť, barvy a zvuky se v přírodě vyskytují, ale slova ne*“ (G. Benn).“ Představy modernistických umělců jsou v tomto ohledu v rozporu s našimi tvrzeními – berme je však spíše jako revoltu proti tradicionalismu a konzervativismu. Ve skutečnosti všichni tušíme, že umění je univerzální nástroj vyjadřování a sdělování pocitů – dlouhá tisíciletí zkušeností civilizovaného lidstva to potvrzují. Umění je v tomto smyslu jazykem pocitů a tedy i qualií. Je pak logické, že budeme-li se chtít něco více napříště dozvědět o qualiích, měli bychom prostředky kognitivních věd zkoumat umění.

Bibliografie

- CHAITIN, G. (2007). *Metamaths*. London: Atlantic Books.
 DAMASIO, A. (2004). *Hledání Spinozy*. Praha: Dubbuk.
 EDELMAN, G. (2010). *Širší než obloha*. Praha: Paseka.
 GEHLEN, A. (1972). *Duch ve světě techniky*. Praha: Svoboda.
 HUME, D. (1972). *Zkoumání lidského rozumu*. Praha: Svoboda.
 LAKOFF, G. & NÚÑEZ, R. (2000). *Where Mathematics Comes from*. New York: Basic Books.