

Kognitivní kontrarevoluce?

Cognitive Contrarevolution?

Jaroslav Peregrin

Filosofický ústav
Akademie věd České Republiky
Jilská 1, 110 00 Praha 1
&
Filozofická fakulta
Univerzita Hradec Králové
Rokitanského 62, 500 03 Hradec Králové
jaroslav.peregrin@uhk.cz

Abstrakt/Abstract

Ve standardních výkladech moderních dějin studia mysli ve dvacátém století se dočteme, že zatímco kolem poloviny tohoto století ovládl studium mysli zpozdilý behaviorismus, v šedesátých letech nastoupila „kognitivní revoluce“, která nadvládu behaviorismu smetla a otevřela cestu ke skutečně nepředpojatému a adekvátnímu studiu mysli. V tomto textu se chci nad tímto standardním výkladem zamyslet a zpochybnit ho: konkrétně chci poukázat na to, že behaviorismus nebyl ve všech ohledech tak zpozdilý, jak by se z tohoto pohledu mohl jevit; a že „kognitivní revoluce“ neznamenal jenom vymanění se z nepřiměřených metodologických okovů, ale i otevření prostoru pro to, čemu budu říkat „magická teorie mysli“ a co podle mého názoru nepatří ani do vědy, ani do rozumné filosofie.

Standard accounts of the modern history of the study of mind in the twentieth century tell us that while around the half of the century the study came to be dominated by inept behaviorism, the sixties witnessed the onslaught of the "cognitive revolution", which wiped out the dominance of behaviorism and opened the way for the truly unprejudiced and adequate study of mind. In this text I want to reconsider this account and challenge it: in particular, I want to point out that behaviorism was not in all respects so inept as it may suggest; and that the "cognitive revolution" did not mean merely breaking out of inadequate methodological fetters, but rather also an opening of space for what I will call the "magical theory of mind" and what I think belongs neither to science, nor to reasonable philosophy.

Kognitivní revoluce?

Standardní výklad moderních dějin studia mysli ve dvacátém století je takový, že v průběhu tohoto století ovládl studium mysli behaviorismus, který teoretiky mysli nepatřičným způsobem omezoval, a tak musela přijít *kognitivní revoluce*, která behaviorismus svrhla, jím uvalená omezení odstranila, a otevřela cestu ke skutečně hlubokému studiu mysli. Co to byl, z historického hlediska, *behaviorismus*? V

psychologii je spojován zejména se jmény J. Watsona a B. F. Skinnera¹ a je především, jak to charakterizuje Graham (2010), „normativní teorie o tom, jak je třeba vědecky provozovat psychologii“. Tito behavioristé se snaží z psychologie vypudit vše to, co zavání nevědeckými spekulacemi o neviditelných mechanismech za lidským chováním, o ‚myšlenkách‘, ‚přesvědčeních‘, ‚přáních‘ apod. (které nikdo nikdy neviděl, ale které jsou přesto údajným předmětem psychologie), a ponechat v ní pouze studium objektivně měřitelného a hodnotitelného lidského chování. Vedle tohoto psychologického behaviorismu ovšem existuje i behaviorismus filosofický, spojený především se jménem G. Ryla², kterému jde hlavně o odpověď na otázku *co to je mysl*.

Podívejme se nyní na instruktivní příklad způsobu, jak bývá kognitivní revoluce obvykle prezentována (Plháková, 2006, s. 227):

„Kognitivní psychologie se zrodila v 50. a 60. letech v USA a posléze zdomácněla také v Evropě. V první polovině 20. století v USA dominoval behaviorismus, který odrazoval psychology od zkoumání nepozorovatelných poznávacích procesů. Během 50. a 60. let se situace postupně začala měnit. Do USA pronikla teorie kognitivního vývoje Jeana Piageta. George Miller publikoval své výzkumy paměti, práce Noama Chomského přispěla k rozvoji lingvistiky. Jerome Bruner zkoumal formování pojmů, objevily se první pokusy o počítačovou simulaci lidské inteligence. Zakladatelé kognitivní psychologie zdůrazňovali, že chtějí-li psychologové pochopit chování, musí zkoumat subjektivní mentální události. Vnitřní manipulace s psychickými událostmi má na chování výrazný vliv, takže výlučné zaměření na vnější projevy poskytuje pouze neúplný obraz o tom, proč se lidé chovají tak, jak se chovají.“

Takovýto pohled, obávám se, vyznívá následujícím způsobem: lidské chování je (více či méně) výsledkem „vnitřní manipulace s psychickými událostmi“. Jsou tu tedy jakési „poznávací procesy“ a „subjektivní mentální události“, které jsou sice nepozorovatelné, nicméně jsou skutečné. Behavioristé nám sice zakazovali se jimi zabývat, ale my jsme se dnes již – díky kognitivní revoluci – dokázali tomuto zákazu vzepřít a prozkoumáním tohoto subjektivna jsme získali potřebný obraz o *příčinách* lidského chování.

Obávám se, že tohle je způsob, kterým mnoho psychologů i filosofů myslí „kognitivní revoluci“ skutečně vidí; tento způsob ale vzbuzuje mnohé otázky. Proč nám behavioristé zakazovali ony „subjektivní mentální události“ zkoumat? Protože jsou „neviditelné“? Ale ve vědě se setkáváme se spoustou věcí, které jsou – z toho či onoho důvodu – neviditelné (ať to jsou třeba síly, které jsou neviditelné svou podstatou, nebo atomy, které jsou pro nás neviditelné, protože jsou příliš malé, nebo vzdálené galaxie, které nevidíme, protože jsou příliš daleko) a nikoho by ani nenapadlo zakazovat je zkoumat. Abychom mohli říkat, že tu *jsou*, i když jsou neviditelné, musíme samozřejmě

¹ Viz Watson (1919), Skinner (1938, 1953).

² Viz Ryle (1949).

mít nějaké způsoby, jak je *detekovat* (jak objektivně prokazovat jejich existenci) – ale pokud tomu tak je, není s jejich neviditelností žádný zásadní problém.

Jak můžeme detekovat *mysl*? Co nás vede k závěru, že nějaká entita *myslí*? Jak na to spektakulárně poukázal Alan Turing (1950), nemůže to být nic jiného než to, že se tato entita určitým způsobem („rozumně“) chová – že s ní můžeme rozmlouvat, že jsou její projevy určitým způsobem koherentní, že reaguje na vnější podněty určitými specifickými druhy způsobů³. Situace, kdy chování prohlašujeme za projevy neviditelné mysli se tedy zdá být nikoli nepodobná té, kdy fakt, že předměty padají k zemi, připisujeme působení neviditelné přitažlivé síly.

Je tu ale podstatný rozdíl. Přitažlivost nepovažujeme za nějakou věc, která by kolovala někde „v žilách“ zeměkoule, ale v podstatě prostě jenom právě za její schopnost přitahovat předměty. Analogicky bychom museli říct, že *mysl* není nic jiného než schopnost lidského organismu produkovat určitý druh chování (chování, které vnímáme jako „inteligentní“, „smysluplné“ apod.). To je ale v podstatě to, co říkali behavioristé – ti odmítali připustit, aby bylo chování vykládáno jako pouhý projev něčeho vnitřního a na něm nezávislého⁴. Tím vědec prostě překračuje mandát, který mu dávají jeho data.

Ale nejsou ona chybějící data, která nám dovolují jít za hranici, kterou vytyčili behavioristé, dodávána *introspekci*? Není tomu tak, že *mysl*, kromě toho, že můžeme studovat její komplikované projevy, také *vidíme*, když se, podíváme do sebe? Je třeba zdůraznit, že takové „dívání se do sebe“, neboli *introspekce*, nebylo po té, co bylo jako nepřijatelná metoda odmítnuto behavioristy, nikdy fakticky rehabilitováno, a to ani „kognitivními revolucionáři“. Je totiž zřejmé, že prohlašovat to, co člověk údajně vidí „v sobě“, za *data* je zásadně problematické (prohlášení jediného člověka nepotvrzené z žádných jiných zdrojů se normálně za bernou minci nebere a v případě introspekce je potvrzení z nezávislého zdroje principiálně nemožné). Tohle byl po behavioristech ochoten zpochybnit málokdo, a je tedy třeba vidět, že o čemkoli mohly být spory mezi behavioristy a „kognitivními revolucionáři“, o introspekci to rozhodně nebylo. (To je třeba si uvědomit zejména v souvislosti s výše citovaným výrokiem, že „zakladatelé kognitivní psychologie“ chtěli „zkoumat subjektivní mentální události“.)

Podívejme se, jak je obvykle definován obor *kognitivní psychologie*. Jenom několik příkladů: Podle Wikipedie je to „podobor psychologie zkoumající vnitřní mentální procesy“. Učebnice *Cognitive psychology: a student's handbook* uvádí, že tento obor se „zabývá vnitřními procesy, které jsou součástí toho, jak chápeme své prostředí a rozhodujeme se, jaké činy by mohly být vhodné.“⁵ Co tu znamená slovo

³ To se stalo předmětem jeho od té doby vášnivě diskutovaného návrhu, že to, zda daná entita *myslí*, lze měřit tím, čemu se po té, co to Turing navrhl, začalo říkat Turingův test - viz Saygin, Cicekli & Akman (2000).

⁴ Gilbert Ryle tohle odmítl jakožto „dogma ducha ve stroji“ (viz Ryle, 1949, s. 22). Podle něj máme tendenci vidět v čemkoli, co funguje příliš složitě na to, abychom to dokázali jednoduše pochopit, nějakého „ducha“, který tím hýbe.

⁵ Viz Eysenck & Keane (2005).

„vnitřní“? Zdá se být zřejmé, že nejde o vnitřek *lebky*, to jest o procesy, které se odehrávají v mozku (jejich studium se také obvykle považuje za neurologii, nikoli za psychologii), ale o jakýsi imaginární vnitřek naší *mysli*, který obývají myšlenky a jiné mentální obsahy.

Nahlédněme do autoritativní Sternbergovy učebnice kognitivní psychologie, která vyšla i v českém překladu. Tam se říká, že „kognitivismus“ je „přesvědčení, že většina způsobů lidského chování může být pochopena, zjistíme-li, jak lidé myslí“. To je ale tvrzení, se kterým by – v nějakém jeho smyslu – souhlasil asi každý. Námitkou behavioristů je ovšem to, že zjistit, jak lidé myslí, nemůžeme jinak, než že pozorujeme, jak se chovají, a tím nás takováto tvrzení vedou do nebezpečné blízkosti bludného kruhu. Sternberg konstatuje, že „většina behavioristů se vyhýbala pohledu do ‚černé skříňky‘ lidské mysli a zaměřovala se pouze na pozorovatelné chování“. Problém je v tom, že tato formulace bere za hotovou věc, že do lidské mysli lze nahlížet nějak příměji (jestliže to behavioristé mohli *odmítat*, tak to muselo být alespoň v principu možné), než skrze to chování – což je právě to, co behavioristé popírali.

Zdá se mi tedy, že poměrně běžný výklad „kognitivní revoluce“ jako události, kdy revolucionáři odstranili bariéry, kterými nám behavioristé svévolně zahradili přístup k legitimnímu předmětu psychologických a filosofických výzkumů, totiž „vnitřní manipulaci s psychickými událostmi“, je značně problematický. Je nepochybné, že svržení behaviorismu znamenalo odstranění některých zcela nejapných metodologických omezení, které na studium mysli behavioristé uvalili, a v tomto smyslu skutečně otevřelo prostor pro nové a plodné cesty výzkumů. Mám ale podezření, že se s vaničkou příliš restriktivního behavioristického paradigmatu vylilo i dítě rozumně střízlivého přístupu ke studiu mysli.

Primární intuice

My všichni máme nějaké základní intuice ohledně našeho vlastního vědomí. Myslím, že je lze přibližně shrnout do následujících bodů:

- Každý z nás má mysl, která existuje nějakým způsobem diametrálně odlišným od toho, jak existuje jeho tělo.
- Každý si je své mysli vědom (= má vědomí), může do ní nahlížet.
- V mysli jsou k vidění různé myšlenky, pocity, přání apod.
- Mysl je s tělem nějak integrálně propojena.

Shrňme tyto body pod hlavičku *primární intuice* (ohledně mysli).

Pokud se přijmou nekriticky, vedou tyto intuice k představě dvou úrovní existence: zatímco naše těla existují v časoprostoru, kde podléhají kauzálním zákonům, naše mysli existují nějakým jiným způsobem (určitě ne v prostoru – ač obvykle říkáme, že mysl máme „v hlavě“, nikdo si asi nemyslí, že by se někde uvnitř hlavy dala najít

komůrka, ve které se jako ryby prohánějí myšlenky, nápady, přání, představy ...). Tohle je obrázek, jehož explicitní formulací se do dějin filosofie zapsal René Descartes⁶. Podle něj je člověk součtem dvou substancí: těla, které je tzv. *res extensa* („věc rozprostraněná“), to jest je to fyzická věc mezi ostatními fyzickými věcmi našeho fyzického světa, a mysli, která je *res cogitans* (tj. „věc myslící“) a existuje nějak mimo tento svět.

Problémy takového *dualismu* však vplynuly na povrch velmi rychle (nejkřiklavějším z nich je asi problém toho, jak by dvě substance, vyskytující se ve zcela mimoběžných rovinách existence, dokázaly ‚spolupracovat‘ a vytvořit takovou jednotu, jakou je člověk; viz o tom podrobněji níže) a mezi lidmi, kteří se dnes myslí do hloubky zabývají (ať už jsou to vědci nebo filosofové), by už bylo těžko hledat někoho, kdo by byl ochoten se k takovému karteziánskému dualismu explicitně hlásit. Téměř všichni se shodují, že hovořit o mysli jako o nějaké věci, která existuje stejně bytostně jako tělo či jeho části, je nanejvýše nějaký druh metafory – mysl je, podle téměř obecně přijímaného názoru spíše něco takového jako činnost mozku. To, jak mnozí současní filosofové a kognitivní vědci k mysli fakticky přistupují, však naznačuje, že implicitně nějaká forma karteziánského dualismu jejich uvažování o mysli více či méně determinuje⁷.

Moderní věda nám svět předvádí jako komplikovanou síť prostoročasových entit propojených neustálými kauzálními interakcemi (to jest interakcemi, které se řídí zákony příčin a účinků, jaké odkrýváme v rámci přírodních věd). Spousta opovržlivého toho už bylo napsáno na téma „mechanisticko-redukcionistického“ obrazu, který nám tak věda „vnucuje“⁸ – já mám ale pocit, že věda sama je v tom dost nevině. Ta buduje určitý obraz světa, takový obraz, jaký jí vychází z metod, se kterými ona pracuje a které mají většinou co dělat s měřením, vážením apod., ale nezdá se mi, že by nám ona sama vnucovala nějakou redukci v tom smyslu, že by nám bránila přemýšlet o světě i jiným způsobem. (Nějací ti *fachidioti*, kteří by k tomu měli tendence, se samozřejmě najdou, to se mi ale zdá být spíše okrajová záležitost.) Vědecký obraz světa je jeden možný obraz; je sice v mnoha ohledech mimořádně pozoruhodný a v některých i zásadně důležitý, ale rozhodně ne jediný možný: básníci, filosofové, kněží i ostatní mohou předkládat svoje vlastní obrazy, které se mohou s tím vědeckým všelijak překrývat, doplňovat, ale mohou s ním i soupeřit.

Jistě můžeme udělat třeba to, že vědecký obraz světa odsuneme někam do pozadí (možná jenom pro nějaké specifické účely) a upřednostníme nějaký jiný – třeba nějaký animistický obrázek, na kterém je svět kolbištěm vesměs oduševnělých entit. (Takový

⁶ Viz. Robinson (2011).

⁷ Faktem je, že přijde-li explicitně na přetřes povaha mysli, jsou téměř všichni ti, kdo se myslí vážně zabývají, monisty a materialisty; *de facto* se však myslí zabývají tak, jako by to byla svébytná, nehmotná věc, kterou můžeme (svým ‚vnitřním zrakem‘) pozorovat docela tak, jako vnějším zrakem pozorujeme vnější svět.

⁸ Tento názor se často objevuje u autorů blízkým synkretismu tzv. *new age* (viz např. Bohm, 1980, či Capra, 1984), ale rezonuje i v díle mnohých filosofů.

obrázek nás pak může třeba užitečně vést k respektu k přírodě ...) Problémem ovšem může být, že onen obrázek, který nám podává věda, je v nás v některých ohledech zakořeněný mnohem hlouběji, než by se nám mohlo zdát a než si jsme ochotní připustit, a že zcela se od něj oprostít může být docela těžké. Co ale nemůžeme v žádném případě, je vzít nějaké kousky toho, co se nám líbí více než ten vědecký obraz, a ty do toho vědeckého implantovat. Ten vědecký obraz totiž tvoří velmi integrální a uzavřený celek, do kterého nejde jen tak něco cizorodého zasadit.⁹

Představme si například, že bychom se k tomuto obrázku rozhodli přidat nějaké substanciální, ale nečasoprostorové myslí či duše. Pokud to nemá být obohacení pouze ornamentální, pokud mají být tyto myslí v rámci výsledného obrázku nějak skutečně činné, budou muset být nějak propojeny s těly, nějak vnímat, co se s těmito těly děje a naopak nějak řídit jejich chování. Tímto způsobem jsme ovšem rovnýma nohama ve všech těch problémech, na kterých ztroskotal karteziánský dualismus. Půjdeme-li zpátky po kauzálním řetězci od nějakého činu dané osoby (třeba stisknutí spouště pistole), budeme muset dříve nebo později na takovou mysl či duši narazit. Ale jak a kde? Stisk spouště je zřejmě způsoben stahem určitých svalů, a ten je zase způsoben pohybem určitých elektrických potenciálů po nervových drahách směrem od mozku do inkriminované ruky. Budeme-li hledat příčinu toho, proč se tyto potenciály vydaly z mozku do této ruky, budeme asi muset rozplétat fantasticky komplexní interakce mezi neurony mozku, při nichž potenciály tékají mezi nimi, a pokud by se nám je rozplést podařilo, najdeme asi příčiny některých z nich v pohybech nějakých potenciálů do mozku zvenčí, od smyslových receptorů. Atd.

Jenomže pokud má být ve hře netriviálně mysl, pak musí tento řetězec příčin a účinků někde nějakými ‚tajnými dvířky‘ uhýbat pryč z našeho časoprostoru do onoho odlišného rozměru, ve kterém mysl existuje, a pak se zase vracet. Jak a kde by (v mozku?) mohla taková tajná dvířka existovat? Jak by vůbec mohla vypadat? Descartes měl tip, že by něco takového mohlo být lokalizováno v té části mozku, které se říká *šišinka*; nicméně další vývoj zkoumání mozku a zákonů našeho světa nám čím dál tím víc zmenšuje prostor, v němž by šlo dát představě takových, transdimenzionálních vrátek‘ vůbec nějaký smysl.

Zdá se mi, že problém je v tom, že mnoho lidí, kteří se zabývají myslí, se chce jak držet vědy, tak si podržet obraz myslí jako nějaké věci, kterou lze prohlížet, zkoumat a testovat nezávisle na zkoumání hmoty, protože její povaha je nějak bytostně nehmotná. A zdá se mi, že nadšení, se kterým bývá obvykle opěvována ‚kognitivní revoluce‘, není vyvoláváno jenom tím, že nás tato revoluce zbavila nemístné metodologické svěrací kazajky, do které studium myslí zavřeli někteří přehnaně horliví behavioristé, ale také tím, že nám dovolila (nebo se zdála dovolit) se bez výčitek svědomí vrátit k onomu nekoherentnímu způsobu uvažování o lidské myslí, které nám sice může intuice vnucovat, které ale bylo behavioristy (a dalšími) po právu vykázáno za dveře.

⁹ Tím samozřejmě nechci říct, že by se tento obraz sám nevyvíjel.

Magické teorie mysli

V lidovém podání se primární intuice ohledně mysli zkrátka velmi snadno transformuje do čehosi, čemu budu říkat *magická teorie mysli*. Tato teorie vychází z představy, že hmota je ‚mechanická‘, a aby vzniklo něco ‚nemechanického‘, jako je člověk (a potažmo jakýkoli druh *významu*, protože význam nemůže vzniknout jinak než skrze člověka a jeho mysl), je třeba aby byla oplodněna něčím ‚nehmotným‘ či ‚nadpřirozeným‘, co do hmotného světa prosákne z jiného rozměru existence. Lidská mysl se pak jeví jako něco, čím musí být pouhý ‚mechanický‘ organismus oplodněn, aby vznikl ‚skutečný‘ člověk, a takto spirituálně oplodněný organismus se pak dále stává prostředníkem obcování ‚nadpřirozeného‘ s hmotou, místem oplodňování mechanických věcí významy. (Tady je také kořen oněch nářků nad tím, že by svět a my lidé v něm měl být vykládán jenom jako *mechanismus* – nedodá-li se k mechanismu nějaký virtuální magický šém, který z něj udělá *více* než jenom hmotu, nemůžeme mít skutečného člověka.)¹⁰

Proč mluvím o *magii*? Magickými nazývám věci, které něco způsobují, aniž by tak činily standardními, popsitelnými mechanismy, jimiž normálně příčiny způsobují účinky v našem reálném světě. Svět funguje standardně tak, že se věci dějí z nějakých příčin, které jsou v principu (i když třeba zdaleka ne vždy v praxi) popsitelné a klasifikovatelné. Způsobí-li třeba tenisová raketa to, že se míček vydá takovou či onakou rychlostí tím či oním směrem, je to jenom proto, že se ona sama pohybuje a nějakou rychlostí a nějakým směrem a tento míček odrazí. Raketa, která by způsobila pohyb míčku, aniž by se sama hnula, a obecněji věc, která by vyvolala nějaký účinek bez toho, aby ho v běžném slova smyslu zapříčinila, je magická; tak jako deštník Rubeuse Hagrida, který způsobí, že má Dudley Dursley najednou prasečí ocásek, aniž mu tento ocásek narostl nebo byl implementován způsobem, jakým se taková věc normálně může objevit.

Z tohoto hlediska není teorie mysli, která vidí mysl jako něco, co přesahuje kauzální vlastnosti lidského mozku či lidského organismu, a přitom to je schopno způsobovat věci v reálném světě, také ničím jiným než magií. Člověk, to je velmi složitý komplex biologických tkání, řízený nervovou soustavou s centrem v mozku; nicméně v podání magické teorie mysli kromě toho získává ještě jisté další kvality díky přítomnosti nehmotné substance jménem *mysl*. Přítomnost této substance nedovoluje, abychom ho vnímali jako (byť i fantasticky složitý) mechanismus a abychom to, co může způsobovat, redukovali na kauzální potenciál tohoto organismu jako takového. Je to tahle nehmotná, neprostorová substance, která proudí našimi vnitřky a která z nás činí ne pouze mechanické roboty, ale skutečné lidi, schopné, na rozdíl od robotů, produkovat významy. A byl-li behaviorismus zčásti motivován snahou nás od tendencí pouštět se do magických teorií mysli jednou pro vždy odříznout, pak mnoho lidí vidí v kognitivní revoluci znovuotevření cesty do těchto filosoficky neblahých krajín.

¹⁰ Představa člověka, jako pouhého – byť fantasticky složitého – mechanismu je zřejmě pro mnoho lidí – včetně některých filosofů – skutečným strašákem. Já osobně to vnímám opačně: fascinuje mě, co všechno dokáže takový mechanismus, jakým jsem já a jakým jsou jiní lidé, prožívat, vymýšlet a podnikat!

Co když si ale uvědomíme, že primární intuice nelze přijímat natolik nekriticky, aby nás vedly k magické teorii mysli? Co když doceníme fakt, že magická teorie mysli *není* slučitelná s našimi názory na zbytek světa? Slova „vnitřní“ a „mentální“ mohou být samozřejmě nadále použitelná jako metafory – pak je ale navýsost důležité mít na paměti jejich omezenou platnost. A především je třeba přemýšlet o tom, *pro co* to jsou metafory.

Mysl jako černá skříňka

Fakt, že standardně pozorujeme jenom ‚vstupy‘ a ‚výstupy‘ jiného člověka, to jest jeho chování či jeho projevy, a jenom na základě nich soudíme, že jde o myslící bytost, je zcela zřejmý. Samozřejmě to vnímáme nějak tak, že ‚vstupy‘ (to jest to, co daný člověk vnímá) a ‚výstupy‘ (co dělá) spolu nějak souvisí. (A faktem rozhodně je, že se nám to nejeví tak, že by tu byla nějaká přímá a nekomplikovaná souvislost, jakou chtěli, jak uvidíme dále, hledat Watson a Skinner – máme tendenci si tuto souvislost naopak představovat jako zprostředkovanou pomocí velmi složitého ‚soukolí‘ přesvědčení, přání atd., které jsou ‚vstupy‘ všelijak iniciovány a které naopak vedou k ‚výstupům‘.) Takovýhle způsob chápání jiných lidí je tím, co Daniel Dennett nazývá *intencionálním postojem*¹¹ a co mnoho současných filosofů připisuje speciálnímu modulu lidské mysli, tzv. *teorii mysli*¹². V souvislosti se spleť přesvědčení, přání atd., kterou máme tendenci ‚vidět‘ v myslících bytostech se také někdy hovoří o ‚lidové psychologii‘¹³.

Ať tak či onak, je třeba mít na paměti, že postulujeme-li něco jako ‚vnitřní strukturu‘, která zprostředkovává vztah mezi ‚vstupy‘ a ‚výstupy‘ lidského organismu, pak to nečiníme na základě toho, že bychom ji nějak pozorovali, ale děláme to z pozice, z jaké odhadujeme vnitřek černé skříňky. Můžeme pak takovou černou skříňku otevřít a konfrontovat to, co je v ní, s obrázkem, který jsme si takto vytvořili? V jistém smyslu můžeme – můžeme samozřejmě zkoumat, jak je vztah mezi ‚vstupy‘ a ‚výstupy‘ zprostředkován nervovou soustavou a mozkiem. Tak ale rozhodně dostaneme obrázek, který se nebude krýt s tím, který si děláme ‚intuitivně‘ – s oním obrázkem, jehož základními prvky jsou přesvědčení, přání atd.

Můžeme pak samozřejmě říci, že obrázek, který si o ‚vnitřku‘ člověka děláme z intencionálního postoje, je jenom velmi hrubou, neumělou, a proto velmi nepovedenou verzí toho, co potom ve skutečné podobě poznáváme, když studujeme mozek. Z tohoto pohledu je ‚lidová psychologie‘ jenom naivní verzí neurologie, stejně tak jako je ‚lidová fyzika‘, to jest soubor principů, podle nichž se řídíme v našem každodenním obcování s fyzickým světem, jenom naivní verzí teorií, které vypracovali Newton, Einstein a další.

Dramatický rozdíl mezi tím, jak lidské mysli chápeme intuitivně, a tím, co nám neurofyziologové říkají o lidských mozcích, nám ale vnucuje představu, že když

¹¹ Viz Dennett (1987).

¹² Viz např. Carruthers a Smith (1996).

¹³ Viz Ravenscroft (2010).

uvažujeme o mysli, uvažujeme přece jenom o něčem jiném než o mozku. A odtud je už jenom krok k nějaké té „mentální substanci“. To je pochopitelně velmi podporováno pocitem, že tuto substanci přímo „vidíme“ v akci, když nahlédneme do sebe. Výsledkem ale bývá magická teorie mysli: implicitní teorie nehmotné, nečasoprostorové substance, která je nicméně kauzálně účinná, to jest ač je nedetekovatelná, je schopná organismus naplnit kauzálními silami, které nemá sám ze své podstaty.

Myslím, že rozumná cesta chápání mysli vede mezi těmito dvěma extrémů: mezi názorem, že onen vír přesvědčení, přání a dalších „mentálních“ entit, které jakoby „pozorujeme“ z intencionálního postoje (případně prostřednictvím introspekce), je jenom naším neumělým odhadem toho, co dělá mozek příslušného organismu, a mezi názorem, že se týká něčeho esoterického, co nemá s mozkiem nic společného. Myslím, že to můžeme vidět tak, že „lidová psychologie“ se týká čehosi, co je podloženo činností mozku, co je ale zachycováno na takové úrovni abstrakce, že to není přeložitelné do žádného neurofyziologického povídání o mozku¹⁴. Relativní nezávislost mysli na mozku tak *de facto* spočívá v tom, že je mozek tak složitý, že některé relevantní aspekty jeho fungování nelze než uchopit jazykem, který není reálně převoditelný do jazyka neurofyziologie (tak jako např. fungování šachového programu nemůžeme vysvětlit v jazyce strojového kódu počítače, na kterém běží, i když v poslední instanci není tento program ničím jiným než takovým strojovým kódem). Relativní závislost je věcí toho, že jakékoli projevy organismu nemohou být ničím jiným než důsledky příčin, které se zpravidla odehrávají v mozku.

Slasti a strasti behaviorismu

Jak už jsem řekl, byl nástup behaviorismu podle mého názoru do velké míry motivován snahou o to, aby se z teorie mysli vyloučilo to, co jsem výše nazval *magií*, a aby byla dovedena do podoby podniku začlenitelného do hlavního proudu vědy, ve kterém dnes přestalo být místo pro karteziánské „přemítavé substance“. Ve své filosofické podobě, zejména u Gilberta Ryla, se soustředí na snášení důvodů pro to, že vidět mysl jako nějakou věc, jako nějakou substanci, je jenom jakási škodlivá iluze.

Rylovou klíčovou myšlenkou bylo to, že vidět mysl jako nějakou další substanci vedle těla je *kategoriální chyba*: mysl podle něj není žádná věc, na kterou bychom se mohli dívat jako na nějakou další složku člověka vedle jeho složek tělesných: je to spíše způsob, jak tělesné složky člověka fungují dohromady. (Podobně jako, v Rylově slavném příkladu, *univerzita* není nějaká další budova vedle jednotlivých fakult, kolejí, menz apod. – je to spíše způsob, jak tyto budovy fungují dohromady.)

Behaviorismus v psychologii, kde je dílem především Watsona, Skinnera a dalších, s tím rylovským do nějaké míry souvisí, má ale také na něm nezávislou agendu,

¹⁴ Tento návrh v podstatě není ničím jiným než *anomálním monismem*, jak ho známe od Donalda Davidsona (1970; 2001).

kteřá ho mnohem více vede k praktickým, metodologickým důsledkům. Myslím, že ho můžeme, s velkým zjednodušením, charakterizovat pomocí tří základních pilířů:

- 1) objektivně popsatelné jsou pouze projevy mysli, tedy její vstupy (stimuly) a výstupy (reakce) (takže teorie mysli je v jistém podstatném smyslu teorií těchto projevů);
- 2) mysl bychom měli popisovat, mechanisticky, jakožto deterministické propojování stimulů s reakcemi; a
- 3) popis propojování stimulů s reakcemi v mysli musí být natolik přímočarý, aby nebylo třeba postulovat žádnou netriviální skrytou strukturu mysli.

Mám pocit, že je to bod 3), který je hluboce problematický a díky kterému si behaviorismus vysloužil tak špatnou reputaci. Myslím, že toho bylo třeba se skutečně zbavit – velmi rychle se ukázalo, že popisovat třeba jazykové projevy člověka prostřednictvím nějakých jednoduchých propojení stimulů a reakcí je prostě nemožné (na tom je ostatně založena Chomského kritická recenze Skinnerovy knihy *Verbal behavior*¹⁵, která byla jednou z rozbušek kognitivní revoluce).

Současně mám ale pocit, že s body 1) a 2) je to mnohem komplikovanější a že chápeme-li kognitivní revoluci, jak se to často děje, jako chvályhodné odvrhnutí i jich, může to způsobit více škody než užítu. Spor může být o to, zda můžeme, v rámci naší teorie, postulovat nějaké komplikované modely toho, jak jsou propojeny pozorovatelné vstupy a výstupy organismu; nebo také o to, zda můžeme takovým modelům připsat status popisu něčeho, co skutečně existuje někde vevnitř tohoto organismu. Ve své slavné knize o kognitivní revoluci píše Baars (1986):

„Kognitivní psychologové si nemyslí, že učení se je nutně věcí asociování pozorovatelných stimulů s reakcemi, že základní formou učení je podmiňování nebo že role dědičných faktorů v rámci lidského konání a zkušenosti je jenom malá. ... Jeden z těch, se kterými jsme o tom dělali rozhovor, to nazval osvobozením – povolením užívat teoretickou představivost bez předem daných omezení.“

Čertovým kopytkem se mi zdá být termín „používat teoretickou představivost“. Mám-li ho číst ve smyslu „budovat teoretické modely propojení vstupů s výstupy“, pak s tím rozhodně nelze než souhlasit: bránil-li behaviorismus tomuto (a je nepochybné, že tomu tak do velké míry bylo), pak ho rozhodně bylo třeba odvrhnout. Mám-li ho ale číst tak, že do teorie mysli znovu vpouští – byť implicitně – myšlenku mentální substance, kterou je třeba zkoumat zcela nezávisle na zkoumání pozorovatelných vstupů a výstupů organismu, pak bychom si měli dobře rozmyslet, zda nás kognitivní revoluce nevede z deště pod okap.

¹⁵ Viz Chomsky (1967).

Mysl a počítač

Důležitým trumfem, který „kognitivní revoluce“ dostala do ruky, bylo to, že se konečně objevil způsob, jak začalo být možné o „vnitřku mysli“ uvažovat vědecky respektabilním způsobem. To nebylo důsledkem pokroků ve studiu mozku (neurofyziologie nám toho v šedesátých letech ještě o detailech fungování mozku mnoho říci schopna nebyla), ale důsledkem rozvoje počítačů. Člověka, který se počítači zabýval a který sledoval, jak se vypořádávají s úkoly, které do té doby dokázaly řešit jenom inteligentní bytosti, nemohlo nenapadnout, že lidská mysl by mohla být také ‚něco jako‘ počítač – a počítač se tak stal klíčovým modelem pro to, jak o fungování mysli uvažovat.

„Kognitivní revoluce“ se tak velmi rychle sblížila s rodícími se teoriemi umělé inteligence.¹⁶ Počítač je samozřejmě něco, co může propojovat vstupy a výstupy velmi komplikovaným způsobem, a co tak jde zcela proti prvnímu z výše uvedených principů behaviorismu. Na druhé straně není vůbec zřejmé, proč by musela jít proti ‚mechanistickému paradigmatu‘, které behaviorismus nastolil. Počítač přece není nic jiného než soustava nějakých složitě propojených elektrických obvodů, které vstupy s výstupy propojují sice možná neobyčejně komplikovaně, ale pořád ve zřejmém smyslu mechanisticky. (Uvědomme si, že fakt, že počítač pracuje na bázi elektroniky, není zásadní – v podstatě by ho bylo možné udělat i ze zcela mechanických částí.)

Vraťme se ještě k behaviorismu: již před jeho revolučním svržením byly v jeho lůně odstartovány evoluční procesy směřující k jeho modifikaci a zejména k uvolnění oněch striktních metodologických omezení, jaká jsou ve zkratce vyjádřena v bodě 3). Mám tu na mysli zejména tzv. neobehaviorismus badatelů, jako byli Tolman nebo Hull: podle nich je sice mysl věcí propojení stimulů s reakcemi (to jest je v tomto smyslu ‚mechanická‘), ovšem není důvod předpokládat, že funguje na nějakém jednoduchém principu propojování – vztah mezi stimuly a reakcemi může být velmi komplikovaný (metafory mysli jako ‚mapy‘ ap.).¹⁷ Proč se pak stal model počítače symbolem revolučního svržení behaviorismu, a nikoli jeho evolučního zdokonalení?

Odpověď podle mne odhaluje velmi problematický rys „kognitivní revoluce“: v rámci této revoluce totiž není počítač obvykle chápán jako *mechanické* zařízení, je totiž brán jako prostředek zpracovávání *informací*, přičemž *informace* je chápána jako něco svou podstatou nevměstnatelného do ‚mechanistického‘ paradigmatu. Obávám se ovšem, že tady jsme svědky dalšího pokusu o import magie do vědy a filosofie. Počítač je zcela zřejmě zařízení, na kterém není nic nemechanického a přitom je z mechanistického paradigmatu vyjmut proto, že se některým věcem, které se v něm dějí, dá magické jméno – *informace*.

Draaisma (1998) o tom píše:

¹⁶ Klíčovou roli v tomto ohledu sehráli zejména Newell a Simon (1972).

¹⁷ Viz Tolman (1932); Hull (1943). Souhrnně o tom pojednává Smith (1990).

„Hull a Turing – a potažmo behaviorismus a umělá inteligence – koncipovaly odlišně také analogii mezi mozkem a strojem. Pro Hulla znamenaly mechanické modely analogii lidského nervového systému, protože jejich reakce na podněty vykazovaly v obou systémech identické schéma; model a organismus vykazovaly stejné chování. Pro Turinga byl rozum strojem, neboť mozek byl speciálním Turingovým strojem, systémem, jehož logické vlastnosti se shodovaly s vlastnostmi počítače. Shoda v reakcích, která byla pro Hulla podstatou jeho „psychických strojů“, byla pro Turinga odvozeninou. Ne nadarmo kladl Turing důraz na schopnost počítačů zpracovávat informace a ukládat je. Zavedení metafory počítače do psychologie je dáno především použitelností této představy. A zvláštním směrem psychologie, který z toho bude těžit, neměl být behaviorismus, ale kognitivní psychologie.“

To je zajímavý a leccos vysvětlující postřeh. (Jakkoli mám pocit, že to, co autor říká o Turingovi, platí mnohem více o protagonistech pozdější „kognitivní revoluce“ než přímo o něm.) Zlom v přístupu myslí nastává nikoli v okamžiku, kdy se odmítne představa nutně jednoduchého, ‚reflexivního‘ propojení ‚vstupů‘ a ‚výstupů‘ organismu a začne se spekulovat o nějaké komplexní ‚vnitřní struktuře‘ (protože to by kognitivní revoluci spustili už Hull a Tolman), ale až tehdy, když se v souvislosti s vnitřní strukturou začne hovořit o *informacích*, *logice* atd.

Jenomže co to je *informace*? Myslím, že tohle slovo můžeme chápat v širším a užším smyslu. V tom širším je to jakýkoli ‚signál‘, který může dostat organismus a nějak ho využít pro své účely. V tomto smyslu je informací třeba světlo dopadající na sítnici oka člověka nebo pohyb pavučiny, který vnímá svou nohou pavouk. V užším smyslu je informací něco, co je sdělováno a zpracovááno velmi specifickým způsobem, zpravidla nějaká propozice, tj. nějaký význam věty. V tomto užším smyslu je informací třeba, že *svítí slunce* nebo že *pavouci nepatří mezi hmyz*; příjemcem informace v tomto smyslu může být jediné rozumná, jazykem disponující bytost.

Co to pak může znamenat, že stroj pracuje s *informacemi*? Chápeme-li toto slovo v tom širším smyslu, pak s informacemi pracuje každý jen trochu složitější stroj. Auto monitoruje stav benzínu v nádrži i teplotu vzduchu; bankomat přijímá čísla platebních karet. Chápeme-li to slovo v užším slova smyslu, je situace poněkud složitější. Pak můžeme říci, že zpracovává informace tehdy, když disponuje něčím jako rozumem či inteligencí – když je schopen informaci přijmout *jakožto informaci*, to jest hospodařit s ní jako rozumná bytost (to jest uvést ji do souvislosti s jinými informacemi, odstraňovat nekonzistence, vyvozovat důsledky a informace komplexně využívat pro své vyrovnávání se se světem). V případě tohoto užšího chápání termínu *informace* je tu ovšem ještě jedna alternativa: můžeme říci, že stroj pracuje s informacemi tehdy, když pracuje s větami (či nějakými jejich kódy), které *pro nějakého člověka* vyjadřují informace. V tomto posledním smyslu tedy k tomu, abychom řekli, že stroj pracuje s informacemi, stačí, aby byl naprogramován tak, aby například z vět *Jestliže prší, je*

mokro a *Prší* ‚vyvodil‘ větu *Je mokro* (to jest aby, když dostane na vstupu ty první dvě, vydal na výstupu tu třetí).

Máme tedy tři smysly, ve kterých může stroj „zpracovávat informace“. Ten nejužší se zdá být nepoužitelný, protože v jeho případě může „zpracovávat informace“ jedině skutečně inteligentní stroj, a takové rozhodně v době „kognitivní revoluce“ k dispozici nebyly (a je minimálně diskutabilní, zda je něco takového k dispozici dnes)¹⁸. Ale ty další dva jsou ve skutečnosti ještě nepoužitelnější. Pokud bychom se přiklonili k tomu prvnímu, pak by „zpracování informací“ necharakterizovalo žádnou specifickou třídu strojů – téměř každý stroj nějaké informace – v tomto smyslu slova – zpracovává. A podobně by tomu bylo i v případě toho posledního smyslu: i v jeho případě by v podstatě téměř každý stroj mohl být nahlížen jako „zpracovávající informace“, neboť v tomto případě by příliš nezáleželo na tom, co tento stroj sám dělá, ale jak my jeho vstupy a výstupy interpretujeme, a jako kódování nějakých informací je možné interpretovat v podstatě cokoli.

Zdá se mi tedy, že jediný způsob, jak dát *netriviální* smysl pojmu stroje *zpracovávajícího informace*, je vrátit se k onomu prostřednímu smyslu a říci, že stroj zpracovává informace tehdy, když pracuje s něčím, co svou komplexitou připomíná lidský jazyk, a dělá s tím podobné věci, jaké s jazykem dělá člověk. (To je samozřejmě velmi vágně řečeno, ale tady není místo to rozvádět.) Nakonec je to tedy věc určitého (specifického) druhu *komplexity* – a není vůbec jasné, proč by měl Hull zůstat před jejími dveřmi a za ně by mohli jenom kognitivní revolucionáři.

Jak už jsem ale řekl, mám pocit, že *informace* se nejčastěji používá jako další magické slovo: řekneme-li v tomto smyslu, že nějaká věc, třeba nějaký zvuk nebo nějaká řada klikyháků na papíře, *je informací* (nebo *nese informaci*, nebo *vyjadřuje informaci*, ...) pak tím *de facto* říkáme, že tato věc má nějaké vlastnosti či schopnosti, které přesahují vlastnosti či schopnosti, které má v důsledku toho, čím je: to jest, že kromě například schopnosti vyvolat v někom akustický vjem (se všemi následnými důsledky v jeho mozku atd.) má schopnost svého posluchače *informovat* v tom smyslu, že způsobí něco v jeho mysli, která *není* jeho mozkiem.

Revoluce nebo kontrarevoluce?

Vraťme se ke standardnímu pohledu na kognitivní revoluci, tj. k pohledu, ze kterého ji vidíme jako skutečnou *revoluci* hodnou toho jména – jako na převrat v rámci moderního vývoje studia mysli, který nás osvobodil z okovů behaviorismu. Chci poukázat na to, že je zde možný i opačný pohled: Na úsvitu moderní filosofie i psychologie se čím dál více prosazuje poznání, že zachovat v rámci našeho obrazu světa, který nám předkládá věda a který se – ať chceme nebo nechceme – postupně stává i způsobem, jakým svět skutečně vnímáme a chápeme, něco takového, jako je descartovská „přemítavá

¹⁸ Tzv. Loebnerovu cenu určenou pro první počítač, který projde Turingovým testem (<http://www.loebner.net/Prizef/loebner-prize.html>), se dosud nikomu nepodařilo získat.

substance“ prostě nelze. Behavioristé a někteří další filosofové a psychologové tuto myšlenku revolučním způsobem implementovali, „kognitivní revoluce“ však tuto filosofickou revoluci zvrátila.

Je přitom třeba znovu zdůraznit fakt, že mysl v podobě nějaké esoterické substance není – oproti našim primárním intuicím – něčím, s čím bychom se ve světě skutečně bezprostředně setkávali. Bezprostřední zkušenost s myslí jiných lidí nemáme – bezprostředně můžeme být svědky jedině toho, jak se chovají (v krajním případě toho, co se děje v jejich mozcích). A bezprostřední zkušenost se svou vlastní myslí, které se nám, jak se může zdát, dostává prostřednictvím introspekce, musíme brát s velkou rezervou – jelikož chybí intersubjektivní korektiv, který je k dispozici v případě všeho toho, co jsme ochotni označovat za objektivní fakta, dává v této oblasti stěžejní smysl rozlišení mezi fakty a fantaziemi.

Behaviorismus byl jedním z pokusů postavit se do cesty přirozené, ale neudržitelné tendenci chápat mysl jako substanci. Tyto pokusy byly podle mne vpravdě revoluční – šly proti lidsky hluboce zakořeněnému způsobu, jak mysl vnímat. Ano, behaviorismus šel do extrémů a vliv, který měl ve své ortodoxní podobě na rozvoj studia mysli, byl v mnohém neblahý. V jeho lůně se ale již rodily méně ortodoxní varianty, které zdaleka tak neblaze nepůsobily. A na druhé straně: kognitivní revoluce se sice vypořádala s behavioristickými ‚okovy‘ (které ovšem již samy od sebe povolovaly), avšak za cenu, že odsunula stranou behavioristický pokus o zablokování magických teorií mysli.

Gardner (1985, 12) píše:

„Behaviorismus odpovídal mnoha potřebám vědecké komunity, včetně některých, které byly zcela legitimní: nespokojenost s přijímáním introspektivní evidence jako dané, bez jakékoli možnosti vědecké kontroly a jakékoli možnosti odmítnutí; nespokojenost s vágními a obecnými pojmy jako vůle a účel a touhou vysvětlit lidské chování s použitím těch samých konstruktů, které byly aplikovány (se zjevně velkým úspěchem) na chování zvířat. A v kontextu potíží, které vyvolalo spoléhání se na introspektismus (...) behaviorismus vypadal v prvních desetiletích století jako závan svěžího povětří. Není divu, že se tak rychle uchytil a rychle upoutal nejbystřejší duchy jedné generace badatelů. Avšak z pohledu zpět byla cena, kterou jsme zaplatili za přísné přilnutí k behaviorismu, příliš vysoká. Dokud byl behaviorismus u vesla – to jest během dvacátých, třicátých a čtyřicátých let – přicházely otázky o povaze lidského jazyka, plánování, řešení problémů, představivosti a podobně na přetřes jenom pokradmu a s obtížemi, pokud byly vůbec tolerovány.“

Otázkou nyní je, jak máme, z pohledu zpět, hodnotit „kognitivní revoluci“? Nebyla cena za ni příliš velká? Není samozřejmě pochyb, že pokrok, který za poslední desetiletí udělalo studium mysli, budování jejích modelů i umělá inteligence, je enormní. Není také pochyb, že k tomuto pokroku přispěla „kognitivní revoluce“ (i když legitimní

otázkou je, jak moc zásluh připadá jí a jak a co by se bylo událo i bez ní). Otázkou ovšem je, zda tento příspěvek vyváží invazi naivně mentalistických teorií mysli, které se jí v patách vyrojily a které uvažování o mysli zamořily. Lidem, kteří budují systémy umělé inteligence, nemusí na teoriích mysli, o které se opírají, tak moc záležet, ale jde-li nám o *filosofii* mysli, je pro nás právě tato otázka klíčová.

Závěr

Myslím, že „kognitivní revoluce“ uskutečnila některé blahodárné věci, především nás zbavila předsudečných a brzdicích metodologických omezení výzkumu mysli. Také vedla k mnohým novým impulsům ohledně vymyšlení komplexních mechanismů propojení vstupů a výstupů organismů jakožto ‚černé skříňky‘, které nám do propojení těchto vstupů a výstupů zjednaly nové vhledy. To do velké míry souviselo s přijetím počítače jako paradigmatu modelování mysli a její činnosti.

Současně si ale myslím, že je otázkou, zda tohle dostačuje k tomu, abychom ji skutečně nazývali *revolucí* – jak jsme viděli, behaviorismus sám v sobě skrýval potenciál k evolučnímu vývoji s podobným výsledkem¹⁹. Myslím, že mnozí příznivci „kognitivní revoluce“ považují za její největší výtěžek to, že nás zbavila představ mysli jakožto něčeho mechanického a že nám dovolila přitakat naší intuitivní tendenci vnímat mysli jakožto něco poněkud esoterického, co lze studovat nějak příměji než prostřednictvím studia jejích projevů. Proto je tu i jedno nebezpečí, které je s kognitivní revolucí svázáno, a to je nebezpečí resuscitace karteziánského dualismu a magické teorie mysli. Z hlediska toho, že tohle s sebou „kognitivní revoluce“ také fakticky přinesla, je možné ji považovat i za svého druhu kontrarevoluci – za zmaření revoluční snahy behavioristů a dalších teoretiků mysli systematicky zablokovat magickou teorií mysli.

Literatura

Baars, B. J. (1986): *Cognitive Revolution in Psychology*. Guilford Press, New York.

Benjamin, L. T., Jr. (2007): *A Brief History of Modern Psychology*. Blackwell, London.

Bohm, D. (1980): *Wholeness and the Implicate Order*. Routledge, London.

¹⁹ Benjamin (2007, s. 215) píše: „Historická analýza naznačuje, že tu žádná revoluce nebyla, přestože někteří kognitivní psychologové s oblibou líčí přerod jejich oboru jako těžce vybojovanou válku proti silám temné strany, neboli Johnu ‚Darth Vaderovi‘ Watsonovi a jeho nohsledům“.

- Capra, F. (1984): *The Turning Point: Science, Society, and the Rising Culture*. Bantam Books; český překlad *Bod obratu: Věda, společnost a nová kultura*, Maťa/Dharmagaia, Praha, 2002.
- Carruthers, P. and P. K. Smith (1996): *Theories of Theories of Mind*. Cambridge University Press, Cambridge.
- Davidson, D. (1970): „Mental Events.“ In *Experience and Theory*, eds. L. Foster a J. W. Swanson, The University of Massachusetts Press; reprinted in Davidson (1980), s. 207–243; slovenský překlad „Mentálne události“ v Davidson: *Čin, myseľ, jazyk*, Archa, Bratislava, 1997, s. 35–64.
- (2001): *Subjective, Intersubjective, Objective*. Clarendon Press, Oxford; český překlad *Subjektivní, objektivní, intersubjektivní*, Filosofia, Praha, 2004.
- Dennett, D. (1987): *The Intentional Stance*. MIT Press, Cambridge (Mass.).
- Draaisma, D. (1998): *De metaforenmaschine*. Historische Uitgeverij, Utrecht; český překlad *Metafory paměti*, Mladá fronta, Praha, 2003.
- Eysenck, M. & T. Keane (2005): *Cognitive Psychology: a Student's Handbook*. Psychology Press, London.
- Gardner, H. (1985): *The Mind's New Science (A History of the Cognitive Revolution)*. Basic Books, New York.
- Graham, G. (2010): „Behaviorism.“ In *Stanford Encyclopedia of Philosophy*, ed. E. Zalta. Dostupné na:
<<http://plato.stanford.edu/archives/fall2010/entries/behaviorism/>>.
- Hull, C. L. (1943): *Principles of Behavior: An Introduction to Behavior Tudory*. Appleton-Century, New York.
- Chomsky, N. (1967): „A Review of B. F. Skinner's *Verbal Behavior*.“ In *Readings in the Psychology of Language*, eds. L. A. Jakobovits a M. S. Miron, Prentice-Hall, s. 142–143.
- Newell, A. and H. A. Simon (1972): *Human Problem Solving*. Printice Hall, Englewood Cliffs.
- Plháková, A. (2006): *Dějiny psychologie*. Grada, Praha.
- Ravenscroft, I. (2010): „Folk Psychology as a Theory.“ E. Zalta (ed.): *Stanford Encyclopedia of Philosophy*, Dostupné na:
<<http://plato.stanford.edu/archives/fall2010/entries/folkpsych-theory/>>.
- Robinson, H. (2011): „Dualism.“ In *Stanford Encyclopedia of Philosophy*, ed. E. Zalta. Dostupné na: <<http://plato.stanford.edu/archives/win2011/entries/dualism/>>.
- Ryle, G. (1949): *The Concept of Mind*. University of Chicago Press, Chicago.
- Saygin, A. P., I. Cicekli & V. Akman (2000): „Turing Test: 50 Years Later.“ *Minds and Machines* 10: 463–518.

- Skinner, B. F. (1938). *The Behavior of Organisms*. Appleton-Century-Crofts, New York.
- (1953): *Science and Human Behavior*. Macmillan, New York.
- Smith, L. D. (1990): „Metaphors of Knowledge and Behavior in the Behaviorist Tradition.“ In *Metaphors in the history of philosophy*, ed. D. E. Leary, Cambridge University Press, Cambridge.
- Sternberg, R. J. (1999): *Cognitive Psychology* (2nd edition). Harcourt, Fort Worth; český překlad *Kognitivní psychologie*, Portál, Praha, 2002.
- Tolman, E. C. (1932): *Purposive Behavior in Animals and Men*. Century, New York.
- Turing, A. (1950): „Computing Machinery and Intelligence.“ *Mind* 59: 433–460.
- Watson, J. B. (1919). *Psychology from the Standpoint of a Behaviorist*. Lippincott, Philadelphia.