
Polemika

Ještě o vědomí, behaviorismu a zdravém rozumu

Jaroslav Peregrin —

Filosofický ústav AV ČR, v. v. i., Praha
peregrin@flu.cas.cz

Nevím, jak dlouho může čtenáře zajímat, když se dva filosofové přou o nějaké speciální téma. Z tohoto důvodu jsem váhal, zda mám v diskuzi s kolegou profesorem Sousedíkem pokračovat ještě veřejně, nebo už ji přenést do soukromého prostoru mezi námi dvěma. Nakonec jsem ale usoudil, že v této chvíli jde přece jenom o věci, které až zas tak speciální nejsou, a které by tedy snad mohly být pro širší filosofickou veřejnost zajímavé; ještě jednou proto pokračuji veřejně.¹

Sousedík píše: „Představme si, jak Peregrin ve svém příspěvku navrhuje, robota, který se bude ve všech *ohledech* chovat jako člověk (to jest včetně všech projevů emocí atd.). Peregrin soudí, že v takovém případě bude prý zřejmé, že onen robot „nejen má vědomí, ale navíc se my na něj dokonce ani *nebudeme schopni* dívat tak, jako by ho neměl“. S tímto míněním se nemohu osobně ztotožnit.“²

K tomu, bych rád podotkl, že mnou uvedený soud není nějakou apriorní spekulací, ale předpovědí opřenou o – podle mne – dost pádnou empirickou evidenci. Jeden ze známých aforismů, které je možné najít ve Wittgensteinových *Filosofických zkoumáních*, zní: „Můj postoj k němu je postojem k duši. To, že má duši, není můj názor.“³ Wittgenstein zde poukazuje na to, že beru-li někoho za oduševnělou či myslící bytost, není tomu tak proto, že bych si z jeho projevů *vyvodil* názor, že tomu tak je – tento názor si vytvářím spon-

1 Reaguji zde na Sousedíkův text „Behaviorismus a zdravý rozum“ (*Filosofický časopis*, 66, 2018, č. 3, s. 446–448), který byl jeho reakcí na můj příspěvek „Vědomí ve světle ‚zdravého rozumu‘“ (tamtéž, s. 442–446).

2 Sousedík, S., Behaviorismus a zdravý rozum, c.d., s. 447.

3 Wittgenstein, L., *Filosofická zkoumání*. Přel. J. Pechar. Praha, FLÚ AV ČR 1993, s. 225.

tánně a nereflektovaně. Beru-li tedy někoho za myslící bytost, pak to podle Wittgensteina není závěr žádné dedukce, ale víceméně mimovolně zaujímaný a nevyhnutelný postoj. A právě tomuto Wittgensteinovu vhledu se v poslední době dostává překvapivě silné podpory, jež vychází z výzkumů vědných v rámci kognitivní psychologie.

Většina psychologů se shoduje na předpokladu, že máme nějakou vrozenou schopnost vnímat lidi jakožto opatřené myslí; nejčastěji se o této schopnosti hovoří, možná trochu matoucí způsobem, jakožto o *teorii mysli*.⁴ Různí protagonisté se ovšem neshodují na detailech: někteří mají za to, že teorie mysli je něco jako samostatný modul v rámci mozku/mysli, zatímco jiní to popírají; někteří se domnívají, že jde doslova o teorii, zatímco jiní to vidí jinak, třeba jako něco založeného na empatii atd.⁵ I přes tyto rozdíly se nicméně společně shodují na tom, že naše tendence připisovat lidem mysl má genetický základ, a není tedy věcí našeho úsudku či rozhodnutí. A protože takové přisouzení mysli nemůže vycházet z ničeho jiného než z toho, jak ten, komu mysl připisujeme, vypadá a jak se chová, zdá se mi opodstatněné předpokládat, že bude-li se něco v těchto ohledech s člověkem zcela shodovat, nepřipisovat tomu mysl prostě ani nedokážeme.

Píše-li tedy Sousedík „Nevidím důvodu, proč bych měl celku vzniklému sestavením [součástí robota – J.P.] připisovat více vědomí než odstředivému regulátoru“,⁶ pak je tu třeba oponovat a zdůraznit, že to, na co jsem poukázal, se netýkalo důvodů. Uvedl jsem totiž, že se domnívám, že pokud tu budou roboti nerozpoznatelní od člověka, pak se na ně nedokážeme dívat jinak než jako na myslící bytosti – ne kvůli důvodům, ale proto, že „brát něco za myslící“ je postoj, který zaujímáme mimovolně. A pokud už jde o důvody, přiznám se, že se mi nepodařilo dát rozumný smysl Sousedíkovu argumentu s odstředivým regulátorem. Zdá se mi, že ten argument zní: odstředivému regulátoru bychom jistě nechtěli připisovat vědomí, odstředivý regulátor je stroj, tedy žádnému stroji nelze připisovat vědomí – což je ale natolik zřejmé *non sequitur*, že takhle to Sousedík jistě nemyslel. Nejsem si tedy vůbec jistý, o jaký argument má jít.

4 Carruthers, P. – Smith, P. (eds.), *Theories of Theories of Mind*. Cambridge, Cambridge University Press 1996.

5 Jako první s touto myšlenkou zřejmě přišel již v sedmdesátých letech minulého století primatolog David Premack (viz např. Premack, D. – Woodruff, G., Does the Chimpanzee Have a Theory of Mind? *The Behavioural and Brain Sciences*, 1, 1978, No. 4, s. 515–526). Ve filosofii se podobným směrem ubíraly úvahy Daniela Dennetta (*The Intentional Stance*. Cambridge, Mass., MIT Press 1987), ale nejvíce tento pohled zakořenil v kognitivní psychologii (viz např. Garton, A. F., *Exploring Cognitive Development: The Child as Problem Solver*. Oxford, Blackwell 2004; nebo Siegal, M., *Marvelous Minds: The Discovery of What Children Know*. Oxford, Oxford University Press 2008).

6 Sousedík, S., *Behaviorismus a zdravý rozum*, c.d., s. 447.

Jinak se s kolegou Sousedíkem zřejmě neshodneme na tom, co je *poznatek*. Já vycházím z předpokladu, že my lidé jsme odedávna vyvíjeli způsoby, jež nám měly umožňovat, abychom se něco nového dozvídali o našem světě (a o nás samotných jako součástech tohoto světa) a abychom uměli rozlišovat mezi tím, co je skutečným poznatkem, a tím, co jako poznatek jenom vypadá. Místem, kde se sedimentovaly výsledky těchto našich snah, se stávala naše věda – takže mi připadá rozumné se i v otázkách po povaze poznatků a poznání inspirovat tím, co se ustálilo ve vědě. A zdá se mi, jak jsem o tom psal ve svém předchozím vstupu do této diskuse, že věda nám říká, že poznatkem může být jenom to, co je potvrditelné z nezávislých zdrojů; což ovšem nedovoluje propůjčit zkoumání své vlastní mysli – nakolik je chápána jako neprodyšně soukromá – status skutečného poznání.

Musíme jít ale ještě dál. Ty nejzákladnější poznatky, ke kterým věda dospívá, jsou *kauzální* povahy – říkají, že nějaký děj *zapříčňuje* jiný děj. Abychom ovšem mohli něco takového konstatovat, musíme být schopni detekovat oba tyto děje (nezávisle na sobě) a předvést mechanismus, který kauzálním způsobem vede od onoho prvního k tomu druhému. Z toho vyplývá, že děje, které se odehrávají v mysli, se nemohou stát součástí kauzálního vysvětlení. (Těmi se mohou stát například děje v mozku, které je možné objektivně detekovat, avšak nikoli něco, co se údajně děje v něčí mysli a k čemu nemáme žádný bezprostřední přístup.)

Sousedíkův příklad kauzálního vysvětlení prostřednictvím záměru pojava tého v mysli, který se týká Evy, jež pojala záměr utrhnout ze stromu jablko – a v důsledku toho ho skutečně utrhla, tuto podmínku zřejmě nesplňuje. Zda někdo skutečně má nějaký záměr, nijak detekovat nemůžeme (na rozdíl například od toho, co se mu děje v mozku, nebo od zvuků, které vypouští z úst). Tím nechci říci, že předpoklad, že někdo má nějaký úmysl, nám nemůže sloužit jako druh vysvětlení – v běžné mezilidské interakci nám k tomu bezpochyby často slouží –, jen to není onen druh kauzálního vysvětlení, s jakými pracuje věda.

Ve vysvětlení, které uvádí Sousedík, totiž nejde o *příčinu* toho, co Eva udělala, ale o *důvod*, jaký k tomu měla. (Diskuse o vztazích mezi důvody a příčinami jsou zejména v současné filosofii společenských věd velmi intenzivní.) Zásadní rozdíl mezi příčinami a důvody spočívá v tom, že příčiny jsou deterministické – příčina má *nevyhnutelně* svůj důsledek –, zatímco důvody ne. I přesto, že mám důvod něco udělat, tak to udělat nemusím – vztah zde není kauzální, ale normativní. To, že důvody mohou být v běžném hovoru o tom,

7 Viz např. Risjord, M., Reasons, Causes, and Action Explanation. *Philosophy of the Social Sciences*, 35, 2005, No. 3, s. 294–306; nebo Setiya, K., Reasons and Causes. *European Journal of Philosophy*, 19, 2011, No. 1, s. 129–157.

co lidé dělají, prostředky vysvětlování, spočívá v tom, že předpokládáme, že lidé jsou racionální neboli že skutečně dělají to, pro co mají dobré důvody.

Sousedík tedy navrhuje vysvětlit fakt, že Eva utrhla ze stromu jablko, faktem, že „pojala (ve své mysli pouze privátně dostupný) záměr utrhnout si jablko a (protože její mysl samotná není s to tento záměr uskutečnit) užila své paže jako nástroje svého druhu a jablko její pomocí utrhla“.⁸ Jak to ale vím? Co se odehrává v Evině soukromé mysli nejenže nevím, ale ani to vědět nemohu. Proč se tedy domnívám, že Eva pojala ten záměr? Inu jistě právě proto, že to jablko utrhla. Ale tady už se začínáme pohybovat v jakémisi podezřelém kruhu: chtěli jsme vysvětlit fakt, že Eva utrhla jablko, faktem, že pojala příslušný záměr, ale teď vidíme, že naše přesvědčení, že Eva pojala ten záměr, je zcela odvozeno z našeho přesvědčení, že utrhla to jablko. Mohu zjistit, že Eva pojala záměr utrhnout jablko, nějak nezávisle na tom, že ho skutečně utrhne? Například tak, že nám Eva řekne, že ho chce utrhnout? Obávám se, že nikoli. K tomu bychom totiž potřebovali s jistotou vědět, že slova, která Eva vypouští z úst, znamenají to, co podle nás znamenají, že Eva nelže atd., což s patřičnou jistotou nemůžeme vědět nikdy.

Sousedík ale, zdá se, tohle vše ví. Na otázku, co bych asi jeho vysvětlení řekl já, si sám odpovídá naprosto správně: „Rekl by pravděpodobně, že se v něm dovolávám činitele (totiž pouze privátně dostupného Evinu záměru), o němž mám (v nejlepším případě) jen velmi problematické, ve skutečnosti spíše vůbec žádné poznání.“⁹ Ano, to je to, co říkám; a rozdíl mezi námi je pak zřejmý v tom, že zatímco já v tom vidím demonstraci toho, že tady nemůže jít o nic takového, jako je vědecké vysvětlení (a připomeňme, že zde se náš spor odvinul od mého tvrzení, že pojem soukromé mysli nepotřebujeme pro vědecké vysvětlování), Sousedík v tom zjevně nevidí problém.

Na závěr Sousedík píše, že jsem o tom, v čem spočívá moje verze behaviorismu „dosud mnoho neřekl“.¹⁰ Tím mě, příznávám se, poněkud zaskočil: o tom byla má knížka, kterou Sousedík recenzoval. Jenom v kostce shrnuji: myslím si, že má-li se behaviorismus stát udržitelným, musíme k němu přidat historickou, vývojovou dimenzi. Pokud to neuděláme, existuje mysl jenom v „oku pozorovatele“, je prostředkem, pomocí něhož pozorovatel „dává smysl“ chování ostatních. Lidská společnost ale fungují na bázi naprosté vzájemnosti – každý je interpretem a interpretovaným zároveň. V knize tedy konstatuji: „Právě takovéto zrušení opozice mezi divákem a hercem je předmětem vývoje, který vede od subjektivní konstrukce k intersubjektivní, institucionální realitě. Vezmou-li něčí konstrukci za svou ostatní a opatří-li ji nadstavbou

8 Sousedík, S., Behaviorismus a zdravý rozum, c.d., s. 448.

9 Tamtéž.

10 Tamtéž.

institucionalizovaných kritérií pro to, kdy něco v rámci této konstrukce ‚objektivně existuje‘, a kdy ne, může vzniknout nová skutečnost. K takovému vývoji ovšem může dojít jedině v rámci *společnosti*, která se vyvíjí – jinými slovy, naše mysl by se nemohla stát tím, čím je, nebýt tohoto společensko-historického kadlubu.“¹¹

Redakce Filosofického časopisu se obrátila na profesora Sousedíka s dotazem, zda hodlá – obdobně jako v minulém případě – reagovat na polemický příspěvek profesora Peregrina. Dostalo se nám však záporné odpovědi: prof. Sousedík se nejen domnívá, že vše podstatné bylo již mezi nimi řečeno, ale gentlemansky také ponechává právo posledního slova kolegovi, jehož k diskusi sám ponejprv vyzval.

Polemika je tedy v této fázi ukončena, nicméně má-li někdo z kolegů zájem v ní pokračovat a toto dle mého názoru zajímavé a nosné téma dále rozvíjet, je to samozřejmě na stránkách Filosofického časopisu možné. Oběma dosavadním diskutujícím děkuji.

James Hill, šéfredaktor FČ