

## Kauzalita činitele: představení hlavních argumentů knihy

Petr Dvořák

*Filosofický ústav  
Akademie věd České republiky  
Jilská 1, 110 00 Praha  
petr.dvorak@flu.cas.cz*

Článek shrnuje základní argumenty v knize P. Dvořáka *Kauzalita činitele. Úvod do analytické diskuse o svobodě vůle* (Togga 2020). Kniha obhajuje jednu dříve též nazývanou kauzalitu činitele, coby nejlépe schopnou uchovat naše běžné intuice týkající se významu pojmu svoboda vůle a důsledků pro morální odpovědnost. V první části knihy uvádí základní pojmy a klíčový argument pro inkompatibilismus. Hlavního rivala, kompatibilismus v podobě tzv. standardní teorie (D. Davidson), odmítá ve třetí části. Hlavní teze, která leží v pozadí v obhajované hájence kauzality činitele, spočívá v tom, že kontingentní (sebe-)determinace představuje určitý střed mezi deterministickou a indeterministickou kauzalitou. Druhá část knihy pak hraje pomocnou úlohu: hájí nějaký typ dualismu (ať už vlastností či substanční dualismus) proti nereduktivnímu fyzikalismu tím, že se zaměřuje na tzv. problém vyloučení.

Klíčová slova: svoboda, determinismus, kauzalita činitele, inkompatibilismus, kompatibilismus

V následujícím pojednání představíme hlavní argumenty knihy *Kauzalita činitele. Úvod do analytické diskuse o svobodě vůle*<sup>1</sup> v uzavřené podobě, která, doufáme, nevyžaduje znalost knihy samé. V naší prezentaci oproti knize zaměníme posloupnost jejích částí a část, která se v knize objevuje jako v pořadí druhá a týká se filozofických předpokladů kauzality činitele v oblasti filozofie zvané filozofie mysli, si necháme až na konec. Nejprve

---

<sup>1</sup> Dvořák, P. (2020): *Kauzalita činitele. Úvod do analytické diskuse o svobodě vůle*. Togga, Praha.

představíme hlavní pojmy a ukážeme neslučitelnost běžného robustního (inkompatibilistického) pojetí svobody vyžadujícího alternativní možnosti s determinismem. Dále představíme údajnou neslučitelnost tohoto pojetí, které kromě existence alternativních možností vyžaduje i to, aby svobodný činitel determinoval své chtění a jednání, s popřením determinismu. Na základě obou argumentů se ukáže, že činitele je třeba chápat jako nedeterministického, ale zároveň takového, který nekoná náhodně, ale plně determinuje své chtění a jednání, tj. že nemůže být indeterministický (kdy indeterminismus zahrnuje alespoň zčásti náhodný vznik). Takový činitel musí být obdařen schopností kontingentní determinace, kterou představujeme ve druhé části tohoto pojednání v kontrastu s dnes v analytické filozofii dominantní tzv. „standardní teorií“. Poslední část naší prezentace věnujeme odmítnutí nereduktivního fyzikalismu, jenž coby teorie mysli představuje filozofické pozadí „standardní teorie“, a poukazu na racionální oprávněnost nějaké formy dualismu, kterou jako teorii mysli předpokládá kauzalita činitele, tak jak ji zde chápeme.

## PRVNÍ ČÁST KNIHY:

### Inkompatibilismus a problém spojený s popřením determinismu

První část knihy vysvětluje základní pojmy determinismu, svobody, inkompatibilismu. Uvádí určitou verzi typického argumentu pro inkompatibilismus (Argument následku) a logické dilema, podle něhož nejen determinismus, ale ani jeho popření neumožňuje existenci inkompatibilistické svobody (Standardní argument). Ukazuje, že námitka, podle níž popření determinismu u vzniku jednání automaticky znamená, že dané jednání vzniká náhodně a je tudíž v rozporu s tím, aby bylo inkompatibilisticky svobodné, je lichá. Absence determinismu příčin neimplikuje absenci determinace příčinami. Vysvětluje, že řešení spočívá v zavedení kauzality činitele, který je kontingentně (tzn. nedeterministicky) determinující příčinou svobodných rozhodnutí.

Determinismus je názor, který říká, že je-li dána minulost a zákony přírody, v každém okamžiku  $t$  existuje jen jediná možná budoucnost. Jinak řečeno, to, jaké události budou existovat v libovolném budoucím okamžiku, je v tomto okamžiku  $t$  již dané. Přesněji:

(D) Nachází-li se v  $t_0$  svět ve stavu  $P$  a platí-li zákony přírody  $L$ , potom je pro každou událost v každém okamžiku  $t_1$  následujícím po  $t_0$  nutně dáno, zda se vyskytne či nikoliv.<sup>2</sup>

Ponecháme-li stranou mikrosvět a zákony kvantové fyziky, zdá se, že makrosvět, který obýváme, popisují deterministické zákony. Na odhalení těchto zákonitostí je založen úspěch vědy. O deterministickou povahu těchto zákonů se opírá fakt, že věda dokáže přesně předpovídat mnohé jevy. I predikovatelnost našeho jednání a jeho vysvětlení udáním důvodů (motivů) vyžaduje pravidelnost a podle některých svědčí o tom, že i ono je výsledkem deterministického působení prostředí, našeho charakteru, dědičnosti atd.

Zdá se, že nutnou podmínkou toho, aby jednání bylo svobodné, je možnost nekonat to, co se koná, možnost učinit opak. V okamžiku bezprostředně před rozhodnutím či konáním existuje ve vztahu k dané minulosti *vícero* možných budoucností, tj. neexistuje jediná událost, která se uskuteční. Takže se zdá být zřejmé, že svoboda jednání není slučitelná s tím, aby jednání bylo výsledkem deterministického působení. Proto, vyskytne-li se alespoň jedno svobodné jednání, nemůže skutečnost existovat tak, jak tvrdí determinismus. Jinými slovy, determinismus musí být nepravdivý. A naopak, je-li pravdivý determinismus, neexistuje žádné svobodné jednání a, jelikož jen za takové jednání může být jeho činitel odpovědný, žádný činitel není za nic morálně odpovědný. To je pochopitelně velmi nepříznivý důsledek, který odporuje běžným názorům i naší praxi. Buď tedy platí determinismus, nebo existují jednání, která jsou svobodná. Determinismus a svoboda rozhodování a jednání jsou vzájemně neslučitelné, nekompatibilní. To je teze o inkompatibilismu.

Uveďme určitou verzi tzv. Argumentu následku (*Consequence Argument*).<sup>3</sup> Výrazem „jednání“ v daném argumentu rozumějme rozhodování či z něho plynoucí jednání:

2 Stavem  $P$  míníme souhrn všech fakt (případně událostí) existujících v daném čase. Množinou přírodních zákonů  $P$  míníme všechny zákony fyzikální, chemické, biologické a jiné, které jsou deterministické, tj. umožňují v každém okamžiku z relevantních fakt odvodit pro libovolný budoucí okamžik jediný fakt.

3 Existují různé verze. Základní pocházejí od Petera van Inwagena. Srov. Van Inwagen (1983, kap. 3).

1. Determinismus je pravdivý (předpoklad)
2. Tedy, každá událost je determinována minulostí  $P$  a zákony přírody  $L$
3. Tedy, jednání  $x$  činitele  $S$  je determinováno  $P$  a  $L$
4.  $S$  by mohl jednat jinak než  $x$  pouze tehdy, kdyby  $S$  měl kontrolu nad  $P$  nebo  $L$
5.  $S$  nemá kontrolu ani nad  $P$ , ani nad  $L$
6. Tedy,  $S$  nemůže jednat jinak než  $x$
7. Tedy,  $S$  není svobodný ve vztahu k jednání  $x$

Protože platí pro každého činitele  $S$  a každé jednání  $x$ :

8. Tedy, jestliže platí determinismus, teze o svobodě jednání je nepravdivá.

Premisa 1 je předpoklad, premisa 2 vychází z definice determinismu. Premisa 3 vztahuje tvrzení ve druhé premise na jednání. Premisa 4 stanoví podmínku existence alternativních možností jednání.<sup>4</sup> Premisa 5 konstatuje výše zmíněnou nemožnost měnit minulost a zákony přírody. Premisa 6 vyvozuje důsledek, neexistenci možnosti jednat opačně. Ta je nutnou podmínkou svobody, proto platí závěr 7. Zobecnění na libovolné jednání libovolného činitele vede k popření svobody (závěr 8).

Nabízí se možnost odmítnout závěr o neexistenci svobodného jednání tak, že popřeme první premisu argumentu, předpoklad o platnosti determinismu. Neplatí-li determinismus, pak ne každá událost je určená minulými událostmi a zákony přírody v tom smyslu, že ji lze z těchto faktorů dedukovat. Události, které takto nejsou vyvoditelné, jsou (svobodná) lidská jednání. Aby v daném okamžiku existovala možnost jednat jinak, nemohou zde být *žádné* faktory, z nichž by šlo dedukovat, že se realizuje jedna z těchto možností. Říkejme této pozici indeterminismus. Nyní však mnozí uvažují takto: Pokud bezprostředně před vznikem události neexistují žádné determinující faktory, pak je vznik události čistě náhodný. Jak jsme již poznamenali výše, náhodný děj se vymyká kontrole. A svoboda je druh kontroly nad dějem, nad událostí, již je v našem

4 Již víme, že v případě svobody jednání se jedná o podmínku nutnou. O tom argument hovoří. V případě svobody rozhodování by se jednalo o podmínku nutnou i postačující.

případě určité jednání. Proto žádné jednání nemůže být svobodné. Vzniká tedy následující logické dilema známé v literatuře jako Standardní argument:

1. Buď je determinismus pravdivý, nebo je determinismus nepravdivý (a platí indeterminismus).
2. Jestliže platí determinismus, nemáme svobodu jednání.
3. Jestliže platí indeterminismus, nemáme svobodu jednání.
4. Tedy, nemáme svobodu jednání.

Pravdivost třetí premisy tohoto argumentu se opírá o výše zmíněnou implikaci „Pokud bezprostředně před vznikem události neexistují žádné determinující faktory, pak je vznik události čistě náhodný“. Domníváme se, že aby tato implikace byla pravdivá, je nutno předpokládat, že oněmi „determinujícími faktory“ se míní příčiny. Je jistě pravda, že pokud bezprostředně před vznikem události neexistují žádné příčiny, tzn. pokud vznik události nic nepůsobí, pak je vznik události čistě spontánní a náhodný. To ale z popření determinismu nijak neplyne. Musíme rozlišit *determinující* faktor ve smyslu něčeho, co působí, že účinek má určitý charakter, určitou povahu či podobu, tj. determinuje jej, a *deterministický* faktor, tj. něco, z čeho lze dedukovat, že účinek nastane, ať už příslušný faktor determinuje či ne, tj. má kauzální charakter nebo ne.<sup>5</sup> Determinismus tedy postuluje, že každá událost plyne z deterministických faktorů, indeterminismus to popírá. Determinující faktor může být deterministický, ale nemusí. Je to právě vznik události při absenci determinujících faktorů, co lze chápat jako náhodu. „Neexistují žádné deterministické faktory“ (to říká indeterminismus, nakolik jej chápeme jako negaci determinismu) tedy neimplikuje „neexistují žádné determinující faktory“. Popření determinismu vede k následující formulaci zmíněné implikace:

---

5 Minulý stav světa a přírodní zákony samy o sobě nepůsobí budoucí stavy světa. To, co působí budoucí stav světa, jsou všechny existující příčiny a jejich kauzální schopnosti, jejichž působení popisují přírodní zákony. Ty působí budoucí fakta, která souhrnně nazýváme „stavem světa“. Minulé stavy světa a přírodní zákony tedy samy o sobě nejsou determinující faktory, ale lze je označit za deterministické faktory (platí-li determinismus).

Determinismus nechceme zúžit na kauzální determinismus, v němž by deterministické kauzální faktory byly zároveň determinujícími. To proto, že existuje např. teologický determinismus, podle něhož budoucí události plynou z toho, že je Bůh předem poznává. Boží poznání se ale nepovažuje za příčinu těchto událostí.

„pokud bezprostředně před vznikem události neexistují žádné deterministické faktory, pak je vznik události čistě náhodný“. A ta pravdivá není: I když absentují deterministické faktory (kauzální či nekauzální), mohou existovat nedeterministické determinující faktory, nedeterministické příčiny. Nakolik nedeterministická příčina nedeterminuje všechny aspekty účinku, ale jen některé, lze hovořit o náhodnosti. Například nějaká částice udělí zrychlení jiné částici, ale přesná hodnota daného zrychlení touto příčinou určena není a je dílem náhody. Nebo nestabilní jádro atomu způsobí radioaktivní rozpad, ale kdy přesně, není určeno. Pokud však nedeterministická příčina determinuje účinek plně, tj. všechny jeho aspekty, pak taková determinace náhodu na této rovině determinace vylučuje. Činitel determinuje všechny aspekty svého jednání. Proto zde nelze hovořit o náhodě. Činitel je zároveň nedeterministickou příčinou svého jednání. To znamená, že své jednání nedeterminuje deterministicky, tj. jeho vlastní determinaci nelze vyvodit z jiných faktorů, kauzálních či nekauzálních.

Souhrnně je tedy k námitce z náhodnosti vůči těm, kdo popírají determinismus, třeba říci toto: Máme zde dva různé pojmy náhody: náhodu coby absenci determinace a náhodu jakožto absenci determinismu. My zde chápeme náhodu v prvním uvedeném smyslu. Náhodu v druhém pojetí nepovažujeme automaticky za náhodu. Absence determinismu je náhodou jen tehdy, je-li zároveň absencí determinace. U námitky z náhodnosti lze tedy rozlišit dvě podoby. Jedna je skutečná a relevantní, druhá nikoliv. Skutečná námitka směřuje proti indeterministické kauzalitě, v níž hraje roli náhoda, protože determinace toho, který z možných účinků vzniká, nepochází (zcela) od příčiny, od ní pochází jen pravděpodobnost, s níž náhodně vznikne ten či onen účinek. Toto je relevantní námitka vůči libertarianismu, který pojímá kauzalitu jako vztah událostí (*event-causal libertarianism*), nikoli proti libertarianismu založenému na kauzalitě činitele (viz níže).<sup>6</sup>

6 Libertariáni jsou inkompatibilisté, kteří zároveň přijímají, že existuje svoboda chtění a jednání, jež a) implikuje existenci alternativních možností chtění a jednání, b) činitel je posledním zdrojem (příčinou) svého chtění a jednání a nic jiného nebo že dané chtění a jednání nemá příčinu (kdyby bylo zdrojem něco jiného nebo nic, pak by činitel nebyl posledním zdrojem svého jednání, neměl by nad jeho vznikem plnou kontrolu). Protože jsou inkompatibilisté a zároveň přijímají svobodu, musí odmítat determinismus.

Naproti tomu zde máme námitku, která tvrdí, že pokud účinek nevzniká deterministicky a není jej tudíž možné vysvětlit kontrastivně (proč vzniká právě tento a proč nevzniká jiný), je náhodný a činitel tedy nad ním nemůže mít kontrolu, vládu. Existuje-li kontrastivní vysvětlení, pak účinek vzniká nutně a opačný účinek je vyloučen jako nemožný. Tato námitka je však lichá, protože je založena na falešném dilematu „buď nutný, nebo náhodný“, přičemž pravdivá je (tautologická) disjunkce „buď nutný, nebo nahodilý“, rozuměj: „buď je, nebo není vyloučena možnost vzniku opačného účinku, než je ten, který vzniká“. Jak je zřejmé z našeho výkladu, „náhodný“ implikuje „nahodilý“, ale naopak to neplatí.<sup>7</sup>

Je také jasné, proč inkompatibilisté a libertariáni jako O'Connor odmítají, že jediné možné vysvětlení je vysvětlení kontrastivní, proč se realizuje *A* a proč ne *ne-A*. Nekontrastivní vysvětlení je sice slabší, ale u nedeterministické kauzality ani jiné být nemůže. Kontrastivní vysvětlení totiž předpokládá, že existují kauzální postačující podmínky, jinak řečeno, předpokládá deterministickou kauzalitu.

Ve snaze vyhnout se třetí premise (ukázat její nepravdivost) někteří argumentují tak, že rozliší kauzalitu událostí od kauzality činitele. V prvním případě do kauzálního vztahu vstupují události nebo stavy. To znamená, že roli příčiny a účinku hrají jen události či stavy. Ve druhém případě roli příčiny hraje činitel a účinkem je událost nebo stav. Vysvětleme na příkladech: zapnutí vypínače způsobilo, že se rozsvítila žárovka. Zapnutí vypínače a rozsvícení žárovky jsou události. Událost je vlastně změna vlastností nějakého objektu. Tato změna implikuje určité časové trvání, minimálně tedy dva časové okamžiky (před změnou a po změně). Místo události můžeme uvažovat stav, tj. skutečnost, že daný objekt má příslušnou kauzálně relevantní vlastnost, aniž by se zde výslovně zohledňoval časový rozměr. Stav, že vypínačem prochází proud, působí, že svítí žárovka. Příkladem kauzality událostí v oblasti, která nás zajímá, může být vztah mezi motivem a jednáním: Petr pocítil touhu uhasit žízeň a uvědomil si, že když si vezme z lednice pivo a napije se, žízeň uhasí. Máme zde dvě mentální události: vznik touhy a vznik přesvědčení, jež jsou motivy pro jednání a které společně vyvolají intenci

7 Proti nutnosti kontrastního vysvětlení v souvislosti s kauzalitou činitele srov. Connor & Jacobs (2013, s. 173–192, srov. s. 187).



tímto způsobem jednat. Komplexní mentální událost, že se Petr rozhodne určitým způsobem jednat, tj. intence vzít si z lednice pivo a napít se, je pak účinkem. Totéž je možno popsat v termínech kauzální relace mezi stavy: Petr je v určitém mentálním stavu, to znamená, že má určité mentální vlastnosti (touhu a přesvědčení), a to působí, že má jinou mentální vlastnost (intenci jednat). Zastánci rozlišení kauzality událostí a kauzality činitele nyní řeknou, že právě svobodné záměrné jednání takto popisovat nelze, tj. nejde ho popisovat v termínech událostí či stavů, jak jsme to právě učinili. Událost či stav intence jednat tak, jak jsme řekli, nevyvolají události či stavy (to že Petr nabyl či má určité motivy), ale činitel sám, Petr. To, co mají tito teoretikové na mysli, je skutečnost, že účinek, Petrovu mentální vlastnost, nevyvolává fakt, že Petr nabyl či má jiné mentální vlastnosti, jako kdyby činitelem byla sama tato vlastnost či příslušný stav (konfigurace, uspořádání) Petrovy mysli či mozku, protože takové působení zbavuje Petra aktivní role. To, že Petr nabyl či má určité mentální vlastnosti (motiv) jej sice ovlivňuje, aby vyvolal intenci, tj. jiný mentální stav, ale přímo tento nově vzniklý mentální stav nepůsobí. Nepůsobí vlastnost či stav Petrovy mysli, ale Petr sám. Kdyby tomu tak bylo, tj. kdyby působila vlastnost či stav Petrovy mysli, tak by Petr byl pasivním pozorovatelem dění, které probíhá bez jeho aktivního přičinění, a nikoliv původcem a aktérem. Podle nich buď událost působí jinou událost nutně, deterministicky, nebo událost vzniká indeterministicky, což zahrnuje náhodu. To, že událost vzniká indeterministicky, znamená, že tato událost buď vůbec není determinována příčinou, nebo se determinace příčiny nevztahuje na některý aspekt události – účinku, případně je dílem náhody fakt, že příčina v jednotlivém případě účinek způsobí. Ve všech těchto případech zde nemůže být kontrola nad účinkem, takže jednání, má-li být svobodné, nemůže být takto působeno. Naproti tomu, zavedeme-li zvláštní typ kauzality, kdy jednání (což je událost) působí činitel, a nikoliv nějaká událost či stav, v němž se činitel nachází, pak takový účinek, tj. vznik jednání, může být kontrolován, protože činitel nepůsobí ani deterministicky, ani náhodně (aby to vylučovalo kontrolu).

Domníváme se, že zavedení rozlišení mezi kauzalitou událostí a kauzalitou činitele není potřebné. Ve skutečnosti je totiž ontologicky základní kauzalita činitele, která platí nejen ve zvláštním případě, kdy je



činitelem rozumem obdařený člověk působící svobodné jednání, ale ve všech případech příčinnosti. Činitelem je vše, co působí ve smyslu aristotelské účinné příčiny: oheň, který ohřívá, magnet, který přitahuje magnetickou silou, elektron, který naráží do jiného, rostlina, která provádí látkovou výměnu s okolím či fotosyntézu. Každý takový činitel má nějakou vlastnost či vlastnosti, díky nimž takto může působit. Tyto vlastnosti mu udělují kauzální schopnost vyvolat určitý účinek, který za určitých kauzálních podmínek buď manifestuje nutně, nebo indeterministicky. Proto lze také druhotně kauzální vztahy zachycovat v jazyce stavů nebo událostí. To však s vědomím, že daný popis není z ontologického hlediska, tj. z hlediska, jak se věci ve skutečnosti mají, základní. Můžeme říci, že fakt, že oheň má vysokou teplotu (stav ohně), způsobil jiný fakt, že voda vře (stav vody). Je to ale proto, že oheň díky tomu, že má příslušnou teplotu, realizuje svou schopnost ohřívát a za podmínky, že se voda nachází v nádobě v jeho blízkosti, v ní vyvolá var. Říkáme-li, že působící příčinou je stav, tj. to, že objekt má určitou vlastnost, lze tomu rozumět buď tak, že působí objekt na základě toho, že má uvedenou vlastnost, nebo tak, že působí daná vlastnost objektu a vyvolává jinou vlastnost na témže či jiném objektu, bez přičinění těchto objektů. První interpretace je správná, druhá nesprávná. Ale právě tato druhá interpretace je základem „standardní teorie“, o níž bude řeč níže. Odmítáme-li zde kauzální působení stavů, týká se to pouze této druhé interpretace. Můžeme tedy říkat, že mentální stav vyvolal jiný, ale musíme si být vědomi, že tím míníme to, že činitel způsobil svůj mentální stav díky tomu, že se nachází v jiném mentálním stavu, a to buď nutně (deterministicky) nebo indeterministicky. V prvním případě nejde o svobodné působení: například když se činitel rozhodl deterministicky na základě nezvladatelného strachu. U takového deterministického působení se zdá jazyk kauzality událostí či stavů případněji, protože se jím jakoby správně signalizuje, že činitel hrál v produkci účinku pasivní, ale nikoliv aktivní roli. Ve skutečnosti ale činitel hraje aktivní roli aspoň v tom smyslu, že je to on, kdo působí. To, čeho je zbaven, je možnost učinit jiné rozhodnutí, či nerozhodnout vůbec, takže ztrácí nad vznikem intence kontrolu. Nemíníme tím však, že by vlastním činitelem byl příslušný mentální stav a ten vyvolával jiný stav, ať už deterministicky či indeterministicky, přičemž činitel by byl

jen trpným pozorovatelem dění. Toto je model spojený se „standardní teorií“, kterou podrobujeme kritice ve třetí části knihy.

### TŘETÍ ČÁST KNIHY: Kritika „standardní teorie“ a kauzalita činitele jakožto schopnost kontingentní determinace

Ve třetí části knihy nejprve představujeme, co míníme „standardní teorií“, a uvádíme důvody v její prospěch. Poté nabízíme alternativu, totiž kontingentní determinaci chtění a jednání činitelem jako odpověď na námitky proti „standardní teorii“. Přitom se inspirujeme zejména dílem současného anglického analytického filozofa Thomase Pinka.<sup>8</sup>

V analytické filozofii se od poloviny šedesátých let 20. století rozšířil vlivem D. Davidsona výklad záměrného jednání, které je svobodné, jenž lze nazvat pro svůj vliv „standardní teorií“.<sup>9</sup> Teorie má původ u T. Hobbesa, D. Huma a dalších raně novověkých anglických filozofů.<sup>10</sup> Svoboda se zde chápe jako dobrovolnost: jednání je určeno vnitřním motivem činitele (touhou a poznáním, že dané jednání vede k naplnění předmětu tužby), který určuje jednání. Ono „určuje“ ve „standardní teorii“ znamená kauzální působení ve smyslu deterministické (eficientní) kauzality. Existenci samotného vnitřního motivu lze také chápat jako produkt deterministické kauzality. Jak vidíme, z teorie přímo plyne kompatibilismus, tedy názor, že existence svobodného jednání je slučitelná s determinismem.

Jednání je podle „standardní teorie“ záměrné, pokud je působeno mým příslušným mentálním stavem, motivem; je dobrovolné, pokud jedním tak, jak chci, tj. ve shodě s mou touhou a poznáním.<sup>11</sup> Jinak ře-

---

8 Srov. zejm. Pinkův úvod do problematiky svobody vůle: Pink (2004a). Ten doplňujeme studiem, v nichž Pink v něčem posunul svou argumentaci: Pink (2003), Pink (2008) a Pink (2019).

Tyto články v sevržené podobě představují argumentaci, která se detailněji objevuje v knize *Self-Determination*. (Pink, 2016).

9 Tak také např. Velleman (1992, s. 461–481). Pro počátek „standardní teorie“ u Davidsona a její spojení s nereduktivním fyzikalismem založeným na supervenienci mentálních stavů na fyzických srov. Davidson (1963, s. 685–700).

10 Hobbes (1841, 5. díl, zejm. s. 367 a 372–373). K expozici Hobbesovy teorie v dialogu s Bramhallem srovnej Russell (2011). Ke kritice pak Pink (2004b, s. 127–153).

11 V tomto pojetí záměrné a dobrovolné jednání u člověka rozsahově splývají, pojmy se liší nanejvýš obsahově. Tam, kde se rozezná jako motiv intence, která se liší od touhy a přesvědčení, tam lze

čeno, je zde „správný“ kauzální řetězec od motivů k jednání.<sup>12</sup> Sebevláda, sebedeterminace, jinak řečeno, svoboda mého jednání, není nic jiného než tato dobrovolnost jednání. Tu ruší nátlak, působení vnější příčiny. Brání se tak kauzálnímu vlivu mých vlastních motivů na mé fyzické jednání. Buďto vnější příčina fyzickému jednání brání, nebo fyzické jednání působí proti vnitřním motivům, tj. je s nimi v nesouladu.

Je jasné, že působí-li ve standardní teorii motivy jednání deterministicky, je inkompatibilistická svoboda tohoto jednání vyloučena (viz argument následku výše). Je vyloučeno, aby činitel za splnění stejných kauzálních podmínek (s totožnou množinou motivů) jednal odlišně, než ve skutečnosti jedná. Neexistují alternativní možnosti jednání, resp. existují jen tehdy, uvažujeme-li odlišné motivy. Neplatí absolutně, že činitel, jedná-li, mohl jednat také odlišně, ale jen relativně k odlišnému chtění: činitel, jedná-li, mohl jednat také odlišně, *kdyby chtěl* (tj. kdyby měl jinou motivaci, odlišnou dvojici touha/poznání, která by působila intenci jednat odlišně).<sup>13</sup>

Coby alternativu stavíme proti „standardní teorii“ kauzalitu činitele, kterou chápeme jako polyrealizovatelnou schopnost kontingentní determinace. Vysvětleme, co to znamená, a jak se liší od běžné nutné a indeterministické kauzality v přírodě. Odlišme pojmově a obecně vztah schopnosti vůči jejímu realizujícímu aktu, a to jen jednoho druhu nebo více druhů, podle typu schopnosti. Odlišme dále vztah činitele vůči schopnosti, kterou realizuje. Je-li touto schopností specificky schopnost svobodné volby, pak první vztahuje schopnost vůči různým aktům (aktuálním jednáním) *A*, *B*, *C* včetně nulového aktu, druhý pak vztahuje činitele ke schopnosti svobodné volby vztahem determinace (aby realizovala akt *A*, *B*, *C* nebo nerealizovala nic). První vztah u této specifické

---

některá jednání chápat jako nezáměrná, ale dobrovolná, ovšem pouze tehdy, nepůsobí-li touha/přesvědčení intenci nutně, což není případ „standardní teorie“, jak ji zde prezentujeme.

12 O tom, co znamená „správný“ bude ještě řeč níže v souvislosti s námitkou z deviantních kauzálních řetězců.

13 Z Humovy definice svobody plynou hypotetické alternativní možnosti, tj. existuje možnost, za předpokladu, že je tu opačný motiv:

Svobodou můžeme tedy mínit pouze moc jednat či nejednat podle určení vůle – tj. rozhodneme-li se zůstat v klidu, můžeme, a rozhodneme-li se pro pohyb, můžeme rovněž. Tato hypotetická svoboda náleží nyní podle všeobecného souhlasu všem, kdo nejsou ve vězení či v řetězcích. Hume (1996, sekce 8, O Svobodě a nutnosti, část I, § 23, s. 135–136).

schopnosti udává polyrealizovatelnost schopnosti (*multi-wayness*), tedy to, že schopnost může být za týchž okolností determinována více způsoby, může být realizována kontrárními akty, nebo vůbec nerealizována. Druhý vztah charakterizuje *kontingentní* determinovatelnost schopnosti činitelem. Činitel determinuje schopnost k činnosti tak, že by byl býval mohl determinovat k odlišné specifické činnosti či nedeterminovat k žádné činnosti. Kontingentní determinace je redukovatelná na realizaci polyrealizovatelné schopnosti. Lze rozlišit trojí druh polyrealizovatelné schopnosti za týchž kauzálních okolností: (i) tu, která umožňuje realizovat různé akty, přičemž se vždy nutně některý akt realizuje, (ii) schopnost, která umožňuje realizovat různé akty včetně možnosti, že se žádný nerealizuje, (iii) schopnost, která umožňuje realizovat jen jediný akt, nebo žádný akt. Schopnost svobodné volby je polyrealizovatelná ve smyslu (ii).

Schopnost svobodné volby se neodlišuje od běžné kauzální schopnosti tím, že je polyrealizovatelná. Polyrealizovatelná je rovněž indeterministická kauzální schopnost. To, co charakterizuje schopnost svobodné volby, je právě kontingentní determinace. Akty této schopnosti jsou determinované činitelem, ale na rozdíl od aktů (deterministické) kauzální schopnosti nejsou determinované nutně, tj. k jedinému specifickému výsledku (aktu), přičemž by zároveň neexistovala možnost jej nerealizovat.<sup>14</sup> Neplatí tedy námitka z náhody. Ta by platila tehdy, kdyby činitel nedeterminoval, který z množiny různých výsledků se realizuje. Tak je tomu u indeterministické kauzality, u níž je determinace aktu záležitostí náhody. Zatímco tedy kauzální schopnosti v přírodě se vyznačují buď nutnou determinací, nebo indeterminací, schopnost volby má uvedený zvláštní charakter.<sup>15</sup>

14 Samotná možnost determinace jen k jedinému výsledku ještě nezakládá nutnou determinaci, jak plyne z toho, co bylo řečeno výše. Determinace, která buď determinuje k jednomu výsledku nebo k jeho nerealizaci, je rovněž kontingentní a příslušná schopnost, která je takto determinována, je (slabě) polyrealizovatelná ve smyslu (iii) výše. U nutné determinace musíme vyloučit logickou možnost nerealizace výsledku (je-li přítomna příčina a jsou splněny všechny kauzální podmínky). Schopnost deterministické kauzality je tedy monorealizovatelná.

15 Indeterministická schopnost nemusí znamenat, že činitel ji vůbec nedeterminuje. Činitel zpravidla determinuje jen částečně, tj. (nutně) determinuje to, že se realizuje určitý typ aktu, přičemž to, který ze specifických kontrárních aktů se realizuje, determinováno není, je to dílem náhody. Např. srážka jedné částice s jinou determinuje její akceleraci. Přesná míra akcelerace je však dílem náhody. Nebo částečná determinace představuje jen různé rozložení pravděpodobností, s nimiž se realizují kontrární možnosti.

Nyní uvedeme námitky proti „standardní teorii“, z nichž vyplyne, jakou jinou teorií zmíněný rozšířený výklad nahradit. Většina těchto námitek vychází z toho, že „standardní teorie“ neodpovídá tomu, co zakoušíme jako aktéři. Ukazuje se, že teorie, která chápe svobodu rozhodování robustně, jako polyrealizovatelnou schopnost kontingentní determinace jednání, je dobře zakotvena ve zkušenosti samé.

Námitka mizejícího činitele (*disappearing agent objection*) je formulovatelná proti teoriím, které staví na kauzalitě stavů nebo událostí a chápou jednání jako sérii stavů či událost působenou mentálním stavem či mentální událostí. Zvláště silná je proti libertarianismu událostí (*event libertarianism*), protože tam navíc přichází do hry náhoda místo deterministické kauzality, jak je tomu ve „standardní teorii“. <sup>16</sup> Platí-li, že jednání působí mentální stavy, pak jaká je role činitele jako takového? Předpokládá se, že činitel je totožný se souborem jeho mentálních stavů, ale to je určitý redukcionismus, který je těžko přijatelný. V každém případě tato pozice odporuje zkušenosti. Já, coby nositel mentálních stavů, nejsem totožný s žádným jednotlivým mentálním stavem nebo jejich souborem. Pozoruji-li, že nějaká nezvladatelná touha či závislost coby (prožívaný) mentální stav působí moji intenci a jednání, pak vím, že to nejsem já, že já si přeji jednat jinak. Hovořím-li v takovém případě v jazyce kauzality stavů či událostí, pak proto, že nejsem skutečným původcem svého jednání coby činitel, třebaže jsem to já, kdo koná, a to s nutností blížící se deterministické kauzalitě. Vnitřní zkušenost vydává přesvědčivé svědectví o tom, že činitel je zdrojem svého svobodného rozhodování a jednání, a nikoliv pasivním pozorovatelem dění, resp. toho, jak jeho mentální stavy působí jednání.

Další námitky proti standardní teorii, totiž námitky z deviantních kauzálních řetězců, si byl vědom již Davidson.<sup>17</sup> Problém spočívá v tomto: motivy způsobí jednání, které ale není záměrné a dobrovolné. Obtíž je tedy v tom, že tyto motivy nepředstavují kauzální postačující podmínku pro dobrovolné a záměrné jednání. Tento kauzální řetěz od motivů k jednání není „správný“, ale je deviantní. Přesněji řečeno, motiv je správný, ale způsob, jímž je jednání způsobeno, nikoliv. Davidson přiznává, že

<sup>16</sup> Srov. např. Pereboom (2014, s. 59–69), Pereboom (2017, kap. 1). Dále Velleman (1992).

<sup>17</sup> Davidson (2001, s. 79).

udat obecná kritéria správnosti způsobu působení, aby jednání bylo záměrné, není možné.

Závažnou námitkou je, že „standardní teorie“ neodlišuje dostatečně aktivní a pasivní motivaci. Ve „standardní teorii“ je vznik intence deterministicky způsoben dvojicí motivů touha/poznání. Proto je stejně pasivní jako samy tyto motivy. To však zcela odporuje naší zkušenosti. Zakoušíme, že zatímco nad motivy a poznáním nemáme přímou kontrolu, rozhodnutí čili tvorba intencí jednat tak a tak za účelem dosažení cíle, což předpokládá volbu cíle jako takovou (že jej chci svým jednáním dosáhnout), je něco, co se nám nepřihází, ale co činíme my sami, co je v naší moci. Zkrátka řečeno, tvorba intence jednat k dosažení cíle je jednáním, a to takovým, nad kterým máme největší kontrolu. Realizace vnějšího jednání, které je touto intencí působeno, již může narazit na vnější překážky v podobě externích příčin, takže se nemusí zdařit. Proto jej máme ve své moci méně. Dobrovolnost se týká pouze vnějšího jednání, které je působeno intencí, nakolik je v souladu s tímto motivem. O dobrovolnosti či nedobrovolnosti nemá smysl hovořit u vnitřního aktu rozhodnutí (tvorby intence k jednání), protože ten není působen žádnými motivy ve smyslu aristotelské eficientní kauzality. Není zde předchůdná intence pojmut intenci k jednání. To ovšem neznamená, že rozhodnutí samo není vůbec motivováno. Je motivováno normativními vlastnostmi cíle, resp. tím, že je poznám. Mentální stav poznání těchto normativních vlastností ale nepůsobí jako eficientní příčina rozhodnutí. Touto eficientní příčinou aktu rozhodnutí, vytvoření intence k jednání, jsem já sám, mé „já“. Normativní vlastnosti cíle působí jinak, jazykem aristotelské filozofie se říká, že působí jako finální příčina. Jako něco, co na mne klade nárok. Metaforicky se hovoří o tahu (v případě účinné příčiny pak o tlaku).

Standardní teorie směšuje dvě různé schopnosti do jedné, totiž motivační schopnost cíle a schopnost činitele k rozhodnutí se pro jednání. Jedna věc je schopnost cíle ovlivnit činitele, jeho přitažlivost, aby se pro něj rozhodl, druhá pak schopnost činitele rozhodnout se. Z pohledu „standardní teorie“ je zde jediná schopnost, a to deterministická kauzální schopnost motivů (touhy/poznání a intence), tedy mentálních stavů. Motivace i sebedeterminace činitele se vysvětlují jediným mechanis-

mem. Toto vysvětlení je sice jednoduché (pracuje s minimem různých objektů a výkladových principů), uniformní (ve smyslu toho, že rozličné interakce v přírodě se vysvětlují stejně) a pracuje se známou a neproblematickou schopností, tj. deterministickou kauzální schopností. To jsou jistě nepochybné teoretické výhody. Na druhou stranu se zdá, že v případě schopnosti motivovat a sebedeterminovat se jde evidentně o dvě různé schopnosti: Mají odlišné nositele i odlišný způsob působení. Cíl, který motivuje, je myšlenkový předmět, mentálně uchopený věčný stav, jenž v realitě dosud neexistuje. Zdrojem sebedeterminace je činitel sám, a nejen některý jeho mentální stav. Ten je jistě důležitý, protože právě proto, že jej činitel má, či lépe řečeno v něm je, může realizovat svou schopnost k jednání. Působení je také odlišné: cíl motivuje, více či méně přitahuje, klade nárok na činitele. Tuto stránku cíle poznává činitel rozumem. Motivovat může cíl jen jako rozumem nazřený, ale není to mentální stav sám, co motivuje, nýbrž jeho poznávaný obsah, věčný stav s danými normativními vlastnostmi. Z vnitřní zkušenosti víme, že jako činitelé jsme pasivními recipienty, jsme podrobeni tomuto působení cíle. Vnitřní zkušenost nás ovšem také informuje o tom, že jsme to my sami, naše „já“, jež je zdrojem rozhodnutí pro nějaké jednání (zdrojem intence) a jeho následné fyzické realizace. Zde nejsme pasivními recipienty, nýbrž aktivními aktéry. Zastánci „standardní teorie“ jsou nuceni tyto vnitřní zkušenosti odmítnout jako iluze.

Mezi pasivní a aktivní motivací (za aktivní motivaci považujeme tvorbu intence) existuje také rozdíl v kritériích jejich racionality. Je racionální po něčem toužit, nakolik je to lákavé, atraktivní, vzbuzující touhu. Naproti tomu jako kritérium racionality rozhodnutí (vzniku intence) atraktivita (normativní vlastnosti) cíle nestačí. Je zde zapotřebí další kritérium, totiž dosažitelnost cíle jednáním. Pokud by cíl nebyl jednáním dosažitelný, pak rozhodnout se pro jeho získání daným jednáním není racionální. Naproti tomu je zcela racionální toužit po věčných stavech, o nichž vím, že je nemohu svým jednáním realizovat. Z toho plyne, že motiv, který je pasivní, nad jehož vznikem nemám přímou kontrolu, a aktivní motiv, který vytvářím (svobodně) sám, mají zcela jiný charakter. Jednomu jsem podřízen a jeho vznik nemám ve své moci, druhý je



podřízen mně a mám jej ve své moci.<sup>18</sup> Teorie, která druhý chápe jako první, nemůže být správná.

Poslední z hlavních námitek proti „standardní teorii“ a jejímu pojetí svobody jako dobrovolnosti pochází ze specifické oblasti zkušenosti, a to z mravní intuice. Všeobecně rozšířené pojetí mravní odpovědnosti za jednání, které Derk Pereboom nazývá odpovědnost v základním smyslu záslužnosti (*basic desert sense*), nehledě na jeho následky či něco jiného, předpokládá svobodu v tom smyslu, že jednání kauzálně deterministicky nevyplývá z ničeho, nad čím by činitel neměl kontrolu, jako je tomu ve „standardní teorii“.<sup>19</sup>

## DRUHÁ ČÁST KNIHY: Filozofie mysli, kritika nereduktivního fyzikalismu a důvody pro dualismus

Výše jsme představili první a třetí část knihy. Druhá část knihy, kterou jsme zatím vynechali, je věnována filozofii mysli a předpokladům v rámci této disciplíny, na nichž kauzalita činitele staví. Kauzalita činitele není kompatibilní s fyzikalismem. Hrozbou je zejména jeho v analytické filozofii často zastávaná *nereduktivní* podoba, která se nachází v pozadí výše představené „standardní teorie“. Ta uznává existenci mentálních stavů, takže je *ontologicky* neredukuje na stavy fyzické (proto „nereduktivní“). Redukuje je však kauzálně, a tak *explanačně*, jak uvidíme.

Kauzalita činitele rovněž vyžaduje určitý typ dualismu. Neplatí-li nereduktivní fyzikalismus (i jiné formy fyzikalismu), zůstávají dveře dualismu otevřeny. Snažíme se obhájit racionální přijatelnost dualismu,

18 Srovnáváme zde touhu po předmětu coby motiv vznikající spontánně a rozhodnutí jednat tak, abych předmět získal (intence). Ve scholastické teorii, kterou nalezneme např. u T. Akvinského jsou přítomny tři elementy: (i) přitažlivost předmětu, (ii) rozhodnutí pro předmět (záměr) a (iii) rozhodnutí jednat k dosažení cíle (intence), jež vede k jednání. Pouze první, tj. přitažlivost předmětu, nemáme ve své moci, druhé a třetí máme.

19 Pereboom píše, že pro inkompatibilistu, jímž je i on sám, je nutnou podmínkou mravní odpovědnosti za jednání následující: Jednání je svobodné ve smyslu vyžadovaném mravní odpovědností pouze tehdy, když nevzniká jako produkt deterministického procesu, v jehož historii se nacházejí kauzální faktory mimo kontrolu činitele. D. Pereboom, *Living Without Free Will*, Cambridge University Press, Cambridge 2001, s. 3. Pereboom se domnívá, že toto kritérium je logicky nezávislé na kritériu, aby činitel mohl jednat jinak, než jedná, tj. na tom, že je schopnost svobodného jednání polyrealizovatelná. Toto druhé kritérium pro mravní odpovědnost odmítá a souhlasí zde s kompatibilitou jako H. Frankfurt a J. M. Fischer.

pokud ne substančního, tak alespoň slabšího dualismu, totiž dualismu vlastností. To je nutné kvůli zamezení tomu, aby se kauzalita činitele nebrala vážně jenom proto, že vyžaduje nějaký dualismus a veškerý dualismus byl považován za „mysterianismus“ bez jakéhokoli rozumového oprávnění. Kniha prezentuje jen nástin takové argumentace, protože filozofie mysli není její hlavní zaměření.

Nejprve tedy shrňme argumentaci proti fyzikalismu. Domníváme se, že fyzikalismus neplatí v žádné verzi. Jak známo, reduktivní fyzikalismus (teorie identity) se neslučuje s vícečetnou realizovatelností (*multiple realizability*).<sup>20</sup> Teorie identity, podle níž je druh mentálního stavu totožný s druhem stavu mozku, záhy po své formulaci čelila klíčové námitce: Druhově jeden mentální stav, např. bolest u různých druhů organismů, je biologicky realizován velmi odlišnými způsoby. Fenomén mozkové plasticity odhaluje, že jeden mentální stav může být realizován vícero konfiguracemi vlastností mozku, tedy vícero mozkovými stavy, a to i u jediného druhu organismu, člověka. Pak je zřejmé, že daný typ mentálního stavu nemůže být totožný s jedním typem mozkového stavu (s jedním typem konfigurace či biologické realizace). Proto mezi stavy fyzickými a mentálními není identita (tzv. typová identita).

Naproti tomu nereduktivní fyzikalismus naráží na problém vyloučení. Pokud jsou mentální stavy odlišné od fyzických a platí princip kauzální uzavřenosti fyzického, nemohou být mentální stavy kauzálně činné, což je důsledek odporující zkušenosti, protože se nám zdá evidentní, že naše mentální stavy hrají kauzální roli vůči dalším mentálním stavům i stavům fyzickým (např. chci vyvolat vzpomínku a ta se vybaví, nebo chci pohnout prstem a ten se pohne). Proto je třeba odmítnout princip kauzální uzavřenosti fyzického, který je nutnou podmínkou fyzikalismu. Vyjádřeme přesněji: Nereduktivní fyzikalismus stojí na dvou níže uvedených tezích (2. a 3.) a tezi o supervenienci mentálních stavů na fyzických (5.). Problém vyloučení je tento: Kauzalita mentálních stavů (1. teze) není slučitelná se zbylými dvěma tezemi níže (2. a 3.), platí-li zároveň předpoklady o nemožnosti přeurčení a superveniencie mentálních stavů na fyzických (4. a 5). Proto je nereduktivní fyzikalista nucen opustit tezi 1., což se nám však, jak bylo řečeno, jeví jako nepřijatelné:

20 Putnam (1975, kap. 14, 18, 19, 21). Také J. Fodor v článcích z let 1965–1974.

1. Mentální stavy jsou kauzálně aktivní (působí fyzické stavy i mentální stavy).
2. Princip kauzální uzavřenosti fyzického: každý fyzický stav má postačující fyzickou příčinu.
3. Mentální stavy jsou odlišné od fyzických stavů.
4. Nemožnost ontologického přeurčení: každý stav nemůže mít více než jednu ontologickou postačující podmínku (ať už je determinace konstituční či kauzální).<sup>21</sup>
5. Závislost mentálních stavů na fyzických (silná supervenience): pro výskyt každého mentálního stavu je nutně postačující výskyt některého fyzického stavu.

Určitý problém vzniká u 5. teze. O jaké nutnosti zde hovoříme? To, zda uvedený operátor nereduktivní fyzikalista bere jako nomologickou nutnost, tedy korelaci ve všech možných světech, v nichž platí tytéž relevantní přírodní zákony, či jako metafyzickou nutnost, tedy korelaci ve všech možných světech, závisí na tom, zda je instanciace fyzických vlastností slučitelná s tím, aby platily odlišné fyzické zákony popisující kauzální interakce těchto vlastností a jejich determinační schopnosti. Pokud ano, lze nomologickou nutnost odlišit od metafyzické a nereduktivní fyzikalista je s to pomocí ní vyjádřit svou tezi o supervenienci. Pokud ne, pak je příslušná modalita ve fyzikalistově tezi o supervenienci mentálních stavů na fyzických výlučně modalitou metafyzickou. Kniha se ve druhé části podrobně věnuje těmto možnostem, ale pro celkový argument ve prospěch kauzality činitele to není rozhodující, proto zde z důvodu úspory místa tyto úvahy nereprodukuje.

Supravenienci mentálního na fyzickém může přijmout i dualista, třebaže na rozdíl od nereduktivního fyzikalisty nemusí. Rozdíl mezi nimi nespočívá v síle modality modálního operátoru, zda jde o nutnost nomologickou nebo metafyzickou.<sup>22</sup> Ať už tedy ten či onen přijímá no-

---

<sup>21</sup> Zde jsme formulaci oproti předchozí diskusi zobecnili na ontickou postačující podmínku a nemožnost mít více než jednu takovou podmínku.

<sup>22</sup> Například J. van Cleve ve Van Cleve (1990, s. 222), chybně tvrdí, že rozdíl spočívá v tom, že zatímco nereduktivní fyzikalisté jsou vázáni k přijetí metafyzické supervenience (jež je základem slabé emergence), dualisté k přijetí nomologické, která je základem silné emergence (van Cleve není necessarista s ohledem na přírodní zákony).

mologickou nebo metafyzickou supervenienci, hlavní rozdíl mezi nimi spočívá v tom, že zatímco nereduktivní fyzikalista míní, že existenci všech mentálních vlastností lze vyvodit čistě na základě znalosti, které fyzické vlastnosti existují a které fyzické a psychofyzické zákony platí. Dualista to popírá. Psychofyzické zákony stanoví determinační vztah mezi fyzickými a mentálními vlastnostmi: které fyzické vlastnosti jsou postačující ontologickou podmínkou pro ty či ony mentální vlastnosti. Toto vyvození všech mentálních fakt z fyzických umožňuje kauzální uzavřenost fyzického a „překlad“ z fyzických vlastností do mentálních prostřednictvím psychofyzických zákonů. Naproti tomu podle dualisty žádné takové vyvození všech mentálních fakt z fyzických čistě na základě fyzických a psychofyzických zákonů není možné. To proto, že mentální vlastnosti samy hrají kauzální roli a působí některé fyzické vlastnosti, takže neplatí princip kauzální uzavřenosti fyzického. Existují tedy indeterministické psychofyzické zákony ještě jiného druhu: v nich některá mentální vlastnost představuje ontologickou nutnou podmínku pro vznik určité fyzické vlastnosti. Tyto zákony nelze odvodit z fyzických a psychofyzických zákonů prvního druhu. To znamená, že se u mentálních vlastností objevují nové kauzální schopnosti, které nejsou převoditelné na kauzální schopnosti čistě fyzických vlastností. „Vynoření se“ mentální vlastnosti a tím i kauzální schopnosti, kterou uděluje nositeli, na základě fyzických vlastností a s nimi spojených kauzálních schopností jeho částí popisuje pojem *emergence*.

Emergence je inverzním vztahem k redukci.<sup>23</sup> O emergenci mentální vlastnosti z fyzické báze tedy hovoří jak nereduktivní fyzikalisté, tak dualisté, a to dualisté vlastností i emergenční substanční dualisté. O emergenci se obecně zpravidla hovoří v kontextu mereologického vztahu částí a celku. Protože je celek uspořádaný a má i jistou dynamiku v čase, často se také hovoří o *systemu* a jeho částech. Vlastnosti částí, které tvoří tento celek či systém, nazýváme nižší. Udělují těmto částem určité kauzální schopnosti, které popisují tzv. nižší zákony. Sem patří nerelační vlastnosti jednotlivých částí i relační vlastnosti, které má více částí najednou,

23 Užitečné představení historie pojmu emergence ve filozofii, jenž má původ v díle tzv. britských emergentistů od J. S. Milla (*A System of Logic*, 1843) po C. D. Broada (*The Mind and Its Place in Nature*, 1925), včetně určitého specifického pojetí vztahu emergence a superveniencie nalezeneme u McLaughlin (1997, s. 25–43).

tzn. jejich vazby a uspořádání. Díky těmto nižším vlastnostem a relacím částí získá souhrn částí schopnost působit emergenci (synchronně determinovat) vyšší vlastnosti tohoto celku, která mu uděluje novou kauzální schopnost. Emergující vlastnost je tedy vlastnost složeného celku či systému a nikdy ne vlastní části tohoto celku. Jde o vlastnost jiného druhu, než jsou vlastnosti částí a je jimi determinována: obvykle se říká, že emergující vlastnost supervenuje na nižších vlastnostech částí, relačních i nerelačních, ale některé koncepce emergence chápou tuto determinaci jako kauzalitu.

Uveďme definiční kritéria emergenčního vztahu. Mentální vlastnost  $M$  nositele emerguje z fyzických vlastností jeho částí  $F_1, F_2, F_3, \dots$ , právě tehdy, když

- a)  $M$  supervenuje na  $F_1, F_2, F_3, \dots$
- b) Pro libovolnou mentální vlastnost  $M$  není fakt, že nositel bude mít  $M$ , v principu vyvoditelný z (i) fakt, že jeho části mají  $F_1, F_2, F_3, \dots$ , (ii) fyzických zákonů popisujících kauzální interakce fyzických vlastností částí nositele a (iii) zákonů o determinaci mentálních vlastností fyzickými.

To je definice tzv. silné (někdy zvané radikální) emergence, která se liší od slabé emergence v podmínce b). Zastává ji emergentní dualista a není slučitelná s fyzikalismem. Podmínka u slabé emergence, která naopak slučitelná s fyzikalismem je, zní takto:

- b)\* Pro libovolnou mentální vlastnost  $M$  je sice fakt, že nositel bude mít  $M$ , vyvoditelný z (i) fakt, že jeho části mají  $F_1, F_2, F_3, \dots$ , (ii) fyzických zákonů popisujících kauzální interakce fyzických vlastností částí nositele a (iii) zákonů o determinaci mentálních vlastností fyzickými, ale toto vyvození je nečekané či překvapivé.

Nečekanost či překvapivost je zpravidla dána praktickou neproveditelností vyvození kvůli extrémní složitosti, množství faktorů a proměnných či chaotickému chování systému. To podstatné, co říká podmínka b), je toto: známe-li v přítomném okamžiku všechny fyzické vlastnosti všech

částí systému, a máme-li k dispozici fyzické a psychofyzické determinační zákony, takže známe i všechny mentální vlastnosti v daném okamžiku, nejsme s to pro následující okamžiky predikovat, jaké bude mít systém fyzické vlastnosti ani jaké bude mít vlastnosti mentální. K tomu bychom potřebovali další typ zákonů (byť takové zákony budou indeterministické), řídící kauzální interakce mentálních vlastností. Takové zákony by byly fundamentální, tj. nevyvoditelné z fyzických a psychofyzických zákonů. V rozlišení silné a slabé emergence jde právě o tuto fundamentalitu.

Alespoň některé mentální vlastnosti jsou obdařeny kauzální schopností zasahovat do fyzického řádu a skrze něj díky determinaci mentálního fyzickým také do mentálního řádu. Tato schopnost tedy působí od vyšších vlastností k nižším, tak řečeno se shora dolů. Proto se nazývá *downward causation*, sestupná kauzalita. Podmínka b) tedy u silné emergence platí proto, že mentální fakta mají tuto zvláštní kauzální schopnost, která není redukovatelná na kauzální a determinační schopnosti nižších (fyzických) vlastností částí systému. V tomto smyslu jde tedy o novou schopnost: novou v porovnání s těmi, které částem systému udělují fyzické vlastnosti, ať už působí jednotlivě nebo dohromady. Důsledek je jasný: věda, která odhaluje jen fyzické zákony (fyzika, chemie, biologie), si nemůže nárokovat úplnost vysvětlení existence všech fakt. Silný emergentismus (dualismus) tedy odporuje fyzikalismu, nikoliv však naturalismu, protože existence a působení mentálních vlastností není *sui generis*, nýbrž vysvětluje se z přirozených příčin a přírodní věda může postihovat i je. Výklad toho, jak mentální stavy vznikají a působí, tedy není v principu mimo přírodovědecké vysvětlování.

To, co jsme doposud řekli, je slučitelné jak s dualismem vlastností, tak se substančním dualismem. Dualismus vlastností předpokládá substanční monismus. Tedy mysl, „já“, musí být identické s mozkem nebo celým organismem. Pak mozek či celý fyzický organismus je nositelem jak fyzických, tak také mentálních vlastností, nachází se jak ve fyzických, tak mentálních stavech, působí jiné fyzické stavy a synchronně determinuje stavy mentální. Tedy například v  $t_1$  mozek ve fyzickém stavu  $F1$  působí jako nepostačující, nutná, neúplná, nedeterministická příčina toho, že se v  $t_2$  očitne ve fyzickém stavu  $F2$ . Mozek v mentálním stavu  $M1$

je také nepostačující nutná, neúplná a nedeterministická příčina toho, že se v  $t_2$  ocitne ve fyzickém stavu  $F2$ . Tím, že je mozek jak v  $F1$ , tak  $M1$  *současně*, je nepostačující, nutná, úplná a nedeterministická příčina toho, že se v  $t_2$  ocitne ve stavu  $F2$ .

Někdo by mohl namítat, že tak radikálně odlišnou kauzální schopnost nemůže nést biologický organismus či jeho část, mozek, nýbrž že zde musí být rovněž radikálně odlišný nositel daných mentálních stavů a jimi udělené schopnosti, mysl, netělesné „já“.<sup>24</sup> V knize sumarizuji některé současné argumenty, proč přejít k substančnímu dualismu. Zmiňuji argumenty Plantingy, Lowa, Haskera a T. O'Connora. Z úsporných důvodů uveďme jen argument J. E. Lowa. Ten se pokusil ukázat, že nositel mentálních stavů, „já“, nemůže být identický s celým organismem, ani s libovolnou jeho částí. Zní takto:<sup>25</sup>

1. Já jsem subjektem všech a výlučně mých mentálních stavů (mé mentální stavy by nemohly existovat, pokud bych já nebyl).
2. Ani mé tělo jako celek, ani jeho libovolná část nemůže být subjektem všech výlučně mých mentálních stavů (většina mentálních stavů může existovat, i kdyby mé tělo i jeho části nebyly).
3. Já nejsem identický s celkem mého těla ani jeho libovolnou částí.

První premisu považuje Lowe za evidentní, proto hlavní váha argumentace směřuje ve prospěch 2. tvrzení, že na těle jako celku ani jeho libovolné části nezávisí, že mám všechny a výlučně mé mentální stavy. Platí princip, že žádná entita nemůže být považována za subjekt určitých mentálních stavů, pokud by dané stavy mohly existovat, i kdyby daná entita neexistovala. Je jasné, že *mé* mentální stavy by nemohly existovat, aniž bych já existoval. Mentální stavy musí vždy mít subjekt, kterému náleží a žádný jiný; v opačném případě by šlo o jiné mentální stavy (daný subjekt je pro tyto stavy esenciální). Ale mnohé, možná všechny mentální stavy by mohly existovat, i kdyby mé tělo či libovolná jeho část, kterou aktuálně má, nebyla. To znamená, že moje tělo jako celek nemůže být

---

24 Např. podle Morelanda a Craiga implikuje dualismus vlastností kauzalitu událostí, a tak není slučitelný s kauzalitou činitele. Moreland & Craig (2003, s. 241).

25 Lowe (2008, s. 96).



nositelem všech a pouze mých vlastních mentálních stavů. Proto mé „já“ nemůže být identické s mým tělem.

Lowe připouští, že kdyby byl mozek jako celek zničen, byly by zničeny všechny mé mentální stavy. Neexistuje však žádná část mozku taková, např. jeden určitý neuron, že kdyby byla zničena, zanikly by tím všechny mé mentální stavy.

Loweův argument se snaží dokázat substanční dualismus, podle něhož „já“ coby nositel mentálních stavů, není identické s mým tělem či jeho částí. Lowe se však nehlásí k jeho známé karteziánské podobě, nýbrž zastává emergenční substanční dualismus, kdy je substance mého „já“ ontologicky závislá na mém organismu. Svou variantu dualismu nazývá „nekarzeziánský substanční dualismus“ (v angličtině NCS, *Non-Cartesian substance dualism*). Podle Lowa se karteziánský dualismus mylí v tom, že „já“ má výlučně mentální vlastnosti, že je schopné existovat nezávisle na těle a že jednoduchost „já“ nutně implikuje nerozlehlost. U Lowa je „já“ odlišné od těla a libovolné jeho části a jednoduché (shoda s Descartem), ale také rozlehlé, fyzické, mající nejen mentální, ale i fyzické stavy, např. instanciuje vlastnost „být 180 cm vysoký“.<sup>26</sup> Podobně jako socha není identická se svým materiálem, protože má odlišné podmínky persistence, ale to ji nečiní nefyzickou.<sup>27</sup> Socha nemůže bez materiálu existovat. Sdílí s ním řadu vlastností: zaujímá stejné místo, má tentýž tvar, velikost, hmotnost. Analogie samozřejmě není dokonalá, protože socha na rozdíl od substančního „já“ není jednoduchá.

Lowe se domnívá, že jeho dualismus má coby vysvětlení teoretické výhody oproti konkurenčním výkladům.<sup>28</sup> Na rozdíl od karteziánského dualismu nemá potíže s párováním, na které upozornil Kim.<sup>29</sup> Oproti reduktivnímu fyzikalismu uchová svébytnou roli mentální kauzality a její intencionální charakter (tím, že ji neredukuje na neintencionální fyzickou kauzalitu).

26 Některé fyzické stavy jsou připsatelné jen tělu, ty tedy „já“ mít nemůže.

27 Socha má odlišné podmínky persistence (přetrvávání) než kus materiálu, z něhož je vytvořena: Socha zanikne v okamžiku, kdy materiál, např. kus bronzu, výrazně změní tvar. Materiál sám existuje dál. Naproti tomu táž socha přetrvává i výměnou některé své části (např. ruky), materiál ne, resp. již to není týž materiál. Srov. Lowe (2006, s. 5–23, s. 9).

28 Tamtéž, s. 11nn.

29 Kim (2001, s. 30–43). Tuto námitku zde z úsporných důvodů nepředstavujeme.

## Literatura

- Davidson, D. (2001): „Freedom to Act.“ In *Essays on Actions and Events*, D. Davidson. 2. vyd., Clarendon Press, Oxford, 2001, (článek vyšel poprvé v roce 1973), s. 63–81
- Davidson, D. (1963): „Actions, Reasons, and Causes.“ *The Journal of Philosophy* 60 (23): 685–700.
- Dvořák, P. (2020): *Kauzalita činitele. Úvod do analytické diskuse o svobodě vůle*. Togga, Praha.
- Hobbes, T. (1841): *The Questions concerning Liberty, Necessity, and chance*. In *The English Works of Thomas Hobbes of Malmesbury*, ed. W. Molesworth, John Bohn, London, 1841, druhý reprint Scientia Verlag, Aalen, 1966.
- Hume, D. (1996): *Zkoumání o lidském rozumu*. Přeložil J. Moural. Svoboda, Praha.
- Kim, J. (2001): „Lonely Souls: Causality and Substance Dualism.“ In *Soul, Body, and Survival: Essays on the Metaphysics of Human Persons*, ed. K. J. Corcoran, Cornell University Press, Ithaca, s. 30–43.
- Lowe, E. J. (2006): „Non-Cartesian Substance Dualism and the Problem of Mental Causation.“ *Erkenntnis* 65 (1, Prospects for Dualism: Interdisciplinary Perspectives): 5–23.
- Lowe, E. J. (2008): *Personal Agency: The Metaphysics of Mind and Action*. Oxford University Press, Oxford.
- McLaughlin, B. P. (1997): „Emergence and Supervenience.“ *Intellectica* 25 (2): 25–43.
- Moreland, J. P. & Craig, W. L. (2003): *Philosophical Foundations for a Christian Worldview*. InterVarsity Press, Downers Grove, IL.
- O'Connor, T. & Jacobs, J. (2013): „Agent Causation in a Neo-Aristotelian Metaphysics.“ In *Mental Causation and Ontology*, S. C. Gibb, E. J. Lowe & R. D. Ingthorsson, Oxford University Press, Oxford, s. 173–192.
- Pereboom, D. (2001): *Living Without Free Will*. Cambridge University Press, Cambridge.

- Pereboom, D. (2014): „The Disappearing Agent Objection to Event-Causal Libertarianism.“ *Philosophical Studies* 169 (1): 59–69.
- Pereboom, D. (2017): „Responsibility, Agency, and the Disappearing Agent Objection.“ In *Le Libre Arbitre. Perspectives contemporaines*, ed. J.-B. Guillon, Collège de France, Paris. Dostupné z: <https://books.openedition.org/cdf/4942#entries>.
- Pink, T. (2003): „Action and Self-Determination.“ *Intellectica. Revue de l'Association pour la Recherche Cognitive* (36–37): 247–259.
- Pink, T. (2004a): *Free Will. A Very Short Introduction*. Oxford University Press, Oxford.
- Pink, T. (2004b): „Suarez, Hobbes and the Scholastic Tradition in Action Theory.“ In *The Will and Human Action From antiquity to the present day*, eds. M. W. F. Stone & T. Pink, Routledge, London, s. 127–153.
- Pink, T. (2008): „Intentions and Two Models of Human Action.“ In *Reasons and Intentions*, ed. B. Verbeek, Routledge, London, s. 153–179.
- Pink, T. (2016): *Self-Determination. The Ethics of Action*. 1. díl. Oxford University Press, Oxford.
- Pink, T. (2019): „Freedom, Power and Causation.“ *Organon F* 26 (1): 1–28.
- Putnam, H. (1975): *Mind, Language and Reality: Philosophical Papers*, 2. díl. Cambridge University Press, New York.
- Russell, P. (2011): „The Free Will Problem [Hobbes, Bramhall and Free Will].“ In *The Oxford Handbook of Philosophy in Early Modern Europe*, eds. D. M. Clarke & C. Wilson, Oxford University Press, New York 2011, s. 424–444.
- Van Cleve, J. (1990): „Mind-Dust or Magic? Panpsychism Versus Emergence.“ *Philosophical Perspectives* 4 (*Action Theory and Philosophy of Mind*): 215–226. DOI:10.2307/2214193.
- Van Inwagen, P. (1983): *An Essay on Free Will*. Oxford University Press, Oxford.
- Velleman, J. D. (1992): „What Happens When Someone Acts?“ *Mind, New Series*, 101 (403): 461–481.

## Abstract

### ***Agent Causation: Introducing the Chief Arguments of the Book***

The paper summarizes chief arguments in the author's monograph *Agent Causation: Introduction to the Analytic Debate on Free Will* (Togga 2020). The book defends one particular theory, that of agent causation, as best suited to preserve our common intuitions regarding the meaning of free choice and its consequences in the sphere of moral responsibility. In the first part it introduces the terms of the debate and a key argument for incompatibilism. The main competitor, compatibilism of the so-called standard theory (D. Davidson), is rejected in the third part of the book. A central thesis underlying the view of agent causation defended in the book is that contingent (self-) determination is a kind of mean between deterministic and indeterministic causation. The second part of the book serves a subservient purpose defending dualism of some kind (either property dualism or substance dualism) against non-reductive physicalism via the so-called exclusion problem.

Key words: freedom, determinism, agent causation, incompatibilism, compatibilism

Dvořák, P. (2021): „Kauzalita činitele: představení hlavních argumentů knihy.“ *Filosofie dnes* 13 (1): 3–28. Dostupné z [www.filosofiednes.ff.uhk.cz](http://www.filosofiednes.ff.uhk.cz).