

Vysoká škola ekonomická v Praze

Fakulta informatiky a statistiky

Katedra informačních technologií

Studijní program: Aplikovaná informatika

Obor: Kognitivní informatika

**Svoboda vůle a přístup k informacím o
vlastních intencích**

DIPLOMOVÁ PRÁCE

Student : Bc. Ondřej Havlíček

Vedoucí : Mgr. Juraj Hvorecký, Ph.D.

Oponent : Prof. MUDr. Jiří Horáček, Ph.D.

2012

Prohlášení:

Prohlašuji, že jsem diplomovou práci zpracoval samostatně a že jsem uvedl všechny použité prameny a literaturu, ze které jsem čerpal.

V Praze dne 21. června 2012

.....

Ondřej Havlíček

Poděkování

Na tomto místě chci vyjádřit své díky Mgr. Juraji Hvoreckému, Ph.D. za vřelý přístup a cenné rady při vedení mé práce. Velké poděkování patří také MUDr. Oldřichu Vyšatovi, Ph.D. za odborné konzultace a poskytnuté zázemí pro výzkum. Významnou měrou se o vznik této práce zasloužili Prof. Ing. Aleš Procházka, CSc., Michal Hájek, Ing. Jan Kremláček, Ph.D., Simone Kühn, Ph.D. a RNDr. Ivana Malá, CSc. Děkuji také všem účastníkům mého experimentu za jejich ochotu.

Abstrakt

Diplomová práce se zabývá problematikou svobodné vůle a věrohodností introspektivního přístupu k záměrům jednání. Tradiční problém svobodné vůle, odpovědnosti a determinismu se dostává v současnosti do popředí především díky rozvoji kognitivních věd. Na základě jejich poznatků mnoho autorů považuje svobodnou vůli za iluzi. Cílem práce je diskutovat tradiční konceptualizaci tématiky a ukázat její problematičnost. Snažím se namísto ní navrhnout alternativní, kompatibilistické pojetí svobody a odpovědnosti, které je vnitřně koherentní a v souladu s poznatky vědy. Jádrem práce je replikace a modifikace nedávné experimentální studie zkoumající svobodné rozhodování a introspektivní přístup k vlastním intencím. Účastníci experimentu prováděli reflexivní akce, kterým mylně přisuzovali záměrnost.

Cíle teoretické části práce bylo dosaženo na základě studia rozsáhlé zahraniční literatury. Metodou je filosofická analýza pojmů s důrazem na současné vědecké poznatky. Metodou empirické části je kognitivně-psychologický experiment zahrnující měření a analýzu reakčních časů a evokovaných potenciálů.

Nejprve je provedena diskuze pojmů relevantních k problému svobodné vůle. Poté je podán výklad nejdůležitějších experimentů zpochybňujících tradiční intuice o lidské mysli a svobodě vůle. Následuje podrobný popis mého experimentu, jeho výsledků a implikací. Nakonec jsou vysloveny závěry ohledně povahy svobodné vůle a odpovědnosti a některé implikace pro trestní právo a běžný život.

Ačkoli se v teoretické části opírám o poznatky množství autorů, přispívám ve značné míře svými vlastními názory a argumenty. Významným vlastním přínosem je pak především empirická studie, jejíž realizace si vyžádala mimo jiné vytvoření prezentační aplikace, provedení experimentálních úloh s relativně velkým počtem účastníků a matematické a statistické zpracování získaných dat.

Klíčová slova

Svobodná vůle, odpovědnost, determinismus, agence, vědomí, jáství, introspekce.

Abstract

This diploma thesis deals with the problem of free will and the reliability of an introspective access to action intentions. The traditional questions of free will, responsibility and determinism are currently receiving a lot of attention due to the advances in cognitive sciences. Because of various scientific findings, many authors claim that free will is just an illusion. The purpose of this study is to discuss traditional conceptualization of this topic and point out its problems. As an alternative, I am trying to suggest a compatibilistic theory of freedom and responsibility, which is internally coherent and compatible with scientific evidence. The core of my thesis is a replication and modification of a recent experimental study, investigating voluntary decision-making and an introspective access to one's own intentions. The participants of the experiment performed reflexive actions which they misjudged as intentional.

The goal of the theoretical part of my thesis has been accomplished on the basis of examining a voluminous foreign literature. The method involves a philosophical analysis of concepts with an emphasis on contemporary scientific findings. In the empirical part of my work, results were obtained in a cognitive-psychological experiment involving reaction time and evoked potential data acquisition and their analysis.

First, a discussion of concepts relevant to the problem of free will is undertaken. An explication of the most significant experiments that question traditional intuitions about the human mind and free will is then given. A detailed description of my experiment follows, including its results and implications. Finally, conclusions about the nature of free will and responsibility are made together with certain suggestions for both the criminal justice system and everyday situations.

Although I base the theoretical part of my thesis on the results of many authors, I contribute to it to a substantial degree with my own opinions and arguments. The empirical study is the most significant contribution of mine as its realization demanded i.a. a development of a stimulus presentation software, execution of the experimental tasks with a relatively large number of subjects and a mathematical and statistical evaluation of the acquired data.

Keywords

Free will, responsibility, determinism, agency, consciousness, the self, introspection.

Obsah

1	Úvod	1
1.1	Problém svobodné vůle	1
1.2	Struktura a cíle práce	2
1.3	Metoda dosažení cílů	3
1.4	Předpoklady a omezení	3
1.5	Výstupy a přínos práce	4
2	Základní pojmy	6
2.1	Svobodná vůle	7
2.2	Determinismus	9
2.2.1	Kauzální determinismus	10
2.2.2	Genetický a environmentální determinismus	11
2.2.3	Příčinnost	13
2.3	Indeterminismus	14
2.4	Možnost	17
2.5	Odpovědnost	20
2.6	Jáství	22
3	Filosofické problémy ve světle vědeckých poznatků	24
3.1	Kauzální potence vědomí	24
3.1.1	Rychlost vědomí	25
3.1.2	Libetův experiment	26
3.1.3	Dekódování rozhodnutí v lidském mozku	28
3.1.4	Epifenomenalismus	31
3.2	Sense of agency	32
3.2.1	Fenomenologie agence	33
3.2.2	Magické myšlení a hypotéza postdikce	34
3.2.3	Retrospektivní konstrukce soudu volního rozhodnutí	35
3.2.4	Integrovaný model	37
3.3	Důsledky výzkumů	37
4	Retrospektivní konstrukce soudu volního rozhodnutí, replikace a modifikace	39
4.1	Úvod	39
4.2	Metoda	40
4.2.1	Participanti	40
4.2.2	Technické aspekty	41
4.2.3	Procedura	43
4.3	Analýza a výsledky	47

4.3.1 Behaviorální data	48
4.3.2 Evokované potenciály	55
4.4 Diskuze a shrnutí	62
5 Jak svobodní jsme	68
5.1 Ontologická svoboda	68
5.2 Morálně relevantní svoboda.....	69
5.3 Implikace pro trestní právo	71
5.4 Implikace pro běžný život.....	72
6 Závěr	73
Terminologický slovník.....	75
Seznam literatury.....	76

1 Úvod

„*Skutečná magie*, jinak řečeno, znamená tu magii, která není skutečná, zatímco magie, která je skutečná, ta která může být opravdu provedena, *není skutečná magie*.“

Lee Siegel (1991, s. 425)

1.1 Problém svobodné vůle

Pravděpodobně všichni zdraví lidé zažívají minimálně v některých okamžicích svých životů silný pocit, že jsou svobodnými autory svých rozhodnutí. Pocit, že se rozhodli ze své svobodné vůle a pouze oni sami jsou za své akce odpovědní. Na předpokladu svobodné vůle je založeno mnoho společenských institucí. Je důležitý například pro právní systém, který jedince trestá za nezákonné činy, pro které se zřejmě rozhodnout nemuseli. Naopak shovívavě se chová k lidem, kteří jednali pod vlivem nátlaku či manipulace a nerozhodli se tak pro svůj čin zcela svobodně. Předpoklad svobody vůle je také důležitý pro morální hodnocení. Chválíme lidi pro některá jejich rozhodnutí, a haníme pro jiná. Lidská svoboda je důležitá i pro libertariánské ekonomické teorie. A především, svobodnou vůli lze považovat za zásadní předpoklad smysluplnosti lidské existence na tomto světě.

Přesto si různí myslitelé již od starověku kladou otázku, jak je to doopravdy s pociťovanou lidskou svobodou. Tato otázka má různé podoby: Jak můžeme být svobodní, když svět se řídí fyzikálními zákony, když Bůh zná všechny naše budoucí činy, když je má povaha dána mými geny a prostředím? Princip problému je vždy stejný: otázka slučitelnosti svobodné vůle a různých podob determinismu.

Problematika svobody vůle je ve filosofii rozebírána již po minimálně dvě tisíciletí. Co je však zásadním způsobem nové, je rozvoj v oblasti kognitivních věd, především neurověd a informačních věd. Různé experimenty ukazují ve shodě s předpokladem determinismu, že již lze vypočítat predikce pro některá jednoduchá lidská rozhodnutí dříve, než si je toho rozhodnutí člověk sám vědom. Introspekce, dle které je vědomí kauzální příčinou rozhodnutí, tak může být v tomto ohledu zcela klamavá.

Zatímco dříve šlo pouze o problém filosofů, díky vědeckému pokroku se o tomto problému začíná dovídat i širší veřejnost, která si tyto vědecké výsledky nemusí umět vždy dobře vyložit. Ale ani sami vědci pracující na zmíněném problému neoplývají vždy dostatečnými filosofickými znalostmi potřebnými pro správnou interpretaci svých výsledků. V takovém neporozumění se skrývá jisté riziko negativních společenských důsledků. Zda a jak velké, to mohou ukázat psychologické a sociologické studie, přičemž některé z nich již tím směrem skutečně ukazují.

Ať už ale svobodná vůle existuje či ne (a ať už tím pojmem myslíme cokoli), vyvstává další významná otázka: Jak tento fenomenální prožitek svobody našich akcí vůbec vzniká? Jde o důležitou součást širšího problému vzniku vědomí a jáství; součást, která byla do nedávné doby přehlížena. Teprve v současnosti se na tuto otázku začíná zaměřovat empirická věda i filosofie.

Z těchto všech důvodů považuji za více než kdy dřív nezbytné kombinovat vědecký přístup s filosofickou reflexí. Snažit se o dialog mezi přírodními vědami (kam lze v této oblasti počítat především neurovědy, kognitivní psychologii a informační vědu) a obory humanitními (především samozřejmě filosofií, ale nezbytně i právem) a o vzdělávání veřejnosti.

O tematiku svobodné vůle – jednu z nejpálčivějších a nejdůležitějších otázek lidské existence – jsem pojal zájem již před přibližně pěti lety a nyní jsem se rozhodl zvolit si ji jako téma této diplomové práce.

1.2 Struktura a cíle práce

Práce sestává jak z teoretické diskuze, tak z empirického výzkumu. Teoretická část se zaměřuje na představení a diskuzi dle mého názoru nejdůležitějších aspektů problematiky s akcentem na současný stav filosofického bádání a nejdůležitější vědecké poznatky. Cílem práce přitom primárně není nalézt odpověď na otázku, zda svobodná vůle existuje. Ačkoli jde o velmi intuitivní pojem, není totiž zcela jasné, co jím myslíme. Můžeme ale objasňovat významy relevantních pojmů a s přispěním empirických poznatků zkoumat hranice svobody: co vůbec můžeme svobodnou vůlí myslet. Snažím se proto o kompatibilistickou¹ konceptualizaci této problematiky tak, aby byla vnitřně konzistentní, ve shodě s empirickými výsledky, a v co

¹ Kompatibilismus je teze, že svobodná vůle je kompatibilní s determinismem, více dále.

největší shodě s běžným, intuitivním pojetím svobody, odpovědnosti a podobně. Na základě toho se snažím podat nástin odpovědi na to, jaký druh svobody je možný, jaký nikoli, a jaké jsou případné implikace.

Součástí mé práce je také experimentální studie. Ta se snaží ověřit výsledky nedávného výzkumu, který má k problematice svobodné vůle značnou relevanci. Konkrétně hypotézu, že soud o provedení intencionálního rozhodnutí může být nepravdivě zpětně vykonstruován na základě výsledku provedené akce. To znamená, že nemáme úplně pravdivý a přímý introspektivní přístup k informacím o vlastních intencích.

1.3 Metoda dosažení cílů

Problém svobodné vůle je klasickým filosofickým problémem v tom, že vyvstává ze složité spleti pojmů, které zdánlivě vedou ke kontradikcím (McKenna 2009). Metodou řešení takového problému je nalezení takové konceptualizace problému, která je vnitřně koherentní. Navíc musí korespondovat s empirickou realitou, takže nesmí být v rozporu s výsledky empirických studií.

Dosažení cílů v teoretické části práce předcházelo studium především zahraniční *peer-reviewed* literatury. Na základě této rešerše je provedeno představení vybraných nejdůležitějších filosofických aspektů, relevantních pojmů a poznatků ve vědě a jejich diskuse. Jako klíčový filosofický zdroj byla vybrána práce amerického filosofa Daniela Dennetta pro svoji aktuálnost, respekt k vědě a pro představení koncepce kompatibilismu, kterou považuji za nejpropracovanější a pro kterou také argumentuji.

Způsob realizace praktické části je experimentální studie zahrnující behaviorální a elektrofyziologická měření provedená na třiceti zdravých participantech během specifické úlohy. Metodologie této studie je podrobně popsána dále (sekce 4.2).

1.4 Předpoklady a omezení

Přístup v práci spočívá na množině více či méně explicitně vyjádřitelných filosofických předpokladů.

Mým hlavním východiskem je *naturalismus*. V rovině *ontologie*, tedy „toho, co existuje“, jde o tezi, že „vše co existuje, náleží do přírodního (přirozeného) světa“ (Jacobs 2009), což znamená, že předpokládám, že ve světě není nic „nadpřirozeného“. Ontologický naturalismus implikuje odmítnutí např. dualismu a idealismu. V rovině *epistemologie*, tedy způsobu poznávání „toho co je“, tzv. *naturalizovaná epistemologie* „zdůrazňuje použití metod, výsledků a teorií empirických věd. Odlišuje se od přístupů, které zdůrazňují apriorní konceptuální analýzu nebo trvají na teorii poznání, která je nezávislá na konkrétních vědeckých detailech o tom, jak mysl či mozek fungují.“ (Wrenn 2005)

Důležitým východiskem je *metafyzický anti-esencialismus*, minimálně v psychologické a společenské doméně. Jde o antitezi k přesvědčení, že pojmy z těchto domén jsou realizací nějaké (skutečně ontologicky existující) obecné formy, která je určena souborem samostatně nutných a společně dostačujících atributů (ať už je lze objevit a vyjádřit či ne). Dle anti-esencialismu neexistují žádné ontologické esence či podstaty těchto pojmů, ať už jde o „židli“, „hromadu“, „člověka“, „vědomí“ či „svobodnou vůli“. Z tohoto pohledu bádání nad tím „co je to *skutečně* svobodná vůle“ postrádá smysl. Stejně tak není zcela možná klasická (aristotelská) definice těchto pojmů.

Z hlediska filosofických směrů je práce orientována především na anglo-americkou analytickou tradici, především na filosofii Daniela Dennetta.

Významným omezením empirické studie jsou technické prostředky, které byly k dispozici pro elektrofyziologické měření. Dále, participantům nebylo za účast v experimentu placeno, což omezilo množství nasbíraných dat.

1.5 Výstupy a přínos práce

Hlavním výstupem práce je provedená experimentální studie, resp. její popis a výsledky v praktické části práce. Tato část má přibližnou strukturu vědeckého článku a je určena především vědcům a filosofům zabývajícím se problémem svobodné vůle, vědomí, rozhodování, fenomenologií agence či introspektivními iluzemi. Věřím, že tento výstup je velmi přínosný i přesto, že jde v zásadě o replikaci předchozí studie. Především v základu vědecké metodologie je požadavek opakovatelnosti vědeckých výsledků. V ideálním případě by všechny výsledky měly být podrobeny ověření v dalších nezávislých replikacích, zvláště pokud se jedná o výsledky s významnými implikacemi, jako je i tato. To se bohužel zvláště v této oblasti vědy

příliš často neděje. Původní studie je poměrně nedávná (2009) a žádné replikace nebyly ještě provedeny, jak mi bylo potvrzeno její hlavní autorkou. Studie byla provedena s relativně malým počtem subjektů (12, plus dalších 12 v modifikované verzi experimentu), proto je obzvláště důležité provést ověření na širším vzorku participantů (30). Přínosem oproti původní studii je začlenění elektrofyziologického měření (EEG) během experimentu a následná analýza evokovaných potenciálů, což má za cíl přinést objektivnější zhodnocení behaviorálních výsledků studie.

Dalším výstupem je zpracovaná teoretická část práce. Přínos vidím v tom, že toto téma, pokud vím, nebylo dosud v českém prostředí adekvátně zpracováno z hlediska kombinace vědeckého a filosofického přístupu. Cílovou skupinou je primárně spíše odborná veřejnost, doufám ale v přínos i pro širší veřejnost, protože se domnívám, že je důležité, aby o této problematice měla povědomí.

2 Základní pojmy

Problematika svobodné vůle je po teoretické stránce obzvláště složitá, protože zahrnuje velké množství důležitých pojmů (konceptů), které je samozřejmě pro účely této práce potřeba definovat. Základním problémem ale je, že význam konceptů psychologické či společenské domény vzniká spontánně (neúčelově), vyvíjí se a je nutně „fuzzy“. S tím souvisí, že dané pojmy často nelze ani definovat tak, aby výsledné definice byly koherentní, odpovídaly empirickým datům, a nebylo možné proti nim najít intuitivní protipříklady.² Ačkoli se analýze těchto konceptů věnují filosofové již dlouhou dobu, stále se ukazuje, že snad u všech konceptů, kromě těch účelově zavedených a přesně definovaných (jako jsou např. koncepty matematiky) není ani možné k žádné dokonalé definici dojít. Vhodnější se ukazuje být popis takovýchto pojmů pomocí nějaké teorie a vztahů k ostatním pojmům v rámci dané teorie (Margolis a Laurence 2011). Přesto bude užitečné pro porozumění problematice uvést alespoň operativní definice pojmů a následně je vysvětlit podrobněji.

Domnívám se, že mnoho nejasností ohledně otázky svobodné vůle vyvstává právě z jazykových problémů. Největší problém vidím v implicitním esencialismu mnoha autorů. K mnoha pojmům přistupují jakožto k refererujícím k nějaké rigidní entitě, ačkoli pojmy představují spíše vágní modely, které používáme, abychom byli schopni mluvit o nějaké části (doméně) reality. Příkladem jsou kromě samotné „svobodné vůle“ především pojmy „vědomí“ či „já“, jak je podrobněji diskutováno dále. Někteří autoři požadují, abych byl já či mé vědomí zdrojem akcí, aby to byla *má* akce. Ale co je to „já“ nebo „vědomí“, o kterém mluví?

Významným jazykovým problémem je směšování slovníků různých úrovní popisu³ či explanace a různých úrovní kompozice. Může dojít k tzv. kategorické chybě jako v klasickém příkladu návštěvníka v Oxfordu: „Vidím tu knihovnu a mnoho budov s učebnami, ale kde je *univerzita*?“ Je důležité uvědomit si, jaké typy vlastností lze smysluplně⁴ predikovat jakým pojmům. Nelze například navazovat psychologické vlastnosti na fyzikální či biologické entity a naopak. „Mozek znal mé rozhodnutí dříve než já.“ nebo „Mé vědomí ovlivňuje realitu na kvantové úrovni.“ jsou příklady (oblíbených) vět bez jasného smyslu.

² Viz například různé definice svobodné vůle v (O'Connor 2011).

³ Tyto úrovně jsou samozřejmě také vágní, ale dle rostoucí úrovně abstrakce lze uvažovat např. úrovně fyziky, biologie, psychologie či sociologie.

⁴ Tak, aby predikování vlastnosti vracelo pravdivostní hodnotu.

Kvůli povaze našeho jazyka, který je uzpůsobený spíše pro běžné užití, než pro nízkoúrovňové exaktní fyzikální či neurovědné popisy, se výše uvedeným problémům nedá snadno vyhnout. Je možné, že se jich sám dopouštím. Často to nevádí a je možné takové výrazy používat metaforicky, ale je potřeba dát pozor na způsob jejich použití v argumentaci.

Nyní již přistoupím k diskusi hlavních relevantních pojmů. Pro definice bylo využito slovníku vytvořeného odborníky z vědy i filosofie přímo za účelem sjednocení terminologie v problematice svobodné vůle (Haggard et al. 2010) a dalších zdrojů.

2.1 Svobodná vůle

Pokud se někdo ptá, zda existuje svobodná vůle, přijde mi jako jediná správná odpověď další otázka: „Co myslíte svobodnou vůlí?“ Základní pojem „svobodná vůle“ lze sice definovat jako „schopnost vykonávat svobodné akce nebo konat svobodně“ (Haggard et al. 2010). Co to ale je svobodná akce? To již záleží na celkové teorii svobodné vůle a některých dalších teoriích, které zastáváme, ať už explicitně či implicitně.

Je možné vymezit několik základních filosofických pozic podle postoje jejich zastánců k pravdivosti teze o determinismu a o svobodě lidské vůle:

		Vědomé bytosti svobodnou vůli	
		mají	nemají
Realita determinována	je	Měkký determinismus Kompatibilismus	Tvrký determinismus (Inkompatibilismus)
	není	Libertariánství (Inkompatibilismus)	Tvrký indeterminismus (Inkompatibilismus)

Obrázek 1: Základní filosofické pozice v problematice svobodné vůle

Nejintuitivnější tezí zřejmě je, že případná platnost determinismu vylučuje svobodnou vůli. Tento postoj je nazýván jako inkompatibilismus. Naproti tomu je dle některých autorů determinismus se svobodnou vůlí slučitelný, ne-li dokonce pro ni nezbytný. Tento směr je označován jako kompatibilismus či také měkký determinismus.

Někteří inkompatibilisté zůstávají ohledně existence svobodné vůle agnostičtí. Obecně ale lze v rámci inkompatibilismu rozlišit tři podskupiny. Dle metafyzických libertariánů máme svobodnou vůli, a *proto* realita není kompletně determinována. Dle tvrdých deterministů svobodnou vůli nemáme právě proto, že realita determinována je. Další autoři se pak hlásí k pozici tvrdého indeterminismu (jinak také tvrdého inkompatibilismu), tedy že svobodnou vůli mít nemůžeme, i když realita determinována není.

Je tedy jasné, že inkompatibilisté a kompatibilisté musejí za svobodnou akci považovat něco jiného. Podle inkompatibilistů je *nutnou* podmínkou svobodné akce, aby „kombinace minulosti – až do okamžiku akce – a přírodních zákonů ponechávala agentovi dvě či více možností“ (Haggard et al. 2010), přičemž mohou existovat další nutné podmínky. K tomuto tzv. principu alternativních možností přidává asi nejvýznamnější proponent libertariánství Robert Kane další nutnou podmínku, kterou je tzv. ultimátní odpovědnost. Zjednodušeně řečeno, agent musí být zodpovědný za svou vlastní povahu (*character*) (Kane 1998). Viz sekce 2.5 Odpovědnost.

Inkompatibilistickou svobodnou vůli bych si dovolil označit za ontologickou svobodnou vůli. Ve struktuře reality musí být kromě „principu“⁵ determinismu (a případně ještě principu náhody) ještě jakýsi princip svobody, který umožňuje agentovi (který je brán zjevně esencialisticky) být *causa sui*, nezapříčiněnou příčinou jeho akcí či jeho povahy. Je nutno říci, že někteří zastánci libertariánství (např. zmíněný Kane) se snaží podat i naturalistické vysvětlení jen za pomoci náhody, například využitím teorií kvantové mechaniky (viz 2.3 Indeterminismus).

Kompatibilisté naproti tomu kladou následující *dostačující* podmínky: „Jakákoli záměrná akce vykonaná na základě informované, racionální deliberace, přičetnou osobou a bez donucení, je svobodnou akcí, i když byla deterministicky zapříčiněna.“ (Haggard et al. 2010)

Je vidět, že se v případě kompatibilismu jedná o slabší pojem svobodné vůle. Pochopitelnou kritikou inkompatibilistů je, že nejde o *skutečnou* svobodu. Kompatibilistický filosof Daniel Dennett často cituje⁶ dle mého názoru výstižnou analogii s divadelní magií: „*Skutečná magie*, jinak řečeno, znamená tu magii, která není skutečná, zatímco *magie*, která je skutečná, ta která

⁵ Nutno přiznat Rudolfu Carnapovi, že není příliš jasné, co může být výrazem „ontologický princip“ myšleno. Ačkoli ani mně nedává tento výraz jasný smysl, bývá používán tímto způsobem.

⁶ Např. (Dennett 2003, s. 7)

může být opravdu provedena, *není skutečná magie*.“ (Siegel 1991, s. 425) Předestřu, že se s tímto postojem také ztotožňuji; inkompatibilistická, absolutní, ontologická svobodná vůle nemůže být podle mého názoru skutečná. Namísto toho se musíme snažit nalézt „nemagické“ vysvětlení v rámci naturalismu.

Pokud jsem inkompatibilistickou svobodnou vůli označil za ontologickou, mohl bych se pokusit o analogické označení kompatibilistické svobody za epistemologickou svobodu. Cítíme se svobodní, protože nemůžeme vědět, jak se v budoucnu rozhodneme. Pokud bychom to nějakou vyspělou metodou zjistili, stane se to novým vstupem do rozhodování a opět nebudeme mít jistou znalost svého budoucího rozhodnutí.⁷ Naproti tomu, kdyby byl pravdivý fatalismus⁸, nějak bychom zjistili scénář celého svého života a nemohli se řízení osudu nijak zbavit, jistě bychom se příliš svobodní necítili. Jiný způsob, jak nahlížet na kompatibilistickou svobodu jako na epistemologickou, lze nalézt u Dennetta (2004). Jsme podle něj tím svobodnější, čím lépe se dokážeme díky schopnosti predikce vyhnout nepříznivým budoucím událostem, a čím lépe se díky vzdělání dokážeme bránit manipulaci.

Další možnou distinkcí je rozlišení vědomé a ne-nutně-vědomé svobodné vůle. Vědomá svobodná vůle je intuitivní koncept, dle kterého jsou svobodné akce způsobené naším vědomím. Vědomou svobodnou vůli lze považovat za inkompatibilistickou. Zároveň vyžaduje speciální teorii vědomí, dle které má vědomí kauzální potenci. Že je teorie vědomé svobodné vůle značně problematická, ukazuji v sekci 3.1 Kauzální potence vědomí. Naproti tomu kompatibilistické teorie kauzálně činné vědomí vyžadovat nemusí.

2.2 Determinismus

Problém není pouze s vymezením toho, co myslíme svobodnou vůlí. Problematický je i pojem determinismu (a s ním spojených pojmů jako nutnost a příčinnost), který by měl dle inkompatibilistů rozhodovat o možnosti existence svobodné vůle. Lze ho definovat například jako tezi,

⁷ Například kdyby pokročilejší varianta metody popsané v sekci 3.1.3 (Dekódování rozhodnutí v lidském mozku) zjistila, že za *současného stavu* světa (především mého mozku) provedu za 10 vteřin akci A, mohl bych provést akci B, což by znamenalo, že tato zpětnovazební smyčka zneplatní pravdivost předchozí predikce.

⁸ Konkrétně jeho varianta nazývaná někdy finalismus, dle které všechny děje osudově spějí k určitým konečným cílům, čemuž nelze přes jakoukoli lidskou snahu zamezit.

že „v každém okamžiku existuje přesně jedna fyzikálně možná budoucnost.“ (Van Inwagen 1983, s. 3) Různých definic determinismu ale existuje mnoho, stejně jako existuje několik příbuzných druhů determinismu na různých úrovních popisu. Například determinismus kauzální, genetický, sociální, teologický, logický, psychologický či technologický. Jejich společným znakem je určenost věcí jistými příčinami. Nejčastěji se determinismem myslí determinismus kauzální, kterému v podstatě odpovídá i výše uvedená definice.

2.2.1 Kauzální determinismus

Již starověké civilizace vyzorovaly zjevný řád v přírodních dějích a začaly si utvářet kromě nadpřirozených vysvětlení i určitý pojem přírodního zákona (Hawking a Mlodinow 2011). První zastánci determinismu, o kterých víme, byli Leukippos a jeho žák Démokritos. Viz např. Leukippův výrok: „Nic se neděje nazdařbůh, nýbrž vše z důvodu (logos) a nutně.“⁹ V novověku formuloval filosof a vědec G.W. Leibniz tzv. princip dostatečného důvodu: vše, co je, má dostatečný důvod proč je to tak, a ne jinak (Hofer 2010). Postupem času se tak vyvinula celá moderní věda, která by bez existence nějakých kauzálních vztahů ve světě ani neměla co zkoumat. Z náhodných či singulárních událostí nelze tvořit obecné teorie.

Ačkoli nám může připadat jasné, že se svět řídí určitými deterministickými fyzikálními zákony, z filosofického hlediska to není tak jednoduché. Determinismus je totiž v podstatě metafyzická teze. Dle Davida Huma (1739) nelze pravdivost determinismu nijak dokázat, můžeme jen poukazovat na asociace mezi jevy v naší mysli.

Otevřenou otázkou je ale kauzální determinismus i ve vědě. Dle některých interpretací kvantové mechaniky jsou fyzikální zákony na nejnižší úrovni pouze probabilistické, viz 2.3 Indeterminismus.

Koncept determinismu silně závisí na dalším ve vědě otevřeném konceptu, kterým je čas. Standardně přijímané fyzikální teorie jsou deterministické časově symetricky (Hofer 2010). Takové pojetí by znamenalo, že současný stav věcí „fixuje“ všechny stavy budoucí, ale i minulé. Pokud by ale byla realita na hypotetické elementární úrovni diskrétní spíše než spojitá, mohl by hypoteticky jeden současný stav být způsoben i více minulými stavy. Determinismus by tak byl jen „jednosměrný“ ve směru (termodynamické) šipky času.

⁹ Zlomek B 2 z Aetia (Svoboda 1962, s. 121)

Časově symetricky chápal determinismus i Pierre-Simon Laplace. V myšlenkovém experimentu uvažoval entitu („Laplaceův démon“, v podstatě ztotožnitelný s Bohem), která by dokázala pojmout znalosti o všech fyzikálních zákonech, stejně jako o stavu a všech vlastnostech všech částic (či jiných elementárních komponent) vesmíru. Taková entita by pak měla znalost všech budoucích i minulých stavů světa. Musela by bohužel být „nekonečná“ a stát mimo náš svět (Hoefler 2010).

Tento myšlenkový experiment ilustruje časté tendence spojovat determinismus a predikovatelnost. Je nutné uvědomit si, že neschopnost predikovat nějaký jev neznamená, že tento jev není determinován. Tento fakt ilustruje např. deterministický chaos či jiný komplexní systém, kde je množina příčin a síť kauzálních vztahů příliš složitá pro praktickou možnost predikce. Nemožnost predikce může být dokonce principiální, nikoli pouze praktická – jako možná v případě Heisenbergovy relace neurčitosti – a přesto mohou být tyto jevy ontologicky deterministické.

Dalším omylem je směšování determinismu a fatalismu (konkrétně finalismu). Zatímco kauzální determinismus znamená (v časově jednosměrném pojetí) determinaci od příčin k následkům, „tlak odspodu“, finalismus předpokládá naopak metafyzický „tah seshora“ k nějakému cílovému stavu či účelu.

2.2.2 Genetický a environmentální determinismus

Nemusíme se zabývat ani tak metafyzickou a potenciálně spornou otázkou, jako je kauzální determinismus. Komplikaci pro naše intuice o svobodné vůli představují i jiné, (popisně) vyšší úrovně determinace. Nejvýraznější z nich jsou determinace genetická (či biologická) a sociální (obecněji environmentální).

Kým jsme, je dáno z velké části dědičností, naší genetickou informací. S pokrokem genetiky se daří nalézat další a další geny kódující různé morfologické a fyziologické vlastnosti. Silně geneticky podmíněné jsou však i vlastnosti duševní. „V současnosti existuje rozsáhlá evidence, která podporuje závěr, že individuální rozdíly ve většině, ne-li ve všech, spolehlivě měřených psychologických vlastnostech, normálních i abnormálních, jsou podstatně ovlivněné genetickými faktory.“ (Bouchard 2004, s. 148) Tento nedávný přehled například uvádí vliv dědičnosti na osobnostní rysy v rozmezí 40 až 50 %. U inteligence dosahuje míra dědičnosti u 26letých jedinců až 88 %. U komplexnějších rysů je tato míra samozřejmě nižší, u religiozity

například pouze 11 % u 16 letých dívek a 22 % u chlapců. Jiná studie (Waller et al. 1990) ale našla přibližně 50% vliv dědičnosti v této oblasti. U všech obdobných studií silně záleží na použité metodice a uvedená čísla je nutné brát s rezervou. Lze však říci, že čím složitější fenomén a čím je člověk starší, tím je vyšší vliv prostředí.

Dle sociálního konstruktivismu a příbuzných směrů je vliv prostředí a sociálních fenoménů dokonce zásadní určující silou, zatímco vliv genetiky je zanedbáván (Sokal a Bricmont 1999). Stejně jako čistý genetický determinismus je i toto pouze krajní pozice v tzv. „*nature-nurture*“ debatě. V té jde právě o určení vlivu dědičnosti (*nature*) a prostředí (*nurture*) na různé lidské vlastnosti. Je sice nepochybné, že výchova a vstupy z prostředí po celý život zásadním způsobem ovlivňují jak lidskou osobnost, tak i veškeré další jednání. Ale tvrzení, že je člověk při narození *tabula rasa*, je ve světle výše uvedených poznatků neudržitelné.

Naše osobnost je určena z většiny naším mozkiem, především architekturou jeho neurální sítě. Proto bychom mohli přenést nature-nurture debatu právě na tuto úroveň. Vývoj mozku je geneticky určen v mnoha ohledech, ale jeho komplexita je mnohem větší, než jakou je genetická informace schopna kódovat. Pouze geny proto jistě nestačí. V dětství se vytvoří mezi neurony velké množství synapsí – podstatně více, než máme v dospělosti – a tyto synapse jsou pak vyřizovány a přetvářeny podle vstupů z prostředí. Na vývoji nervové soustavy se tak podílí jak dědičnost, tak prostředí.

Prostředí tak hraje při výstavbě organismu značnou roli. Ovlivňuje například do jisté míry expresi různých genů. Na základě genotypu a prostředí je tak utvářen náš fenotyp. Ten společně s kulturou určuje způsob, jakým žijeme, a tím naše prostředí. Adaptace na toto prostředí následně ovlivňuje genotyp dalších generací. Mluví se o tzv. koevoluci genů a kultury (např. Feldman a Laland 1996). Tyto vlivy jsou složitě propleteny. Dokonce samotná hranice mezi tím, co lze považovat za vrozené a co za získané, není ostrá, protože pojem vrozenosti je vágní (Samuels 2004). Jednoznačné řešení nature-nurture debaty je tak kvůli této komplexitě a vágnosti obtížné, ne-li nemožné.

Každopádně, je nepochybné, že za svoji genetickou výbavu nejsme nijak zodpovědní. Nejsme příliš odpovědní ani za svou výchovu a vstupy z prostředí. Zdá se, že na této úrovni popisu determinace genetikou, prostředím a případnými stochastickými jevy (viz 2.3 Indeterminismus) vyčerpávají zdroje vlivů na osobnost člověka a jeho konání (Cashmore 2010).

2.2.3 Příčinnost

Když mluvíme o svobodě či odpovědnosti, používáme nutně koncept příčinnosti. Abych byl svobodný a zodpovědný za akci, musím být *já* příčinou dané akce. Nesmí být její příčinou něco jiného. Problematický je zde ale jak pojem jáství, tak příčiny. Domnívám se, že jde opět o problém úrovní popisu. Nejprve k pojmu příčiny:

Výše bylo pojednáno o kauzálním determinismu, který zjevně implikuje příčinnost. Jde o kauzální vztahy na hypotetické elementární úrovni a popis příčinnosti je zde poměrně jednoznačný. Nebýt toho, že byl vesmír v čase t ve stavu A, nebyl by vesmír v čase $t+1$ ve stavu B. Stav A je příčinou stavu B. Kauzální determinismus je však kvůli své metafyzické podstatě „silnější teze než teze, že každá událost má příčinu nebo množinu příčin.“ (Haggard et al. 2010). můžeme používat nejen ve spojení s kauzálním determinismem, ale opět na více úrovních popisu a abstrakce a do různé hloubky kauzálního řetězu. Obvykle dle toho, jak je to právě nejpraktičtější. Rostoucí míra abstrakce ale může do pojmu příčinnosti přinášet jistou vágnost. Například na výše popsané biologické a environmentální úrovni se již objevují značné komplikace ve vyjasnění, co je příčinou čeho.

Pojem příčinnosti však používáme nejčastěji na úrovni běžného života. Je mnohem praktičtější říci, že příčinou mého strachu je divoký pes přede mnou, než říci, že mám excitovanou amygdalu a vyplavený adrenalin kvůli dopadu určité skupiny fotonů na mou sítnici atd. V některých případech se ale mohou projevit problémy s vágností takto užívaného pojmu příčinnosti, někdy dokonce se složitými důsledky. Lze to ilustrovat na klasické právní úloze, ve verzi dle (Dennett 2004):

Všichni na pouštní základně nesnáší Freda a chtějí ho zabít. Než je Fred vyslán na výpravu přes poušť, Tom otráví jeho lahev vody. O tom ale neví Dick, který poté vyleje vodu z lahve a naplní ji pískem. Následně Harry proděraví lahev, aby „voda“ po cestě vytekla. Fred skutečně zemře. Kdo zapříčinil jeho smrt?

Žádná správná odpověď na této úrovni popisu zřejmě neexistuje. Analýza kauzálních řetězců na nižší úrovni ale není možná, protože na té nejsou definovány pojmy jako „osoba“. Důvodem nezdaru může být to, že intuitivně při kauzální analýze hledáme *nutné* příčiny. Tento případ podle mě nedokážeme rozsoudit právě proto, že akce Toma, Harryho a Dicka jsou *dostatečné* příčiny smrti Freda, ale žádná z nich není příčinou nutnou. Pokud podmínka nutnosti selže, používáme další heuristiky, jako je například temporální priorita: Můžeme mít *pocit*, že větší odpovědnost za Fredovu smrt má Tom, ačkoli by Fred nutně zemřel i bez Tomovy akce.

Pokud pak máme dostatečně složité jevy, můžeme se dostat ke kontra-intuitivním závěrům, že dané jevy nemají *žádnou* příčinu. Jako příklad Dennett uvádí 1. světovou válku. Zavraždění Ferdinanda bylo jistě postačující příčinou, ale ne nutnou, protože válka mohla¹⁰ vypuknout i bez jeho zavraždění. Co lze považovat za příčinu války pak podle mě záleží na úhlu pohledu a schopnosti kontrafaktuální představivosti, ale žádná „správná“ odpověď opět neexistuje. Komplex příčin nejde na této úrovni smysluplně popsat; za tím účelem bychom opět museli na elementární úroveň, kde ale neexistují žádné vraždy a války a popis příčin ve fyzikálních termínech by nebyl příliš užitečný. Dennett proto navrhuje provokativní tezi, že existují události bez příčiny, i v případě, že platí determinismus. Determinismus je totiž „doktrína o dostatečnosti“ (Dennett 2004, s. 83), kdežto „zapříčinění obecně předpokládá nutnost“ (Dennett 2004, s. 84). Na každodenní morálně relevantní úrovni popisu (kde je něco jako morálka a odpovědnost definováno) tak lze i v deterministickém vesmíru říci, že jedinou (nutnou) příčinou mé akce bylo v dané situaci mé rozhodnutí.

2.3 Indeterminismus

Indeterminismus neznámá, že by nic nebylo determinováno, ale pouze, že není determinováno zcela vše; že existují indeterministické děje a že v daný moment existuje více fyzicky možných budoucností. Inkompatibilisté, konkrétně metafyzičtí libertariáni, se snaží nalézt v nějaké formě indeterminismu záchranu pro možnost existence svobodné vůle a morální odpovědnosti. Již starověcí epikurejci se pokoušeli nalézt mezeru v řádu příčin a následků, která by umožnila vznik „novosti“. Podle Lucretia je to zajištěno náhodným odchýlením pohybu atomů (Cashmore 2010).

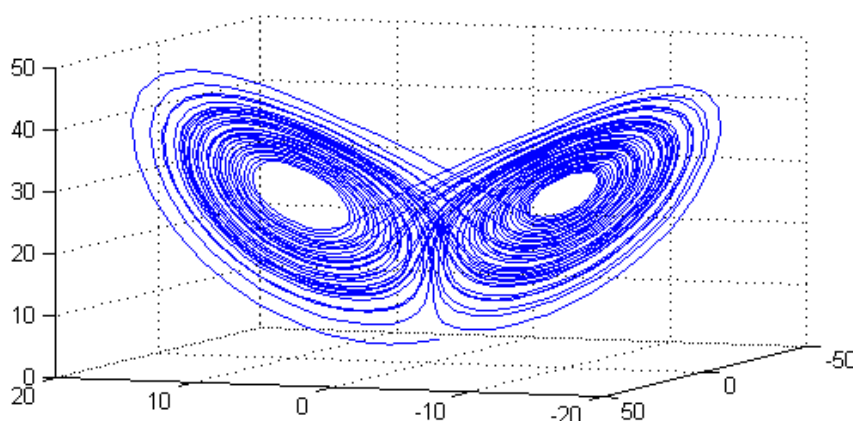
Indeterminismem se obvykle myslí, že determinismus je doplněn náhodnými ději, které způsobují nepredikovatelné, stochastické chování. Je opět důležité rozlišit indeterminismus ontologický a epistemologický. Ontologický indeterminismus by vyžadoval existenci ontologické náhody. Kdežto epistemologický indeterminismus znamená pouze to, že nějaké jevy jsou tak komplexní, že je nelze predikovat. Proto se jeví jako náhodné a lze je popisovat maximálně probabilisticky. Chci dále ukázat, že zatím nemáme jasný důkaz existence ontologického indeterminismu.

¹⁰ Viz epistemická možnost v sekci 2.4 Možnost

Uvažme klasický příklad hodu kostkou. Pokud bychom měli dostatečně výkonný počítač napojený na kameru, dokázal by predikovat hozené číslo již ve chvíli, kdy kostka opustí naši ruku. Výsledek hodu lze označit za pseudo-náhodný, stejně jako například náhodně generovaná čísla v počítačích.

Jacques Monod rozlišuje praktickou náhodu a faktickou náhodu. Za příklad praktické náhody bere právě hod kostkou, protože neumíme spočítat výsledek hodu. Kdežto faktická náhoda je „zkřížení dvou zcela na sobě nezávislých příčinných závislostí.“ (Monod 2008, s. 115) Jako když se dva staří přátelé náhodou potkají po letech v cizí zemi. Jak ale víme, že jde skutečně o „nezávislé příčinné závislosti“? Jde o stejný problém jako s výše zmíněným zapříčiněním. Ani u hodu kostky ani u setkání přátel nejsme schopni uvést žádnou *nutnou* příčinu na naší makroúrovni. Kompletní popis světa – řekněme před sto lety – však jistě mohl být *dostatečný* pro určení výsledku hodu i setkání přátel. Laplaceův démon by oboje „vypočítat“ dokázal. Ani Monodova faktická náhoda (pokud je vůbec kvalitativně odlišná od praktické) tak není v rozporu s determinismem.

Dalším často uváděným argumentem pro indeterminismus je chaos. Teorie chaosu je rozvíjející se matematická disciplína, která studuje chování velmi komplexních nelineárních systémů s velkou citlivostí na počáteční podmínky. Chování těchto systémů je z praktických důvodů nepredikovatelné. Lze je popsat pouze pravděpodobnostně, např. pomocí tzv. atraktorů systému:



Obrázek 2: Lorenzův „podivný atraktor“ s parametry $\rho = 28$, $\sigma = 10$, $\beta = 8/3$

Linie znázorňuje vývoj Lorenzova systému v čase. Linie vyplňuje určitý podprostor ve fázovém prostoru, ale nikdy se sama neprotíná.

Přestože vývoj chaotických systémů nelze predikovat, jsou tyto systémy brány za zcela deterministické. Mluví se přímo o deterministickém chaosu. Indeterminismus je zde pouze epistemologický. Dokonce, jejich pravděpodobnostní popis je umožněn jen díky podkladovému determinismu (Stewart 2009, s. 379).

Vůbec nejčastějším útočištěm indeterminismu je pole kvantové mechaniky.¹¹ Obvykle se má za to, že určité děje mikro-úrovně, jako například radioaktivní rozpad, jsou ontologicky náhodné a popsitelné pouze pravděpodobnostně. Celá věc je však složitější. Je potřeba rozlišit rovnice kvantové mechaniky od jejich interpretací. Schrödingerova rovnice popisuje vývoj vlnové funkce systému zcela deterministicky. Indeterminismus vstupuje do hry při hypotetickém kolapsu této vlnové funkce, kdy systém nepredikovatelně přejde z vlnové superpozice do jediného vlastního stavu. Kolaps je však součástí jen některých interpretací kvantové mechaniky, například nejznámější kodaňské interpretace. Indeterminismus je obsažen také v Heisenbergově relaci neurčitosti, dle které nelze určit některé páry fyzikálních vlastností současně. Neurčitost je to ale pouze epistemologická. Některé interpretace kvantové mechaniky, například Bohmova, zachovávají determinismus zavedením tzv. skrytých proměnných na pozadí indeterministických kvantových jevů. Hypotézu skrytých proměnných zastával také Einstein, který pronesl slavný výrok o tom, že Bůh nehraje kostky. Již to, že například zmíněný radioaktivní rozpad má stále Poissonovo rozdělení a nikoli při některých pozorováních jiné, naznačuje existenci skrytých proměnných. Tzv. Bellovy nerovnosti však zdá se vylučují možnost interpretace kvantové mechaniky, která by zároveň obsahovala skryté proměnné a byla tzv. lokální. Existují i lokální teorie obcházející Bellovy nerovnosti pomocí právě teorie chaosu (Stewart 2009). Problematika je stále otevřená.

Pomocí kvantové mechaniky se snaží vysvětlit problém vědomí a svobodné vůle několik teorií. Dle Ecclese (1989) by kvantová mechanika dokonce mohla zajišťovat dualistickou interakci mezi tělem a duší. Fyzik R. Penrose a anesteziolog S. Hameroff se snaží nalézt více naturalistické vysvětlení. Dle nich může uvnitř neuronů (konkrétně mikrotubulů) být zachována kvantová koherence a může tam probíhat dokonce nenáhodný „sebe-kolaps“ (tzv. objektivní redukce) vlnové funkce (Hameroff a Penrose 1996). Téma je velmi kontroverzní, debata stále probíhá, ale tyto teorie obecně nejsou příliš uznávány, např. právě kvůli příliš rychlému procesu tzv. kvantové dekoherence v „teplém a vlhkém“ mozku (Tegmark 2000).

¹¹ Výklad kvantové mechaniky v tomto odstavci je podán především dle (Hofer 2010) a (Stewart 2009). Místy se ale prezentují partikulární názory a proto jsou uvedeny příslušné reference.

At' už jsou ale kvantové jevy na mikro-úrovni náhodné epistemologicky či ontologicky, zdá se, že na makro-úrovni determinismus přesto platí. Šlo by ho označit jako „adekvátní determinismus“ (Doyle 2012). Jde zhruba o tezi, že jakékoli kvantové vlastnosti jsou na makro-úrovni zanedbatelné a vyruší se kvůli zmíněné dekoherenci (Pavlík 2004), (Stewart 2009). Nejdůležitější v debatě o indeterminismu a svobodné vůli je pak to, že není příliš jasné, jak by hypotetická náhoda vůbec mohla přispět ke svobodnosti vůle. Libertariáni se snaží vypracovat takové teorie (viz 2.5 Odpovědnost). Mně osobně však přijde pro svobodnou vůli daleko důležitější determinismus než indeterminismus. Akce, která má příčiny v mé kauzálně determinované zkušenosti, mých přáních a podobně, je z mého pohledu mnohem svobodnější a především *moje*, než akce, která má příčinu v náhodném (kvantovém) ději.

2.4 Možnost

Modálním protipólem výše diskutované nutnosti je možnost. Tento pojem používáme ve svých životech zcela běžně. Například v kontrafaktuálním uvažování typu „kdybych se býval učil, zkoušku bych býval udělal“. Libertariáni existenci alternativních možností (*I could have done otherwise*) považují za nutnou podmínku svobody a odpovědnosti, a proto předpokládají, že svět není deterministický. Jak by ale koncept možnosti byl slučitelný s případným determinismem, pro jehož minimálně „adekvátní“ formu výše argumentuji? Co ospravedlňuje naše každodenní používání tohoto pojmu? A především, jaká možnost je potřebná pro morálně relevantní svobodnou vůli?

Problematika modalit je vysoce komplikovaná jak po stránce epistemologie, tak ontologie. Dle (Vaidya 2011) lze rozlišit např. tyto druhy možnosti:

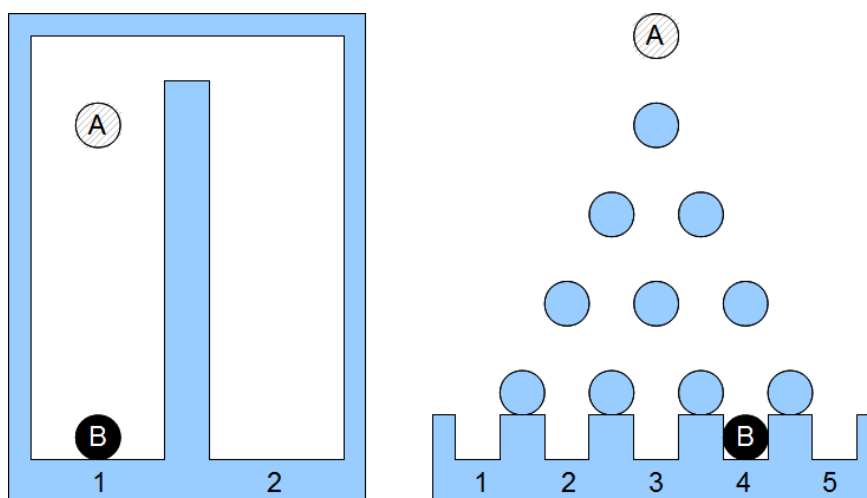
- epistemická = znalosti nevylučují X
„Je možné, že kolem Pluta obíhá čajová konvička.“
- logická = X je konzistentní v systému S
„Itálie vyhraje světový pohár“ je možný výrok v rámci výrokové logiky.“
- konceptuální = X není vyloučeno množinou všech konceptuálních pravd
„Je možné, že Země je placatá.“
- metafyzická = X je pravdivé v metafyzicky možném světě
„Je možné, že se nějaké fyzikální částice pohybují rychleji než světlo.“

- fyzická (fyzikální) = X je logicky konzistentní s fyzikálními zákony
„Je možné, aby vlak (za určitých podmínek) jel rychlostí 150 km/h.“

Pro debaty o svobodné vůli lze ponechat stranou možnost logickou a konceptuální. Možnost, kterou požadují libertariáni, je možnost ontologická. Ačkoli je to předmětem debat, zde uvedená metafyzická možnost je podle mě spíše možností konceptuální a ne ontologickou. Proto ponechám stranou i možnost metafyzickou a z naturalistického hlediska budu identifikovat ontologickou možnost s možností fyzikální. Aby částice cestovala rychleji než světlo, musí to umožňovat fyzikální zákony, ať už v současnosti známé nebo neznámé. Pokud přijmeme kauzální determinismus, množina ontologických možností a ontologicky možných světů se značně zúží: existuje jen jeden ontologicky možný svět: ten současný (a jeho vývoj v čase). Ačkoli je to silně proti-intuitivní, nemohl jsem se více učit na zkoušku, protože jsem tehdy nechtěl (a nemohl chtít) kvůli daným fyzikálním zákonům a danému stavu věcí v té době. Bylo to stejně *fyzikálně* nemožné jako (v dané situaci) přeměnit olovo ve zlato úderem kladiva, avšak více představitelné (viz dále). Dle některých fyzikálních teorií existuje skutečně více fyzikálně možných (možných, protože *skutečných*) světů, ale ty se od sebe odlišují fyzikálními zákony. I malá změna zákonů by zamezila vzniku života v takovém světě (Hawking a Mlodinow 2011), natož aby tam existovala ta samá osoba učící se na zkoušku. Jediné ontologické možnosti tak mohou_(epistemologicky) platit na kvantové úrovni, což je – jak bylo řečeno výše – na makroúrovni většinou zanedbatelné.

Domnívám se, že běžné uvažování o možnostech má povahu nikoli ontologických, ale epistemologických možností. Něco považujeme za možné, když nemáme informace, že by tomu něco bránilo. Jenomže nejsme Laplaceovi démoni, takže nám mnoho informací chybí. (Především, nemáme přímý přístup ke všem informacím v nás samotných, protože vědomé procesy jsou pouze špičkou ledovce v moři nevědomých procesů.) Když mluvíme o tom, že něco je (nebo mohlo být) možné, myslíme tím, že to je myslitelné, v čemž se implicitně spoléháme na svou schopnost představit si takovou událost (čemuž pomáhá implicitní pozadí vlastních znalostí a zkušeností). Jsme v tom poměrně dobří v jednoduchých záležitostech běžného života. V nich posuzujeme jako možné ty situace, ve kterých naše představivost nenarazí na žádný *fyzikální* rozpor. Představivost ale těžko vezme v úvahu pozice všech částic v blízkém vesmíru (především v mém mozku) a jejich kauzální historii, aby našla v představě mé osoby (v té době) usilovně se učící na zkoušku fyzikální rozpor. Naproti tomu v přeměně olova na zlato úderem kladiva (v konkrétní situaci) najdeme rozpor poměrně snadno, ale i to je závislé na našich znalostech. Proto se nám zdá, že některé události jsou více možné než jiné.

Pokusil jsem se znázornit rozdíl mezi ontologickou a epistemologickou možností na následujícím obrázku:



Obrázek 3: Epistemologická možnost

Předpokládejme uzavřené systémy bez vnějších intervencí, ve kterých působí homogenní gravitace, materiály nejsou pružné atp. Na obrázku vlevo zřejmě uznáme, že není možné, aby kulička z pozice A bývala spadla¹² do pozice 2. Let kuličky proti gravitaci by byl *příliš velkým* porušením fyzikálních zákonů. Přesněji, příliš proti-intuitivní. Přesto otevřená přepážka provokuje naši mysl k uvažování, že to třeba *nějak* možné je, což demonstruje naše posuzování možnosti spíše podle představitelnosti, než podle fyzikálních zákonů. Na obrázku vpravo však intuice povede už mnoho lidí k závěru, že kulička bývala mohla spadnout do jakékoli z pozic, protože *neznají* fyzikální detaily tohoto systému. Také, překulení se na druhou stranu, než je dáno detailním tvarem objektů, není intuitivně *tak velkým* proviněním proti fyzice. Ontologickým „proviněním proti realitě“ jsou oba případy ve skutečnosti ve stejné míře. Možnost je zde v obou vyobrazených případech pouze epistemologická.

Epistemologická možnost je vysoce užitečný, až nepostradatelný koncept našich myslí a našich životů. Umožňuje nám uvažovat nad různými alternativními scénáři ať už v budoucnosti nebo v minulosti, které můžeme hodnotit na *škále* možnosti, jakožto míře jejich myslitelnosti a představitelnosti podle našich omezených intuic, znalostí a zkušeností. Není to ale typ možnosti požadovaný pro svobodnou vůli metafyzickými libertariány. V *přesně stejných* okolnostech jsem nemohl konat jinak, než jsem konal (*could not have done otherwise*), stejně jako

¹² bývala spadla = v tom konkrétním čase, tedy za stejných podmínek

kulička nemohla v daných okolnostech spadnout do jiné pozice (*could not have fallen elsewhere*).

Dle kompatibilisty Dennetta je takové uvažování o přesně stejných okolnostech triviální (je jasné, že jsem nemohl konat jinak) a nezajímavé (Dennett 2004). Zajímá ho, zda jsem mohl konat jinak v množině *dostatečně blízkých* světů – takových, které se odlišují od současného jen v některých *myslitelných* detailech. Je to sice vágní vymezení, které se opírá o epistemologickou možnost a představitelnost/myslitelnost, ale odpovídá díky tomu vágnosti běžně užívaného pojmu možnost. Dennettův pojem možnosti tak znamená spíše obecnou možnost, schopnosti, *kompetence*, určitého systému. Tento pojem je pro naše uvažování o svobodě v praktické rovině užitečnější než ontologická možnost. Umožňuje nám uvažovat nad obecnými možnostmi. Například, nemohl jsem porazit Garryho Kasparova v šachu v žádném z dostatečně blízkých představitelných světů, protože k tomu nemám kompetenci. Ale býval jsem mohl pomoci babičce přes přechod, protože jsem nebyl spoután řetězy a kompetenci k tomu jistě mám. Stačila „jen“ jiná konfigurace stavu mého mozku. Protože to zřejmě nebylo možné ontologicky, za svůj čin (resp. absenci činu) nejsem zodpovědný v jakémsi ultimátním, metafyzickém smyslu. Míra mé kompetence se ale stává jednou z podmínek *praktické* morální odpovědnosti.

2.5 Odpovědnost

Protože je svobodná vůle relativně vágní pojem, často se definuje přes pojem odpovědnosti (McKenna 2009). Svobodná vůle je to, co lidem umožňuje konat tak, že nesou morální odpovědnost za své činy. Pojem odpovědnosti ale není o mnoho jednoznačnější. Je to otázka nesmírně široké disciplíny filosofické etiky. Pro potřeby této debaty můžeme zjednodušeně rozlišit opět dva druhy odpovědnosti: ultimátní a praktickou.

Robert Kane považuje za nutnou podmínku svobodné vůle tzv. ultimátní odpovědnost (Kane 1998). Abych byl morálně zodpovědný, musím být dle Kanea ultimátním zdrojem svých rozhodnutí. Regres příčin musí být ukončen u *mě*. To znamená, že souhrn předchozích na mě působících vlivů *nepostačoval* k zapříčinění *mého* rozhodnutí. Toho má být dosaženo vložením (kvantové) náhodnosti dovnitř tzv. praktického rozumu (*faculty of practical reasoning*). Kvůli náhodnému vlivu pak nebudou předchozí vlivy dostačující podmínky mého rozhodnutí. Kane uznává, že v mnoha případech (ne-li většině) se rozhodujeme deterministicky, ale přesto

jsme odpovědní, protože akce vycházejí z naší osobnosti či povahy (*character*). Ultimátní odpovědnost je odpovědnost právě za vlastní osobnost. Toho je dosaženo při tzv. sebe-utvářejících akcích (*self-forming actions*) a pouze u nich je důležitý indeterminismus. Jedná se o těžká dilemata, kdy máme dobré důvody pro více směrů konání. Náhoda způsobí, že zvolíme právě jeden ze směrů, což ovlivní naši osobnost pro další rozhodování, ať už deterministická či nikoli. Na námitku, že za náhodné vlivy nejsme odpovědní, odpovídá Kaneova teorie tzv. pluralitní racionality. Jsme odpovědní za výsledný směr konání, protože jsme pro něj měli důvod, protože jsme ho chtěli. Kdyby náhoda rozhodla pro druhý směr, budeme za něj odpovědní též, protože i ten jsme chtěli.

Nelze se ubránit dojmu, že se Kaneovi nepodařilo podat zcela uspokojivé vysvětlení odpovědnosti, která má být „ultimátní“. Dennett ukazuje, že Kaneova teorie pro své závěry vlastně „nepotřebuje indeterminismus, který inspiroval její vznik.“ (Dennett 2004, s. 126) Závěry by pak šlo pojímat jako podporu kompatibilismu a ne libertariánství. Dennett dále poskytuje rozsáhlou kritiku Kaneovy teorie. Argumenty jsou svým rozsahem bohužel nad rámec této práce. Stručně řečeno, zahrnují mimo jiné kritiku Kaneova esencialismu ohledně sebe-utvářejících akcí a jáství jako takového. Jinou kritiku ultimátní odpovědnosti podává také Galen Strawson (1994).

Je možný jiný druh odpovědnosti? Zjevně ano, protože ho běžně aplikujeme, když chválíme a haníme lidi za jejich akce, ačkoli se v nich ve skutečnosti „ultimátní odpovědnost“ v Kaneovo pojetí neskrývá. Může se zdát, že každodenní „praktická“ odpovědnost není skutečnou odpovědností, protože ji nelze nijak metafyzicky založit. Je to ale dle mého pohledu podobný případ jako se „skutečnou magií“.

Zásadní přehodnocení teorie morální odpovědnosti přinesl P.F. Strawson (1975). Podle něj je chybou hledat založení morální odpovědnosti v nějakých racionálních či metafyzických kritériích. Morální hodnocení vychází z tzv. reaktivních postojů a jejich role ve společnosti. Například někým poškozená osoba má obvykle osobní reaktivní postoj ukřivdění či zlosti. Okolí má reaktivní postoj rozhořčení a samotný škůdce může v reflexi získat postoj provinění (McKenna 2009). Bez ohledu na možné závěry filosofů či zjištění neurověd, těchto postojů se nemůžeme lehce zbavit (Pizarro 2011). Jsou hlubokou součástí naší psychologie a našich sociálních životů. Narůstá empirická evidence, že se moralita (a stejně tak i koncept odpověd-

nosti) spontánně vyvinula v průběhu genově-kulturní evoluce a stále se vyvíjí.¹³ Morální odpovědnost je čistě praktické, spontánně (nezáměrně) vyvinuté kritérium (viz sekce 5.2). Co však možná bude potřeba ve světle nových poznatků přehodnotit, jsou některé dílčí principy trestní odpovědnosti (viz sekce 5.3).

2.6 Jáství

Jak již bylo uvedeno výše, za zcela klíčový pojem v debatě o svobodné vůli považují koncept Já, jáství, osoby. Jáství běžně připisujeme atributy jako „zdroj či příčina akcí“, „svoboda“, „morální odpovědnost“, „vědomí“. Co nebo koho tímto Já ale myslíme?

Dle lidově psychologické intuice je Já často ztotožňováno s vědomím i celou myslí, a toto souhrnné psychično je chápáno dualisticky¹⁴, jako bezrozměrná substance (*res cogitans*) odlišná od mozku, těla a jiné rozprostraněné hmoty (*res extensa*). Tato představa však není z drtivé většiny explicitně uznávána ani filosofy (z „apriorních“, logických důvodů) ani vědci (z empirických důvodů). Vážným problémem je (ne)možnost interakce dvou odlišných substancí. Dále, pokud budeme uvažovat o Já substancialisticky jako bodově velkém, jak si je Já vědomo věci? Lze použít metaforu Já jako vnitřního agenta dívajícího se na plátno tzv. karteziánského divadla, kde se promítají děje okolního světa. Pak bychom ale potřebovali další plátno a další Já uvnitř prvního Já atd. Pokud je Já substancí, co jsou jeho substanciální vlastnosti? David Hume argumentoval, že žádné takové nejsou. Já je podle něj soubor (*bundle*) neustále se proměňujících složek bez žádného podkladového substanciálního jádra. Množí se empirická evidence, že jednotné Já je spíše¹⁵ složitý psychologický a sociální konstrukt (Hood 2012).

Přes tento filosofický a empirický pokrok a oficiální odmítání dualismu mnoho autorů stále implicitně i explicitně uvažuje o Já jako o substancí. Chtějí, aby Já bylo *causa sui*, ultimátně odpovědné, sebe-utvářející a dokonale svobodné. „Jinými slovy, chtějí zázrak.“ (Dennett

¹³ Literatury k tomuto tématu je velké množství v oblastech jako evoluční psychologie, sociobiologie, behaviorální ekonomie a podobně. Základní přehled lze nalézt např. v (Dennett 2004).

¹⁴ Tuto teorii se v západní filosofické tradici poprvé pokusil rozpracovat René Descartes, proto se jí říká karteziánský dualismus.

¹⁵ Nechci použít slovo „pouze“, protože tím Já není o nic méně fascinující. Ba co víc, takto je *skutečné*.

2008, s. 250) Pokud budeme brát Já substanciálně a zároveň naturalisticky (což je rozpor, ale často zastáváný), tak obzvlášť ve světle současných poznatků (3.1 Kauzální potence vědomí) musíme dojít k závěru, že *Já nedělá věci, ale věci se mu dějí*. Tedy že svobodná vůle neexistuje. To je však karteziánská chyba pojmání Já jako bodové substance, s ostatními atributy „přivěšenými“ na tento bod. „Pokud se učiníte dostatečně malými, můžete brát za externí prakticky cokoli.“ (Dennett 1984, s. 143) Klíčem je naopak pojímat jáství dostatečně široce, jako *konstituované* mnoha atributy a relacemi. „Když vyhostíme onoho vnitřního agenta a rozložíme jeho úlohy nejen přes celý mozek, ale i tělo a ‚okolní‘ kulturní zásobárnu ... nemusíme vyhostit svobodnou vůli!“ (Dennett 2008, s. 254-5)

Na fyzické úrovni nenajdeme žádné místo, v mozku ani nikde jinde, kde by sídlilo jáství. Je důležité uvědomit si, že jde o široký koncept vysoké, každodenní úrovně popisu. Díky tomu je i v deterministickém světě možné připisovat jáství vysoko-úrovňové atributy, jako zdroj důvodů (5.2), příčina svých akcí (2.2.3), morální odpovědnost (2.5) a tedy i svoboda.

Navrhované pojetí jáství je částečně kontroverzní v tom, že za konstitutivní součást Já považuje celý mozek, včetně nejen vědomých, ale i nevědomých procesů, což nám může připadat kontra-intuitivní (viz 3.3). O vztahu vědomí, nevědomí, jáství a agence¹⁶ pojednává následující kapitola.

¹⁶ agence = vlastnost agenta, bytosti, která koná

3 Filosofické problémy ve světle vědeckých poznatků

Současné kognitivní vědy dospěly do stádia, kdy dokáží svými výzkumy lidské mysli zásadním způsobem ovlivňovat filosofické debaty, včetně témat, která byla dříve považována za metafyzická – a tedy neempirická – jako je právě svobodná vůle či vědomí.

3.1 Kauzální potence vědomí

Jak bylo předestřeno v předchozí kapitole, introspektivně vnímáme, že je to vědomá část našeho Já, která je svobodným původcem našich akcí. Odborně řečeno, cítíme, že vědomí má kauzální potenci, že je schopno kauzálně působit na svět. Tato teze je sice intuitivní, ale značně problematická.

Opět, jako v případě jáství, se dostáváme k problému dualismu. Descartes bral celou mysl jako transparentní, vědomí tak bylo její celou podstatou. Vědomí je pak jako imateriální substance vyloučena z podléhání kauzálnímu determinismu. Zároveň má ale schopnost kauzálně působit na materiální substance. Jak by mohla taková interakce probíhat? Byly navrženy různé teorie, od idealistických, které ruší materiální svět a veškerá kauzace pak probíhá jen v mysli, po epifenomenální, dle kterých je vědomí jen „hvizdem doprovázejícím činnost lokomotivy“ bez jakékoli kauzální účinnosti.

Problém se možná skrývá v samotném pojmu vědomí, který je silně vágní. Například dle zastánců eliminativního materialismu nejde o vědomí vůbec koherentně mluvit, a kvůli pokroku neurověd potká tento termín podobný osud jako „duši“ či „flogiston“ (Churchland 1988). Jevy jako vědomí a jáství jsou však tak komplexní povahy, že i kdyby se podařilo je skutečně převést do nízko-úrovňového slovníku fyziky a neurověd, bude velmi nepraktické, ne-li nemožné, o nich na tak nízké úrovni mluvit či přemýšlet. Proto je často zastáván metodologický dualismus. Ten spočívá v ontologickém monismu, ale pro popis jevů z praktických důvodů používá dualistický slovník.

Z hlediska přírodních věd jsou veškerá rozhodnutí produktem mozku, který je deterministickým fyzikálním systémem. Ve shodě s přírodovědným pohledem je i narůstající množství experimentů, zaměřujících se na naše vědomí a rozhodování. Některé významné z nich jsou

dále podrobněji popsány. Nejprve jsou uvedeny experimenty problematizující běžnou představu, že vědomí je potřebné pro konání a že uvědomění předchází akci.

3.1.1 Rychlost vědomí

Jak rychle probíhá uvědomování si událostí ve světě? Může nám připadat, že vědomí má přímý, prakticky okamžitý přístup ke světu. A také, že nejprve si něco uvědomíme, a teprve poté reagujeme svojí akcí. To však z hlediska složitosti nervových procesů, které k tak komplexnímu mentálnímu procesu (vědomí něčeho) vedou, není vždy možné. Rychlost šíření nervového vzruchu je řádově pouze v desítkách metrů za sekundu a množství kognitivního zpracování potřebného pro vznik vědomého prožitku je obrovské. Kvůli zpoždění vědomého prožitku za objektivními ději ve světě se dá říci, že vlastně žijeme v minulosti. Dle současných výzkumů jde v případě jednoduchých vjemů (bliknutí světla) o cca 80 milisekund, přičemž vědomí je zpětně přetvářeno tak, že tuto latenci zatím ne příliš známými způsoby kompenzuje (Eagleman a Sejnowski 2000).

Některé události se dějí tak rychle, že k vytvoření uvědomění nestihne dojít během naší akce, a dochází k němu až později nebo vůbec. Příkladem je reakce na tenisový míček, který může letět rychlostí 200 km/h a přesto je úspěšně odpálen. K vědomému zpracování v takovém případě nemůže stihnout dojít a hráč si svůj odpal skutečně uvědomí spíše až v době, kdy od něj míček letí pryč.

Castiello et al. (1991) provedli experiment, ve kterém došli k podobnému závěru. Účastníci měli za úkol dotknout se toho ze tří cílů (kolíků), který se zrovna rozsvítí. Měli zároveň ohlásit (pronést „Tá!“), když si uvědomí rozsvícení cíle. Ale v některých případech, jakmile účastník započal pohyb rukou k cíli, byl náhle rozsvícen jiný cíl místo původního. Účastníci měli v takovém případě změnit pohyb ruky k nově rozsvícenému cíli a opět ohlásit uvědomění si změny cíle. Reakční čas zahájení pohybu ruky k primárnímu cíli (ať už se potom změnil nebo ne) byl kolem 300 ms. Vokální reakční čas pro primární cíl byl 375 ms. Překvapivě, reakční čas změny pohybu k novému cíli, poté co byl zahájen pohyb k primárnímu cíli, byl jen kolem 100 ms, ale vokální reakční čas na tuto změnu cíle byl 420 ms. Z rozdílu v časech lze vyvodit, že změna pohybu byla zahájena dříve, než došlo k uvědomění si změny cíle.

V některých případech ale k uvědomění nedojde vůbec, ačkoli provedeme odpovídající akci. Účastníci v jiné studii (Taylor a McCloskey 1990) reagovali správně na primární stimul, který byl po 50 ms překryt jiným stimulem, ale byli si vědomi pouze druhého.

Z tohoto lze vyvodit, že vědomí přichází až po delším neurálním zpracování. Domnívám se, že vědomí je konstituováno různými druhy komplexních nervových procesů podobně, jako je teplo konstituováno pohybem molekul (ačkoli tato analogie nemusí být úplně přesná). Ačkoli jsou kladeny určité podmínky na povahu těchto procesů, mohou být velmi rozmanité, a tím pádem jsou různé i druhy vědomých stavů (Edelman 2005). Nedá se přímo říci, že vědomí (jako jakási samostatná věc, substance) má kauzální potenci, „dělá věci“. Spíše, tyto nervové procesy odpovídající vědomí jsou kauzálními příčinami našich akcí. Než bude tento problém kauzální účinnosti vědomí více probrán (v sekci 3.1.4), je potřeba se zmínit o dvou významných experimentech.

3.1.2 Libetův experiment

Americký vědec Benjamin Libet provedl v sedmdesátých a osmdesátých letech dvacátého století sérii experimentů zabývajících se časovým aspektem vědomí. Experiment z roku 1983 lze považovat za nejvýznamnější z nich. Stal se zcela zlomovým pro výzkum svobodné vůle; dalo by se říci, že soustavné vědecké zkoumání této oblasti v podstatě odstartoval.

Účastníci sledovali bod vykreslovaný osciloskopem pohybující se kruhově po číselníku (tzv. Libetovy hodiny). Bod se pohyboval poměrně rychle; jeden oběh bodu trval 2,56 vteřiny. Číselník měl 12 hlavních značek a mezi nimi ještě 12 vedlejších. Ciferník tak byl rozdělen na 24 dílků; na jeden dílek připadalo 107 ms. Zároveň bylo prováděno měření elektrické aktivity motorických a premotorických oblastí mozkové kůry pomocí EEG.

Účastníci měli za cíl nechat oběhnout bod alespoň jednou dokola a poté „kdykoli budou cítit, že se jim chce“¹⁷ (Libet et al. 1983, s. 625) rychle zavřít svoji dlaň. Tento čas byl zaznamenán pomocí EMG¹⁸ a bral se jako čas „0“. Účastníci byli konkrétně instruováni „nechat nutkání k akci kdykoli se samo objevit bez žádného plánování dopředu nebo soustředění se na to, kdy

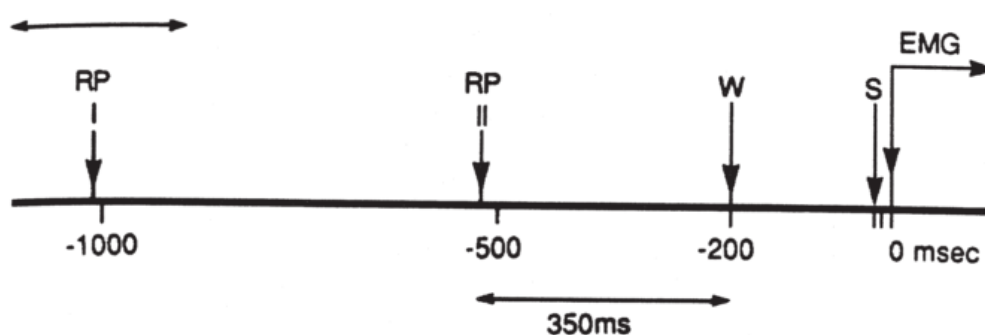
¹⁷ „at any time thereafter when he felt like doing so“

¹⁸ Elektromyograf, zařízení snímající aktivitu svalu

akci vykonat¹⁹ (Libet et al. 1983, s. 625), aby byla dle autorů zajištěna spontaneita a svobodnost akce. Už zde bychom mohli namítat, že spontaneita nemusí znamenat svobodnost. Pro tu bychom mohli naopak vyžadovat více předchozí deliberace. Zde ale šlo především o to, aby byla akce vyvolána endogenně a nikoli exogenně.

Po účastnících bylo vyžadováno, aby si po každém ze série pokusů vybavili pozici na hodínách, kdy u nich poprvé došlo k „uvědomění si vlastního ‚chtění‘ vykonat danou vlastní akci“,²⁰ tzv. čas „W“. (Libet et al. 1983, s. 627). V instrukcích byl tento prožitek (*experience*) popsán zároveň slovy „nutkání“ (*urge*), „intence“ (*intention*), či „rozhodnutí“ (*decision*), ačkoli účastníci se prý sami nejvíce ztotožňovali se slovy „chtění“ (*wanting*) nebo „nutkání“ (*urge*). Je vidět, že varieta termínů je značně široká; především slovo „nutkání“ bychom nemuseli zrovna asociovat se svobodností akce.

Nejdůležitější výsledky studie ilustruje následující obrázek:



Obrázek 4: Posloupnost událostí volní akce v Libetově experimentu. Zdroj: (Libet 1989)

RP I je začátek přípravného motorického potenciálu, v případech, kdy účastník uvedl nějaké plánování akce; RP II když k plánování nedošlo. W je čas pociťovaného chtění vykonat akci. EMG je čas zaregistrování akce elektromyografem. S je pociťovaný čas aplikace externího podnětu na kůži v kontrolní podmínce experimentu, relativně ke skutečnému času aplikace (bod 0).

Čas W byl hlášen průměrně 200 ms před časem nula. Co však bylo považováno za překvapivé, bylo zjištění, že přípravný motorický potenciál²¹ byl pomocí EEG zaznamenán průměrně

¹⁹ „to let the urge to act appear on its own at any time without any preplanning or concentration on when to act“

²⁰ „conscious awareness of 'wanting' to perform a given self-initiated movement“

²¹ Tzv. Bereitschaftspotential (BP) nebo Readiness Potential (RP) je elektrofyziologický korelát zahájení volní motorické akce pravděpodobně v suplementární motorické kůře, ačkoli to je stále předmětem diskuse (Bareš 2011).

už 350 ms před časem W. To znamená, že vědomý prožitek chtění vykonat akci nemohl být kauzální příčinou samotné akce. Ta již byla zahájena předtím na nevědomé úrovni a vědomý prožitek chtění či nutkání k akci byl vytvořen až poté.

Libet sám ale odmítal interpretaci, že by svobodná vůle byla vyloučena, že by vědomí nemohlo řídit naše akce. Navrhl, že tyto nevědomě zahájené nervové procesy lze vědomě vetovat. Taková interpretace mi přijde velmi zvláštní. Jednak lze předpokládat, že i ostatní naše akce mají příčiny v s-vědomím-nekorelujících (což je asi přesnější vyjádření než „nevědomých“) nervových procesech. Mohlo by to znamenat, že naše tělo je neustále nevědomě nuceno k různým akcím, které musíme neustále selektovat vetováním většiny z nich. To je skutečně zvláštní model svobody. A jednak, i veto je akce. Pak bychom se mohli dostat do kruhu, kde by akce vetovaly akce vetující akce. Lze očekávat, že i vetující akce má nevědomé příčiny. To však Libet odmítá, protože veto je dle něj jiný druh akce, totiž *řídící* akce. Žádná evidence prý nevyklučuje možnost, že by se „řídící proces mohl objevit bez vyvolání předchozím nevědomým procesem“ (Libet 1999, s. 53). Domnívám se, že se zde Libet nedokáže oprostít od uvažování o vědomí jako o jakési karteziánské samostatné „věci“, substanci, možná kvůli dle jeho názoru špatným důsledkům plynoucím z opačného pohledu. Podle Libeta by nebylo možné považovat nikoho odpovědným za nevědomě zahájené akce, což je podle mne nepřesný závěr založený na chybném substančním pojetí vědomí a jáství.

Nejslavnější Libetův experiment vzbudil pochopitelně mnoho emotivních reakcí a kontroverzí. Kritiky a reinterpretace se objevují i po dvaceti letech (např. Zhu 2003). Domnívám se, že není třeba se jimi zde zabývat. Na jejich základě bylo provedeno mnoho replikací i jinak designovaných experimentů, které dospívají k podobným závěrům. Pokud chceme provádět kritiku, měli bychom kritizovat nejaktuálnější a nejlépe designované experimenty. Dle mého názoru nejvýznamnější z poslední doby je uveden dále.

3.1.3 Dekódování rozhodnutí v lidském mozku

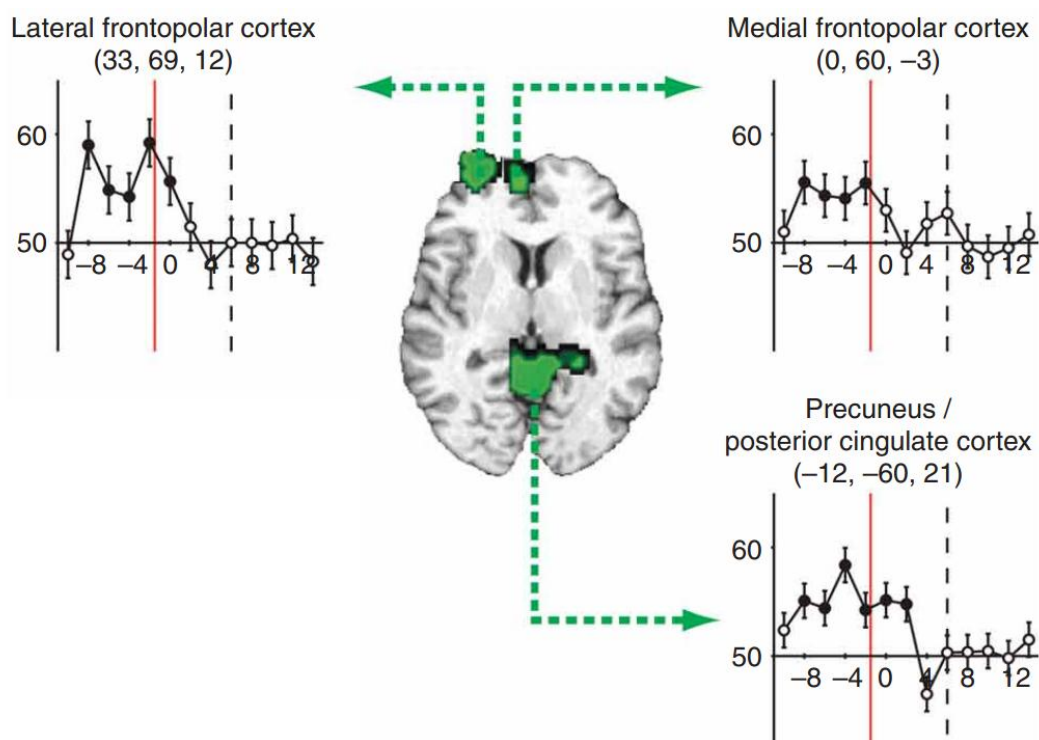
Častým terčem kritik byl poměrně krátký časový rozdíl mezi přípravným motorickým potenciálem a časem W. Je proto možné odvolávat se na různé možné nepřesnosti v časových soudech účastníků a podobně. Úloha také nezahrnovala rozhodnutí mezi více možnostmi, ale pouze ohledně času stisku: „when-decision“ namísto „what-decision“ (Brass a Haggard 2008). Dalším omezením Libetova experimentu byla možnost zjišťování pouze přípravného

motorického potenciálu, který je informací až o poměrně pozdní mozkové aktivitě předcházející pohyb. Tato omezení překonává nedávná studie z Německa (Soon et al. 2008).

Podstatným rozdílem oproti předchozím studiím s EEG bylo použití funkční magnetické rezonance (fMRI) a pokročilých metod analýzy dat. Participantů ve skeneru sledovali na obrazovce řetěz po sobě jdoucích písmen. Každé bylo zobrazeno na 500 ms. Po nějaké chvíli, když se jim chtělo²², se měli rozhodnout mezi stiskem levé a pravé klávesy a ihned poté ji stisknout. Následně měli určit písmeno, které bylo právě na obrazovce, když učinili své rozhodnutí. Časová přesnost 500 ms je dostatečná, protože fMRI snímání bylo prováděno každé dvě vteřiny.

Autoři se snažili analýzou dat z fMRI určit, pro kterou klávesu se účastník rozhodne. Konkrétně, snažili se zjistit míru prediktivní informace v různých částech mozku v různém čase před stiskem i po stisku klávesy. Bylo použito speciální statistické metody rozpoznávání vzorů. Dekodér se učil rozpoznávat vzorec aktivace v určité malé oblasti mozku a na základě tohoto vzorce vydal predikci, zda člověk později stiskne levou či pravou klávesu. Hranice náhody je 50 %. Dle očekávání, statisticky významně šlo výsledek určit během provádění akce z primární a ze suplementární motorické kůry. Klíčové však bylo, že dvě mozkové oblasti kódovaly výsledek rozhodnutí ještě před samotným provedením tohoto rozhodnutí. Jednou oblastí je frontopolární kůra, další oblast zahrnuje precuneus až posteriorní cingulární kůru. Především ve frontopolární oblasti byla statisticky významná prediktivní informace přítomna již 7 vteřin před provedením rozhodnutí. Když se vezme v úvahu pomalá hemodynamická odezva při fMRI snímání, lze podle autorů uvažovat až o 10 vteřinách. Nejdůležitější výsledky shrnuje následující obrázek:

²² Opět bylo použito formulace „when they felt the urge to do so“ (Soon et al. 2008, s. 543). Zde to však není tak na škodu, protože cílem zkoumání není „when-decision“, ale „what-decision“.



Obrázek 5: Míra prediktivní informace o rozhodnutí v různých částech mozku. Zdroj: (Soon et al. 2008)

Na ose y je přesnost predikce v procentech, na ose x čas v sekundách relativně k okamžiku motorické akce. Levá plná vertikální čára značí, kdy nejdříve si byli účastníci vědomi svého rozhodnutí. Pravá přerušovaná čára označuje začátek dalšího pokusu. Plné body značí statistickou významnost na hladině 0,05.

Predikci rozhodnutí 7 - 10 vteřin dopředu již není možné snadno zpochybnit odvoláváním se na nepřesné časové odhady a jiné s tím spojené problémy, které byly relevantní v Libetově experimentu. Dalším významným rozdílem je získání predikce konkrétního „what“ rozhodnutí, nikoli pouze nespecifické přípravné aktivity.

Prediktivní informace je poprvé získána z prefrontálních oblastí. Dle autorů se lze domnívat, že právě zde dochází k začátku tvorby rozhodnutí, a nikoli až v suplementární motorické kůře (SMA), jak bylo dříve předpokládáno (Eccles 1982). SMA kóduje spíše jen výsledný motorický plán.

Experiment byl již úspěšně replikován (Bode et al. 2011) s lepším vybavením (7T skener), které umožnilo získat podrobnější data. Rozšířenou zpětnou vazbou od účastníků se také minimalizoval vliv vědomého uvažování ještě před rozhodnutím. Autoři i zde místo termínu „kauzální vztah“ raději volí „vliv“ na rozhodnutí o 7 vteřin později. I přes tuto korektní opatr-

nost autorů se táží, jaký jiný vztah než kauzální by mohl vysvětlit statisticky významné predikce?²³

Výsledky experimentu nejsou příliš překvapivé z hlediska předpokladu ontologického kauzálního determinismu. Jsou však v silném rozporu s lidově-psychologickou intuicí svobodné vůle, dle které vědomí vybírá a způsobuje naše akce. Neznamenají samozřejmě neexistenci kompatibilistické svobodné vůle, jsou však velmi problematické pro teze o karteziánské substancialitě, kauzální potenci vědomí a ontologické svobodě vůle.

Zřejmou námitkou zastánců kauzální potence vědomí či ontologické svobody může být, že ani zde nejde o skutečná svobodná rozhodnutí, zejména kvůli jejich jednoduchosti. Dá se však (snad nepříliš spekulativně) očekávat, že s rozvojem pokročilejších metod půjde v budoucnu predikovat z mozkové aktivity i složitější rozhodnutí. Pochopitelně ale jen do určité, zatím neznámé, míry, která je dána epistemologickými limity komplexity mozkových dějů, dějů v prostředí, minulostí agenta a podobně.

3.1.4 Epifenomenalismus

Dá se z těchto experimentů definitivně usoudit, že je vědomí epifenomémem? Že je způsobené neurálními procesy, ale samo na ně nemůže působit? Může se zdát, že výsledky experimentů tomuto závěru nasvědčují, ale není to tak jednoduché. Už jen to, že mohu mluvit o svém vědomí, lze brát jako projev toho, že mé vědomí má příčinný vliv. Hlavní otázka je, co vlastně myslíme vědomím. Na teorii vědomí pak závisí odpověď na tyto palčivé otázky.

Především, epifenomenalismus je pozice aplikovatelná pouze v případě, že zastáváme teorii dualismu substancí, který jsem výše odmítl, anebo tzv. dualismu vlastností. Dle druhé jmenované teorie existuje sice jen jedna (fyzická) substance, ale má dva druhy vlastností: fyzické a mentální. Epifenomenalismus by zde znamenal, že mentální vlastnosti nemají kauzální účinnost. Zda je rozlišení vlastností korektní a zda jsou mentální vlastnosti redukovatelné na fyzické, je předmětem stále probíhajících debat. Není zde prostor pro podrobnou diskuzi, uvedu

²³ Pro vyslovení nepochybného závěru o kauzální determinaci by bylo samozřejmě potřeba dosáhnout prediktivní úspěšnosti 100 %. Pro postulování kauzální souvislosti dvou jevů ale stačí nalezení statisticky významné predikce nad úroveň náhody. Neschopnost dosáhnout 100% prediktivní síly je už záležitost čistě epistemologická, daná nedokonalostí použitých metod a komplexitou dějů.

pouze, že existují jiné teorie vědomí kompatibilní s uvedenými empirickými výsledky, které odmítají epifenomenalismus.

Například dle biologického naturalismu Johna Searla patří mentální vlastnosti sice do jiné ontologie (zde spíše ve významu používaného *slovníku*: do slovníku první osoby), ale jsou kauzálně redukovatelné na fyzické. Vědomí je kauzální vlastností mozku, jako je trávení kauzální vlastností žaludku. Dle Searla můžeme legitimně nahlížet na systém z více úrovní popisu: Mentální a fyzická kauzace „nejsou dva nezávislé příběhy, ale popisy na různých úrovních jednoho spojitého kauzálního systému.“ (Searle 2007, s. 331)

Dennett také odmítá epifenomenalismus, jakožto inkohrentní pozici. Zmatení dle něj vychází z mylných intuicí ohledně toho, co to vědomí je. Navrhuje tzv. *multiple drafts model*, dle kterého vědomí tvoří ty procesy, které zrovna vyhrají v kompetenci o řízení chování a vyvoláním určitých důsledků „vstoupí do paměti“. Ale jde o proces distribuovaný v prostoru i čase. „Protože [vědomí] v žádném konkrétním okamžiku ‚opravdu nic nedělá‘, může se zdát, že je naprosto epifenomenální společník, černý pasažér.“ (Dennett 2004, s. 246) Že si jsme vědomí „pružin“ svých akcí je dle Dennetta spíše výjimkou. Přesto tvrdí, že je v rámci jeho teorie vědomí omyl říci, že je nějaká „mentální aktivita nevědomě zahájena a pouze později vstupuje do vědomí.“ (Dennett 2004, s. 242) Vědomí tedy „dělá věci“, ale není potřeba vždy. Mozek pochopitelně optimalizuje procesy pro co nejlepší efektivitu. Vědomé rozhodnutí sahá od první deliberace, že teď budu odpalovat tenisové míčky, nebo že v tomto experimentu budu mačkat tlačítka, po celou dobu hry nebo experimentu, ale „přímé uvědomění“ partikulárního rozhodnutí v každém okamžiku není potřeba. Celé toto konání a rozhodování, ať už je v něm míra vědomí různých procesů jakákoli, je *moje* konání. Mnoho kritiků Dennettovo vysvětlení nepovažuje za příliš uspokojivé, protože podle nich vědomí spíše zrušil, než vysvětlil. Na to jim obvykle odpovídá, že jde o nejlepší možné vysvětlení. Vědomí podle Dennetta sice není „skutečná magie“, ale „hrst triků“, avšak dává nám vše, co potřebujeme. Tedy morálně relevantní svobodu a odpovědnost.

3.2 Sense of agency

Ať už se to má s povahou svobodné vůle a vědomí jakkoli, vynořují se jiné přímo související otázky. Jak je fenomenální prožitek agence, tzv. *sense of agency* (SoA) vytvořen? Co jsou nutné podmínky pro to, abych se považoval za původce svých akcí? Kvůli čemu se prožitek

svobodné vůle dostavuje? Jaké jsou psychologické a neurální mechanismy tohoto prožitku? Experimentální i teoretické zkoumání SoA se začalo více rozvíjet až v posledním desetiletí. Do hry kromě vědy vstupuje i filosofie. Ta se snaží o konceptuální rozbor fenoménu agence. Problematiku SoA stručně představím, protože je to relevantní k mému dále popsanému experimentu (s. 39).

3.2.1 Fenomenologie agence

Americký filosof Shaun Gallagher, inspirován výzkumy se schizofreniky (Graham a Stephens 1994), zavedl následující distinkci:

- Sense of Agency (SoA, prožitek či vědomí agence): „Vědomí (*sense*), že já jsem ten, kdo způsobuje či produkuje danou akci. Například, vědomí, že jsem ten, kdo způsobuje pohyb něčeho, nebo že jsem ten, kdo produkuje určitou myšlenku ve svém proudu vědomí.“ (Gallagher 2000, s. 15)
- Sense of Ownership (SoO, prožitek či vědomí vlastnictví): „Vědomí (*sense*), že já jsem ten, kdo prochází daným prožitkem. Například, vědomí, že mé tělo se hýbe, bez ohledu na to, zda jde o volní pohyb.“ (Gallagher 2000, s. 15)

Sense of agency se dá dále diferenciovat na složku pocitu a soudu:

- Feeling of Agency (FoA, pocit agence): „nepojmový, nízko-úrovňový pocit bytí agentem nějaké akce“ (Synofzik et al. 2008, s. 222)
- Judgment of Agency (JoA, soud agence): „pojmový, interpretativní úsudek o mém bytí agentem“ (Synofzik et al. 2008 s. 222)

Konceptuální rozlišení je důležité pro relevantní teoretické diskuze, design experimentů i jejich korektní interpretaci. V některých případech je však takové rozlišení problematičké. Pro označení prožitku svobodné vůle se přikláním k synonymnímu používání termínů „sense of agency“ či „prožitek agence“. Ve většině experimentů (ne-li ve všech) je zkoumán spíše soud agence. Reflektivně interpretujeme své akce jako intencionálně a svobodně způsobené. Kdežto běžně ve svých životech cítíme spíše pocit agence.

3.2.2 Magické myšlení a hypotéza postdikce

Podle amerického psychologa Daniela Wegnera je vědomá svobodná vůle iluzí (Wegner 2002). Domněnku, že naše vědomí způsobuje fyzické akce ve světě, přirovnává k magickému myšlení. To je kognitivní klam spatřování kauzálních souvislostí tam, kde žádné ve skutečnosti nejsou, avšak tato souvislost je pro nás nejlepší vysvětlení. Příkladem může být zlepšení zdravotního stavu příbuzného po mé modlitbě.

Wegner přichází s teorií zdánlivé mentální působnosti (*apparent mental causation*), dle které jsou akce způsobené nevědomě a prožitek vědomé vůle vzniká na základě tzv. postdikce: je zpětně vykonstruován na základě výsledku akce. Konkrétně, Wegner postuluje tři podmínky (Wegner 2008):

- Konzistence: Předchozí myšlenky musí být konzistentní s výsledkem akce
- Priorita: Předchozí myšlenky musí předcházet akci
- Exkluzivita: Nesmí být přítomny jiné možné příčiny akce

Jeho teorie vychází z několika zajímavých experimentů, které disociují zapříčinění akce a prožitek agence. Například, v jednom experimentu (Pronin et al. 2006) byla jedna skupina účastníků vedena k negativním myšlenkám ohledně druhé osoby. Tu měli zranit pomocí vo-oodoo panenky. Cílová osoba poté schválně udala, že cítí bolest hlavy. Ti účastníci, kteří měli předtím myšlenky kompatibilní s výsledkem akce, si připisovali častěji kauzální vliv na zdravotní stav druhé osoby. Podobně, fanoušci sledující osobu střelící na koš měli pocit, že oni částečně způsobili výsledky střelby. Naopak, v jiném experimentu (Wegner a Erskine 2003) měli účastníci za cíl potlačit myšlenky o svých intencích během rutinního úkolu (namotávání vlákna, kroužkování samohlásek). Ti pak necítili takovou intencionálnost (*intentionality*) či prožitek vůle (*experience of will*) ohledně svých akcí, na rozdíl od účastníků, kteří se na své intence soustředili.

Wegner se také snaží podat evoluční vysvětlení toho, proč máme vědomý prožitek vůle. Vědomí není dle Wegnera (2008) zcela kauzálně impotentní; pouze nezpůsobuje akce přímo. Místo toho nám vědomí prý poskytuje jakousi sociální emoci autorství, na které můžeme zakládat odpovědnost. Zajímavá mi v tomto kontextu přijde také teorie, že vědomí je jakýmsi korelátem ideomotorického učení z důsledků akcí (Dennett 1984).

Wegnerova teorie jistě nebude úplně kompletní. Podmínka konzistence předchozích myšlenek a výsledků akcí je částečně zpochybněna studií (Kühn, Brass 2009), ve které žádné intence

před akcí nebyly v mysli vědomě přítomny, ale byly spíše zpětně vytvořeny až na základě provedené akce. Sám Wegner toto zřejmě uznává. Je spoluautorem studie demonstrující, že „pocit agence“²⁴ nemusí vyžadovat předchozí intencionální, cílově orientované myšlenky, pouze očekávání konkrétních důsledků akce (Aarts et al. 2005). Nejen „intence“, ale i „očekávané důsledky“ se zjevně počítají za „předchozí myšlenky“ (viz 4.4). Existují teorie, že v běžných cílově orientovaných situacích jsou prediktivní informace o důsledcích akcí generovány motorickým systémem a že podobný základ v motorických plánech mají i intence (viz 3.2.4). V této problematice je každopádně mnoho konceptuální nejasnosti i mnoho prostoru pro empirický výzkum.

3.2.3 Retrospektivní konstrukce soudu volního rozhodnutí

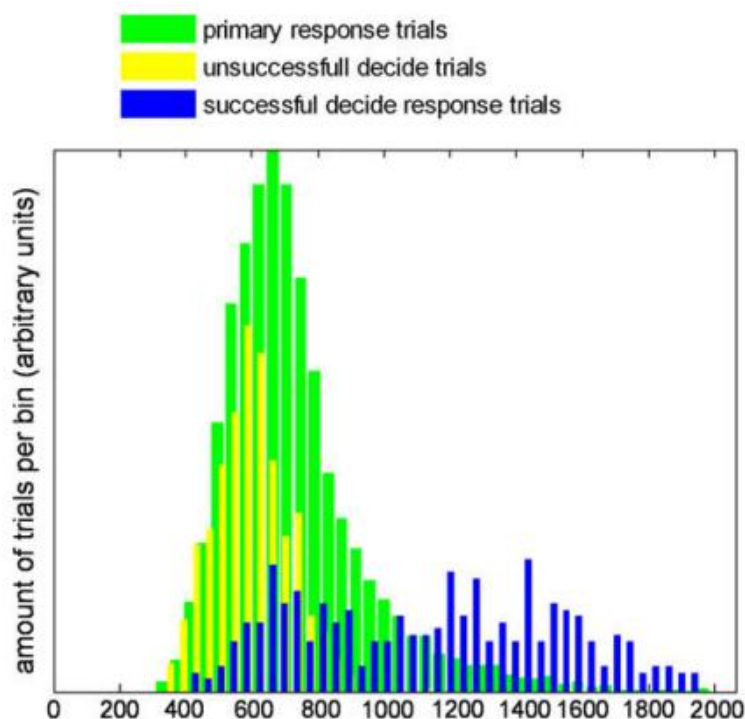
Zatímco ve Wegnerovo experimentech šlo obvykle o zmatení v připsání autorství mezi akcemi vykonanými mnou či někým jiným. V této studii (Kühn a Brass 2009) autoři zkoumali, zda může zmatení vyvstat i u akcí pouze dané osoby, bez cizího vlivu. Konkrétně, zda jsou účastníci schopni rozlišit mezi volním rozhodnutím znovuzahájit a neschopností zastavit předpřipravenou akci. Design experimentu je zde popsán pouze stručně, protože podrobně je popsán v kapitole o mé replikaci experimentu (s. 39).

Účastníci měli za cíl reaktivně odpovídat stiskem tlačítka na zobrazení primárního stimulu. V některých případech byl ale zobrazen odlišný stimul znamenající zastavení akce. Zastavit akci se subjektům podařilo (kvůli nastavení experimentu) přibližně v 50 % případů. V jiných případech byl zobrazen stimul znamenající rozhodnutí. V tom případě měl účastník zastavit svou impulzivní odpověď a rozhodnout se, zda chce tlačítko stisknout, či ne.²⁵ Poté měl indikovat, zda se skutečně pro akci či její inhibici vědomě rozhodl.

Předpokladem bylo, že uvědomění si jiného stimulu, inhibice akce a případná reinitializace akce, by mělo trvat déle, než jen impulzivní odpověď na primární stimul. Analýza reakčních časů (Obrázek 6) ale poskytla překvapivý výsledek.

²⁴ Ekvivalentně jsou použity termíny „sense of agency“, „prožitek osobního zapříčinění“, „pocit autorství“ a „připsání autorství“.

²⁵ „whether-decision“ dle (Brass a Haggard 2008)



Obrázek 6: Distribuce reakčních časů dle typu události. Zdroj: Kühn a Brass (2009)

Na ose x je reakční čas v milisekundách, na ose y počet událostí spadajících do jedné třídy histogramu. Zeleně jsou reakce na primární stimul. Žlutě události, kdy účastníci indikovali, že nedokázali provést inhibiční akce v situaci rozhodnutí. Modře jsou klíčové události, kdy účastníci indikovali úspěšné rozhodnutí.

V podstatné části případů, kdy účastníci indikovali, že se vědomě rozhodli pro stisk tlačítka, to nemohli zjevně v tak krátkém čase stihnout. Introspekce účastníků v těchto případech zřejmě vedla k retrospektivnímu vytvoření soudu o záměrnosti akcí, a to nejspíše interpretativně na základě výsledků akcí. V tomto soudu je zahrnut i soud o úspěšném vetování impulzivní primární odpovědi, ale i tento soud o vetování je mylný.

Zajímavým otevřeným problémem je, proč dochází k tomuto introspektivnímu klamu pouze v některých případech. Dále (v sekci 4.4) se snažím podat možné vysvětlení tohoto jevu. Klemm (2010) na základě zmíněného experimentu namítá i proti jiným typům experimentů, že introspekce není důvěryhodným zdrojem pro hodnocení, zda byla nějaká akce svobodná. Jakým jiným způsobem ale zjišťovat prožitek agence (SoA), než jako JoA pomocí interpretační introspekce? Explicitní vyjádření pocitu o agenci (FoA) již vyžaduje reflexi a konceptualizaci, takže dochází také spíše k vytvoření soudu (JoA). Proto experimentálně zjišťovat přímo pocit agence není snadné.

3.2.4 Integrovaný model

Současný stav výzkumu fenoménu sense of agency naznačuje, že pro vytvoření SoA je důležitá jak postdiktivní složka (vjem výsledku akce), tak prediktivní (Sato 2009). Predikce se zakládá na intenci provést akci, která je pravděpodobně realizována jako motorický plán akce obsahující očekávané důsledky akce (Haggard 2008). Britský neurovědce Patrick Haggard na základě svých (např. Haggard et al. 2002) i jiných studií navrhuje model vysvětlující prožitek intencí, agence a částečně i jáství: „Za prvé, intence predikují akce a cíle akcí, ke kterým odkazují. Za druhé, sensorický vjem akce a jejích důsledků spouští rekonstrukci intence konat. Tyto procesy váží dohromady různé komponenty tohoto prožitku a vytvářejí uvědomění si akce jako celku, smrštěné v čase relativně k podkladovým nervovým procesům. Toto nakonec vytváří reprezentaci sebe jakožto agenta, který je schopen plynulého volního chování.“ (Haggard 2008, s. 941)

3.3 Důsledky výzkumů

Ačkoli výzkum v této oblasti stále pokračuje, několik pravděpodobných závěrů již můžeme na základě doposud získaných výsledků vytvořit. Nejprve ohledně vědomí. Většina mozkových procesů nemá korelát ve vědomém prožívání. Pouze některé komplexní správně organizované procesy odpovídají vědomým prožitkům. Jak přesně, tomu zatím nerozumíme. Je zcela smysluplné předpokládat, že tyto s-vědomím-korelující procesy mají fyzikální příčiny v složitých nevědomých procesech. Nemělo by proto být příliš překvapivé, když dojde k nalezení neurálních determinant lidských rozhodnutí, které předchází samotné uvědomění rozhodnutí. Svobodná vůle tím není popřena, je pouze naturalizována, „odkouzlena“, zbavena schopnosti porušovat fyzikální zákony. Je potřeba jiný, kompatibilistický pohled na to, co za svobodnou vůli můžeme považovat.

Výzkum v oblasti sense of agency nám již začíná dávat jisté odpovědi, jak vědomý prožitek svobodné vůle vzniká. Může být skutečně generován vazbou mezi (nevědomě vytvořenými) „záměry“ a percepční zpětnou vazbou. Podrobněji tomu rozumíme u motorických akcí, ale svobodné autorství cítíme i u vlastních myšlenek, aniž bychom se pozastavovali nad tím, odkud k nám vlastně přicházejí. Existují však i modely podobné modelům pro motorickou SoA, které vysvětlují pocitované autorství myšlenek, i proč někteří schizofrenici naopak pociťují některé myšlenky jakožto cizí (Gallagher 2000).

Sense of agency je klíčovou součástí našeho jáství a sebe-vědomí. Porozuměním SoA tak můžeme časem lépe pochopit, jakým mechanismem je konstruováno samotné Já, které si je vědomo sebe sama, a navíc se cítí být svobodným autorem svých akcí. Dochází k „odkouzlení“ i tohoto Já. Je čím dál jasnější, že není žádnou metafyzickou substancí či zdrojem nezapříčiněných příčin. V tom vidím pozitivní důsledky pro možnost vysvětlení svobodné vůle. Kdyby totiž bylo Já skutečně jakousi dualistickou substancí, nezbylo by nám, než akceptovat epifenomenalismus. Skutečně bychom byli jenom bezmocnými pozorovateli svých životů. Takto ale můžeme pochopit, že Já je velmi široké: Je tvořeno z mnoha různě „fuzzy“ složek. Významnou složkou Já je pochopitelně mozek, se různými²⁶ v něm probíhajícími procesy, vědomými i nevědomými. Proto i když jsou akce zahájeny na nevědomé úrovni, jsou to pořád akce, které pochází z tohoto Já; jsou to *mé* akce. Já koná dle své vůle. Na svobodu (resp. míru svobody) této vůle je pak potřeba pohlížet naturalisticky, kompatibilisticky, nikoli „magicky“ a metafyzicky.

²⁶ Protože je jáství fuzzy koncept, záleží, jak široce ho zrovna chceme uvažovat, a tedy i jaké procesy započteme. Mohu například považovat za součást svého Já schopnost tvořit gramaticky správné věty, ale už nikoli schopnost regulovat krevní tlak.

4 Retrospektivní konstrukce soudu volního rozhodnutí, replikace a modifikace

4.1 Úvod

Většina z nás se podobně jako René Descartes domnívá, že je jim jejich mysl transparentní. Máme pocit, že introspekce nám umožňuje důvěryhodně vnímat svou vlastní mysl, jako nám oči dovolují vnímat okolní svět. Dnes již uznáváme, že nám smysly prezentují pouze zlomek skutečné bohatosti externího světa. Stejně tak již víme, že z veškerých procesů v našem mozku představuje vědomí jen špičku ledovce. Přes tyto znalosti není snadné se ubránit dojmu spolehlivosti a bezprostřednosti našeho přístupu k mysli. A to ani přes evidenci, že zcela systematicky podléháme mnoha kognitivním klamům a introspektivním iluzím. Zajímavým příkladem takové iluze (takřka meta-iluze) je to, že považujeme cizí introspektivní výpovědi za více náchylné k chybám, než své vlastní (Pronin 2009). O sobě samých si myslíme, že známe původ obsahů své mysli, včetně původu svých rozhodnutí.

Dle (Wilson a Dunn 2004) jde v případě introspekce spíše o proces tvořící určité osobní narace. Proces tvorby těchto narací však může být externě ovlivněn různými způsoby. Dle Wegnera (viz 3.2.2 Magické myšlení a hypotéza postdikce) podléháme introspektivní iluzi kontroly svých akcí a vědomé svobodné vůle. Vytvoříme si kauzální naraci mezi naší intencí provést určitou akci a skutečně realizovaným důsledkem akce, pokud jsou splněny podmínky priority, konzistence a exkluzivity. Svou teorii založil na sérii experimentů, při kterých byly manipulovány jak předchozí myšlenky účastníků, tak výsledky akcí. Byla navržena i poměrně radikální teorie, dle které je hlavním zdrojem našich teorií o vlastních interních mentálních stavech pozorování vlastního chování (Bem 1967). Tyto teorie o nás samých mohou být stejně přesné, jako ke kterým dospěje cizí pozorovatel (Wilson a Stone 1985).

Wegnerovy experimenty většinou zkoumaly připsání autorství za akce způsobené někým jiným. Nedávná studie (Kühn a Brass 2009) publikovaná v žurnálu *Consciousness and Cognition* zjistila, že si lidé dokáží retrospektivně přisuzovat intenci vlastní akce, ačkoli žádnou intenci k akci neměli. Toto zjištění podporuje teorie, dle kterých je introspektivní soud o agenci závislý na výsledku akce. Dle autorů také zpochybňuje Wegnerův princip konzistence a možnost vědomého veta. Nakolik oprávněně, bude diskutováno dále.

Jako každou studii s významnými výsledky, je i tuto potřeba podrobit replikaci. Původní studie byla provedena na relativně malém vzorku účastníků, 12 v první a 12 ve druhé, lehce odlišné úloze. Studie ke svým závěrům využila pouze statistické rozdělení reakčních časů (viz sekce 3.2.3). Rozdíl v mé studii je v tom, že replikuje pouze proceduru první, hlavní úlohy, avšak s 30 účastníky. Přináší několik změn v designu, které jsou zaměřené především na možnost ověřit hypotézu nejen pomocí reakčních časů, ale i dodatečným elektrofyziologickým měřením. Změny designu mohou jako sekundární důsledek přinést oslabení velikosti zkoumaného efektu a proto nezvyšují riziko falešně pozitivních výsledků.

4.2 Metoda

4.2.1 Participanti

Experimentu se zúčastnilo třicet participantů, z toho 20 mužů a 10 žen. V původní studii bylo zahrnuto pouze 12 účastníků, proto bylo cílem mé replikace dosáhnout statisticky významnější vzorek populace. Počet účastníků byl stanoven předem podle pravidel vědecké etiky, aby nemohlo docházet k přibírání dalších participantů pro zlepšování případně nevyhovujících výsledků. Několik účastníků (7) však muselo být po vyhodnocení dat vyřazeno, protože se z dat zjistilo, že úlohu i přes podrobné instrukce a názorné vysvětlování před experimentem pochopili nesprávně (4), nebo v jejich případě došlo k výskytu chybějících či nepoužitelných dat (3). Za všechny vyřazené participanty byli zařazeni a testováni náhradníci, aby bylo dosaženo cílového počtu 30 subjektů.

Věk účastníků byl v rozmezí 21 až 31 let, průměrný věk 24,6. Většina účastníků (28) měla dominantní pravou ruku, 2 účastníci měli dominantní ruku levou. Začlenění levorukých účastníků by nemělo mít významný vliv na reakční časy, neboť na polovinu stimulů se odpovídalo levou a na polovinu pravou rukou, a tedy jak levorucí, tak pravorucí účastníci používali svou dominantní ruku v polovině případů. Na záznam EEG tento faktor vliv také nemá, neboť se zjišťovaly potenciály lateralizovaně pouze z motorických oblastí (elektrody C3 a C4, viz dále), kde rozdíly v lateralizaci hemisfér nehrají roli, neboť u leváků i praváků ovládá danou ruku vždy kontralaterální motorická kůra. A především, zjišťovaný evokovaný potenciál P300 má širokou distribuci v obou hemisférách (Soltani a Knight 2000).

Všichni účastníci měli zrak v pořádku, či upravený do normálu brýlemi nebo kontaktními čočkami, a nikdo neměl problém s jasným rozlišením stimulů (ať už druhu písmene či jeho barvy). Nikdo také neměl žádnou medikaci či anamnézu potenciálně ovlivňující výsledky experimentu.

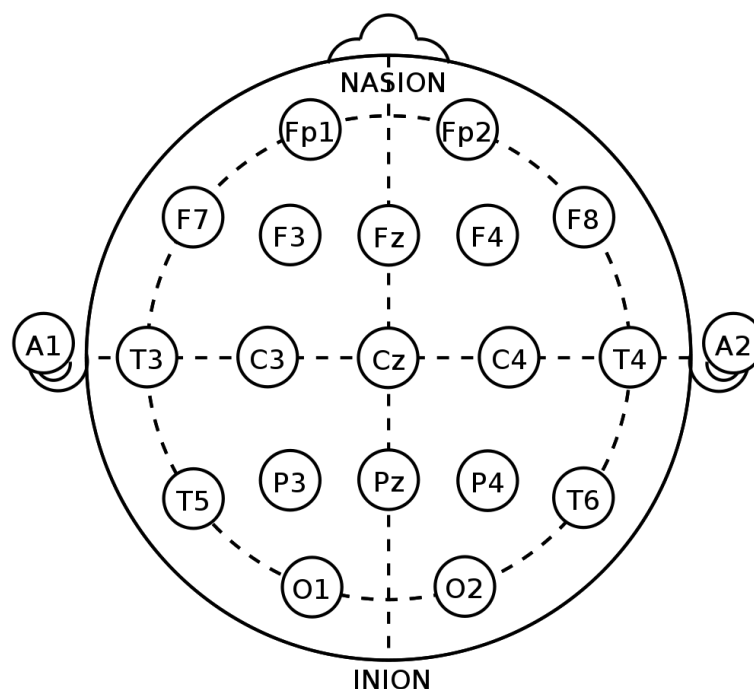
Instrukce k experimentu účastníci obdrželi den předem, aby měli čas k prostudování, navíc jim byly zopakovány a pečlivě vysvětleny znovu před experimentem. Účastníci neznali hypotézu experimentu, bylo jim pouze řečeno, že se výzkum týká rozhodování. Tím měla být zvýšena jistota, že se budou skutečně soustředit na své rozhodování. Tato informace by neměla zkreslit výsledky experimentu ve směru falešně pozitivního nálezu, ale naopak by mohla snížit velikost zkoumaného efektu.

Účastníkům nebyla za účast v experimentu poskytnuta žádná odměna.

4.2.2 Technické aspekty

Účastníci byli umístěni v tiché místnosti, na částečně polohovatelné židli. Poloha byla upravena tak, aby účastník seděl pohodlně, především co se týče opřené hlavy a uvolněných krčních svalů, aby nedocházelo z těchto důvodů k artefaktům v EEG záznamu. Na prezentaci stimulů byl použit CRT monitor s obnovovací frekvencí 75 Hz. Stimuly byly velká písmena M, N, V a W, černé, červené a modré barvy, zobrazené na světle šedém pozadí. Písmena byla vidět pod úhlem přibližně $2,3^\circ$ ze vzdálenosti přibližně 75 cm.

EEG záznam byl nahráván z tří elektrod (C3, C4 a Pz), umístěných v elastické čepici, rozmístěných dle mezinárodního 10-20 systému, viz Obrázek 7. Bylo použito bipolární zapojení s referenční elektrodou Fz. Při zpracování záznamu byl proveden přepočítání dat, aby referencí nebyla Fz, ale elektrody na obou ušních boltcích. Konkrétně, referencí byl průměr hodnot obou uší, $(A1+A2)/2$. V našich experimentálních podmínkách se bohužel nepodařilo dosáhnout optimálních impedancí elektrod pod 5 k Ω . Ty se pohybovaly v rozmezí přibližně 3-17 k Ω , což mělo za následek zhoršenou kvalitu nahrávek. Nahrávání bylo provedeno na EEG zařízení PL-440 (Walter Graphtek CZ s.r.o., Sedlec, Česká republika). Snímkovací frekvence byla nastavena na 256 Hz. Temporální rozlišení EEG záznamu je tak cca 4 ms.



Obrázek 7: Mezinárodní systém 10-20 pro EEG. Převzato z Wikimedia Commons²⁷

Na prezentaci stimulů a záznam odpovědí byla použita vlastní aplikace RetroCon naprogramovaná v jazyce Java, spuštěná v operačním systému Microsoft Windows 2000. Ta počítala také nejdůležitější výstup experimentu, a to reakční časy subjektů, s přesností přibližně na mikrosekundy. Vliv má samozřejmě také obnovovací frekvence monitoru, odezva klávesnice a různá další zpoždění, která nelze kontrolovat. Proto lze říci, že výsledná přesnost reakčních časů je spíše v milisekundách, což je pro další analýzy dostačující.

EEG zařízení mělo samostatné hodiny s odlišným absolutním časem i odlišnou rychlostí běhu času než na počítači s RetroCon aplikací, což vytvářelo ke konci záznamu nepřesnosti v řádu desítek milisekund. Tento stav již nebyl pro potřeby zpracování EEG vyhovující. Bylo proto nutné synchronizovat záznam EEG se záznamem událostí v aplikaci RetroCon. Synchronizace bylo docíleno tak, že aplikace RetroCon při odpovědi účastníka posílala přes COM port signál do speciálně vyrobeného konvertoru, který signál poslal v potřebné podobě do polysonnografického kanálu (Poly1) EEG zařízení. Porovnáním časů pulzů (tzv. triggerů) v Poly1 kanálu a časů z RetroCon aplikace bylo zjištěno, že rozdíl mezi těmito časy je poměrně stálý, avšak postupně se lineárně mírně zvětšující. Z toho bylo usouzeno, že pokud je během zaslání

²⁷Automata124. Commons.wikimedia.org : Electrodes of International 10-20 system for EEG [online]. 29 May 2010 [cit. 2012-05-03]. Dostupný pod licencí Public domain z: <http://commons.wikimedia.org/wiki/File:21_electrodes_of_International_10-20_system_for_EEG.svg>.

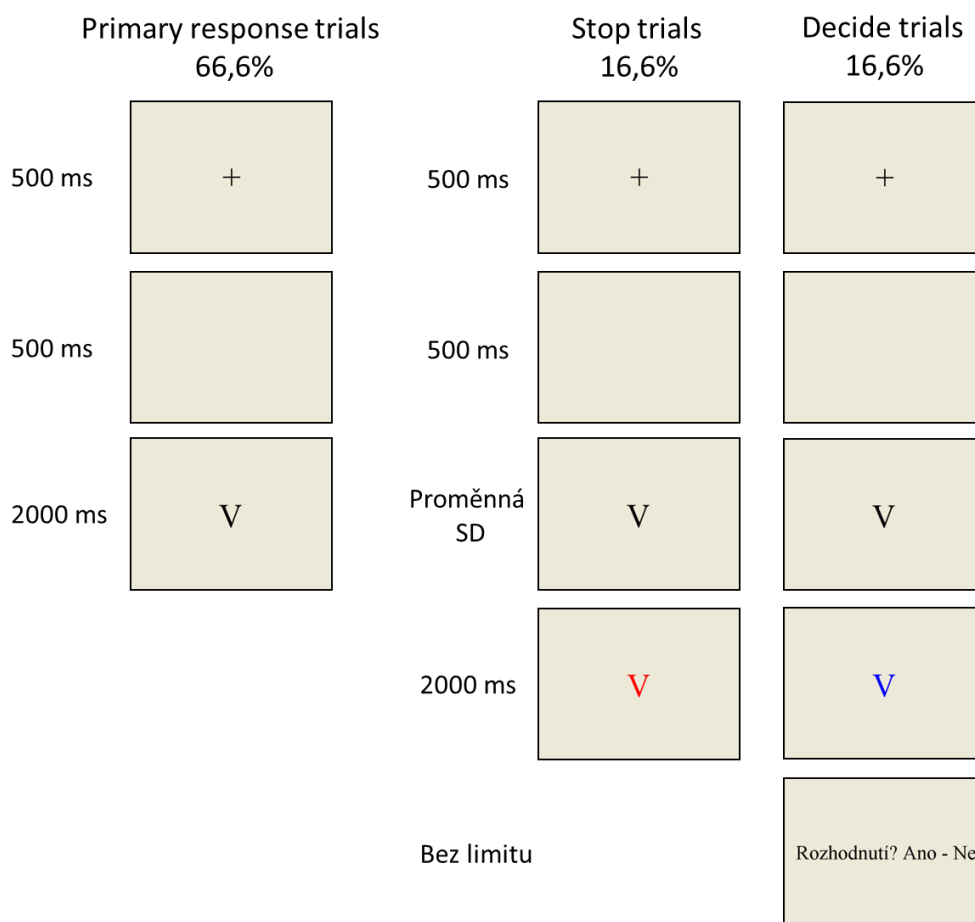
trigger signálu přes COM port nějaká latence, je tato latence stále stejná, díky čemuž lze později v pořádku provést segmentaci a průměrování záznamů. Z uvedeného bylo také patrné, že hodiny v obou zařízeních běží různou rychlostí. Lineární regresí metodou nejmenších čtverců pak byl zjištěn poměr rychlosti běhu obou hodin (cca 0,999444, což při délce záznamu 7 minut vytvoří na konci rozdíl 233 milisekund) a časy poté byly tímto koeficientem upraveny. Získaná přesnost synchronizace byla kolem 5 ms, tedy jednoho až dvou snímků EEG.

4.2.3 Procedura

Experiment replikoval behaviorální proceduru dle Experimentu 1 originální studie (Kühn a Brass 2009, s.13) s některými změnami, které budou popsány dále. Bylo použito Go/Nogo paradigma (spočívající ve stisknutí či nestisknutí tlačítka dle typu zobrazeného stimulu: „Go“ signál vs. „Stop“ signál) rozšířené o „Decide“ signál (spočívající v rozhodnutí se mezi Go a Stop). Jako stimuly byla použita velká písmena M a V, která znamenala stisk levého tlačítka levým ukazovákem, a N a W pro stisk pravého tlačítka pravým ukazovákem (De Jong et al. 1990). Mapování písmen bylo schválně zvoleno poměrně náročně, aby bylo potřeba složitější kognitivní klasifikace, což by mělo mít dle původní studie za cíl optimální variabilitu signální prodlevy (viz s. 45), a dle mého předpokladu i výraznější elektrofyziologický korelát (evokovaný potenciál P300, viz s. 55). V původní studii byl stisk obou tlačítek prováděn stejnou rukou. Užití dvou rukou by nemělo mít vliv na distribuci reakčních časů, neboť každou rukou je prováděna polovina odpovědí, navíc to umožňuje zařazení i participantů s dominantní levou rukou.

Experiment používal design blokové struktury, což znamená, že bylo provedeno více trialů²⁸ za sebou; konkrétně 96 v jednom bloku. Každý trial se navíc skládal z několika částí podle druhu trialu. Strukturu jednotlivých druhů trialů ilustruje Obrázek 8.

²⁸ Pro nedostatek lepšího českého slova použito anglického slova „trial“, znamenající jednotlivou testovanou událost, položku experimentu.



Obrázek 8: Struktura jednotlivých druhů trialů. Zdroj: autor, podle Kühn a Brass (2009)

Na začátku každého trialu byl po dobu 500 ms zobrazen uprostřed obrazovky fixační kříž (ten fixuje zrak účastníků na střed obrazovky, aby byl následně zobrazený stimul zaznamenán co nejrychleji), následovaný 500ms prázdnou obrazovkou. Poté došlo k zobrazení cílového stimulu, tedy některého z písmen M, N, V nebo W, v černé barvě (Go signál). Účastníci byli instruováni co nejrychleji stisknout příslušnou klávesu (levé nebo pravé tlačítko). Stimul byl zobrazen, dokud nedošlo ke stisku klávesy nebo dokud neuplynulo 2000 ms. Poté byla obrazovka na 1000 ms prázdná a následoval další trial. Pokud účastník nestihl odpovědět po dobu 2000 ms, kdy bylo zobrazeno černé písmeno, byl daný trial vyřazen z dalšího zpracování. Stejně tak byl trial vyřazen, pokud účastník odpověděl stiskem špatné klávesy. Pokud stihl odpovědět správnou klávesou, byl zaznamenán reakční čas. Tato situace prosté odpovědi na Go signál je nazývána jako *Primary response* (PR).

V 1/3 případů však došlo po krátké variabilní signální prodlevě (SD, *signal delay*) od zobrazení černého stimulu k jeho přebarvení. V polovině případů (tedy celkově v 1/6) došlo k přebarvení písmene na červenou barvu. To znamenalo Stop signál. V tom případě si účastník musí uvědomit změnu barvy a zastavit svou reflexivní odpověď. Nesmí tlačítko stisknout

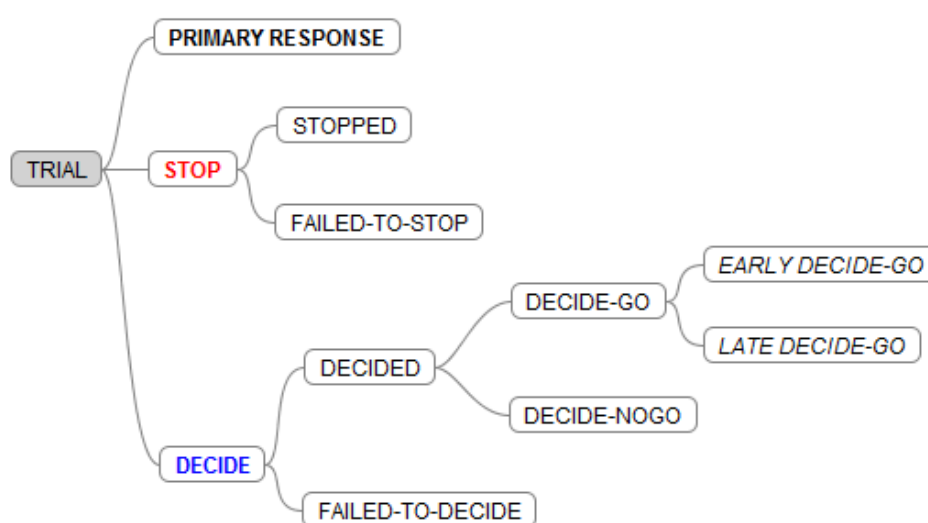
a musí počkat na zmizení stimulu. Pokud si účastník stihl všimnout přebarvení stimulu a skutečně zadržel svou odpověď, byla tato událost klasifikována jako *Stop*. Pokud si účastník nestihl všimnout změny barvy nebo pokud nestihl zastavit svou odpověď a tlačítko stiskl, byla tato událost klasifikována jako *Failed-to-stop*.

V druhé polovině případů (tedy celkově další 1/6) došlo krátce po zobrazení k přebarvení písmene na modrou barvu, což značilo *Decide* signál. V této situaci si měl účastník také uvědomit, že se písmeno přebarvilo, zastavit svou reflexivní odpověď (jakoby se jednalo o *Go* signál) a provést vědomé rozhodnutí, zda opět zahájí akci stisknutí tlačítka, anebo zda tlačítko nestiskne a počká, až písmeno samo zmizí. Po každém *Decide* signálu, ať už participant stiskl tlačítko nebo ne, se zobrazil dotaz „Rozhodnutí? Ano – Ne“. Pokud si účastník byl vědom, že se pro stisknutí nebo nestisknutí skutečně rozhodl, odpověděl levým tlačítkem, jakožto „Ano“. Taková událost byla klasifikována jako *Decide-go*, resp. *Decide-nogo*, dle toho, zda subjekt tlačítko stiskl či ne. Pokud však usoudil, že provedl tuto akci reflexivně bez vědomé deliberace, odpověděl pravým tlačítkem jakožto „Ne“. Tato událost byla označena jako *Failed-to-decide*. Na následnou odpověď, tedy „rozhodnutí o rozhodnutí“ nebyl dán žádný časový limit, na rozdíl od původní studie, kde šlo o 1800 ms. Existuje totiž riziko, že časový tlak na provedení introspekce o záměrnosti rozhodnutí by vedl účastníky k nepřesným odpovědím na tuto klíčovou otázku, což reportovali ve zpětné vazbě účastníci v pilotní studii. Participantů byli instruováni, aby se na zobrazený *Decide* signál rozhodli tlačítko stisknout přibližně v 50 % případů, a ve zbývajících 50 % tlačítko nestiskli. Bylo zdůrazněno, aby si nepřipravovali rozhodnutí dopředu, aby si netvořili žádné strategie, a aby se místo toho rozhodli v každém případě samostatně.

Signální prodleva (SD, čas, za který se původně černě zobrazené písmeno přebarvilo) byla počítána a upravována samostatně pro *Stop* i pro *Decide* signály. Na začátku každého bloku byla stanovena na 300 ms jak pro *Stop* signály, tak pro *Decide* signály, a dále byla upravována schodišťovou metodou (*staircase procedure*). Pokud účastník úspěšně zastavil svou odpověď na *Stop* signál, či pokud po *Decide* signálu odpověděl, že se rozhodl („Ano“), byla SD prodloužena o 20 ms, čímž se úloha stala obtížnější. V opačných případech byla SD zkrácena o 20 ms, čímž se úloha stala jednodušší. Cílem bylo dosáhnout pro každého účastníka takovou prodlevu, aby se udržovala přibližně 50% úspěšnost odpovědí. Toto opatření je důležité jako prevence, aby účastníci nevolili strategii vyčkávání na přebarvení písmena. Pokud by vyčkávali, prodleva by se jim stále přizpůsobovala (prodlužovala by se). Tato informace byla účastníkům sdělena, a byli upozorněni, aby opravdu nevyčkávali. Bylo zdůrazněno, že primárním

cílem je odpovídat co nejrychleji, aniž by udělali příliš mnoho chyb. Přičemž se očekává, že chybu udělají zhruba v 50 % případů, a že to je zcela v pořádku. Signální prodleva byla zvolena jako 300 ms dle původní studie. V mé studii vyšla ze všech trialů průměrná SD pro stop signál 327 ms a pro decide signál 337 ms, což ukazuje, že byla zvolena poměrně správně, resp. mohla být i vyšší, např. 330 ms.

Experimentální design tedy obsahuje celkem šest typů událostí (experimentálních podmínek): PR, Stop, Failed to stop, Decide-go, Decide-nogo a Failed to decide. Událost Decide-go se dá ještě rozčlenit dle reakčního času na *Early decide-go* a *Late decide-go*, dle statistického rozložení reakčních časů Decide-go událostí, viz dále. Schéma těchto událostí ilustruje Obrázek 9.



Obrázek 9: Kategorizace všech typů událostí

Poměr signálů [PR : Stop : Decide] byl zvolen [66,6 % : 16,6 % : 16,6 %]. V původní studii byl použit poměr [75 % : 12,5 % : 12,5 %]. V mé studii bylo rozhodnuto pro větší zastoupení událostí (Decide) především kvůli možnosti porovnat jejich reakční časy s evokovanými potenciály. Pro získání evokovaného potenciálu je potřeba zprůměrovat co nejvíce (přibližně desítky) segmentů EEG daného typu. Z hlediska experimentu jsou klíčové události Decide-go, konkrétně Early decide-go a Late decide-go. Decide události tvoří 1/6 událostí, z toho Decided přibližně 1/12, Decide-go 1/24 a Early/Late decide-go pak kolem 1/48. Posledně jmenovaných událostí je ve skutečnosti ještě o něco méně, protože některé Decide-go události nelze jednoznačně zařadit do Early nebo Late kategorie (viz s. 55). Proto jsem se rozhodl zařadit Decide událostí o něco více. Zároveň jich však nesmí být příliš mnoho, neboť pak by docházelo k oslabování reflexivní PR odpovědi. Zda je poměr událostí stále vyhovující, lze zhruba poznat z výsledného podílu počtu Stopped a Failed-to-stop událostí, který by měl být

kvůli schodišťové proceduře signální prodlevy roven přibližně 1:1. V této studii je tento podíl 56,8 ku 43,2 %. Úloha byla kvůli jinému poměru signálů [PR : Stop : Decide] v této replikaci zřejmě o něco snazší; zkoumaný efekt se i přes to vyskytl (viz s. 50).

Účastníci měli nejprve nejméně dva tréninkové bloky po 30 trialech; v některých případech více, dokud si nebyli dostatečně jisti svojí schopností odpovídat správně vlevo/vpravo dle typu zobrazeného písmene. Poté následovalo pět ostrých bloků po 96 trialech (64 Primary response, 16 Stop a 16 Decide trialů) s krátkou pauzou mezi bloky. Celkový počet trialů tak byl 14 400. Jeden blok trval přibližně 7 minut, celý experiment pak včetně instruktáže, tréninku a připojení EEG elektrod přibližně 1 hodinu pro každý subjekt. Práce na experimentu probíhaly přibližně od května 2011 do dubna 2012.

V původní studii provádělo 12 participantů dvě série pěti bloků po 96 trialech, tedy celkem 11 520 trialů. Méně trialů s více účastníky bylo v mé replikaci provedeno z dvou důvodů. Prvním důvodem bylo ověřit výsledky předchozí studie na širším vzorku populace. Druhým byly praktické důvody. Jak jsem již uvedl, účastníkům nebylo za účast placeno, a již tato hodina byla pro mnohé na hranici času, který byli ochotni experimentu věnovat. Důležitým faktorem byla také únava účastníků při tolika opakováních této stejné úlohy, na což si účastníci často na konci pátého bloku stěžovali. V ideálním designu by samozřejmě bylo nejlepší provést se všemi 30 účastníky v jiný termín druhou sérii experimentů, což však bohužel nebylo z finančních a organizačních důvodů možné.

4.3 Analýza a výsledky

Během experimentů byly nasbírány dva druhy dat. První jsou kvantitativní behaviorální data o výkonu účastníků ve všech trialech, především jejich reakční časy. Tato data byla pořizována aplikací RetroCon a stejný typ dat byl sbírán a vyhodnocován v původní studii. Dále jde nově oproti originální studii o EEG záznam pořízený během plnění úloh.

Část zpracování dat byla provedena již samotnou RetroCon aplikací, část byla provedena v MS Excel, většina pak v programovém prostředí MATLAB pomocí vlastních skriptů, vestavěných funkcí a jedné funkce poskytnuté autory původní studie.

4.3.1 Behaviorální data

Pro každý trial bylo zaznamenáno několik údajů: absolutní čas zobrazení stimulu, absolutní čas přebarvení stimulu (pokud k němu došlo), absolutní čas odpovědi na stimul, reakční čas, čas potřebný k zodpovězení dotazu „Rozhodnutí Ano-Ne“, typ písmene Vlevo/Vpravo, zda účastník odpověděl správnou klávesou a typ události.

Bylo zaznamenáno celkem 14 400 trialů. Z toho dle typu signálu bylo Primary response trialů 9600, Stop trialů 2400 a Decide trialů taktéž 2400.

Úspěšných stop trialů bylo 1364, failed-to-stop trialů bylo 1036. Jejich poměr je 56,8 ku 43,2 %, což se odchyluje od očekávaného poměru 1:1. To naznačuje, že zvolený podíl Stop (a tedy i Decide) signálů k Primary response signálům byl příliš vysoký a úloha tak mohla být o něco málo jednodušší než v původní studii, kde bylo dosaženo poměru 52 ku 48 %. Vyšší poměr by mohl mít za následek slabší primárně zkoumaný efekt v případě decide-go událostí. Očekávaný efekt (bimodální rozdělení) však přesto nastal a lze konstatovat, že poměr signálů zvolený v této studii je v rámci tolerance.

Pro studii nejdůležitějších trialů - decide-go - bylo celkem 865. Decide-nogo trialů bylo 560, a neúspěšných failed-to-decide trialů bylo 975. Poměr decided událostí (decide-go a decide-nogo) k failed-to-decide je 59,4 ku 40,6 %, kdežto v originální studii byl 54 ku 46 %. To je stejný druh odlišnosti jako u Stop trialů a opět indikuje, že úloha byla možná o něco snadnější kvůli zvolenému poměru signálů.

Poměr decide-go a decide-nogo trialů je 60,7 ku 39,3 %, což naznačuje, že rozhodnutí neprovést (inhibovat, vetovat) standardní připravenou akci bylo pro účastníky skutečně obtížné. Tento poměr je zde podstatně vyšší než v původní studii, kde byl 52,5 ku 47,5 %. Dle hypotézy studie dochází v některých případech k retrospektivní konstrukci soudu o záměrnosti akce. Níže uvedená analýza (s. 55) ve shodě s touto hypotézou naznačuje, že z 865 decide-go událostí jich pravděpodobně alespoň 219 nebylo skutečně rozhodnuto a patří do failed-to-decide. Celkový počet failed-to-decide by tak byl 1194. Dále, z oněch 865 decide-go jich bylo pravděpodobně skutečně rozhodnuto 547. To by dávalo celkový počet decided událostí 1107. Zbývajících 99 decide-go událostí nelze podle zvoleného statistického kritéria zařadit. Výsledný poměr decided událostí ku failed-to-decide by tak byl 48,1 ku 51,9 %, což bylo očekáváno kvůli schodišťové proceduře pro signální prodlevu. Toto dále podporuje výsledek následné analýzy a naznačuje, že toto rozčlenění decide-go událostí bylo provedeno správně.

Otevřenou otázkou je, proč získali autoři původní studie poměr blízký 1:1, když i v jejich studii byla část z decide-go událostí ve skutečnosti failed-to-decide.

Ohledně schopnosti účastníků správně odpovědět na písmena M a V levou klávesou a na N a W pravou klávesou bylo zjištěno, že z celkového počtu 14 400 trialů bylo správných odpovědí 11 843, nesprávných 579, a v 1978 případech nedošlo ke stisku klávesy vůbec (stop a decide-nogo). Podíl chyb je 4,7 oproti 4,2 % v originální studii, což není příliš odlišný výsledek.

Základními daty této studie jsou reakční časy účastníků v různých typech výše uvedených událostí. Reakční čas je počítán jako doba mezi zobrazením primárního stimulu (černého písmene) a stiskem příslušné klávesy. Byly počítány pouze události, kde bylo odpovězeno správnou klávesou vlevo či vpravo dle zobrazeného písmene v daném časovém limitu.

Průměry a směrodatné odchylky reakčních časů dle jednotlivých typů událostí uvádí následující tabulka.

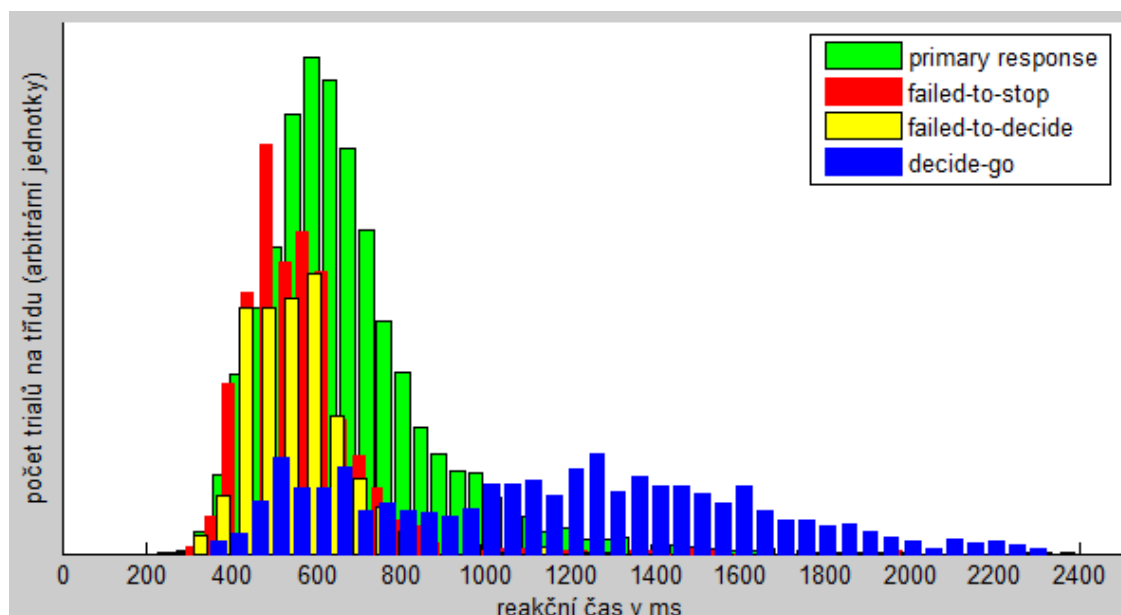
Tabulka 1: Reakční časy dle jednotlivých typů událostí

<i>Typ události</i>	<i>Počet</i>	<i>Průměr (ms)</i>	<i>Směr. odchylka (ms)</i>
Primary response	9572	672,13	208,36
Failed-to-stop	1034	558,20	152,23
Decide-go	865	1158,9	448,81
Failed-to-decide	945	563,74	165,59

Je vidět, že průměry reakčních časů událostí failed-to-stop a failed-to-decide jsou takřka totožné, s podobnou směrodatnou odchylkou. Dvou-výběrový t-test nezamítl nulovou hypotézu o shodnosti průměrů ($t(1977) = -0,775$, $p = 0,44$), což naznačuje, že mají stejnou podstatu. Jde pravděpodobně o stejný kognitivní proces, protože v obou případech jde o reakci na PR signál a o neschopnost participanta reagovat na změněnou barvu stimulu.

Průměrná reakční doba na čisté PR události je také nízká, je však významně odlišná, než u failed-to-stop ($t(10604) = 17,096$, $p < 0,001$) i failed-to-decide událostí ($t(10515) = 15,515$, $p < 0,001$). Možné vysvětlení je, že v části těchto případů účastník vyčkával, zda nedojde k přebarvení stimulu, kdežto jako failed-to-stop a failed-to-decide dopadly ty události, kde účastník s odpovědí nevyčkával. Výsledky mohou značit, že schodišťová procedura nedokázala dostatečně dobře přizpůsobovat signální prodlevy výkonu účastníků a že schodišťový posun signální prodlevy mohl být vyšší než použitých 20 ms.

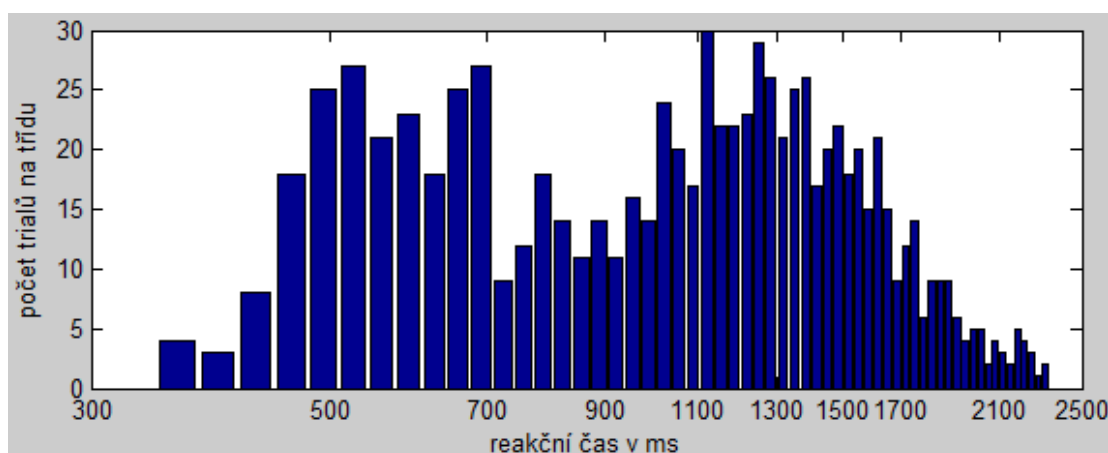
Reakční časy decide-go událostí mají vyšší průměrnou hodnotu, protože inhibice primární odpovědi, rozhodnutí a znovuzahájení odpovědi musí trvat delší čas než pouhá primární odpověď. Tyto reakční časy mají i vyšší směrodatnou odchylku. Důvodem je především to, že z vizuální inspekce jejich histogramu (Obrázek 10) je patrné bimodální rozdělení.



Obrázek 10: Distribuce všech reakčních časů dle jednotlivých typů událostí

Hodnoty tříd byly kvůli řádově odlišným počtům lineárně upraveny pro lepší vizualizaci. Osa y má tak arbitrární měřítko a nevypovídá přímo o počtu událostí, pouze o jejich rozdělení.

Je vidět, že bylo dosaženo stejné distribuce reakčních časů jako v původní studii (Obrázek 6, s.36); lze tak tvrdit, že replikace byla provedena úspěšně. První vrchol bimodálního rozdělení leží zhruba ve stejném místě jako vrcholy ostatních rozdělení, ve kterých šlo o reflexivní odpověď na PR signál. Z histogramů je patrné, že všechna rozdělení vykazují spíše log-normální (LN) rozdělení, než čistě normální, stejně jako v původní studii. Histogram s logaritmicou škálou pouze pro decide-go události (Obrázek 11) poté zjevně vykazuje bimodální normální rozdělení.



Obrázek 11: Distribuce reakčních časů decide-go událostí, logaritmická škála

Bimodalita rozdělení by nasvědčovala, že jde o výsledky dvou různých (zde kognitivních) procesů. To je pro hypotézu celé studie klíčové, proto je potřeba tuto skutečnost důkladně analyzovat. Data byla nejprve zlogaritmována (přirozeným logaritmem) a pak byla modelována jako směs dvou normálních rozdělení:

Rovnice 1: Model bimodálního normálního rozdělení

$$p \times N(\mu_1, \sigma_1^2) + (1 - p) \times N(\mu_2, \sigma_2^2)$$

Parametr p udává proporce ve směsi, tedy jaký podíl dat patří do prvního a jaký do druhého normálního rozdělení. Parametry μ_1 a μ_2 jsou střední hodnoty a parametry σ_1 a σ_2 jsou směrodatné odchylky prvního a druhého normálního rozdělení.

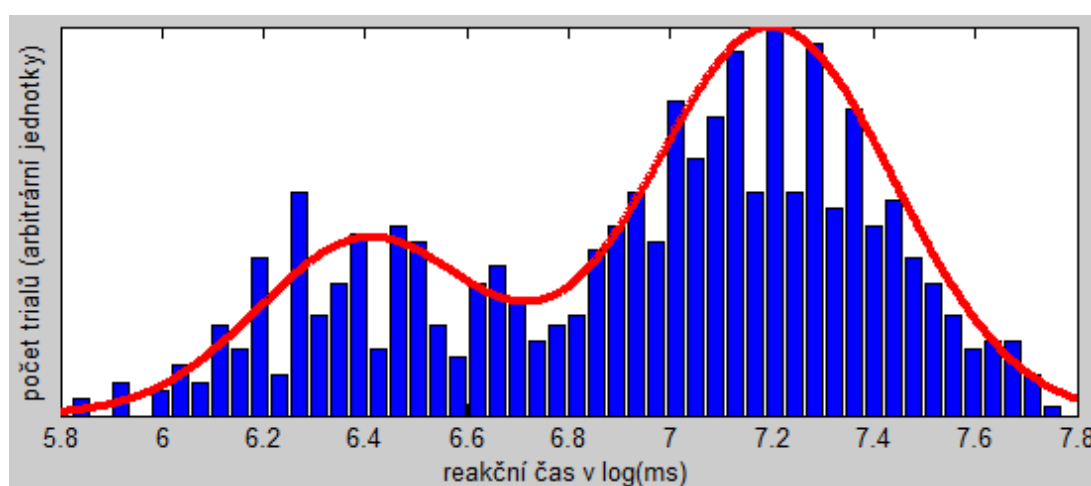
Následně byl proveden odhad parametrů modelu metodou maximální věrohodnosti²⁹. Bylo odhadnuto 5 neznámých parametrů této směsi normálních rozdělení a byly sestrojeny 95% intervaly spolehlivosti. Byly také spočítány charakteristiky složek (medián a střední hodnota) odpovídajících log-normálních rozdělení.

²⁹ Standardní funkce `mle()` programového prostředí MATLAB

Tabulka 2: Odhadnuté parametry modelu decide-go události

Parametr	p	μ_1	μ_2	σ_1	σ_2
Hodnota parametru	0,2918	6,4082	7,2018	0,2163	0,2390
Spodní hranice	0,2478	6,3611	7,1743	0,1835	0,2183
Horní hranice	0,3357	6,4554	7,2294	0,2491	0,2597
Medián LN rozdělení		606,83	1341,9		
Stř. hodnota LN rozdělení		621,16	1380,7		

Grafické znázornění modelu ve formě křivky s těmito odhadnutými parametry lze vidět níže (Obrázek 12).



Obrázek 12: Logaritmované reakční časy decide-go události modelované jako směs dvou normálních rozdělení

Sloupce histogramu jsou naměřené reakční časy, červená křivka je odhadnutý model těchto dat.

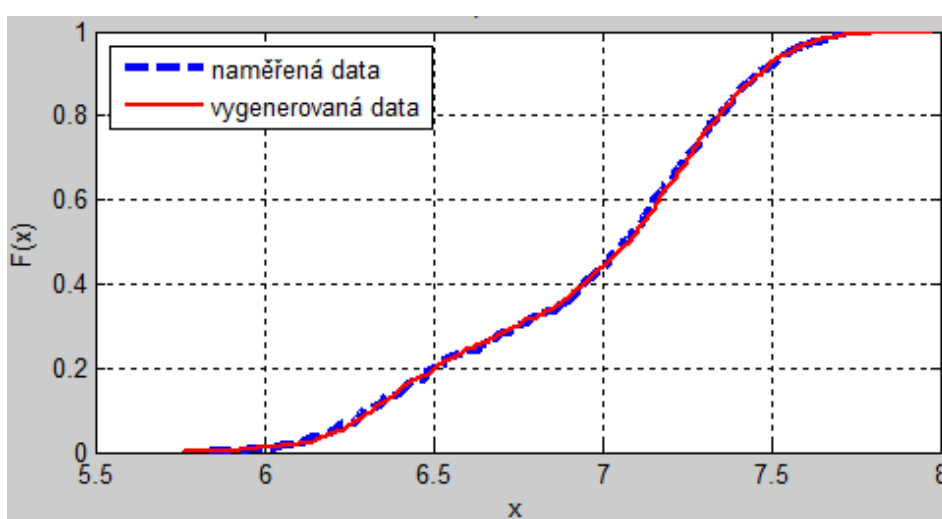
Pro získání měřítka vhodnosti tohoto modelu byl po vzoru původní studie použit dvou-výběrový Kolmogorovův-Smirnovův test (KS-test), který porovnal empirické kumulativní distribuční funkce (CDF, *cumulative distribution function*) naměřených dat a vygenerovaných náhodných dat podle testovaného modelu. Data byla vygenerována za pomoci výše odhadnutých parametrů bimodálního rozdělení. Nejprve byla vygenerována jedna sada pseudo-náhodných dat s normálním rozdělením $N(\mu_1, \sigma_1^2)$ a druhá sada pseudo-náhodných dat s rozdělením $N(\mu_2, \sigma_2^2)$. Tyto dvě sady dat byly pak smíšeny tak, aby každý prvek z prvního rozdělení měl šanci p , že bude vybrán, a naopak každý prvek z druhého rozdělení měl šanci $(1-p)$. Poté byl proveden KS-test testující nulovou hypotézu, že naměřená a vygenerovaná data jsou ze stejného spojitého rozdělení. Alternativní hypotézou bylo, že pochází z různých rozdělení. Použitá hladina významnosti byla 5 %. Protože testování bylo prováděno na ná-

hodně vygenerovaných a náhodně smíšených datech, vše bylo provedeno 10 000krát. Nulová hypotéza byla odmítnuta pouze v 0,51 % případů. Testovou statistikou KS-testu je maximální vzdálenost mezi oběma CDF:

Rovnice 2: Testová statistika KS-testu

$$D_{n,n'} = \sup_x |F_{1,n}(x) - F_{2,n'}(x)|$$

Průměrná hodnota testové statistiky D byla rovna 0,0348 (ve stejných jednotkách jako testovaná data, tedy v $\log_e(\text{ms})$). Obrázek 13 ilustruje tento velmi malý rozdíl mezi CDF původních naměřených dat a náhodně zvolené sady vygenerovaných dat.



Obrázek 13: Porovnání empirických kumulativních distribučních funkcí naměřených dat a dat vygenerovaných na základě statistického modelu

Průměrná asymptotická p-hodnota (pravděpodobnost, že testová statistika vyjde ještě méně příznivě než nulová hypotéza) byla rovna 0,66. To také ukazuje na vhodnost zvoleného modelu.

Model lze otestovat také jedno-výběrovým KS-testem porovnávajícím získaná data a samotný model (Rovnice 1). Nulová hypotéza opět nebyla na hladině významnosti 5 % zamítnuta. P-hodnota byla tentokrát rovna 0,98. Testová statistika D vyšla 0,015654 a test by byl významný až do hodnoty $D = 0,04598$. Takovéto použití Kolmogorov-Smirnovova testu však není vhodné, protože se provádí na stejných datech, ze kterých byly odhadnuty parametry modelu.

Vhodnější je provést ještě chí-kvadrát test dobré shody. Tímto testem byla opět ověřována nulová hypotéza, že data jsou z předpokládaného rozdělení (Rovnice 1). Data byla rozdělena do dvaceti tříd. Počet stupňů volnosti je počet tříd mínus počet parametrů modelu mínus jed-

na, tedy 14. Nulová hypotéza opět nebyla na hladině významnosti 5 % zamítnuta. P-hodnota byla rovna 0,88 a testová statistika 8,233.

Na základě uvedené analýzy lze prohlásit, že uvažovaný model decide-go dat je možné použít.

Pro doplnění replikace byla ještě provedena analýza dat metodou užitou v původní studii. Data byla modelována stejným způsobem, ale byl použit jiný algoritmus metody maximální věrohodnosti³⁰ na odhad parametrů rozdělení. Získané parametry a charakteristiky složek jsou v následující tabulce:

Tabulka 3: Odhadnuté parametry modelu decide-go událostí algoritmem z původní studie

Parametr	p	μ_1	μ_2	σ_1	σ_2
Hodnota parametru	0,3330	6,4450	7,2325	0,2254	0,2107
Medián LN rozdělení		629,55	1383,6		
Stř. hodnota LN rozdělení		645,74	1414,7		

Tyto výsledky se mírně odlišují. Byly proto opět otestovány KS-testem 10 000krát. Nulová hypotéza (rozdělení jsou stejná) byla nyní zamítnuta už v 4,36 % případů, průměrná D hodnota byla 0,0438 a průměrná asymptotická p-hodnota byla 0,43. Chí-kvadrát test dobré shody nezamítl nulovou hypotézu. Testová statistika byla 22,557 a p-hodnota 0,07. Tyto výsledky znamenají, že odhad parametrů pomocí funkce použité v původní studii není tak dobrý, jako výše uvedený.

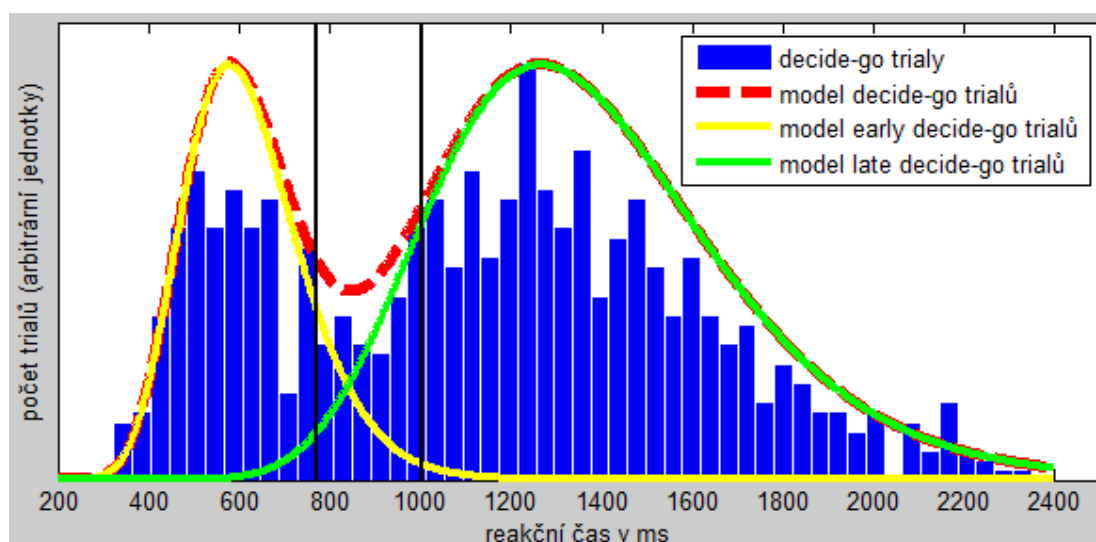
Po nalezení vyhovujícího modelu bylo dále potřeba zařadit decide-go data do tří skupin: První skupina - označená jako *early decide-go* - jsou data, o kterých lze s dobrou pravděpodobností tvrdit, že patří do prvního rozdělení. Druhá skupina byla analogicky označena jako *late decide-go*. Třetí skupina jsou data, u kterých nelze provést zařazení. Bylo potřeba nelézt dvě prahové hodnoty reakčních časů oddělující dané tři skupiny.

Kritérium pro oddělení skupin je nutně arbitrární. Jde zde spíše o praktický požadavek, aby bylo v obou skupinách od každého účastníka dostatečné množství událostí, kvůli možnosti provést kvalitní zprůměrování odpovídajících segmentů EEG. Zároveň by bylo na škodu, kdyby byly ve stejné skupině segmenty EEG odpovídající různým mozkovým procesům. By-

³⁰Funkci laskavě zaslala primární autorka studie Simone Kühn, autorem je Panagiotis Braimakis.

lo zvoleno takové kritérium, že za horní hranici dat, která lze ještě počítat do early decide-go skupiny, byl brán 1% kvantil pravého rozdělení. Podobně, spodní hranice, od které jsou data zařazena do late decide-go skupiny, byl 99% kvantil levého rozdělení. Je tedy pouze 1% pravděpodobnost, že událost zařazená do early-decide go patří ve skutečnosti do late decide-go, a naopak.

Vyšly hodnoty prahů 769,5 a 1003,7 ms. Do early decide-go skupiny bylo zařazeno 219 a do late decide-go skupiny 547 událostí. Nelze dle tohoto kritéria zařadit 99 událostí. Výše provedené porovnání (s. 48) celkového počtu decided událostí (kam byly začleněny late decide-go události) a failed-to-decide událostí (kam byly začleněny early decide-go události) dále obhazuje toto rozčlenění decide-go reakčních časů. Graficky obě rozdělení i prahové hodnoty ilustruje Obrázek 14.

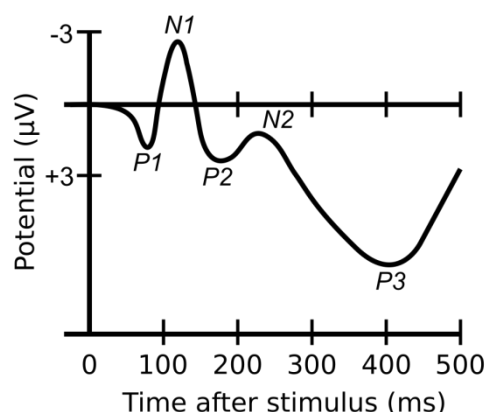


Obrázek 14: Rozdělení Early a Late decide-go trialů

4.3.2 Evokované potenciály

Evokovaný potenciál je elektrické napětí získané elektroencefalografickou metodou ze skalpu, vyvolané jako reakce na externí či interní podnět. Používají se také názvy „kognitivní evokovaný potenciál“ či „potenciál vázaný na událost“ (ERP, *event related potential*). Měřené napětí spojené s touto specifickou mozkovou činností je poměrně nízké, v řádu jednotek mikrovoltů. Naproti tomu běžné napětí EEG je v desítkách mikrovoltů. Pro získání evokovaných potenciálů je tedy potřeba provést zprůměrnění desítek jednotlivých segmentů záznamu. Evokované potenciály mají několik tzv. komponent, což jsou pozitivní či negativní vlny odchylu-

jící se svým napětím od základní linie (*baseline*). Různé komponenty jsou vyvolané různými behaviorálními úlohami. Základní komponenty lze vidět na následujícím obrázku:



Obrázek 15: Komponenty ERP. Převzato z Wikipedia.org³¹

Osa y představuje elektrický potenciál v mikrovoltech; často se uvádí kladné napětí převráceně pod osou x. Osa x představuje čas v milisekundách od stimulu, který evokovaný potenciál vyvolal.

Pro můj experiment je relevantní komponentou P300 či také P3³². Písmeno „P“ značí pozitivní napětí vlny, 300 značí obvyklý čas výskytu vlny v milisekundách od stimulu, který vlnu vyvolal. V případě zrakových stimulů může být dle Bareše (2011) latence (čas od stimulu) komponenty delší (400-550 ms). „P300 je obrazem ukončení poznávací činnosti. Vlna P300 představuje aktivaci mnoha oblastí v neokortexu a limbickém systému v závislosti na stupni volní a mimovolní pozornosti věnované podnětům.“ (Bareš 2011, s. 512) Je vyvolána obvykle použitím tzv. „oddball paradigmatu“. To spočívá v detekci řídkých cílových událostí (oddballů) v sérii událostí běžných. Dle (Picton 1992) amplituda P300 přímo úměrně závisí na řídkosti události a latence na obtížnosti diskriminace. Kognitivně může P300 odpovídat vstupu informace do vědomí či dokončení diskriminačního procesu.

Hypotézou experimentu bylo, že události early decide-go (EDG) mají stejnou podkladovou neurální aktivitu jako failed-to-decide (FD) události. Ta by měla být také podobná pro události failed-to-stop (FS). Ve všech těchto typech událostí participanti dle hypotézy reagovali stis-

³¹ GRINGER. En.wikipedia.org : Components of ERP [online]. 20 February 2009 [cit. 2012-05-03]. Dostupný pod licencí Creative Commons Attribution-ShareAlike 3.0 z: <<http://en.wikipedia.org/wiki/File:ComponentsofERP.svg>>.

³² Konkrétněji, komponenta P300 má dvě subkomponenty, tzv. P3a a P3b, s odlišnými podkladovými kognitivními procesy (Soltani a Knight 2000). Zde je pod názvem P300 myšlena subkomponenta P3b, která následuje krátce po P3a.

kem tlačítka na primární stimul a změnu barvy stimulu si uvědomili příliš pozdě. Tyto tři typy událostí můžeme označit jako „impulzivní“ skupinu. Naproti tomu události stop (S), decide-nogo (DN) a late decide-go (LDG) jsou dle hypotézy reakcí na přebarvený stimul, avšak kognitivní procesy se v těchto třech případech od sebe částečně odlišují. Stop události představují úspěšnou inhibici odpovědi na primární stimul. Decide-nogo taktéž, avšak navíc by měly obsahovat i vědomé rozhodnutí o neodpovědi na přebarvený stimul. Late decide-go události jsou podobné v tom, že také představují inhibici primární odpovědi, ale obsahují vědomé rozhodnutí odpovědět na změněný stimul. Tuto skupinu můžeme nazvat jako „volní“.

Komponentu P300 vyvolává v segmentech průměrovaných k času změny stimulu zaznamenání „odballu“ (změněného stimulu) vždy: ve volní i impulzivní skupině. To bylo z dat částečně potvrzeno. Frekvence „odballu“ i složitost diskriminace je v obou skupinách taktéž stejná. Přesto se lze domnívat, že v impulzivní skupině by mohla být nižší amplituda komponenty, kvůli menšímu množství kognitivního zpracování: to nesouvisí s vědomou inhibicí a rozhodováním, jen s následnou registrací stimulu. Latence by mohla být v impulzivní skupině delší, protože uvědomění změny stimulu nastalo zřejmě později.

Během experimentu bylo nasbíráno přibližně 17 hodin EEG záznamu, který bylo potřeba zpracovat. Elektroencefalografická data byla nejprve oříznuta (dle vizuální inspekce dat a v nich zobrazených triggerů), aby byl vybrán pouze ten úsek, kdy účastník prováděl úlohu, plus několik vteřin před a po úloze kvůli filtrování, které do jisté míry znehodnocuje počáteční a koncovou část dat. Z dat byly vybrány kanály pro elektrody C3, C4, Pz, A1, A2 a polysomnografický kanál. Od kanálů C3, C4 a Pz byla následně odečtena průměrná hodnota z kanálů A1 a A2, aby se ze spojených uší stala reference namísto elektrody Fz.³³ Spojené uši byly zvoleny jako reference, protože se obecně předpokládá, že tyto elektrody obsahují pouze arteficiální (s mozkovými procesy nesouvisející) aktivitu, která je obsažena i ve všech ostatních kanálech. Jejím odečtením by tak měl být zvýšen poměr signálu k šumu. Protože data stále obsahovala kvůli vysokým odporům elektrod velké množství šumu zvláště v pásmu vysokých frekvencí (především 50 Hz síťová aktivita), a protože vlnu P300 generují spíše pomalé mozkové procesy, byl na data (kromě polysomnografického kanálu) aplikován filtr (IIR)³⁴

³³ Původní kanály vyjadřují rozdíl napětí mezi danou elektrodou (včetně ušních) a referencí Fz. Odečtením spojených uší $((A1 - Fz) + (A2 - Fz))/2$ od kanálu např. (C3 - Fz) dostaneme $(C3 - (A1 + A2)/2)$. Tím se referencí stávají spojené uši.

³⁴ *Infinite impulse response*, filtr s nekonečnou impulzní odezvou

eliptický bandpass filtr) propouštějící pouze frekvenční pásmo 0,3 – 30 Hz (Holinger et al. 2000). Filtr byl na každý kanál aplikován dvakrát, jednou ve směru zepředu dozadu³⁵ a podruhé zpět, kvůli eliminaci posunu fáze tímto filtrem.

Záznam byl poté rozdělen na segmenty odpovídající jednotlivým událostem. Kvůli rozdělení bylo potřeba provést výše zmíněnou (4.2.2) synchronizaci časových os a tím umožnit přiřazení behaviorálních dat (časů zobrazení stimulu, časů přebarvení stimulů, typů událostí) k EEG záznamu. Následně byly vybrány segmenty záznamu v rozsahu 100 ms před stimulem až 1000 ms po stimulu. Pro všechny události byla provedena segmentace dle primárního stimulu (časem nula bylo zobrazení černého písmene). U událostí, při kterých došlo k přebarvení písmene (všechny experimentální podmínky kromě PR) byl ještě provedena segmentace, kde byl za čas nula považován čas přebarvení písmene (TC, *time of change*). Tím se zjišťují evokované potenciály vyvolané změnou stimulu. Jako samostatný typ události byly nově brány nejen decide-go, ale i podmnožiny early decide-go a late decide go, kam byly zařazeny události podle výše spočítaných prahových hodnot reakčních časů. Výsledkem je vznik patnácti skupin segmentů pro každého účastníka: PR, Stop, Failed-to-stop, Decide-go, Early decide-go, Late decide-go, Decide-nogo, Failed-to-decide, a to vše kromě PR znovu pro čas změny (TC).

Celkově vzniklo téměř dvacet tisíc segmentů. Od všech byl odečten jejich průměr v předstimulačním úseku (-100 až 0 ms), aby měly všechny segmenty základní linii rovnou přibližně nule. Všechny segmenty byly poté vizuálně zkontrolovány odborníkem³⁶ kvůli případným artefaktům, a arteficiální segmenty byly z další analýzy vyloučeny. Následně byly segmenty stejné skupiny a od stejného účastníka zprůměrněny. Tím byly získány evokované potenciály. Stejný odborník poté ve všech hotových zprůměrovaných úsecích označil vrchol vlny P300, pokud byla tato komponenta přítomna. V kladném případě byla zaznamenána latence komponenty jakožto vzdálenost označeného vrcholu od času stimulu v milisekundách. Podobně byla zaznamenána amplituda komponenty jakožto hodnota (vzdálenost od osy x) označeného vrcholu v mikrovoltech.

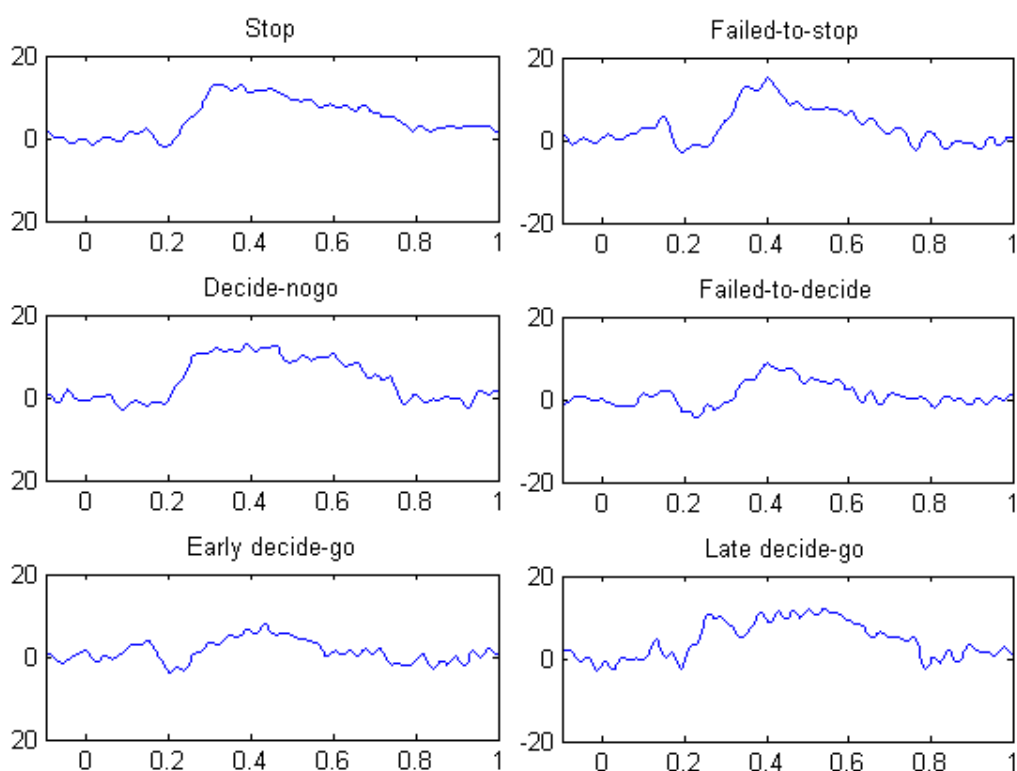
Při analýze dat, která byla provedena najednou (dávkově) až na konci sběru dat bylo bohužel zjištěno, že výsledné zprůměrněné segmenty mají velmi nízké amplitudy; přibližně 1 μV . Vyšlo najevo, že použité zařízení EEG nevyužívá jako referenci odborníkem předpokládanou

³⁵ Kanál je vlastně časová řada dat a filtr tato data prochází vždy určitým směrem.

³⁶ MUDr. Oldřich Vyšata, PhD.

elektrodu Cz, ale Fz, což nebylo uvedeno v manuálu. I v přístroji uložené montáže (schémata bipolárních zapojení pro účely on-line sledování EEG) používaly často přepočítání k Cz, nikoli k Fz. Elektroda Fz ale nebyla plněna gelem a tak způsobila vysoký příjem šumu do signálu. Lze se domnívat, že i po použité filtraci je výsledkem spíše náhodný šum, než skutečný signál. Gel byl aplikován i do elektrody Fz pouze u posledního účastníka, který dosáhl amplitud až 20 μV . Není tak bohužel možné provést intersubjektivní analýzu amplitud a latencí ERP komponent dle jednotlivých typů událostí. Analýza se tak soustředí alespoň na data od třicátého účastníka „A.B“.

Nejdůležitější pro zkoumané hypotézy jsou data zprůměrněná k času změny stimulu. Následující obrázek představuje křivky pro tyto zprůměrněné segmenty. Navíc pro snadnost zobrazení jde o průměry mezi všemi kanály (C3, C4 a Pz). Ostatní vlny pro jednotlivé kanály i průměry k času zobrazení primárního stimulu jsou v příloze A.



Obrázek 16: Evokované potenciály účastníka A.B.

Segmenty jsou zprůměrněné k času přebarvení stimulu. Osa x je v sekundách, osa y v mikrovoltch.

Vrchol, který by mohl odpovídat komponentě P300, je patrný ve všech typech událostí. Jsou vidět zajímavé rozdíly mezi typy událostí, ať už v celkovém tvaru křivky, či konkrétněji v latencích P300. Ve volních událostech dosahuje P300 vrcholu kolem 300 ms, kdežto v impulzivních až kolem 400 ms. Rozdíly v amplitudách lze možná přičíst rozdílu v počtu

průměrovaných segmentů v jednotlivých skupinách. Statistické srovnání amplitud a latencí nebylo provedeno kvůli obtížnému zhodnocení skutečné latence komponenty a kvůli příliš malému vzorku pouze těchto šesti údajů.

Protože nejde snadno srovnat konkrétní parametry, rozhodl jsem se statisticky srovnat celkový tvar křivek: jejich vzájemnou podobnost či rozdílnost. Tato metoda však přináší několik problémů. První otázkou je, jak velký časový úsek pro porovnání zvolit. Jinými slovy, v jak velkém časovém okně je obsažena relevantní informace odlišující typy událostí. Zvolil jsem raději větší úsek, konkrétně celých -100 až 1000 ms, protože počáteční i koncové úseky jsou díky průměrování ve srovnání se středovým úsekem relativně konstantní (blízko nulové hodnoty) a neměly by výsledky významně zkreslit. Druhou otázkou je, jak měřit míru podobnosti křivek. Jako první a nejjednodušší způsob byl zvolen Pearsonův korelační koeficient. Korelační matice pro typy událostí jsou v následujících tabulkách:

Tabulka 4: Korelační matice křivek dle typů událostí, průměr kanálů a Pz

R(a,b)	Průměr kanálů						Pz					
	S	FS	DN	FD	EDG	LDG	S	FS	DN	FD	EDG	LDG
S	1,00	0,83	0,92	0,69	0,73	0,86	1,00	0,90	0,93	0,47	0,51	0,82
FS	0,83	1,00	0,79	0,87	0,88	0,70	0,90	1,00	0,82	0,65	0,66	0,72
DN	0,92	0,79	1,00	0,63	0,65	0,90	0,93	0,82	1,00	0,44	0,43	0,88
FD	0,69	0,87	0,63	1,00	0,81	0,62	0,47	0,65	0,44	1,00	0,79	0,45
EDG	0,73	0,88	0,65	0,81	1,00	0,65	0,51	0,66	0,43	0,79	1,00	0,42
LDG	0,86	0,70	0,90	0,62	0,65	1,00	0,82	0,72	0,88	0,45	0,42	1,00

Tabulka 5: Korelační matice křivek dle typů událostí, C3 a C4

R(a,b)	C3						C4					
	S	FS	DN	FD	EDG	LDG	S	FS	DN	FD	EDG	LDG
S	1,00	0,63	0,81	0,53	0,42	0,78	1,00	0,85	0,89	0,73	0,75	0,89
FS	0,63	1,00	0,64	0,82	0,70	0,49	0,85	1,00	0,81	0,90	0,88	0,78
DN	0,81	0,64	1,00	0,63	0,65	0,83	0,89	0,81	1,00	0,68	0,67	0,92
FD	0,53	0,82	0,63	1,00	0,74	0,54	0,73	0,90	0,68	1,00	0,86	0,70
EDG	0,42	0,70	0,65	0,74	1,00	0,54	0,75	0,88	0,67	0,86	1,00	0,70
LDG	0,78	0,49	0,83	0,54	0,54	1,00	0,89	0,78	0,92	0,70	0,70	1,00

Všechny korelace jsou významné na hladině 0,01, protože všechny křivky jsou si v základu podobné. Je však vidět, že korelace událostí v rámci volní a impulzivní skupiny jsou převážně vyšší než mezi těmito skupinami. Výjimkou je poměrně vysoká korelace mezi S a FS. Pro hypotézy je důležité, že korelace EDG a LDG je relativně nízká. EDG je nejvíce korelováno s FS a FD, což ukazuje na správnost domněnky, že se v EDG událostech účastníci ve skutečnosti nerozhodli pro stisk. LDG je nejvíce korelováno s DG a poté s S, což naopak ukazuje na správnost předpokladu volního rozhodnutí v těchto událostech.

Další použitou metrikou podobnosti křivek je eukleidovská vzdálenost. Ta je definována jednoduše takto:

Rovnice 3: Eukleidovská vzdálenost

$$\sqrt{\sum_{i=1}^n (p_i - q_i)^2}$$

Matice eukleidovských vzdáleností mezi křivkami mají podobné relativní proporce jako výše uvedené korelační matice (vysoce korelované křivky mají nízkou vzdálenost a naopak), proto zde nebudou uvedeny. Byla provedena shluková (*cluster*) analýza křivek algoritmem K-means. Ta zařadila křivky do předem zvoleného počtu shluků pomocí čtverce jejich eukleidovské vzdálenosti. Dle hypotézy existují dva shluky: volní a impulzivní typy událostí. Konečnější je ale provést analýzu pro všechny možné počty shluků, tedy 2 až 5. Výsledné zařazení do shluků ukazuje následující tabulka:

Tabulka 6: Shluková analýza křivek

Kanál	Průměr				C3				C4				Pz			
Počet shluků	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
Stop	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Failed-to-stop	2	3	3	3	2	2	2	2	1	3	3	3	1	3	3	3
Decide-nogo	1	1	1	5	1	3	3	3	1	1	4	4	1	1	1	5
Failed-to-decide	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Early decide-go	2	2	2	2	2	2	2	5	2	2	2	2	2	2	2	2
Late decide-go	1	1	4	4	1	1	4	4	1	1	1	5	1	1	4	4

Čísla ve sloupcích představují číslo shluku, do kterého typ události patří. Shoda v číslech shluků mezi sloupci neznámá, že jde o shluky se stejným centroidem. Pořadová čísla shlu-

ků byla postupně přidávána u typů událostí, které při zvýšení počtu shluků získaly nově samostatný shluk (označené tučně).

Volní a impulzivní shluky byly vytvořeny pro kanály C3 a pro průměr kanálů dle očekávání. V kanálech C4 a Pz je FS přiřazeno do volní skupiny, ale pro tři shluky se již okamžitě vyčleňuje do vlastního shluku. Dle předpokladu FD a EDG sdílí stejný shluk a to nejen pro dva shluky, ale dokonce většinou až pro pět shluků. LDG je naproti tomu poměrně konzistentně zařazeno do shluků společně s volními událostmi S a DN.

Pro porovnání spíše tvaru křivek bez ohledu na amplitudu byly křivky upraveny odečtením průměru a vydělením jednou směrodatnou odchylkou.³⁷ Opět byla provedena shluková analýza:

Tabulka 7: Shluková analýza upravených křivek

Kanál	Průměr				C3				C4				Pz			
Počet shluků	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
Stop	1	1	1	1	1	1	4	4	1	1	4	4	1	1	1	1
Failed-to-stop	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	5
Decide-nogo	1	1	1	1	1	1	1	1	1	1	1	1	1	3	3	1
Failed-to-decide	2	3	3	3	2	2	2	5	2	2	2	5	2	2	2	2
Early decide-go	2	2	2	5	2	3	3	3	2	3	3	3	2	2	4	4
Late decide-go	1	1	4	4	1	1	1	1	1	1	1	1	1	3	3	3

Pro dva shluky vychází zařazení událostí kromě kanálu Pz dle hypotézy. S, DN a LDG spolu opět poměrně konzistentně sdílí zařazení. Avšak EDG již nemá tak robustní sounáležitost s FS a FD. Nicméně EDG a LDG spolu nesdílí shluk v žádném případě, což ukazuje na oprávněnost rozdělení decide-go událostí na tyto dvě skupiny.

4.4 Diskuze a shrnutí

Lze konstatovat, že původní studie byla ve svém jádru úspěšně replikována. Byl nalezen stejný efekt v podobě bimodálního rozdělení reakčních časů v decide-go událostech. Existence daného efektu byla ověřena na poměrně širokém vzorku 30 účastníků, a to i přes některé

³⁷ U výše uvedených korelačních koeficientů nemá tato operace žádný vliv.

změny v designu, které mohly výsledný efekt oslabit. Výše uvedená analýza behaviorálních dat (reakčních časů a počtu jednotlivých typů událostí) podporuje oprávněnost rozdělení decide-go událostí na podskupiny early decide-go a late decide-go. Stisky tlačítka označené účastníky za volně rozhodnuté tak ve skutečnosti pravděpodobně odpovídají směsi dvou různých typů odpovědi. Střední hodnota early decide-go událostí je příliš nízká, než aby za tu dobu bylo možné vykonat vědomou inhibici reflexivní odpovědi a rozhodnutí o reiniciaci stisku tlačítka, jak ukazují podobně krátké reakční časy v ostatních reflexivních typech událostí (PR, failed-to-stop, failed-to-decide). První typ odpovědi (early decide-go) tak ve skutečnosti bude nejspíše reflexivní odpovědí na primární stimul, přičemž soud o jeho záměrnosti byl v těchto případech mylně zpětně vytvořen. Druhý typ odpovědi (late decide-go) byl nejspíše skutečně proveden po úspěšné inhibici primární odpovědi a deliberaci ohledně odpovědi na změněný stimul.

Současná studie si kladla za cíl podpořit tato zjištění i analýzou evokovaných potenciálů získaných elektrofyziologickým měřením z centrální oblasti skalpu. Bohužel byla získána kvalitní využitelná data pouze od jednoho účastníka, proto na jejich základě nelze vytvářet žádné obecné závěry. Lze tvrdit pouze to, že data podporují zkoumanou hypotézu u jednoho účastníka, ale ani u něj nelze vyslovit definitivní závěr. Zjištěné náznaky jsou však velmi zajímavé. Neurální procesy, které jsou podkladem odpovědi v early a late decide-go událostech jsou pravděpodobně skutečně odlišné povahy. Analýzy ukazují, že uvnitř volní skupiny i uvnitř reflexivní skupiny je větší míra podobnosti neurálních procesů než mezi těmito skupinami. Early decide-go vykazuje podobnost spíše s failed-to-decide a failed-to-stop, kdežto late decide-go spíše se stop a decide-nogo. Bylo by zajímavé provést nová měření evokovaných potenciálů na vzorku více účastníků.

Pro rozlišení early a late decide-go událostí postačují zjištěné reakční časy. Otázkou je, jak lze výskyt early decide-go událostí interpretovat. Jednou z možností je lhaní subjektů o provedeném rozhodnutí v části případů. Autoři původní studie nenalezli podporu pro tuto hypotézu na základě provedených psychologickými testy³⁸. Z mého pohledu je tato hypotéza nepravděpodobná spíše proto, že by lhát muselo v podobné míře mnoho účastníků, kteří k tomu navíc neměli žádný důvod.

³⁸ Marlowe–Crown Social Desirability Scale, Paulhus Balanced Inventory of Desirable Responding

Věrohodnější hypotézou je nejistota účastníků při introspekci ohledně záměrnosti jejich odpovědi. Někteří účastníci po experimentu skutečně odpovídali, že odpověď na otázku „Rozhodnutí? Ano-Ne“ pro ně byla obtížná. Za míru jistoty odpovědí lze považovat čas potřebný pro tuto následnou odpověď. V původní studii byla hypotéza o nejistotě účastníků odmítnuta, protože v early decide-go událostech byla průměrná doba následné odpovědi významně kratší (579 ms), než v late decide-go (660 ms). Naproti tomu v mé současné studii byly nalezeny opačné výsledky. V early decide-go případech trvalo „rozhodnutí o rozhodnutí“ průměrně 912 ms, kdežto v late decide-go případech 690 ms. Zjištěný rozdíl je významný ($t(764) = 5,518$, $p < 0,001$). Odlišnost oproti původní studii může být dána odstraněním časového limitu na provedení introspekce.

Různá míra introspektivní jistoty může vysvětlovat, proč se účastníci ohledně záměrnosti svých akcí mýlili jen v části případů. Co ale nejistotu a omyly způsobuje? Jaké ovlivňující faktory přichází v úvahu? Mou hypotézou je, že introspekce účastníků zahrnuje vyvolání vzpomínek na mentální obsahy přítomné v mysli před a při stisku klávesy. Jedním faktorem tak může být míra uvědomění či pozornosti věnovaná vlastním mentálním obsahům, která může ovlivnit proces jejich zapamatování. Například, zda a nakolik jsem si byl vědom přebarvení stimulu nebo míra deliberace³⁹ o mém „whether“ rozhodnutí. Následně musí při introspekci dojít k vybavení předchozích mentálních obsahů z paměti, což je rekonstruktivní proces, který může být ovlivněn externími faktory. Pomocí nich jsou někdy vytvořeny falešné vzpomínky na mnohem komplexnější události, než je záměr provést akci (Loftus a Ketcham 1996). Nepřesnost vzpomínek je nejčastěji zkoumána u dlouhodobé epizodické paměti, ale byla demonstrována i u krátkodobé paměti.⁴⁰ Konfabulace⁴¹ může nastat z různých důvodů

³⁹ Každé rozhodnutí může být zvažováno různě dlouho a komplexně, s různou mírou explicitního uvažování i implicitních vlivů. Dennett uvádí příklad náhlého rozhodnutí (*snap decision*): Jeho žena se ho zeptá, zda by jí po cestě do práce odvezl balíček na poštu. Odpoví takřka okamžitě, že ne, protože by přišel pozdě na konzultaci se svým studentem. Provedl deliberativní rozhodnutí? Dennett (2004, s. 115) dále popisuje velmi dlouhý seznam důvodů vedoucích k jeho rozhodnutí, které přímo při rozhodnutí vědomě nezvažoval, ale mohl je uvést a jeho rozhodnutí jistě *nějak* ovlivnily.

⁴⁰ Jde například o fenomén tzv. rozšíření okrajů (*boundary extension*), kdy si lidé vybavují i po velmi krátkém čase širší okolí fotografie, než jaká jim byla skutečně ukázána. Dalším příkladem je chybné vybavení sémanticky příbuzného slova nepřítomného v prezentovaném seznamu slov (Deese-Roediger-McDermott paradigma).

⁴¹ „Konfabulaci je možné definovat jako výroky či akce, které zahrnují zkreslení vzpomínek.“ (Metcalf et al. 2007, s. 23) Jde například o vytváření nepravdivých nebo nepřesných informací či interpretací vlastních akcí, které ovšem subjekt považuje za zcela pravdivé.

(Metcalf et al. 2007). Jednou z teorií je udržení obrazu své identity, konceptu svého Já jakožto svobodně konajícího. V nejasných případech může dojít k inferenci typu „protože jsem akci provedl, tak jsem to tak jistě chtěl“. Jinou možnou teorií je zkreslení časové posloupnosti událostí. Účastníci mohli mít zpětně dojem, že si nejdříve uvědomili změnu stimulu a až následně zahájili akci, přičemž to mohlo být naopak. Podmínky vybavení z paměti jsou zřejmě ve všech případech stejné. Proto se domnívám, že hlavním faktorem je povaha mentálních obsahů před a při akci, která ovlivňuje jejich zapamatování. Může jít o míru pozornosti a uvědomění změny stimulu, čas uvědomění změny relativně k zahájení akce, či o jiné faktory.

Tyto otázky autoři původní studie nediskutovali, protože hypotézu o nejistotě účastníků na základě rychlosti odpovědí o rozhodnutí zamítli. Vzápětí v článku ale postulují, že by pro účastníky mělo být snazší rozlišit, zda reflexivní akci zastavili či ne, pokud by si sami volili druh vykonané akce. Tím vlastně uznávají, že je rozlišení pro účastníky problematické a že si jím nemusí být jistí. V druhém experimentu stejné studie svoji hypotézu otestovali a zjistili, že v případech, kdy si účastníci mohli zvolit druh odpovědi, byl počet early decide-go událostí menší. Výsledek přímo interpretují tak, že si účastníci byli v těchto případech hodnocením svého rozhodnutí více jistí, ale rychlosti odpovědí o rozhodnutí tentokrát nejsou uvedeny. Vliv míry svobodnosti akce (možnosti zvolit typ odpovědi) na jistotu soudu o záměrnosti rozhodnutí autoři podrobněji nediskutují. Dle mé interpretace pokleslo množství early decide-go událostí proto, že se účastníci více „soustředili“⁴² na svoje mentální obsahy (vjem stimulu, akční plán, pořadí,...), a proto si je dokázali při introspekci věrohodněji vybavit. Větší svobodnost úlohy možná pouze zvýšila míru uvědomění a tedy i zapamatování mentálních obsahů. Bylo by zajímavé provést jinou modifikaci experimentu, která by dávala stejnou míru svobody konání jako základní varianta experimentu, ale zvýšené soustředění na své mentální obsahy by bylo vyvoláno jiným způsobem. Podobně by bylo zajímavé provést variantu úlohy v „dual task“ paradigmatu, kdy by paměť účastníků při každém trialu byla navíc zatížena například zapamatováním si náhodně zobrazených čísel. Jiná vysvětlení výskytu early decide-go událostí a ovlivňujících faktorů jsou samozřejmě možná.

Jaké jsou teoretické implikace získaných výsledků? Autoři původní studie se je snaží aplikovat na Libetovu teorii veta. V early decide-go událostech soud o záměrnosti stisku tlačítka navíc obsahuje i soud o úspěšné inhibici primární reflexivně vyvolané akce. Podle autorů i

⁴² Dennettovsky řečeno, tyto mentální obsahy kvůli své větší bohatosti výrazněji vítězily v kompetici o ovlivnění chování a tedy i o vstup do paměti.

podle Libeta bychom měli mít vědomý přístup k informaci o inhibici (vetu) nevědomě zahájené akce. Vědomé veto v těchto případech ale zjevně nebylo úspěšné, ačkoli si to účastníci mysleli. Dle autorů to „jasně ukazuje proti Libetově (1999) předpokladu, že proces vetování může být vědomě zahájen“ (Kühn a Brass 2009, s. 20). Domnívám se, že to je příliš silné zobecňující tvrzení. V mnoha případech (stop, decide-nogo, late decide-go) účastníci úspěšně vetovali impulzivní odpověď. S určitostí lze tvrdit jedině to, že *v některých případech* nás introspekce ohledně provedení veta klame.

Jak zapadají výsledky studie do Wegnerovy teorie zdánlivé mentální působnosti? Podle autorů získané výsledky zpochybňují Wegnerovu podmínku konzistence předchozích myšlenek s provedenou akcí. Podle nich je totiž předchozí myšlenkou „úspěch v úloze“. Tato interpretace mi přijde zvláštní. Wegner sice formuluje svá kritéria poměrně vágně (čímž se do značné míry snižuje možnost jejich vyvrácení). Za předchozí myšlenky lze brát záměr, akční plán, nebo jenom předpokládané sensorické důsledky akce. Nejsm si ale jist, zda za ně můžeme považovat i tak obecnou abstraktní myšlenku jako „uspějí v úloze“. Podobně vágní je pojem „akce“ a „důsledky akce“.

Obsahem myšlenky před každou akcí při zobrazení černého stimulu je stisknutí tlačítka. Dejme tomu, že se tato akce (reakce na černé písmeno) skutečně uskuteční. Akcí je stisknutí tlačítka a sensorickým důsledkem je proprioceptivní, zrakový a hmatový vjem mého prstu na stisknutém tlačítku. Nastává konzistence myšlenky a akce. Mám dle Wegnera (i dle své zkušenosti) určitý *druh* či *míru* prožitku vědomé vůle. Vědomě a svobodně jsem konal dle instrukcí, viděl jsem černé písmeno a stiskl jsem klávesu, nikdo mě ke stisku nenutil. Vezměme sensorické důsledky akce šířeji: dejme tomu, že písmeno je při mém stisku již modré, nikoli černé. Mám najednou menší míru prožitku vědomé vůle? Na to by následná otázka musela znít například: „Nakolik se cítíte být odpovědní za svůj stisk tlačítka?“ Otázka ale zní: „Uvědomili jste si změnu barvy stimulu, provedli jste inhibici reflexivní odpovědi a deliberaci o reinicializaci stisku?“ Když odpovím „Ano“, ačkoli to dle reakčního času nemůže být pravda (early decide-go události), má odpověď zřejmě nic nemění na mém prožitku autorství akce. Mění se jenom připsání autorství množiny mentálních obsahů před akcí. Podle mého názoru se tak mění spíše pocíťovaný *důvod* akce, a možná *míra* či *druh* prožitku vědomé vůle, ačkoli to otázka nezjišťuje.

Klemm (2010) má zřejmě pravdu, když konstatuje, že jednoduché experimenty nejsou nakonec tak jednoduché. Zkoumané jevy, podkladové teorie i relevantní pojmy jsou komplikované

a často nejasné. Interpretace výsledků je pak velmi obtížná. Není snadné odpovědět, nakolik studie skutečně zkoumá soud o volním rozhodnutí, jak je uvedeno v názvu studie. Lze však říci poměrně jistě, že studie zkoumá retrospektivní přisouzení specifické intence k akci. Studie ukazuje, že si můžeme v některých případech zpětně vykonstruovat nepravdivé záměry konání. Opět se ukazuje, že nám naše mysl není transparentní a introspekce a paměť ohledně vlastních mentálních obsahů není zcela spolehlivá. Podmínky ovlivňující tento jev prozatím nejsou zcela jasné.

5 Jak svobodní jsme

V předchozích kapitolách jsem představil nejdůležitější pojmy relevantní k problému svobodné vůle. Ukázal jsem, že na ně lze nahlížet z více pohledů. Klasické intuitivní pohledy na pojmy jako je svoboda, determinismus, možnost, odpovědnost, osoba (jáství) či vědomí vedou k rozporům jak na čistě konceptuální úrovni, tak v konfrontaci s empirickými výsledky. Na poznatcích současných kognitivních věd jsem ukázal především, že náš koncept vědomého Já jakožto nezapříčiněného zdroje příčin je neudržitelný. Jak můžeme shrnout výše řečené a jaké můžeme vyvodit závěry ohledně svobody lidské vůle? Jak bylo předestřeno za počátku, záleží, co svobodou myslíme.

5.1 Ontologická svoboda

Tradičně je svobodná vůle považována za neslučitelnou s determinismem. Argumenty pro tuto intuitivní inkompatibilitu lze rozdělit do dvou skupin (Vihvelin 2011):

1. Kvůli determinismu nemáme možnost konat jinak.
2. Kvůli determinismu nemůžeme být zdrojem svých akcí.

Domnívám se, že jsem přesvědčivě ukázal platnost různých forem determinismu. Možná_(epistemologicky) existující náhoda na kvantové úrovni je na naší makro-úrovni zanedbatelná; zde platí adekvátní determinismus. Pokud budeme chtít na základě kvantové mechaniky kauzální i adekvátní determinismus přeci jenom zamítnout, (a) náhoda neposkytuje ontologickou svobodu a (b) na vyšší úrovni popisu narazíme na determinismus genetický a environmentální. Pokud přijmeme tradiční intuitivní konceptualizaci problematiky a budeme požadovat „skutečnou“, ontologickou svobodnou vůli, je existence svobodné vůle – domnívám se – výše uvedenými argumenty efektivně vyvrácena. Libertarián Robert Kane (1998) proto předpokládá neplatnost determinismu a snaží se ukázat, jak se nám tím otevřou alternativní možnosti jednání i schopnost být ultimátním zdrojem svých akcí. Dennett (2004) však ukazuje, že Kane poskytuje podporu spíše pro kompatibilistické pojetí svobody, zatímco libertariánské pojetí není udržitelné, ať už je determinismus pravdivý, či nikoli. Zdá se, že pravdu tak mohou mít tvrdí deterministé či tvrdí indeterministé. Ontologická svobodná vůle není možná_(ontologicky), ať už je determinismus pravdivý, či nikoli.

5.2 Morálně relevantní svoboda

Za užitečnější⁴³ kritérium pro posouzení existence a *míry* svobodné vůle považuji místo determinismu spíše morální odpovědnost. Tradičně se naopak bere svobodná vůle za nutnou podmínku morální odpovědnosti. Podle mne je spíše nutné diskutovat oba pojmy současně. Existuje důvod, proč lidem zároveň přisuzujeme svobodu i odpovědnost za jejich činy, ačkoli se v lidech – jak bylo argumentováno – žádná absolutní, „skutečná“ ontologická svoboda ani ultimátní odpovědnost neskrývá. Na základě čeho tedy můžeme lidi i z naturalistického, kompatibilistického pohledu považovat za svobodné i odpovědné?

Jak jsem se pokusil ukázat především v kapitole 2, je podle mne zásadně důležité uvažovat smysluplnost používaných pojmů v závislosti na úrovni popisu, na které se pohybujeme. Podobně uvažuje Dennett (1989), když rozlišuje několik různých postojů, které můžeme při popisu, interpretaci, interakci a předvídání chování libovolného systému zaujmout. Fyzický či fyzikální postoj (*physical stance*) popisuje fyzikální a chemické vlastnosti a vztahy systémů. Můžeme zde nejpřesněji analyzovat kauzalitu a tvořit predikce. Například, jak se bimetalický proužek termostatu ohne při zahřátí a jaký bude následný řetěz příčin ovlivňující teplotu vzduchu. Projektový či návrhový postoj (*design stance*) je úroveň především biologie a inženýrství. Systémům připisujeme určitou funkci, design či účel. Býložravci mají podle běžné intuice oči po stranách hlavy, *aby* spatřili predátora, a nohy pro rychlý útěk. Chápání funkce těchto orgánů stačí pro predikci i vysvětlení útěku zvířete před predátorem, aniž bychom se zabývali fyzikálními detaily fungování a evoluce daných orgánů. Díky znalosti funkce termostatu vysvětlíme stálou teplotu v místnosti i bez chápání podkladové fyziky. Naše vysvětlení a předpovědi ale nemusí být vždy úplně přesné, protože abstrahujeme od množství detailů. Intencionální postoj (*intentional stance*) je úrovní (lidové) psychologie, ale třeba i softwaru. Systémům připisujeme mentální obsahy. Například je možné předpovídat, že se budu při setkání s medvědem *bát* a budu si *přát* dostat se do bezpečí. Lze tak popsat i neživé systémy: Termostat si *přeje* udržet stabilní teplotu, takže když se *domnívá*, že je teplota příliš nízká, *snaží* se o její vyrovnaní. Při interakci s dostatečně komplexním systémem, například s šachovým programem, samozřejmě s lidmi, ale i *se sebou samým*, je intencionální postoj nepostradatelný. Ačkoli přináší problémy v podobě vágnosti pojmů a nepřesnosti vysvětlení a predikce, použití nižších postojů je kvůli obrovské komplexitě systému nepraktické, i když v principu zřejmě možné. Nadstavbou intencionálního postoje je osobní postoj (*personal*

⁴³ z hlediska praktického života ve společnosti

stance). Pomocí něj interpretujeme nějaké systémy jako morálně odpovědné osoby. Jako u všech předchozích postojů, ani zde neexistuje žádné metafyzicky hluboké kritérium. Systémy považujeme za morálně odpovědné agenty, pokud se to z hlediska interakce s nimi vyplácí (Dennett 1984).

Proč je pragmatické považovat za morálně odpovědné dospělé osoby spíše než děti, zvířata nebo termostat? Respektive, proč přisuzujeme dospělým jedincům větší *míru* odpovědnosti, než dětem? Dennett (2004) argumentuje, že společně s našimi kritérii morality se v průběhu biologické a kulturní evoluce vyvíjela schopnost komunikace, především komunikace našich důvodů (*reasons*). Dospělý, zdravý člověk je odpovědný v doslovném smyslu, že dokáže rozumně odpovědět, jaký důvod stál za jeho chováním (Anzenbacher 1994). A především, dokáže chápat cizí důvody a v jejich světle reflektovat důvody své. Což dítě svede v míře malé, termostat v žádné. Důležitou podmínkou odpovědnosti je tak dle Dennetta „*responsiveness to reasons*“ a respekt pro rozum (*reason*) samotný.

Zcela zásadní roli zde hraje výchova a vzdělání: Je potřeba podporovat děti v jejich otázkách „Proč?“, vést je k požadování a oceňování rozumných důvodů. Vzdělaný člověk je méně náchylný k manipulaci a můžeme říci, že i více odpovědný. Lidé dobrovolně a rádi přijímají více odpovědnosti, aby mohli ve společnosti požívat i většího množství práv a svobod. Proto vyspělé společnosti podporují – respektive by měly podporovat – vzdělání, kritické myšlení a výchovu k osobní odpovědnosti. Praktické, skutečné odpovědnosti, nikoli jakési zázračné, metafyzické. Právě proto dle Dennetta nejsou hrozby pro svobodnou vůli metafyzické (např. v podobě determinismu), ale sociální a politické.

Popsaný koncept svobody a odpovědnosti nám může připadat pouze jako redefinice tradičního chápání těchto pojmů, „skutečné“ svobody a „skutečné“ odpovědnosti. Ba co víc, ačkoli jde o redefinici, jasná definice, dle které by šlo veškeré entity ve všech situacích rozdělit na svobodné a nesvobodné, odpovědné a neodpovědné, není podána. Podle mne je uznání, že ani podána být nemůže, velkým pokrokem tohoto přístupu. Není žádné objektivní měřítko či kritérium odpovědnosti, což bohužel způsobuje problémy v soudnictví (5.3). Podobně jako u pojmů „dobro“ či „spravedlnost“ jde o komplexní „fuzzy“ koncepty vysoké úrovně. Na fyzikální úrovni pochopitelně neexistuje žádná svoboda ani odpovědnost; neexistují zde ani osoby. Na intencionální úrovni však existují a pouze zde má smysl se o daných pojmech bavit. Osoba (Já) v širokém pojetí je zdrojem svých důvodů a svých akcí. Nepotřebuje mít ontologickou možnost konat jinak v přesně stejných okolnostech pro praktickou odpovědnost. Má-

me rostoucí odpovědnost a svobodu díky rostoucím kompetencím konání, vzdělání, schopnosti reflexe důvodů, odolnosti vůči manipulaci, schopnosti predikce budoucích událostí včetně důsledků vlastních akcí a společenskému rozvoji. Snažme se, aby tento trend nebyl v budoucnu narušen či zvrácen, ale spíše posilován.

5.3 Implikace pro trestní právo

Přes neexistenci ontologické svobodné vůle je nesporně pragmatické přisuzovat lidem odpovědnost za jejich činy. V opačném případě by byly mezilidské interakce a fungování společnosti takřka nemožné. Nehledě na pragmatičnost, kognitivní mechanismy zodpovědné za naše morální intuice a připisující lidem odpovědnost jsou příliš hluboce zakořeněny v našich myslích (Pizarro 2011). V běžných situacích jsme poměrně snadno schopní přibližně posoudit míru odpovědnosti zúčastněných agentů. V soudnictví však potřebujeme odpovědnost určit obvykle velmi přesně, což se ve složitých případech stává velmi obtížné, ne-li nemožné, protože jednoznačná kritéria odpovědnosti prostě neexistují. Příkladem je 51letý muž, který kvůli epilepsii podstoupil operaci mozku (Devinsky et al. 2010). Následně se u něj vyvinula hypersexualita a další příznaky Kluver-Bucy syndromu. Byl posléze zatčen za stahování dětské pornografie a odsouzen k minimálnímu trestu odnětí svobody. Trest mu byl udělen, protože si zřejmě byl vědom nelegálnosti svého počínání a nedalo se vyloučit, že své činy mohl ovládat.

Rozhodnutí o míře odpovědnosti je v takových případech nesmírně složité. V současné době narůstá debata o vlivu neurověd na právní teorie a soudní praxi. Neurovědní poznatky narůstají a jsou u soudu čím dál více uznávané, což doufáme umožní lepší posouzení podobných komplikovaných případů. Čím dál více je také uznávána teze o neexistenci ontologické svobodné vůle. Toto uznání by pochopitelně nemělo mít za důsledek konec trestání. Měla by se však změnit motivace a způsoby trestání. Za ne příliš smysluplný považují retributivní aspekt trestu. Namísto toho bychom se měli zaměřit na více konsekvenčialistický přístup (Greene a Cohen 2004). V ideálním případě by měl mít trest výchovnou funkci. Odsouzený by měl svůj trest nakonec pochopit a přijmout, jako by řekl „Díky, to jsem potřeboval!“ (Dennett 2004, s. 297) Cashmore (2010) a jiní jím citovaní autoři navrhuji další zajímavou změnu soudní praxe. V první fázi procesu by byla vyloučena přítomnost psychologů a jiných odborníků, kteří běžně svědčí o svéprávnosti obviněného a podobně. Soud by pouze rozhodoval, zda obžalovaný spáchal daný čin. Až následně by panel odborníků soudu radil ohledně vhodného způsobu

nápravného trestu či léčby. Je otázkou, nakolik jsou dané návrhy v současnosti realizovatelné. Odborná debata stále živě pokračuje.

5.4 Implikace pro běžný život

Dle anekdoty vykřikla manželka biskupa z Worcesteru poté, co se dozvěděla o Darwinově teorii: „Bože, doufejme, že to není pravda, ale jestliže ano, modleme se, aby to nevešlo ve všeobecnou známost!“ Problém svobodné vůle je do jisté míry analogický problému evoluce. Teorie evoluce nezpůsobila náhlý vzestup imorálního chování. Došlo však ke změně nahlížení na nás samotné. Změna paradigmatu paradoxně umožnila fundovanější teoretické založení morálky díky rozvoji evoluční psychologie, etologie a podobně. Dá se něco podobného očekávat v případě svobodné vůle? Jak již bylo uvedeno, naše morální intuice a intencionální postoj k ostatním lidem i nám samotným se odmítnutím ontologické svobody pravděpodobně příliš nezmění. Uznání existence kompatibilistické svobodné vůle místo ontologické může dokonce přinést další pokrok v lidské moralitě, především díky zvýšení pokory a uvědomění si důležitosti vzdělání, společenských svobod a jiných důležitých faktorů, na kterých se morálně relevantní svoboda zakládá.

Podobně jako u teorie evoluce, i zde je důležitým předpokladem dostatečné porozumění problému. Proběhly experimenty, ve kterých byla snížena víra účastníků ve svobodnou vůli poté, co četli texty argumentující proti existenci svobodné vůle (Vohs a Schooler 2008), (Baumeister et al. 2009). Účastníci následně více podváděli, když měli šanci, snížila se jejich ochota pomáhat ostatním a zvýšila se jejich agresivita. Na to Dennett (2008) namítá, že čtené texty byly spíše dezinformační a navrhuje jiné znění textů pro budoucí studie. Skutečně, stejné poznatky lze vyložit dvěma způsoby. Buď můžeme prohlásit – jako to činí např. Daniel Wegner – že svobodná vůle je iluze, ale to nevádí, protože máme pořád odpovědnost. Anebo můžeme říci – po vzoru Dennetta – že svobodná vůle je skutečná, jen je to něco jiného, než jsme si představovali, ale na hodnotě našich životů a naší odpovědnosti se nic nemění. Domnívám se, že druhý způsob je vhodnější, neboť první způsob může doopravdy snížit míru naší svobody. Měli bychom se tedy zamyslet nad správnou metodou komunikace poznatků vědy a filosofie laické veřejnosti. Což není lehký úkol jednak kvůli složitosti problematiky a také kvůli dominanci negativní žurnalistiky. Titulky ve stylu „Svobodná vůle je iluze“ bohužel poutají obrovskou míru pozornosti.

6 Závěr

Problematika svobodné vůle je neobyčejně složitá a rozsáhlá. Množství literatury a nových empirických poznatků neustále narůstá. Má práce tak nemůže podat vyčerpávající výklad celého problému. Přesto si kladla za cíl podat co nejlepší přiblížení nejdůležitějších aspektů problematiky. Snažil jsem se ukázat, z čeho pramení zdánlivé kontradikce v tradičním chápání svobodné vůle. Diskutoval jsem proto různá pojetí nejdůležitějších relevantních konceptů, jako jsou svobodná akce, determinismus, náhoda, příčinnost, možnost, odpovědnost, jáství a vědomí. Častou příčinou zmatení ohledně významu daných pojmů je lpění na jejich tradičním metafyzickém a esencialistickém pojetí, zatímco tyto pojmy bylo možné vysvětlit naturalisticky a anti-esencialisticky. S tím souvisí pochopení, že většina zmíněných pojmů je „fuzzy“, neostrá, a použitelná na více úrovních popisné abstrakce (determinismus, příčinnost, náhoda, možnost) či pouze jedné specifické, konkrétně intencionální úrovni (jáství, vědomí, odpovědnost, svoboda). Dále jsem popsal nejvýznamnější vědecké experimenty ohrožující naši intuitivní představu vědomé vůle. Vysvětlil jsem, že zde nejasnosti vychází právě z chybného lidově psychologického chápání vědomí a jáství, dvou pojmů, v kterých bychom měli mít intuitivně vysokou jistotu. Přesto, jak ukazují představené experimenty, se ohledně podstaty vlastní mysli značně mýlíme.

Závěr ohledně netransparentnosti vlastní mysli a nedůvěryhodnosti introspekce přinesla i replikace experimentu, kterou jsem provedl. V původní studii účastníci často indikovali, že se vědomě rozhodli pro svoji akci, ačkoli charakter reakčních časů silně svědčil o tom, že ve skutečnosti nedokázali inhibovat reflexivní akci. Soud o záměrnosti akce byl nejspíše retrospektivně vytvořen na základě výsledku akce. Má replikace zvyšuje validitu předchozích výsledků v několika ohledech. Předchozí výsledky byly ověřeny na větším vzorku participantů. Předpokládaný efekt byl nalezen i přes zeslabení některých podmínek, které by k efektu měly přispívat. Analýza evokovaných potenciálů získaných ze skalpu jednoho z účastníků pomocí EEG dále svědčí pro platnost zkoumané hypotézy, ale na základě dat jedné osoby samozřejmě není možné vyvozovat obecné závěry. Zde je tak prostor pro potenciální další rozšíření práce. Behaviorální data všech účastníků jsou však pro potvrzení hypotézy dostačující a mohu tedy konstatovat, že replikace experimentu byla úspěšná. Uvádím i diskuzi získaných výsledků s ohledem na některé relevantní psychologické teorie.

V závěru práce argumentuji pro nemožnost existence tradičně chápané ontologické svobodné vůle. Tento pojem se jeví jako nekoherentní a odporující empirickým poznatkům. Namísto

toho obhajují kompatibilisticky chápanou morálně relevantní svobodnou vůli. Ta se u člověka vyvinula v průběhu biologické a kulturní evoluce a její vývoj stále pokračuje; je tudíž čistě přírodním fenoménem. Je závislá mimo jiné na výchově k oceňování rozumu a důvodů a na socio-politické úrovni rozvoje společnosti. Dané zjištění vkládá odpovědnost za vlastní svobodu do našich vlastních rukou. Stejně tak je nutná správná komunikace poznatků ohledně svobodné vůle veřejnosti, neboť jejich nepochopení může potenciálně přinést nežádoucí důsledky v podobě sníženého prosociálního chování jedinců.

Domnívám se proto, že cíl práce byl úspěšně dosažen. Po teoretické stránce jsem samozřejmě stál na ramenou obrů a opíral se o jejich díla, především práce filosofa Daniela Dennetta. Struktura i obsah práce a konkrétní argumentace je však ve velké míře založena na mých vlastních názorech a argumentech. Nejsem si také vědom žádné podobné práce v českém jazyce. Významným vlastním přínosem je pak především empirická studie, jejíž realizace si vyžádala mimo jiné vytvoření prezentační aplikace, provedení experimentálních úloh s relativně velkým počtem účastníků a matematické a statistické zpracování získaných dat.

Výsledky experimentu jsou užitečné pro kognitivní psychology a další odborníky zabývající se problémy agence, rozhodování, introspekce, vědomí a paměti. Pojednání o svobodné vůli je pak přínosné na jednu stranu pro humanitně orientované odborníky zabývající se filosofií, právem, ale třeba i ekonomii či sociálními vědami. Na druhou stranu ale přináší zajímavé postřehy i pro přírodní vědce, kterým umožní lépe interpretovat vlastní výsledky a porozumět společenské odpovědnosti při popularizaci svých poznatků. Konečně, domnívám se, že práce je významná i pro širší veřejnost právě proto, aby si dokázala správně vyložit očekávatelné zprávy o neexistenci svobodné vůle a současně chápala, že lidé mají skutečnou, morálně relevantní svobodnou vůli a odpovědnost.

Terminologický slovník

Termín	Význam (zdroj)
Agence	Určující vlastnost agenta. (Haggard et al. 2010)
Agent	Bytost, která koná. (Haggard et al. 2010)
Determinismus	Teze, že v každém okamžiku existuje přesně jedna fyzikálně možná budoucnost. (Van Inwagen 1983)
Elektroencefalografie (EEG)	Metoda snímání elektrického napětí z povrchu skalpu. (vlastní definice autora)
Epifomenalismus	Teze, že ačkoli jsou mentální události způsobené fyzickými událostmi, žádná mentální událost nezpůsobuje fyzickou událost. (Haggard et al. 2010)
Epistemologie	Odvětví filosofie zabývající se poznáním. (vlastní definice autora)
Evokovaný potenciál, Event-related potential (ERP)	Elektrické napětí získané elektroencefalografickou metodou ze skalpu, vyvolané jako reakce na externí či interní podnět. (vlastní definice autora)
Funkční magnetická rezonance, Functional magnetic resonance imaging (fMRI)	Metoda snímání mozkové aktivity na základě změny okysličení krve v mozku. (vlastní definice autora)
Inkompatibilismus	Teze, že existence svobodné vůle není kompatibilní s pravdivostí determinismu. (Haggard et al. 2010)
Kompatibilismus	Teze, že existence svobodné vůle je kompatibilní s pravdivostí determinismu. (Haggard et al. 2010)
Morální odpovědnost	Druh odpovědnosti potřebný pro opodstatněné přisuzování morální viny či zásluhy. (Haggard et al. 2010)
Ontologie	Odvětví filosofie zabývající se podstatou reality, <i>tím, co je</i> . (vlastní definice autora)
Prožitek či vědomí agenty, Sense of agency (SoA)	Vědomí, že já jsem ten, kdo způsobuje či produkuje danou akci. Například, vědomí, že jsem ten, kdo způsobuje pohyb něčeho, nebo že jsem ten, kdo produkuje určitou myšlenku ve svém proudu vědomí. (Gallagher 2000)
Svobodná akce (inkomp.)	Taková akce, kdy kombinace minulosti – až do okamžiku akce – a přírodních zákonů ponechává agentovi dvě či více možností, přičemž mohou být požadovány další nutné podmínky. (Haggard et al. 2010)
Svobodná akce (kompat.)	Jakákoli záměrná akce vykonaná na základě informované, racionální deliberece, přičetnou osobou a bez donucení. (Haggard et al. 2010)

Seznam literatury

- [1] AARTS, H. et al. On the inference of personal authorship: enhancing experienced agency by priming effect information. In *Consciousness and cognition*. 2005. Vol. 14, no. 3, s. 439-58.
- [2] ANZENBACHER, A. *Úvod do etiky*. Praha: Zvon, 1994. ISBN 80-7113-111-3.
- [3] BAREŠ, M. Kognitivní evokované potenciály. In *Česk Slov Neurol N* [online]. 2011. Vol. 107, no. 5, s. 508-517. Dostupné z: <http://www.csnn.eu/ceska-slovenska-neurologie-clanek/kognitivni-evokovane-potencialy-36052?confirm_rules=1>.
- [4] BAUMEISTER, R.F. et al. Prosocial benefits of feeling free: disbelief in free will increases aggression and reduces helpfulness. In *Personality and Social Psychology Bulletin*. 2009. Vol. 35, no. 2, s. 260-268.
- [5] BEM, D.J. Self-perception: An alternative interpretation of cognitive dissonance phenomena. In *Psychological Review*. 1967. Vol. 74, no. 3, s. 183-200.
- [6] BODE, S. et al. Tracking the unconscious generation of free decisions using ultra-high field fMRI. In *PloS one*. 2011. Vol. 6, no. 6, s. 1-13.
- [7] BOUCHARD, T.J. Genetic Influence on Human Psychological Traits. A Survey. In *Current Directions in Psychological Science*. 2004. Vol. 13, no. 4, s. 148-151.
- [8] BRASS, M. a HAGGARD, P. The what, when, whether model of intentional action. In *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*. 2008. Vol. 14, no. 4, s. 319-25.
- [9] CASHMORE, A.R. The Lucretian swerve: the biological basis of human behavior and the criminal justice system. In *Proceedings of the National Academy of Sciences of the United States of America*. 2010. Vol. 107, no. 10, s. 4499-504. Dostupné z: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2842067/>>.
- [10] CASTIELLO, U. et al. Temporal dissociation of motor responses and subjective awareness. In *Brain*. 1991. Vol. 114, no. 6, s. 2639-2655.
- [11] CHURCHLAND, P.M. *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind*. 2. vyd. Cambridge: The MIT Press, 1988. ISBN 0262530740.
- [12] DE JONG, R. et al. In search of the point of no return: the control of response processes. In *Journal of Experimental Psychology: Human Perception and Performance* [online]. 1990. Vol. 16, no. 1, s. 164-182.
- [13] DENNETT, D.C. *Elbow room: The varieties of free will worth wanting*. Cambridge: MIT Press, 1984. ISBN 0262540428.
- [14] DENNETT, D.C. Explaining the“ magic” of consciousness. In *Journal of Cultural and Evolutionary Psychology*. 2003. Vol. 1, no. 1, s. 7-9.
- [15] DENNETT, D.C. *Freedom evolves*. London: Penguin Books Ltd, 2004. ISBN 9780140283891.
- [16] DENNETT, D.C. Some observations on the psychology of thinking about free will. In BAER, J. et al.Eds. *Are We Free? Psychology and Free Will*. New York: Oxford University Press, 2008s. 368. ISBN 0195189639.

- [17] DENNETT, D.C. *The Intentional Stance*. Cambridge: MIT Press, 1989. ISBN 9780262540537.
- [18] DEVINSKY, J. et al. Kluver-Bucy syndrome, hypersexuality, and the law. In *Neurocase*. 2010. Vol. 16, no. 2, s. 140-5.
- [19] DOYLE, R.O. Adequate Determinism. In *The Information Philosopher* [online]. 2012 [cit. 2012-05-14]. Dostupné z: <http://www.informationphilosopher.com/freedom/adequate_determinism.html>.
- [20] EAGLEMAN, D.M. a SEJNOWSKI, T.J. Motion integration and postdiction in visual awareness. In *Science*. 2000. Vol. 287, no. 5460, s. 2036-2038.
- [21] ECCLES, J.C. *Evolution of the Brain: Creation of the Self*. London: Routledge, 1989. ISBN 3492032915.
- [22] ECCLES, J.C. The initiation of voluntary movements by the supplementary motor area. In *Archiv für Psychiatrie und Nervenkrankheiten*. 1982. Vol. 231, no. 5, s. 423-441.
- [23] EDELMAN, G.M. *Wider than the sky: The phenomenal gift of consciousness*. New Haven: Yale University Press, 2005. ISBN 9780300107616.
- [24] FELDMAN, M.W. a LALAND, K.N. Gene-culture coevolutionary theory. In *Trends in Ecology & Evolution*. 1996. Vol. 11, no. 11, s. 453-457.
- [25] GALLAGHER, S. Philosophical conceptions of the self: implications for cognitive science. In *Trends in cognitive sciences*. 2000. Vol. 4, no. 1, s. 14-21.
- [26] GRAHAM, G. a STEPHENS, G.L. Mind and mine. In GRAHAM, G. a STEPHENS, G.L. Eds. *Philosophical Psychopathology*. Cambridge: The MIT Press, 1994s. 91-109. .
- [27] GREENE, J. a COHEN, J. For the law, neuroscience changes nothing and everything. In *Philosophical Transactions of the Royal Society B Biological Sciences*. 2004. Vol. 359, no. 1451, s. 1775-1785.
- [28] HAGGARD, P. Human volition: towards a neuroscience of will. In *Nature reviews. Neuroscience*. 2008. Vol. 9, no. 12, s. 934-46.
- [29] HAGGARD, P. et al. Lexicon of key terms. In *Big Questions in Free Will* [online]. 2010 [cit. 2012-05-16]. Dostupné z: <http://www.freewillandscience.com/FW_Lexicon.pdf>.
- [30] HAGGARD, P. et al. Voluntary action and conscious awareness. In *Nature neuroscience*. 2002. Vol. 5, no. 4, s. 382-5.
- [31] HAMEROFF, S. a PENROSE, R. Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. In *Mathematics and Computers in Simulation*. 1996. Vol. 40, no. 3-4, s. 453-480.
- [32] HAWKING, S. a MLODINOW, L. *The Grand Design*. London: Bantam Books, 2011. ISBN 9780553819229.
- [33] HOEFER, C. Causal Determinism. In ZALTA, E.N. Ed. *The Stanford Encyclopedia of Philosophy* [online]. Spring2010. vyd.2010 [cit. 2012-05-20]. Dostupné z: <<http://plato.stanford.edu/entries/determinism-causal/>>.
- [34] HOLINGER, D.P. et al. Reappraisal of filter effects on P300 voltage and latency. In *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society*. 2000. Vol. 17, no. 3, s. 331-5.

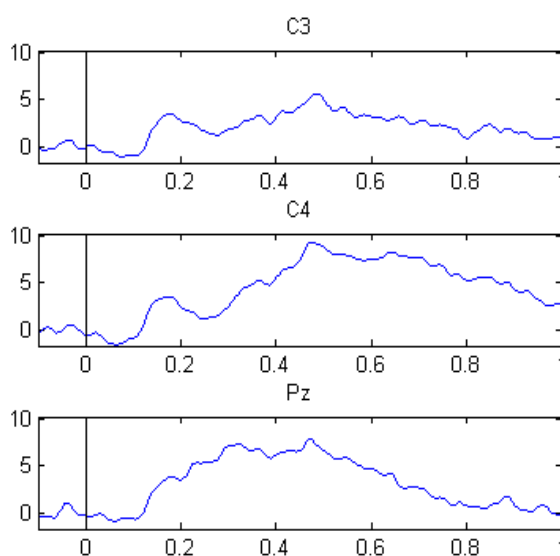
- [35] HOOD, B. *The Self Illusion: How the Social Brain Creates Identity*. [s.l.]: Oxford University Press, USA, 2012. ISBN 019989759X.
- [36] HUME, D. *A Treatise of Human Nature*. Eds. David Fate Norton and Mary J. Norton. Oxford: Oxford University Press, 1739. ISBN 0198751729.
- [37] JACOBS, J. Naturalism. In *Internet Encyclopedia of Philosophy* [online]. 2009 [cit. 2012-05-23]. Dostupné z: <<http://www.iep.utm.edu/naturali/>>.
- [38] KANE, R. *The Significance of Free Will*. New York: Oxford University Press, 1998. ISBN 0195126564.
- [39] KLEMM, W.R. Free will debates: Simple experiments are not so simple. In *Advances in cognitive psychology / University of Finance and Management in Warsaw*. 2010. Vol. 6, s. 47-65. Dostupné z: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2942748/pdf/acp-06-047.pdf>>.
- [40] KÜHN, S. a BRASS, M. Retrospective construction of the judgement of free choice. In *Consciousness and Cognition*. 2009. Vol. 18, no. 1, s. 12-21.
- [41] LIBET, B. Conscious subjective experience versus unconscious mental functions: a theory of the cerebral processes involved. In *Models of Brain Function*. [s.l.]: Cambridge University Press, 1989.
- [42] LIBET, B. Do We Have Free Will? In *Journal of Consciousness Studies*. 1999. Vol. 6, no. 8-9, s. 47-57.
- [43] LIBET, B. et al. Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). In *Brain*. 1983. Vol. 106, s. 623-642.
- [44] LOFTUS, E. a KETCHAM, K. *The myth of repressed memory: false memories and allegations of sexual abuse*. New York: St. Martin's Press, 1996. ISBN 9780312141233.
- [45] MARGOLIS, E. a LAURENCE, S. Concepts. In ZALTA, E.N.Ed. *The Stanford Encyclopedia of Philosophy* [online]. Fall 2011. vyd.2011 [cit. 2012-05-16]. Dostupné z: <<http://plato.stanford.edu/entries/concepts/>>.
- [46] MCKENNA, M. Compatibilism. In ZALTA, E.N.Ed. *The Stanford Encyclopedia of Philosophy* [online]. Winter2009. vyd.2009 [cit. 2012-05-23]. Dostupné z: <<http://plato.stanford.edu/entries/compatibilism/>>.
- [47] METCALF, K. et al. Models of confabulation: a critical review and a new framework. In *Cognitive Neuropsychology*. 2007. Vol. 24, no. 1, s. 23-47.
- [48] MONOD, J. Náhoda a nutnost: Pojednání o přírodní filosofii moderní biologie. In MARKOŠ, A.Ed. *Náhoda a nutnost: Jacques Monod v zrcadle dnešní doby*. Červený Kostelec: Pavel Merhart, 2008s. 27-170. ISBN 80-86818-66-5.
- [49] O'CONNOR, T. Free Will. In ZALTA, E.N.Ed. *The Stanford Encyclopedia of Philosophy* [online]. Summer2011. vyd.2011 [cit. 2012-05-16]. Dostupné z: <<http://plato.stanford.edu/entries/freewill/>>.
- [50] PAVLÍK, J. Dekohorentismus: 4. (a poslední) revoluce ve fyzice ve XX. století. In *E-LOGOS* [online]. 2004. s. 1-90. Dostupné z: <<http://nb.vse.cz/kfil/elogos/science/pav13-04.pdf>>.

- [51] PICTON, T. The P300 wave of the human event-related potential. In *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society*. 1992. Vol. 9, no. 4, s. 456.
- [52] PIZARRO, D. Why Neuroscience Does Not Pose a Threat to Moral Responsibility. In *AJOB Neuroscience*. 2011. Vol. 2, no. 2, s. 1-2.
- [53] PRONIN, E. et al. Everyday magical powers: the role of apparent mental causation in the overestimation of personal influence. In *Journal of personality and social psychology*. 2006. Vol. 91, no. 2, s. 218-31.
- [54] PRONIN, E. The Introspection Illusion. In *Advances in Experimental Social Psychology*. 2009. Vol. 41, s. 1–67.
- [55] SAMUELS, R. Innateness in cognitive science. In *Trends in cognitive sciences*. 2004. Vol. 8, no. 3, s. 136-41.
- [56] SATO, A. Both motor prediction and conceptual congruency between preview and action-effect contribute to explicit judgment of agency. In *Cognition*. 2009. Vol. 110, no. 1, s. 74-83.
- [57] SEARLE, J. Biological Naturalism. In SCHNEIDER, S. a VELMANS, M. Eds. *The Blackwell Companion to Consciousness*. Chichester: John Wiley and Sons Ltd, 2007s. 325-334. ISBN 1405160004.
- [58] SIEGEL, L. *Net of Magic: Wonders and Deceptions in India*. Chicago: University of Chicago Press, 1991. ISBN 9780226756875.
- [59] SOKAL, A. a BRICMONT, J. *Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science*. New York: Picador, 1999. ISBN 0312204078.
- [60] SOLTANI, M. a KNIGHT, R.T. Neural origins of the P300. In *Critical Reviews in Neurobiology*. 2000. Vol. 14, no. 3-4, s. 199-224.
- [61] SOON, C.S. et al. Unconscious determinants of free decisions in the human brain. In *Nature Neuroscience*. 2008. Vol. 11, no. 5. .
- [62] STEWART, I. *Hraje Bůh kostky? Nová matematika chaosu*. Praha: Argo, 2009. ISBN 978-80-257-0024-2.
- [63] STRAWSON, G. The impossibility of moral responsibility. In *Philosophical Studies*. 1994. Vol. 75, no. 1-2, s. 5-24.
- [64] STRAWSON, P.F. Freedom and resentment, and other essays. In *Philosophical Books*. 1975. Vol. 16, no. 1, s. 28-32.
- [65] SVOBODA, K. *Zlomky předsokratovských myslitelů*. Praha: Nakladatelství Československé akademie věd, 1962. ISBN 21-002-62.
- [66] SYNOFZIK, M. et al. Beyond the comparator model: a multifactorial two-step account of agency. In *Consciousness and cognition*. 2008. Vol. 17, no. 1, s. 219-39.
- [67] TAYLOR, J.L. a MCCLOSKEY, D.I. Triggering of preprogrammed movements as reactions to masked stimuli. In *J Neurophysiol*. 1990. Vol. 63, no. 3, s. 439-446.
- [68] TEGMARK, M. Importance of quantum decoherence in brain processes. In *Physical Review E*. 2000. Vol. 61, no. 4, s. 4194-4206.

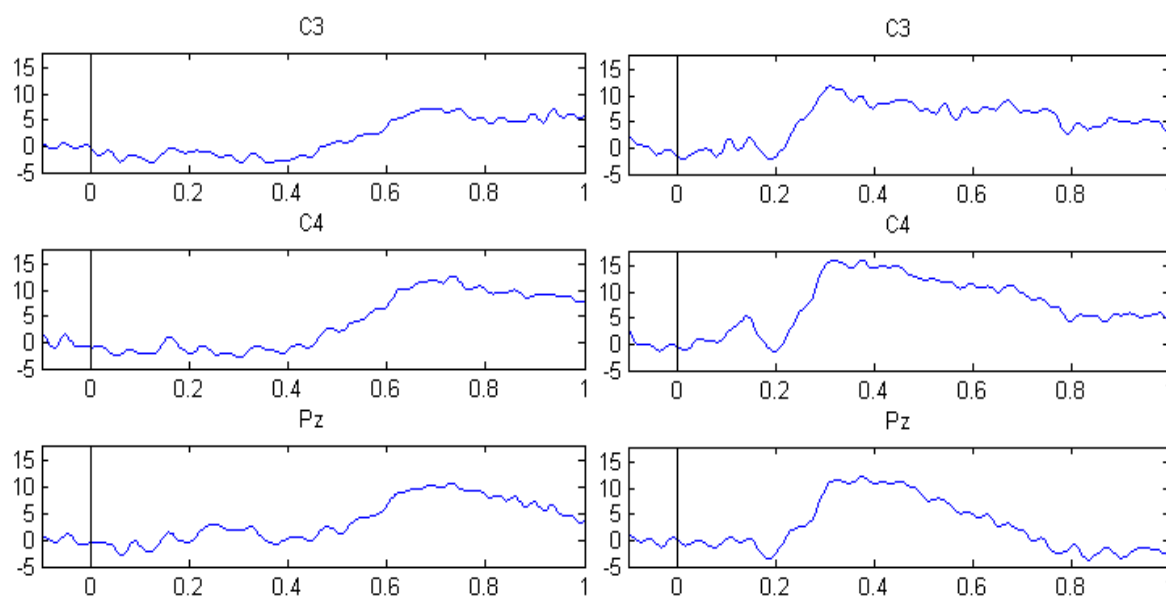
- [69] VAIDYA, A. The Epistemology of Modality. In ZALTA, E.N.Ed. *The Stanford Encyclopedia of Philosophy* [online]. Winter2011. vyd.2011 [cit. 2012-05-22]. Dostupné z: <<http://plato.stanford.edu/entries/modality-epistemology/>>.
- [70] VAN INWAGEN, P. *An essay on free will*. Oxford Oxfordshire: Clarendon Press, 1983. ISBN 0198246242.
- [71] VIHVELIN, K. Arguments for Incompatibilism. In ZALTA, E.N.Ed. *The Stanford Encyclopedia of Philosophy* [online]. Spring2011. vyd.2011 [cit. 2012-06-12]. Dostupné z: <<http://plato.stanford.edu/entries/incompatibilism-arguments/>>.
- [72] VOHS, K.D. a SCHOOLER, J.W. The value of believing in free will: encouraging a belief in determinism increases cheating. In *Psychological science*. 2008. Vol. 19, no. 1, s. 49-54.
- [73] WALLER, N.G. et al. Genetic and environmental influences on religious interests, attitudes, and values: a study of twins reared apart and together. In *Psychological Science*. 1990. Vol. 1, no. 2, s. 138-142.
- [74] WEGNER, D.M. Self is magic. In BAER, J. et al.Eds. *Are we free*. New York: Oxford University Press, 2008. s. 367. ISBN 0195189639.
- [75] WEGNER, D.M. *The Illusion of Conscious Will*. Cambridge: MIT Press, 2002. ISBN 0262232227.
- [76] WEGNER, D.M. a ERSKINE, J.A.K. Voluntary involuntariness: Thought suppression and the regulation of the experience of will. In *Consciousness and Cognition*. 2003. Vol. 12, no. 4, s. 684-694.
- [77] WILSON, T D a STONE, J.I. Limitations of self-knowledge: More on telling more than we can know. In *Review of personality and social psychology*. 1985. Vol. 6, s. 167-183.
- [78] WILSON, Timothy D a DUNN, E.W. Self-knowledge: its limits, value, and potential for improvement. In *Annual Review of Psychology*. 2004. Vol. 55, no. 1, s. 493-518.
- [79] WRENN, C.B. Naturalistic Epistemology. In *Internet Encyclopedia of Philosophy* [online]. 2005 [cit. 2012-05-23]. Dostupné z: <<http://www.iep.utm.edu/nat-epis/>>.
- [80] ZHU, J. Reclaiming volition: An alternative interpretation of Libet's experiment. In *Journal of Consciousness Studies*. 2003. Vol. 10, no. 11, s. 61-77.

Příloha A: Evokované potenciály

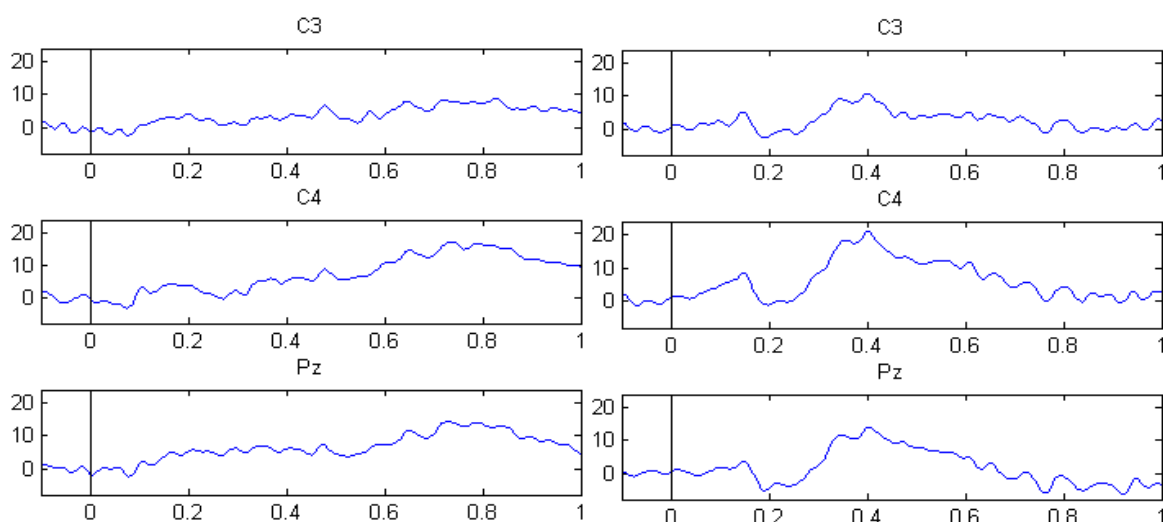
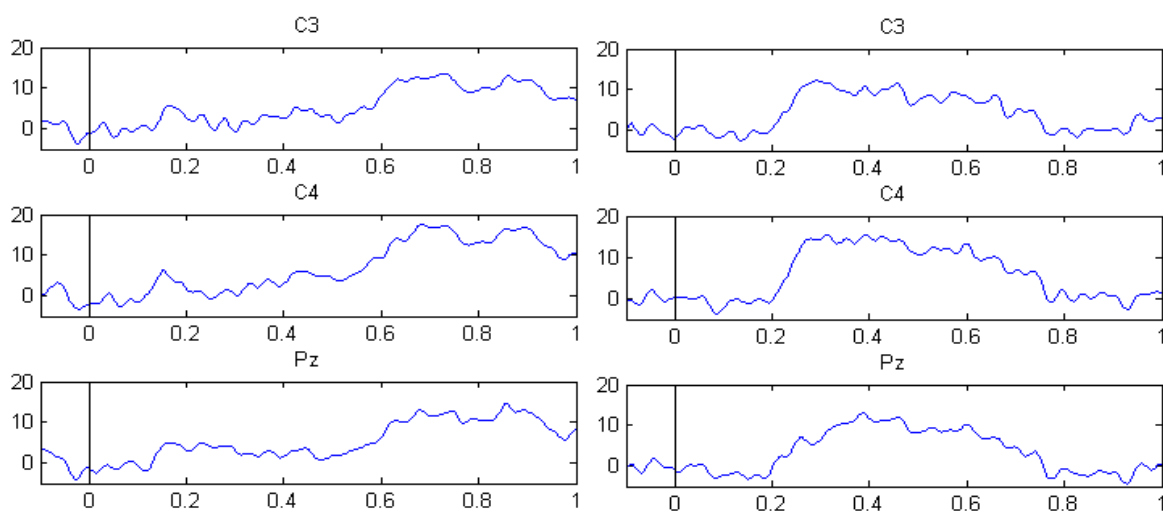
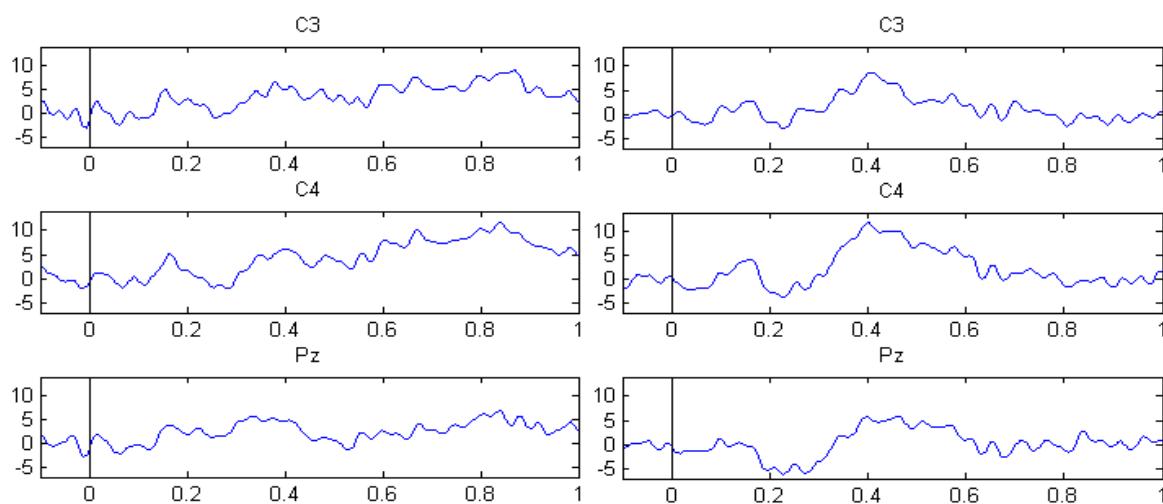
Níže jsou zobrazeny křivky evokovaných potenciálů účastníka A.B. dle jednotlivých experimentálních podmínek (typů událostí). Vlevo jsou vždy potenciály průměrované k času zobrazení primárního stimulu. Křivky na pravé straně jsou průměrované k času změny barvy stimulu, kromě typu události Primary Response (první obrázek), kde ke změně barvy nedocházelo. Na ose x je čas v sekundách, na ose y elektrický potenciál v mikrovoltech.

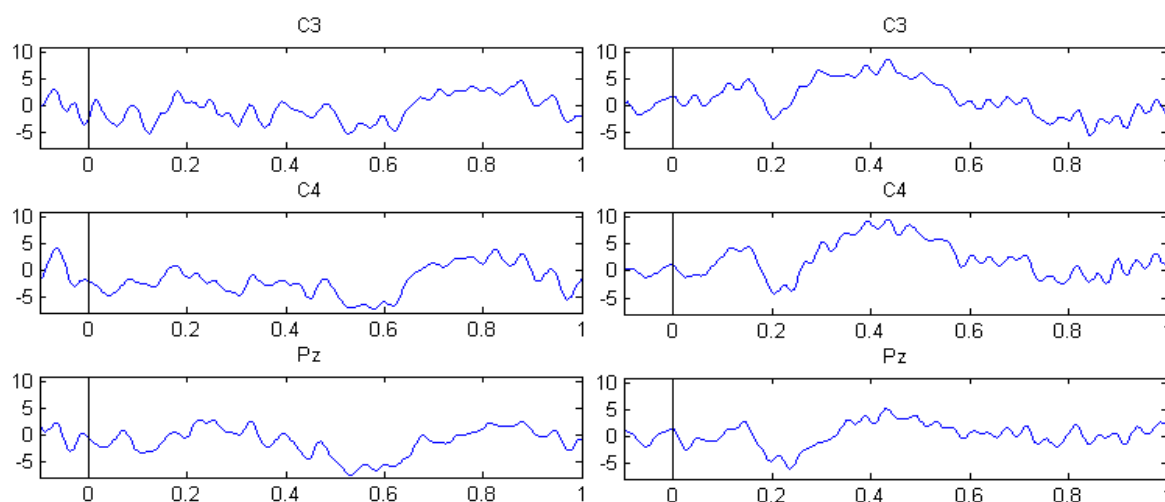


Obrázek 17: Primary Response

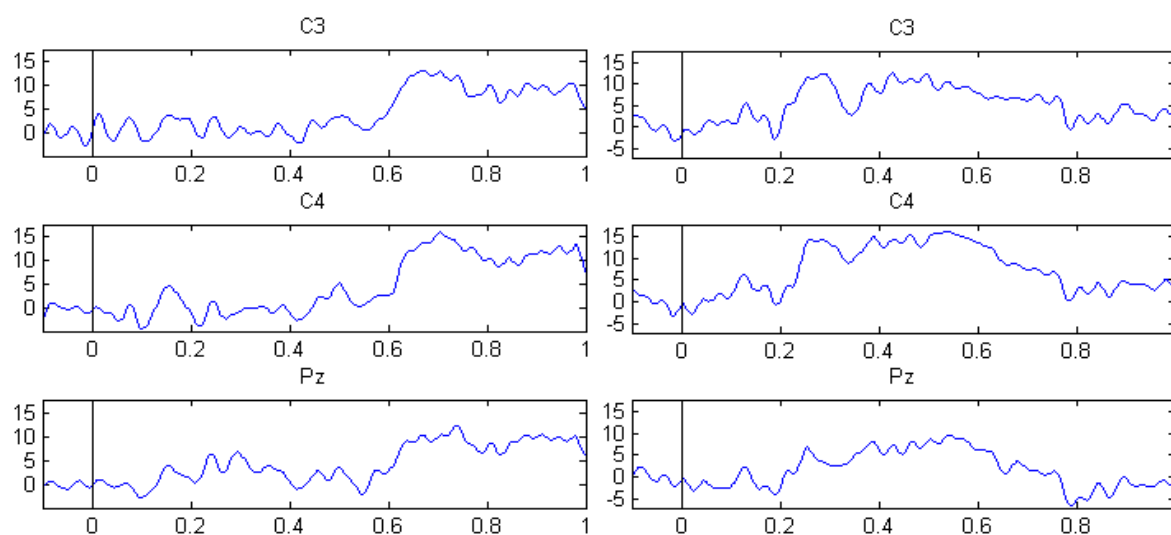


Obrázek 18: Stop

**Obrázek 19: Failed-to-stop****Obrázek 20: Decide-nogo****Obrázek 21: Failed-to-decide**



Obrázek 22: Early decide-go



Obrázek 23: Late decide-go