

Niko Šetar

Univerza v Mariboru

Prihodnost miselnih eksperimentov

Ta članek naslavlja vprašanja povezana s trenutno (sodobno) in prihodnjo rabo miselnih eksperimentov kot metode znanstvenega raziskovanja. V ta namen se najprej obrnemo na splošno veljavnost miselnih eksperimentov, kar naredimo skozi preučevanje kritik, katerih cilj je diskreditacija znanstvene in spoznavne vrednosti miselnih eksperimentov, ter s pregledom odgovorov na tovrstne kritike kot tudi drugih argumentov v prid miselnim eksperimentom, kot so recimo proti-zmote.

Po tem uvodu primerjamo miselne eksperimente z drugimi raziskovalnimi metodami, natančneje empiričnimi eksperimenti in računalniškimi modeli. Poskušali bomo pokazati, da so lahko empirični eksperimenti vsaj toliko zmotni kot miselni eksperimenti, slednji pa imajo tudi določene postopkovne lastnosti, zaradi katerih nam lahko nudijo podatke, ki z drugima dvema omenjenima metodama niso dostopni, kar jim omogoča, da ohranijo svoj status kot veljavna, neodvisna raziskovalna metoda.

Ključne besede: miselni eksperimenti, empirični eksperimenti, računalniško modeliranje, raziskovalne metode

1. Uvod

Miselni eksperiment (v nadaljevanju ME) je raziskovalna metoda, ki se na različnih področjih tako v filozofiji in znanosti uporablja, kadar je izvedba klasičnega, empiričnega eksperimenta (v nadaljevanju EE) nemogoča iz takšnih ali drugačnih razlogov. Teorija ME – sama po sebi precej sodoben pojav – se sooča z definicijo ME v, njihovo strukturo in rabo, kot tudi z odgovarjanjem na kritike, ki trdijo, da ME zastarela ali neveljavna metoda raziskovanja, ki se prej sklicuje na fantazijo kot resničnost.

V danem članku nas predvsem zanimajo te kritike. V odseku 2 si bomo tako najprej ogledali klasične ugovore na ME, od problema sklicevanja na intuicijo pri izvajanju ME, do problemov povezanih z introspekcijo in subjektivnostjo, pa do denimo sklicevanja na vsakdanji jezik. Pri vseh ugovorih bomo povzeli tudi odgovore, ki branijo veljavnost ME. V odseku 3 bomo najprej primerjali ME z EE ter poskušali vzpostaviti zakaj sta EE in ME enostavno dve različni metodi raziskovanja, pri čemer nobena nima absolutne prednosti pred drugo, marveč imata obe svoje prednosti in slabosti. Nato se bomo obrnili k novejšemu konceptu – računalniškemu modeliranju (v nadaljevanju RM). Slednje postaja v zadnjih dveh desetletjih vse bolj popularno in do neke mere začena nadomeščati tako ME kot EE. Skladno s tem se pojavljajo teorije, da bi RM utegnili v celoti nadomestiti ME, saj so natančnejši in bolj objektivni. Poskušali bomo pokazati, da imajo tako kot EE tudi RM svoje prednosti pred ME, a imajo tudi slednji svoje prednosti pred RM, zaradi česar do popolne nadomestitve nikakor ne more priti.

V odsekih 2 in 3 se osredotočamo na znanstveno orientirane ME, kot so npr. Galileova padajoča telesa, Maxwellov demon, ipd., medtem ko bomo v krajšem 4. odseku orisali tudi nadaljnjo funkcijo filozofskih (političnih in etičnih) ME in zakaj je nadomestitev le-teh z RM še bolj neverjetna kot nadomestitev znanstvenih ME. V sklepu ugotavljamo, da ME ostajajo legitimno in uporabno metodološko orodje v svoji klasični vlogi, poleg tega pa v sklopu RM pridobijo še dodatne stranske funkcije.

2. Klasični argumenti proti miselnim eksperimentom

V tem odseku bomo pregledali nekaj klasičnih argumentov proti veljavnosti in uporabnosti ME kot znanstvene metode ter jih poskusili tudi ovreči in pokazati, da ne ogrožajo ne trenutnega statusa, ne prihodnosti ME.

2.1. Problem intuicije

Pogosto se ME očita, da se pretirano zanašajo na intuicijo avtorja oziroma izvajalca. Sorensen (1992) problemu intuicije v okviru ME nadane izraz 'vudu epistemologija,' verjetno najboljše poznan primer takšnega očitka pa je Dennetova (1984) kritika Searlove (1980) Kitajske sobe, kjer ME slavno označi za 'intuicijske črpalke'. Kar Dennet dejansko trdi je, da so ME kot je Kitajska soba formulirani tako, da na dan privlečejo intuitivno smiselne, a končno nepravilne odgovore ali izide, medtem ko pomembnejši odgovori oziroma rezultati ME ostanejo spregledani, ker niso intuitivni. Drugi problem je tudi, da naj bi bile epistemološke in metafizične hipoteze ME zastavljene tako, da jih ne moremo ignorirati ali prilagoditi, kar vodi v krožnost ME, zaradi česar so slednji neveljavni (Virvidakis, 2011) Predpostavke posameznih eksperimentov so pogosto določene tako, da ne izvirajo iz nekega splošnega konsenza, marveč iz subjektivne sodbe ali intuicije avtorja ME. Modalne izjave, ki se pojavljajo kot (pogosto intuitivne) premise ME, po mnenju mnogih filozofov 'visijo v zraku', v smislu, da se njihova možnost ne utemeljuje v ničemer, razen v izjavi sami. Poleg tega se te premise, kot tudi splošni konteksti ME, pogosto navezujejo na možne svetove, ki izhajajo iz intuicij posameznika in so do te mere različni od našega, da implikacije zaključkov niso relevantne za naš svet.

Že Dennett (1984) sam trdi, da je lahko uporaba intuicije za obravnavo kompleksnih scenarijev lahko uporabna saj nam pomaga »videti gozd, ne le dreves«, problem pa je, da je intuicije zelo enostavno (nenamerno) zlorabiti.

Virvidakis (2011) postopa v smeri tega, kako se tovrstni zlorabi intuicij izogniti, oziroma kako bi utegnili potekati odgovorna raba intuicij v ME. Virvidakis izpostavlja, da Kant odgovarja na Leibnizove miselne eksperimente glede teorije prostora kot definirane s strani predmetov, ki se v njem nahajajo. Kant meni, da je narava razmišljanja o prostoru fenomenalne narave, saj gre za prostor kot ga mi izkušamo. Leibnizova redukcija prostora na sistem korelacij in interakcij med predmeti, ki bivajo v prostoru je po njegovem mnenju neupravičena, saj s tem zanemarja epistemske intuicije, ki jih imamo o prostoru brez sklicevanja na stvari, ki se v njem nahajajo. Enako velja za Descartesa (1641), ki je v Meditacijah trdil, da mu nič ni lažje dostopno kot njegov lasten um. Kant (1998) temu ugovarja, češ da Descartes ni dojel, da mu princip *Cogito ergo sum* ne nudi vpogleda v nič povezanega z njegovim umom ali jazom, razen tega, da enostavno biva (obstaja). Za kakršnekoli ugotovitve, ki bi podprle njegovo domnevo, da dejansko pozna lasten um, bi bila potrebna transcendentalna refleksija o naravi sebe in upoštevanje epistemskih pogojev, ki mu omogočajo tovrsten dostop. Ti epistemski pogoji so izpolnjeni skozi dosledno upoštevanje empiričnih spoznanj in izkušenj, ki ustrezno dopolnijo apriorna prepričanja, na katerih slonijo

premise in hipoteze, vsebovane v miselnih eksperimentih. Kantovi pogoji za ustrezno izvajanje ME slonijo na njegovi teoriji modalnosti, ki se lahko povzame kot sledi:

1. Karkoli ustreza formalnim pogojem izkustva (v skladu z intuicijami in koncepti) je možno.
2. Kar je povezano z materialnimi pogoji izkustva (ali zaznave) je dejansko.
3. Česar povezavo z dejanskim se določa v skladu s splošnimi pogoji izkustva, je (obstaja) nujno. (Virvidakis, 2011: 139)

V kolikor so ta tri pravila upoštevana pri vzpostavljanju modalne premise ME, je ta veljaven. Torej, v kolikor so tudi intuicije, uporabljene v ME, skladne s temi pravili, so rabljene odgovorno. Medtem ko je to samo (relativno dober) primer, je glavna poanta tega odgovora (kot tudi Dennettove lastne izjave) zgolj to, da je potrebno preprečiti zlorabo intuicij in intuicijskih črpalk, ter jih v raziskovanju uporabiti odgovorno.

2.2. Problem introspekcije

John Locke (v Sorensen, 1992) zagovarja introspekcijo kot metodo na podlagi samorazpoznavnosti (ang. *self-intimation*) mentalnih stanj – za vsako mentalno stanje obstaja tudi zavedanje, da imamo to mentalno stanje. Leibniz temu principu ugovarja, češ da vsako mentalno stanje zahteva mentalno stanje o zavedanju prvega, in tako naprej do neskončnega regresa, kar pomeni, da je polna introspekcija nemogoča, in je torej naš dostop do lastnih mentalnih stanj nezadosten. Sorensen (1992) izpostavlja, da temu mnenju botruje tudi sodobna psihologija, ki dokazuje, da se ljudje redkokdaj zavedajo lastnih mentalnih stanj, ki so privedla do neke odločitve ali dejanja. To vodi do vprašanja, ali je predmet ME dostopen? Srž tega problema se nanaša na to, da so scenariji, vsebovani v miselnih eksperimentih, mentalna stanja, ki so predmet istega regresa kot introspektivna mentalna stanja.

Tudi če temu ni tako, ostaja vprašanje, kako introspekcija poteka na fizičnem nivoju? ME pogosto zahtevajo nivo intuicije (npr. Brown, 2011), ki za mnoge ni sprejemljiv, ter zahtevajo fizikalistično razlago pojava, v istem smislu kot oko in čutne celice v njem razlagajo pojav vida. Vprašanje delovanja introspekcije se pojavlja tudi izven fizikalističnega dvoma – Auguste Comte (v Sorensen, 1992: 24–25) denimo trdi, da samoopazovanje zahteva razdelitev zavesti osebe na tako radikalen način, da bi šlo za dve osebi v istem umu, kar pa (razen v okviru določenih psiholoških motenj) ni mogoče.

Če je introspekcija dejansko možna, moramo odgovoriti na vprašanja o njeni naravi. Prvo izmed teh je, ali je introspekcija konstantna in zvesta. Sorensen se ponovno obrača na Comteja, ki trdi da

introspekcija vodi do različnih rezultatov pri različnih osebah – je torej intrakonstantna (tj. pri isti osebi konstantna), ne pa interkonstantna (tj. konstantna za različne osebe), le redko pa tudi zvesta, saj praviloma meri subjektovo stališče do propozicije, ne pa česarkoli o propoziciji sami. Na konstantnost ME meče luč dvoma to, da so zgodovinsko mnogi uporabljali iste miselne eksperimente za dokazovanje nasprotnih stališč. Peter Unger (1984) poudarja tudi, da so ME odvisni od družbenega položaja eksperimentatorja. Poleg tega lahko rečemo, da je introspekcija redko celovita, saj je nagnjena k izpuščanju ali dodajanju informacij, ki izhajajo iz naših implicitnih predobstoječih prepričanj. ME seveda prav tako izpustijo in dodajo informacije, ki se zdijo za ME (ne)relevantne, kar pogosto (problematično) izpustijo pa je kakršenkoli konkreten zaključek – predvsem filozofski ME so nagnjeni k temu, da so njihovi sklepi v celoti prepuščeni interpretaciji. Naš poskus odgovora na to kritiko bo precej redukcionističen. Omejili se bomo namreč na vprašanje, ali je ME sploh oblika introspekcije v pravem pomenu besede? Kljub temu, da je znotraj psihologije, kognitivne znanosti in filozofije definicij introspekcije cela kopica, lahko rečemo, da je skupna točka vsem, nekako slovarska definicija ta, da gre za postopek opazovanja in preučevanja lastnih mentalnih stanj. Kadar gradimo ME, načrtno gradimo neko mentalno stanje, za kar seveda potrebujemo zavedanje o tem mentalnem stanju. V kolikor bi potrebovali tudi zavedanje o tem zavedanju, pa zavedanje spet o onem zavedanju, bi Leibnizov argument o neskončnem regresu držal, a temu ni tako. Za izvedbo ME ne potrebujemo tovrstne 'popolne' introspekcije. Sorensenov pomislek o tem, da psihologija nakazuje, da se ljudje redko zavedajo lastnih mentalnih stanj, je v kontekstu ME prav tako odveč. Domneva, da bi lahko neka oseba izgradila ME brez da bi se zavedala mentalnih stanj, ki vznikajo v tem procesu, je absurda – katerega dela postopka se ne bi zavedala? Mar se denimo Galileo ne bi zavedal zakaj si predstavlja padajoča telesa? Ali se ne bi zavedal, da si jih predstavlja? Ali se ne bi zavedal kako naj bi potekalo njuno gibanje? Se naj ne bi zavedal rezultata? Katerakoli izmed teh možnosti bi pomenila, da ME sploh ne more obstajati. Enostavno ker ta obstaja in ker ga je Galileo lahko izvedel, lahko vemo, da se je Galileo zavedal potrebnih mentalnih stanj tekom izvajanja ME.

Tudi Comtejevo (v Sorensen, 1992) stališče o delitvi zavesti je zelo vprašljivo. Smiselno bi bilo, če bi bili ME namenjeni dejanskemu preučevanju podzvestnih procesov ali globokih mentalnih stanj subjekta, ki izvaja ME, a temu ni tako. Gre za preučevanje scenarija (ali mentalnega modela), vsebovanega na mentalnem stanju (vprašanje ali gre za eno ali več mentalnih stanj bomo na tem mestu zanemarili). Razdelitev zavesti ni potrebna, ker ne preučujemo drugih aspektov naše zavesti, ampak uporabljamo zavest za preučevanje nečesa zunanjega tako, da si to predstavljamo znotraj zavesti. Padajoče telo in fizikalne predpostavke za njim niso na noben način del naše zavesti v takšnem smislu, da bi lahko rekli da gre za preučevanje naše zavesti same.

Nazadnje naj omenim še drugi ugovor, ki ga Sorensen (1992) povzema po Comtu, in sicer ugovor iz nekonsistentnosti. Če še malo vztrajamo pri Galileovih padajočih telesih, si je ponovno težko predstavljati, da bi dve osebi, ki izvajata isti ME pod istimi epistemskimi pogoji prišli do drugačnih (subjektivnih) zaključkov. V kolikor do slednjih dejansko pride, gre za napako pri individualni izvedbi ME, bodisi zaradi predsodka, pomanjkanja znanja, ali kakšnega drugega škodljivega faktorja. A to ne pove ničesar o naravi, sestavi ali veljavnosti ME samega, marveč zgolj odseva kompetentnost izvajalca. Druga zgodba je pri ne-znanstvenih miselnih eksperimentih, kjer pa je subjektiven, nekonsistenten rezultat dejansko željen izid – problem vozička denimo ni namenjen temu, da bi vsak izvajalec prišel do istega zaključka.

2.3. Določeni drugi problemi

V tem delu v celoti povzemamo po Sorensenu (1992), ki v svoji knjigi *Thought Experiments* nudi podrobno razdelavo očitkov ME (v naslednjem odseku pa tudi odgovore na slednje).

Najprej je omeniti problem zvestobe. ME merijo med drugim tudi naše prepričanje o nečem, in ne tistega nečesar samega po sebi. Ta prepričanja lahko vodijo v to, da kadar so predobstoječa prepričanja skupna širšemu krogu izvajalcev ME, lahko ME izvajalce prepriča v svojo veljavnost in pravilnost sklepov, čeprav temu dejansko ni tako. Pojavlja se tudi problem pristranskosti tako individualnega izvajalca, kot družbenih pristranskosti, kar lahko vpliva na sklep, ki sledi iz ME. V izrazito katoliški družbi bo, denimo, rezultat ME s slavnim violinistom mnogo pogostejše v nasprotje splava, kot je bil originalni namen, saj bodo pristaši obveznosti nadaljevanja nosečnosti aktivno iskali načine izpodbijanja analogije med violinistom in zarodkom. Še ena težava je, da so nekateri protidejstveni scenariji tako oddaljeni od (tudi možne) realnosti, da bi bilo za to, da se jih sploh upravičeno obravnava, potrebno vzpostaviti širok, podroben, in morda celo nemogoč kontekst. Z roko v roki z vsem tem gre tudi, da ME pogosto poenostavijo stvari do točke absurda – Sorensen pravi, da na tak način, kot bi obliko Francije obravnavali kot popolnoma šest kotno.

Na podlagi vsega opisanega si lahko zastavimo vprašanje, ali je implikacija, da zamisljivost neke situacije pomeni njeno možnost, popolnoma zgrešena? Sorensen problem podrobno analizira – dovolj podrobno, da je njegova obravnava izven obsega tega članka – vendar pa zaključí, da je zgornja implikacija statistično upravičena; pogostejše si bo nekdo zamislil neverjeten, a možen scenarij, kot pa popolnoma nemogočega.

Pogosto se ME sklicujejo tudi na vsakdanji jezik; vendar pa obstajajo polemike tudi glede tega, ali je tovrsten sklic upravičen. Sklic na vsakdanji jezik loči med empiričnimi vzorci, ki opisujejo opazljive ponavljajoče se pojave, in jih upravičeno induktivno posplošijo, ter logičnimi

implikacijami, ki temeljijo na deduktivnem sklepanju iz definicij. Nekateri skeptiki v celoti zanikajo veljavnost sklica na vsakdanji jezik, saj trdijo, da je jezikovno znanje 'navadnega' posameznika omejeno, in je sklic na vsakdanji jezik podobno nesmiseln, kot spraševati slehernika o medicini ali astrofiziki. Sorensen jih odpiše kot radikalne, in se poglobi v zmernejši skepticizem o tem sklicu. Slednji trdi, da je vsakdanji jezik prestavljen preozko ali preširoko. Preozka je na primer definicija angleške besede *bachelor*, ki zaznamuje nekoga, ki je samski po lastni izbiri, a predpostavlja, da je alternativa, da sledi družbenemu standardu, in se poroči. Pri tem izvzema primere samskosti po lastni izbiri iz drugih razlogov, denimo duhovnikov ali pa homoseksualcev v okoljih, kjer je homoseksualnost stigmatizirana ali celo prepovedana. V okviru sklica na vsakdanji jezik pride do 'semantičnega spusta' v miselne eksperimente, pri čemer se vprašanja o stvareh reducirajo na vprašanja o tem, kako govorimo o teh stvareh. V primeru radikalnih predpostavk ME lahko o njih govorimo in jih obdelujemo oz. obravnavamo, vendar pa so brez dejanskega pomena. Obstaja tudi kritika glede informativnosti ME. Carl Hempel (1965) denimo definira, da so lahko ME teoretični, izpeljani deduktivno na podlagi konkretno definiranih principov, ali pa intuitivni. Večina jih po njegovem mnenju leži nekje na tem spektru, a se nagibajo bolj v intuitivno smer – to pomeni, da imajo takšni ME hevristično vrednost pri spoznavanju, ne pa tudi razlagalne vrednosti ali vrednosti v pridobivanju dejanskega znanja. Nadalje trdi, da je miselne eksperimente bodisi mogoče preveriti z empirično eksperimentacijo in jih potrditi, ali pa ne. V kolikor je to mogoče, je ME dejansko nepotreben, saj bi lahko v prvi vrsti izvedli empiričnega; ME tako v splošnem nimajo nikakršne znanstvene vrednosti.

Potrebno je preučiti tudi zmote, ki pritičejo slabim oziroma neveljavnim miselnim eksperimentom. Pogost vzrok neveljavnosti je sklicevanje na mit – slednjega Sorensen označuje kot lažno prepričanje, ki pa ni logična zmota. Biologi pred Darwinom so denimo zagovarjali kreacionizem zaradi korenitega verskega prepričanja in hipotezo podpirali z zapletenostjo živih bitij. ME, ki se sklicuje na mit, bo neveljaven, a Sorensen opozarja, da so tudi EEi lahko osnovani na mitu (tj. lažnem prepričanju), tako so v devetnajstem stoletju izvajale kemijske eksperimente na podlagi hipoteze, da so vsi kemijski elementi sestavljeni iz atomov vodika. ME je lahko neveljaven tudi brez zmote zaradi sklica na mitologijo. Prav tako lahko neveljavnost izhaja iz metodološke zlorabe, kjer lahko denimo filozof uporabi ME kot miselno opornico, ter z njim krožno (ali drugače zmotno) podpre obstoječo hipotezo, ali pa uporabi ME, kjer obstaja bolj zanesljiva metoda dokazovanja ali podpiranja hipoteze, recimo EE, anketna raziskava ali smiseln deduktiven sklep iz zakonov in aksiomov. Pri teh problemih odgovora ne bomo iskali, saj ga bržkone tudi ni. Dovolj je, da sprejmemo, da se ti problemi pojavljajo tudi pri EE in RM, ter niso omejeni na ME.

2.4. O zmotah in protizmotah

Zmotnih ME, kot so recimo ME, ki temeljijo na pristransko zastavljenih hipotezah, zmotno zastavljenih perspektivah, itd., na tem mestu ne bomo vključevali. Razlog za to je precej enostaven: zmoten je lahko prav vsak ME, pa tudi prav vsak EE, se pravi zmotnost ne pove ničesar o veljavnosti ME kot raziskovalne metode na splošno. Prvič jih ne morejo razveljaviti tiste zmote, ki se nanašajo tudi na empirične eksperimente, saj bi morali s tem zanikati tudi veljavnost EE. Drugič, previden raziskovalec lahko preuči, kje se utegnejo pojaviti zmote, in eksperiment osnuje tako, da do zmot ne bo prišlo, s čemer ustvari veljaven ME. Če mu to ne uspe, pa je njegov, in na podlagi tega *samo njegov*, ME neveljaven.

Po drugi strani Sorensen (1992) trdi, da je večina kritik proti ME dejansko protizmot; te so značilnosti ME ali logične strukture, ki na prvi pogled delujejo zmotno, zaradi česar jih mnogi obravnavajo kot zmote, kar pa dejansko niso. Sorensen poglavje odpre s primerom, da se protidejstvena 'kaj če' vprašanja, ki odpirajo miselne eksperimente, pogosto zavračajo na podlagi njihove hipotetične narave, pri čemer implikacija, da so hipoteze neveljavne, v najboljšem primeru temelji zgolj na krožnem argumentu, ki sloni na definiciji hipoteze kot še ne dokazane predpostavke. Seveda naj omenimo da protizmote govorijo v dobro ME in so uperjene proti neutemeljenim kritikam slednjih.

Prvi sklop protizmot so protizmote oddaljenosti, ki se v osnovi sklicujejo na to, da naj bi bila oddaljenost predpostavk, na katerih temeljijo ME, od našega sveta, zmotna. Prvi razlog za tovrstno dojemanje je trivialna interpretacija. V okviru te naj bi bile 'bizarne' predpostavke ME takšne, da so resnične, ko jih interpretiramo trivialno, a sporne, ko jih interpretiramo konkretno. Sorensen pravi, da je potrebno ločiti med trditvijo, da je »dejanje predpostavljanja p bizarno« in idejo, da je »predpostavljen p, ki je bizarna predpostavka« (Sorensen, 1992: 277). Če si predstavljamo bizaren možen svet, v katerem se dogaja nekaj, kar v našem svetu ni mogoče, nismo zagrešili nikakršne zmote – kvečjemu bi jo v primeru, če bi predpostavili, da se nekaj nemogočega dogaja v našem svetu, brez da bi ustrezno prilagodili kontekst.

Naslednja je nemogoča interpretacija. Pri njej večinoma gre za to, da si opazovalec ME napak tolmači, da so njegove predpostavke logično nemogoče, kadar so zgolj praktično nemogoče (o tem smo nekaj povedali že prej, zato na tej točki ne bomo razlagali globlje).

Tretja je interpretacija nedostopnosti, ki se pojavlja večinoma, kadar ME vsebuje izredno nenavadne elemente – pri tem se lahko zanemarja analogijska vloga nenavadnega in nedostopnega scenarija (kot denimo, če bi nedostopnost očitali eksperimentu z violinistom), lahko gre za

mnemonične figure, ki pripomorejo vizualizaciji ME, v končni fazi pa bi lahko šlo tudi za neupravičeno absurde in nedostopne scenarije – a Sorensen poudarja, da je to redek primer, absurdni elementi običajno niso bistveni za končni rezultat eksperimenta.

Sledi retorična verzija oddaljenosti, pri kateri gre v osnovi za lažno obtožbo, da ME vsebuje pristranskost do estetske vrednosti, ki smo jo opisali zgoraj, s čemer se ME razveljavlja na podlagi napačne interpretacije nenavadnih elementov kot nepotrebnih ali neupravičenih. Podoben je tudi ugovor iz nazadovanja, ki ME očita, da lahko s pretiranim poglobljanjem v opisne podrobnosti, ki niso neposredno relevantne za eksperiment, slednjega privedejo do točke, ko gre le še, kot pravi Sorensen, za 'utopičen eskapizem' brez dejanskih implikacij za karkoli v svetu.

Zadnji v tej kategoriji je ugovor iz semantičnega holizma, pri katerem naj bi šlo za jezikovno manevriranje do takšne mere, da je jezik uporabljen v miselnem eksperimentu, z namenom opisa praktično nemogočih in bizarnih scenarijev, tako oddaljen od splošnih komunikacijskih konvencij, da so izjave, ki jih vsebuje, semantično prazne. Redukcija do absurda takšnega primera bi bila, da bi ME opisali tako, da bi z njim preko jezikovnih reinvencij dokazali, da je dva plus dva dejansko pet, tako da bi besedo 'pet' pripisali vrednosti štiri. Sorensen trdi, da je to protizmot zato, ker avtor ME z jezikom ne more možnosti spremeniti v gotovost, v primeru zgornje redukcije bi to pomenilo le, da imamo svet, v katerem sta besedi 'pet' in 'štiri' zamenjani, ne pa tudi v katerem bi veljala kakršnakoli absurda propozicija, kot da ima denimo 'dva in dva' vrednost pet.

Sorensen nadaljuje še s tremi primeri protizmot, ki pa stojijo izven okvira protizmot oddaljenosti. Prva med temi izoliranimi protizmoti je nekaj, kar Sorensen poimenuje »čudno noter, čudno ven« (angl. *strangeness in, strangeness out*) (Sorensen, 1992: 284). Gre za zakoreninjeno pričakovanje, da morajo biti vzroki kvalitativno povezani z učinki, npr. preprost vzrok mora imeti preprost učinek. Nasprotniki ME tako občasno izolirajo čudno predpostavko ME, jo postavijo v (po nepotrebnem) čuden kontekst, nato pa izpeljejo neveljaven, čuden zaključek, ki sledi predpostavki večinoma po principu krožnosti.

Sledi zmota opazovalca, kjer se izvajalcu ME očita, da s tem, da predpostavlja, da je nek zaznavni čut subjekta v eksperimentu v okvari, dela eksperiment neizvedljiv, kajti če subjekt ne zaznava možnega sveta eksperimenta, ga ne more niti opazovalec. Zakaj je to protizmot in ne dejanska zmota je jasno vidno, če si ogledamo vlogo okvarjenega zaznavnega čuta v miselnem eksperimentu obrnjenega barvnega spektra.

Nazadnje Sorensen predstavi protizmoto Kabuki, poimenovano po japonskem slogu gledališča, v katerem odrska tehnika rekvizite premika medtem, ko predstava teče, kar gledalci, ki s slogom niso

seznanjeni, pogosto smatrajo za slabo izvedbo predstave. Gre za zavračanje ME na podlagi 'rekvizitnih' sredstev, ki jih ti uporabljajo – Sorensen pravi, da je takšno zavračanje enako temu, če bi študent fizike zavrnil reševanje naloge, v katerem mora izračunati kot rimskega katapultista, da zadane sovražnikovo obzidje, na podlagi tega, da potovanje po času ni mogoče, ali pa ker ni vojak.

3. Miselni eksperimenti in druge metode

Ta odsek primerja ME z drugimi metodami raziskovanja, ki prav tako temeljijo na posnemanju oziroma modeliranju naravnih pojavov. EE temeljijo na izgradnji fizičnega modela in empiričnem testiranju hipoteze, računalniški modeli (RM) pa na simulaciji naravnega pojava skozi virtualni model. Med znanstveniki na različnih področjih pogosto najdemo pristranskost v prid EE, ki naj bi bili bolj otipljivi, bolj objektivni in bolj zanesljivi, v zadnjem času pa se pojavljajo stališča, ki trdijo, da bodo RM nadomestili ME, saj omogočajo izredno podrobno modeliranje z veliko količino spremenljivk, s katerimi lahko operiramo.

3.1. Empirični in miselni eksperimenti

Da lahko začnemo preverjati, ali so EE res nadrejeni ME, moramo najprej pogledati, ali je ME sploh eksperiment, ter kakšne so podobnosti in razlike med obema kategorijama.

Sorensen (1992) trdi, da so ME eksperimenti v prvi vrsti v taksonomskem smislu, kljub navideznim razlikam med enimi in drugimi – navidezne lastnosti pa niso zadostne za določitev ME kot eksperimentov na podoben način kot fenotip kita ni zadosten za določitev kita kot sesalca, če naj njegov videz primerjamo s tipičnimi primeri kopenskih, štirinožnih, kosmatih sesalcev. Avtor ME primerja z miselnim šahom, pri katerem igralca ne posedujeta dejanske šahovnice, marveč igrata zgolj z besednimi oznakami premikov figur. Pri tem metodologija ni pomembna – ni pomembno namreč, ali obstaja neka Platonska šahovnica, na kateri se udeležujejo poteze obeh igralcev (Brown, 2011), ali pa oba igralca izvajata mentalno simulacijo igre – relevantno je samo, da igra poteka kljub temu, da ne poteka v klasični, fizični obliki. Temu je analogen tudi potek ME.

Tako pri miselnih kot empiričnih eksperimentih, pravi Sorensen, naivni induktivizem ne vzdrži. Rezultat vsakega eksperimenta so »surovi podatki, ki se uporabljajo za preverjanje teorij« (Sorensen, 1992: 233), vsakršna induktivna posplošitev rezultate enega miselnega ali empiričnega eksperimenta na (novo) teorijo pa je zmotna. Naslednji skupni lastnosti Sorensen pravi 'brkljanje' – tako kot se EE pogosto reproducirajo z majhnimi prilagoditvami v instrumentih, spremenljivkah ali postopku, da bi se izpostavile njihove pomanjkljivosti ali vključile nove hipoteze reprodukcije, se tudi v miselnih eksperimentih prilagaja kontekst, spremenljivke, postopek itd. To velja tudi za

filozofske miselne eksperimente: številni utilitaristični in proti-utilitaristični ME, od problema vozička do loterije preživetja, so 'prebrkljane' oblike enega in istega osnovnega eksperimenta, ki v osnovni obliki primerja vrednost enega proti več življenj z različnimi permutacijami. Slednje nakazuje tudi na to, da imajo ME nek standarden format, ki navadno zaobjema kontekst, specifičen problem in akterjevo dejanje v okviru tega problema – tudi EE imajo standarden format, ki obsega med drugim postavitev opreme, vpeljavo in izključevanje spremenljivk, ter sprožitev in spremljanje poteka eksperimenta. Tako EE kot ME vodijo do nepričakovanih učinkov: Louis Pasteur je denimo skozi dokazovanje, da življenje ne nastane spontano v izoliranem okolju, pomotoma dokazal tudi, da je razkroj biološki in ne kemijski postopek. Podobno je Gilbert Harman (1986) pri poskusu podrobnejše razlage Gettierjevega ugovora na teorijo znanja kot upravičenega resničnega prepričanja prišel do tihe premise, da so protidokazi, katerih se ne zavedamo, prav tako ovira pri pridobivanju znanja iz poznane evidence, ki so jo kasneje pograbili mnogi epistemologi (še več, njegov ME se v kontekstu njegovega prvotnega namena smatra kot neuspešen).

Sorensen nadalje trdi, da je večina razlik med miselnimi eksperimenti in empiričnimi eksperimenti, ki bi prve utegnili diskvalificirati, zgolj navideznih. Prva med njimi je, da se ME ne replicirajo, kot se to počne z empiričnimi eksperimenti. Sorensen pravi, da to kratko malo ni res, pri čemer ni potrebna nobena argumentacija, marveč moramo le pogledati »petindvajset let napredovanja Gettierjevih protiprimerov, petnajst let vse bolj zapletenih problemov vozička, ali zadnjih deset let primerov Zemlje dvojčice (Sorensen, 1992: 249). Prav tako ne drži, da ME ne upravljajo s spremenljivkami – zmotna razlika izhaja iz tega, da je upravljanje spremenljivk v miselnih eksperimentih apriorno, kot denimo šahisti upravljajo s spremenljivko nasprotnikove naslednje poteze. Nadalje ne drži niti, da ME niso kvantitativni. Nekateri so eksplicitno kvantitativni, denimo Fisherjev eksperiment (1930), s katerim razlaga uravnoteženo naravno razmerje med številom pripadnikov obeh splov s tem, da si predstavlja variabilna razmerja (tj. svet v katerem je znatno več moških ali znatno več žensk). Spet drugi so kvantitativni, a apriorni, ker se nanašajo na neskončne (ali neskončno majhne) količine, npr. Heronov eksperiment (v Sorensen, 1992: 250), da se predmet na površini brez trenja prične premikati, ko površino nagnemo za poljubno (neskončno) majhen kot.

Razlike, ki dejansko obstajajo med enimi in drugimi eksperimenti gredo v prid miselnim eksperimentom. Noben miselni eksperiment namreč ne zahteva več kot enega eksperimentatorja, kar preprečuje tudi spore glede hierarhije med eksperimentatorji, kot tudi noben ME ne loči med teoretikom in izvajalcem eksperimenta. Rezultati empiričnih eksperimentov so lahko izprijeni zaradi sreče (ali nesreče), če pride do naključne spremembe neodvisne spremenljivke ali česa podobnega – to pri miselnih eksperimentih ni mogoče iz popolnoma očitnih razlogov. ME tudi ne

potrebuje opreme in končno, ME so moralno trivialni. EE so lahko predmet etike, ko so za njihovo izvedbo potrebni živalski ali človeški testni predmeti.

Nadalje želimo pokazati, da je ena izmed podobnosti tudi ta, da EE enako enostavno podležejo nekaterim težavam opisanim v odseku 2.3. kot ME. Zgoraj smo povzeli, da se v ME utegne primeriti, da izvajalec izpelje nepravilne sklepe, a zaradi predobstojećih prepričanj ali gole želje po rezultatih verjame v pravilnost tega sklepa – a to še zdaleč ni omejeno na ME. Šimundić (2013) denimo opisuje 'mučenje podatkov' (ang. *torturing the data*) v medicinskih raziskavah, kar pomeni deljenje rezultatov v različne podkategorije, dokler ne postanejo statistično pomembni za svojo podkategorijo, a so v splošnem nepravilni in praktično popolnoma neuporabni. Poleg tega primera avtorica opisuje tudi druge pristranskosti v zbiranju, obdelavi in interpretaciji podatkov v EE, ki lahko vodijo v neveljavnost eksperimenta. Pojavljajo se tudi implicitne pristranskosti, ki so drugače motivirane, denimo to, da se za raziskave na živalih pogosto uporabljajo samo samci, saj so ti cenejši in pogosto manjši in manj agresivni, kar vodi v kvalitativno nepopolne rezultate, saj se lahko pogosto samice na iste spremenljivke odzivajo popolnoma drugače (Wald in Wu, 2010). Tudi poenostavitev ni omejena na ME, saj je v mnogih laboratorijskih eksperimentih, povezanih z biologijo, fiziko, in ostalimi 'trdimi' znanostmi pogosto nemogoče poustvariti pogoje, v katerih bi se nek postopek ali pojav odvijal v naravi.

3.2. Računalniško modeliranje in miselni eksperimenti

Zgoraj smo omenili, da je računalniško modeliranje (naprej RM) in izvajanje računalniških simulacij na teh modelih metoda, ki se z napredkom računalniške tehnologije vse pogostejše uporablja kot dopolnitev, pa tudi kot nadomestilo EE. Ta za zgodovino znanosti relativno nova metoda ponuja možnost poustvaritve izjemno kompleksnih naravnih pojavov na mikro in makro nivoju, od simulacije trkov delcev in zgradbe molekul, pa do modelov, ki simulirajo delovanje črnih lukenj, razvoj galaksije, itd. Ob tem omogoča tudi vnos velike količine spremenljivk in selektivno upravljanje slednjih. Pomislek, da utegnejo RM zamenjati tako ME kot tudi EE, se tako pogosto pojavlja in na prvi pogled ne izgleda zgrešen.

V kakšnem odnosu pa so ME in RM? El Skaf in Imbert (2012) opredelita ME kot način raziskovanja konceptualnega orodja in razvijanja teoretične podlage, RM pa kot načine zagotavljanja teoretičnih razlag. EE, po drugi strani, le malo pripomorejo k tema dvema funkcijama. Vendar pa avtorja poudarjata, da je osnovna funkcionalna shema vseh treh metod popolnoma enaka, kar opišeta v sledečih korakih (povzeto):

1. Osnovna dejavnost, usmerjena v odgovarjanje na vprašanja.
2. Vključitev scenarija: dejanskega ali izmišljenega, ki vsebuje zakone fizike, pravila, zunanje spremenljivke, ipd. parametre. Scenarij je vrstni opis (ang. *Type-description*), ki vključuje določeno količino podatkov o prikazani situaciji ali pojavu.
3. Razvoj scenarija: preiskava vsebine scenarija skozi analitično logiko, numerično analizo, fizični proces, ali kakšno drugo metodo. Lahko gre za računalniške (RM), mentalne (ME), fizične (EE), ali podobne procese.
4. Rezultati scenarija: relevantni, ne-poljubni podatki, ki lahko vodijo v odgovore na začetna vprašanja.
5. Znanstveni zaključki skozi interpretacijo rezultatov. (El Skaf in Imbert, 2012: 3454-3455)

Na tovrstno funkcionalno izomerijo se osredotoča tudi Arcangeli (2018), ki predvsem poudarja, da se za opredelitev razmerja med ME in RM pogosto pretirano zanašamo na problem identitete, tj. ali so RM vrsta ME, ME vrsta eksperimenta na splošno itd., kar vodi v pristransko preferenco do binarne analize in zmotno stališče da so RM nujno vrsta ME in jih bodo nadomestili kot neke vrste 'naslednja generacija'. Prednost stališča v smeri funkcionalne izomerije je tudi da, da se osredotoča na postopek prej kot na rezultate, kar preprečuje tudi tvorbo pristranske preference do EE nad ME (in RM) kot spoznavno nadrejenih.

Beisbart (2012) pa denimo išče podobnosti med ME in RM s pomočjo Nortonove (1996) argumentacijske hipoteze; tj. miselni eksperiment je logični argument, ki je zanesljiva metoda znanstvenega raziskovanja v kolikor je ta argument zdrav. Po Breisbartovi analogiji je algoritem, ki vodi simulacijo na RM, enostavno serija logičnih izjav v programskem jeziku, končni rezultat pa sklep, ki sledi iz teh izjav. Ponovno gre za vrsto funkcionalne izomerije, le da izhaja iz drugega izhodiščnega pristopa.

Nujno pa potrebujemo tudi opredelitev razlik med RM in ME; namreč če razlik ni in funkcionalna izomerija preraste v funkcionalno identiteto, potem lahko ponovno trdimo, da RM *je* oblika ME, in da lahko torej v celoti nadomesti ME kot njihova sodobnejša različica.

Arcangeli (2018) identificira osnovno strukturno razliko: avtorica trdi, da vsake vrste eksperiment poteka na dveh nivojih: proizvodnja (kar vključuje model, obravnavo oz. postopek in rezultat) in predstavitev (obdelava podatkov in interpretacija obdelanih rezultatov). Medtem, ko je ta razlika pri RM in EE jasna, saj se proizvodnja izvaja v sklopu fizičnega ali računalniškega postopka, je v ME ločevanje med nivojema mnogo bolj zapleteno, saj oboje poteka *in mente* izvajalca. Razlika je zgolj v tem, kdaj se izteče mentalna simulacija in začne obdelava iz nje pridobljenih rezultatov.

El Skaf in Imbert (2013) ilustrirata razlike med vsemi tremi obravnavanimi metodami na primerih poskusov empiričnih in računalniških poskusov Maxwellovega demona. Pred samimi izvedbami sta bila že v prvi polovici prejšnjega stoletja predlagana dva praktična modela. Smoluchowski je denimo predlagal vrata na vzmet, ki se premaknejo, kadar se pritisk nanje poveča zaradi gostote delcev, kar sproži lastno Brownovsko gibanje vratc, zaradi česar spustijo delce v 'napačno' smer. Feynman (Feynman et al., 1977) je predlagal nekoliko bolj zapleten mehanizem z zobatimi kolesom in zaskočno vzmetjo, ki sta nastavljena tako, da se sistem lažje zasuje v 'napačno' smer in prenese hladne (počasne) delce v toplo komoro.

V letu 1997 Kelly in drugi izvedejo EE, ki vključuje kompleksno sintetično molekulo, ki se v sistemu obnaša kot Feynmanov sistem z zaskočno vzmetjo. Molekulo zaradi strukture dela, ki predstavlja zaskočno vzmet, manjši delci lažje zavrtijo v smer urinega kazalca kot v nasprotno smer. Ko so molekulo postavili tako, da so jo počasnejši delci premikali v smer urinega kazalca, so res zabeležili nekaj nepričakovanih energetskih maksimumov, ko so delci prekršili drugi zakon termodinamike in prešli v komoro s hitrejšimi delci, vendar pa se je izkazalo, da povprečna razporeditev ostaja enaka kot v naravnem sistemu. Skordos in Zurek (2003) ustvarita RM Smoluchowskijevega predloga, na podlagi katerega nato izvedeta simulacijo Maxwellovega ME. Zaključek je enak kot pri opisanem EE – čeprav nekatere počasnejše molekule preidejo v drugo sobano in na mikroskopskem nivoju prekršijo drugi zakon termodinamike, je povprečna razporeditev naravna, torej na makroskopskem nivoju zakon ni prekršen. To potrjuje tudi hipotezo, ki jo predpostavljajo Maxwell in kasnejši izvajalci njegovega ME – da je zakon zgolj statistično veljaven, a je statistično veljaven absolutno, se pravi na makroskopskem nivoju ga ni mogoče prekršiti.

Chandrasekharan in drugi (2012) izpostavljajo, da nam RM in ME oboji nudijo nova znanja brez novih empiričnih podatkov, kot tudi, da oboji preverjajo protidejstva (ang. *counterfactuals*) – ME na podlagi konkretnih (mentalno modeliranih) elementov, RM pa na podlagi spremenljivk, vpeljanih v simulacijo – a beležijo pomembne razlike. Glavna izmed teh je, da so RM manj transparentni kot ME in zahtevajo več sprotne vpogleda, nadzorovanja in spreminjanja mehanizma modela in simulacije. Kljub temu pa obstajajo mnoge prednosti RM – ta deluje poljubno, a nudi nove podatke, s pomočjo katerih je mogoče izpopolniti EE. Vzemimo primer preučevanja umetno vzgojenega nevronskega omrežja *in vitro*. V podrobnosti eksperimenta se ne bomo podajali; pomembno je, da sta vzporedni izvedbi istega osnovnega eksperimenta kot EE in kot RM prikazali, kako lahko v RM preverjamo učinek vpeljave poljubnih novih spremenljivk, dokler te ne privedejo do želenega rezultata. To nam omogoči da taiste spremenljivke vpeljemo v EE, ter hitro in zanesljivo potrdimo ali ovržemo ali res privedejo do želenega rezultata. Raziskovanje v RM se začne že pri izgradnji modela, saj je slednjega mogoče dopolniti z različnimi

elementi, ki v preteklih teoretičnih ali empiričnih raziskavah vodijo k želenim rezultatom, ter s tem povečati možnost uspeha tako samega RM kot tudi morebitnega EE, ki ga RM dopolnjuje. Če gre za dodajanje poljubnih elementov, lahko model rekonstruiramo z manjkajočimi elementi in identificiramo tisti specifični element ali kombinacijo elementov, ki je privedla do želenih rezultatov.

RM lahko obravnavajo kompleksnejše naravne pojave in nudijo razlago in vpogled v podrobnosti teh pojavov do mere, do katere ME ne pridejo. Chandrasekharan et al. (2012) opisujejo relacijo med ME in RM s sklepom, da RM omogočajo eksternalizacijo kritičnega mišljenja v znanosti in širijo kognitivno kapaciteto za mentalno simulacijo tako, kot teleskopi širijo zaznavno kapaciteto za videnje. To pomeni, da RM ne nadomestijo ME, temveč jih pomembno dopolnjujejo. ME še zmeraj služijo kot podlaga za osnovno idejo pri izgradnji RM, preden so ti dejansko izgrajeni v računalniškem sistemu, ter sodelujejo pri selektivni vpeljavi spremenljivk – ME predvidi katera spremenljivka bi utegnila pripeljati do želenega rezultata, RM to teoretično preveri na podrobno simuliranem scenariju, z EE dodatno preverimo empirično. ME prav tako še zmeraj igrajo vlogo pri interpretaciji rezultatov RM in EE.

Mnogi so glede vloge ME mnogo bolj optimistični. El Skaf in Imbert (2013) denimo zaključujeta, da dejstvo, da so ME, RM in EE v primeru makroskopskih implikacij Maxwellovega demona z različno metodologijo privedli do istih rezultatov pomeni, da so vse tri metode znanstveno relevantne in funkcionalno izmenljive. Pri tem opozarjata, da funkcionalna izmenljivost ne pomeni epistemološke izmenljivosti; ker ima vsaka izmed zgornjih metod svoj postopek sklepanja in upravičevanja vsaka od njih nudi vpogled v podatke o predmetu raziskave, ki z drugima dvema metodama niso dostopni. Medtem, ko ME lahko še naprej služijo kot samostojna raziskovalna metoda je njihova dodatna uporabnost tudi njihova vloga v pripravi in interpretaciji RM in EE, podobno kot trdijo tudi Chendrasekharan et al. (2012).

Dodatna prednost ME je tudi to, da jih lahko ponovi kdorkoli, tako na nivoju proizvodnje kot na nivoju predstavitve, brez da bi zato potreboval kakršnokoli računalniško ali laboratorijsko opremo – zaradi tega dejstva samega po sebi je miselni ME nenadomestljiv (Arcangeli, 2018).

Najpomembnejša prednost ME pa je verjetno ta, ki jo opisuje Beisbart (2012). Avtor trdi, da so RM kvantitativno široki in kvalitativno ozki (podrobno obdelajo veliko količino spremenljivk in mikro-scenarijev), ME pa kvantitativno ozki in kvalitativno široki (v celoti obdelajo en splošen scenarij). Uporaba obeh metod se zelo razlikuje, ME bo namreč za določene vrste scenarijev vedno bolj smiselna metoda – »noben znanstvenik ne bi izvajal računalniške simulacije za pridobitev rezultata, ki ga lahko nemudoma izpelje v mislih« (Beisbart, 2012: 426). ME je edinstven tudi v tem, da lahko

obravnavajo tako splošne scenarije, da njegovi rezultati privedejo do konceptualnih sprememb, za kar so RM prezapleteni in premalo transparentni; težko si je denimo predstavljati, da bi neko poljubno premikanje spremenljivk v RM lahko privedlo denimo do Galileove teorije o padajočih telesih ali EPR (Einstein-Podolski-Rosen) ugovora na kopenhagensko interpretacijo kvantne fizike (Beisbart, 2012).

Na koncu tega sklopa želimo na kratko pokazati tudi, da tudi RM niso imuni na težave iz odseka 2, enostavno zato ker se tovrstna negativna kritika rada pojavlja vzlic pozitivnim upravičitvam njihove veljavnosti in uporabnosti. Poprej smo omenjali pristranskosti, ki se pojavljajo v ME in EE – tovrstne pristranskosti se pojavljajo tudi v RM. V etiki digitalizacije se sicer primarno ukvarjamo z algoritemsko diskriminacijo, a tudi v znanstvenih RM pride do slednji analognih pojavov. V okviru algoritemske diskriminacije Hacker (2017) opisuje pristranske učne podatke in neenakopravne temeljne resnice. Pri slednjih gre za parametre, vnesene v algoritem, ki izhajajo iz zmotnih posplošitev – v primeru RM bi lahko šlo denimo za vnos osnovnih parametrov, ki temeljijo na zmotnih izsledkih EE; vnesli bi lahko denimo podatke pridobljene skozi študije primarno odraslih samcev laboratorijskih podgan (na primeru EE smo to problematiko omenili zgoraj), medtem ko bi bil naš RM usmerjen v simulacijo nekega pojava na celotni podganji populaciji (tj. tudi samic in mladičev), kar bi lahko pripeljalo v nepopolne rezultate. Pristranski učni podatki, po drugi strani, so podatki, ki jih uporabljamo za učenje programov, ki izvajajo RM. V kolikor so učni podatki pristranski, bo potencialno vsak RM, ki ga izvaja algoritem s temi učnimi podatki, vodil v neveljavne rezultate. Prav tako lahko selektivno vpeljevanje spremenljivk, ki vodijo v želene rezultate, vodi v podobne pojave, kot je prej opisano 'mučenje podatkov' v EE; tj. primeri se lahko, da v RM vpeljemo poljubno kombinacijo spremenljivk, ki vodijo v rezultat, podoben želenemu, ki ga nato interpretiramo tako, da ustreza naši hipotezi.

4. Ne-znanstveni miselni eksperimenti

Nobena izmed podobnosti z EE, ki jih povzemamo po Sorensenu v odseku 2, se ne nanaša izključno na znanstvene miselne eksperimente, marveč tudi na filozofske. Pravzaprav sta obe obliki neodtujljivo povezani, kar lepo ilustrira Shoemakerjev primer (v Sorensen, 1992): Shoemaker (v Sorensen, 1992: 11–12) je v prvi vrsti študiral nevrofiziologijo ljudi, ki so preživeli izgubo ene možganske poloble, v okviru česar je izvajal tako empirične kot (znanstvene) miselne eksperimente, vendar pa je njegov najznamenitejši ME ta, v katerem se vpraša, kaj bi se zgodilo, ko bi gospodu Smithu izrezali desno polovico možganov, in jo presadili v novo telo. Shoemakerjeve pretekle raziskave so pokazale, da bi lahko oba preživela relativno normalno, a vprašanje je, kdo izmed njiju bi bil gospod Smith? Gre za vprašanje identitete, ki je eksplicitno filozofske narave, a je tesno povezano z oziroma celo izpeljano iz znanstvenega ME. Kljub temu pa nas v tem odseku prej

zanima status ME, ki so eksplicitno filozofski, kot sta denimo analogija z violinistom in tančica nevednosti.

Filozofski ME niso nič bolj občutljivi na probleme iz odseka 2.3 kot znanstveni ME, saj uporabljajo isto metodologijo, vsebujejo le drugačne protidejstvene scenarije (Sorensen, 1992; Brown, 2011). Pravzaprav so na nekatere izmed teh problemov manj občutljivi: v mnogih filozofskih ME je denimo subjektivna introspekcija želena, ne pa problematična. Rawlsova tančica nevednosti zahteva preučitev lastnega, subjektivnega stališča do možnih izidov; prav tako denimo analogija z violinistom zahteva preučitev subjektivnega stališča z introspekcijo. Posplošitev in objektivni zaključki sledijo šele na nivoju interpretacije rezultatov.

Vprašanje, ki ostaja je: kakšen je odnos med ME in EE ter RM. Jasno razvidno je, da so mnogi politični in etični ME neizvedljivi v empirični obliki. Tančica nevednosti je denimo praktično neizvedljiva, analogija z violinistom pa bi lahko bila izvedljiva, a bi bila izvedba tovrstnega eksperimenta neetična. Kar pa se tiče ideje filozofskih RM pa je zadeva nekoliko kompleksnejša. Shulzke (2014) vidi distopične video igre kot vrsto političnih RM. Avtor trdi, da lahko skozi video igre odlično simuliramo korupcijo, kriminal, 'zle' institucije, zmotne in zatiralske ideologije ter režime, itd. Dinamika in razsežnost sveta v video igrah omogočata, da lahko simuliramo vsak aspekt družbe z nekim distopičnim elementom; npr. kako tiranija vpliva na življenje v višjih političnih krogih, na podeželju, v *slumih* mest, itd. Vse to omogoča igralcu, da kritično razmisli o svetu, vsebovanem v video igri, in se do njega opredeli.

Vendar pa video igra ni dejanski RM političnega ali etičnega scenarija. Funkcionalna shema igre je drugačna od funkcionalne sheme RM. V RM vnašamo spremenljivke in preverjamo, katere vodijo do želenega rezultata. V video igrah so edine sproti (v razvoju scenarija) vnesene spremenljivke dejanja igralca, ki pa vplivajo na izid igre na enoznačen, vnaprej določen način – izid video igre ne more voditi k novemu znanju. Poleg tega leži težava tudi v tem, da je primarni namen video igre zabava, torej je popolnoma legitimno, da je želeni izid nekega posameznega igralca popolna distopija, tiranija itd. Analogno s tem bi bilo, da v znanstveni RM, ki išče neko sestavino v zdravlilu za raka, vnašamo spremenljivke, za katere upamo, da bodo pospešile razvoj raka.

Prav zaradi tega, ker večina filozofskih ME kliče po introspekciji in subjektivni interpretaciji scenarija, je te scenarije najverjetneje nemogoče simulirati s pomočjo RM, saj ti preveč eksternalizirajo dan scenarij. Poleg tega je težko predstavljivo, kako bi se denimo etični RM izvajal samostojno, kot se znanstveni RM; kaj natanko bi bil učinek simulacije, v kateri voziček Philippe Foot znova in znova povozi enega človeka? Kako bi to doprineslo kakršnokoli novo znanje? Zaradi

njihove odvisnosti od človeškega faktorja introspekcije in subjektivne interpretacije brez težav trdimo, da so etični ME metodološko omejeni na ME.

5. Sklep

V tem prispevku smo pokazali, da raba intuicij ni nujno metodološko slaba in zmotna, čeprav obstajajo mnoga tovrstna stališča. Preprečiti je treba zgolj zlorabo intuicij, kar je dosegljivo z zastavitvijo določenih pogojev za izvajanje ME in raba intuicij – primeroma smo izpostavili Virvidakisov (2011) kantovski pristop, vendar pa ta brez dvoma ni edini, ki lahko pomaga preprečiti intuicijske zmote pri izvedbi ME.

Nadalje smo trdili, da so ugovori proti ME, ki temeljijo na problematiki introspekcije sami po sebi neupravičeni, saj se opirajo na preozko definicijo introspekcije in (zmotno) predpostavljajo, da introspekcija nujno pomeni subjektivnost.

Izpostavili smo nekaj drugih problemov, ki se pogosto očitajo ME, vendar ugotavljamo, da se ti problemi ali njim analogni problemi prav tako enostavno očitajo EE, saj gre za inherentne probleme, ki izvirajo iz človeškega faktorja pri vsaki raziskovalni metodi, in jih je potrebno odpraviti na nivoju raziskovalca, ne na nivoju metode.

Poleg tega smo po Sorensenu (1992) povzeli, da je marsikatera zmota, ki se očita ME, v bistvu proti-zmota oziroma psevdo zmota, se pravi gre za nek veljaven element ME, ki pa se smatra kot zmoten zaradi njegove abstraktne oz. na videz neznanstvene narave, oziroma zaradi pristranskosti kritika proti ME.

Preučili smo relacije med ME in drugimi metodami ter ugotovili, da se EE v veliki meri soočajo z enakimi problemi kot ME. Pri tem ne zanikamo, da obstajajo nekatere težave, ki so specifične zgolj za metodo ME, a na drugi strani obstajajo tudi določene težave, ki so specifične samo za EE. Tako lahko trdimo, da gre za enakovredni metodi, vsaka izmed katerih ima svoje prednosti in slabosti. Pri izvajanju obojih je za pridobitev relevantnih in veljavnih rezultatov potrebna ista mera previdnosti in odgovornosti.

Podobno velja za ME in RM: slednji so bolj primerni za nekatere podrobne, ozko usmerjene znanstvene raziskave, medtem ko so prvi boljše metoda za širše, konceptualno usmerjeno raziskovanje. Prav tako ME igrajo pomembno vlogo pri izgradnji in interpretaciji RM. Poudarek si zasluži tudi ugotovitev, da so RM primerni samo za znanstvene eksperimente, ne pa tudi za politične ali etične, ki ostajajo strogo v domeni ME.

The Future of Thought Experiments

The paper addresses the questions of current (contemporary) and future use of thought experiments as a method of scientific research. To do so we must first address the overall validity of thought experiments, which is done by examining criticisms that aim to discredit thought experiments' scientific and epistemic value and overweighing responses to those criticisms as well as other defences in favour of thought experiments, such as the notion of anti-fallacies.

After doing so, we compare thought experiments to other research methods, namely empirical experiments and computer models. We will argue that empirical experiments are faulty to at least the same degree as thought experiments, while the latter have certain procedural properties that allow them to provide data that is inaccessible by the other two methods and therefore retain their status as a valid, independent research method.

Keywords: thought experiments, empirical experiments, computer modelling, research methods

Literatura

- Arcangeli, M. (2018). »The Hidden Links between Real, Thought and Numerical Experiments« *Croatian Journal of Philosophy*, 18(52), str. 3–22.
- Beisbart, C. (2012). »How can computer simulations produce new knowledge?« *European Journal for Philosophy of Science*, 2, str. 395–434.
- Brown, J. R. (2011). *The Laboratory of the Mind*. New York: Routledge.
- Chandrasekharan, S., Nersessian, N. in Subramanian, V. (2012). »Is This the End of Thought Experiments in Science«. V Frappier, M., Mynell, L. in Brown, J. R. (urđ.): *Thought Experiments in Science, Philosophy, and the Arts*. London: Routledge, str. 239–260.
- Dennett, D. (1984). *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press.
- Descartes, R. (1641). »Meditations on First Philosophy«. *Internet Encyclopedia of Philosophy* (16. junij 2021). URL = <https://yale.learningu.org>.
- El Skaf, R., & Imbert, C. (2012). »Unfolding in the empirical sciences: experiments, thought experiments and computer simulations«. *Synthese*, 190(16), str. 3451–3474.
- Feynman, R. P., Leighton, R. B. in Sands, M. (1977). *The Feynman Lectures on Physics, volume 1*. Boston: Addison-Wesley Publishing.
- Fisher, R. A. (1930). »Sexual Reproduction and Sexual Selection & Natural selection and the sex-ratio«. V Fisher, R. A.: *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.
- Harman, G. (1986). »Moral Explanations of Natural Facts - Can Moral Claims Be Tested Against Moral Reality?« *Southern Journal of Philosophy*, 24 (priloga.).
- Hempel, C. (1965). *Aspects of Scientific Explanation*. New York: Free Press.
- Kant, I. (1998). *Critique of Pure Reason*. Cambridge: Cambridge University Press.
- Kelly, T. R., Tellitu, I. in Sestelo, J. P. (1997). »In search of molecular ratchets« *Angewandte Chemie International Edition in English*, 36 (17), str. 1866–1868.
- Norton, J. D. (1996). »Are thought experiments just what you thought?« *Canadian Journal of Philosophy*, 26, str. 233–366.
- Shulzke, M. (2014). »The Critical Power of Virtual Dystopias« *Games and Culture*, str. 1–20.

- Skordos, P. A. in Zurek, W. H. (2003). »Maxwell's demon, rectifiers, and the second law: Computer simulation of Smoluchowski's trapdoor«. *American Journal of Physics*, 60(10), str. 876–882.
- Sorensen, R. A. (1992). *Thought Experiments*. Oxford: Oxford University Press.
- Šimundić, A. (2013.) »Bias in research«. *Biochemia Medica*, 23 (1), str. 12–15.
- Unger, P. (1984). *Philosophical Relativity*. Minneapolis: University of Minnesota Press.
- Virvidakis, S. (2011). »On Kant's Critique of Thought Experiments in Early Modern Philosophy«. v Ierodiakonou, K. in Roux, S. (urd.): *Thought Experiments in Methodological and Historical Contexts*. Leiden: Brill, str. 127–144.
- Wald, C. in Wu, C. (2010). »Of Mice and Women: The Bias in Animal Models«. *Science*, 327(5973), str. 1571–1572.