Zbornik 21. mednarodne multikonference

INFORMACIJSKA DRUŽBA - IS 2018
Zvezek B

Proceedings of the 21st International Multiconference

INFORMATION SOCIETY - IS 2018
Volume B

Kognitivna znanost

Cognitive Science

Uredila / Edited by
Toma Strle, Olga Markič

http://is.ijs.si

8.–12. oktober 2018 / 8–12 October 2018
Ljubljana, Slovenia

Zbornik 21. mednarodne multikonference

INFORMACIJSKA DRUŽBA – IS 2018

Zvezek B

Proceedings of the 21st International Multiconference

INFORMATION SOCIETY – IS 2018

Volume B

Kognitivna znanost
Cognitive Science

Uredila / Edited by

Toma Strle, Olga Markič

8.–12. oktober 2018 / 8–12 October 2018
Ljubljana, Slovenia

PREDGOVOR MULTIKONFERENCI
INFORMACIJSKA DRUŽBA 2018

Multikonferenca Informacijska družba (http://is.ijs.si) je z enaindvajseto zaporedno prireditvijo osrednji srednjeevropski dogodek na področju informacijske družbe, računalništva in informatike. Letošnja prireditev se ponovno odvija na več lokacijah, osrednji dogodki pa so na Institutu »Jožef Stefan«.

Informacijska družba, znanje in umetna inteligenca so še naprej nosilni koncepti človeške civilizacije. Se bo neverjetna rast nadaljevala in nas ponesla v novo civilizacijsko obdobje ali pa se bo rast upočasnila in začela stagnirati? Bosta IKT in zlasti umetna inteligenca omogočila nadaljnji razcvet civilizacije ali pa bodo demografske, družbene, medčloveške in okoljske težave povzročile zadušitev rasti? Čedalje več pokazateljev kaže v oba ekstrema – da prehajamo v naslednje civilizacijsko obdobje, hkrati pa so notranji in zunanji konflikti sodobne družbe čedalje težje obvladljivi.

Letos smo v multikonferenco povezali 11 odličnih neodvisnih konferenc. Predstavljenih bo 215 predstavitev, povzetkov in referatov v okviru samostojnih konferenc in delavnic. Prireditev bodo spremljale okrogle mize in razprave ter posebni dogodki, kot je svečana podelitev nagrad. Izbrani prispevki bodo izšli tudi v posebni številki revije Informatica, ki se ponaša z 42-letno tradicijo odlične znanstvene revije.

Multikonferenco Informacijska družba 2018 sestavljajo naslednje samostojne konference:

- Slovenska konferenca o umetni inteligenci
- Kognitivna znanost
- Odkrivanje znanja in podatkovna skladišča – SiKDD
- Mednarodna konferenca o visokozmogljivi optimizaciji v industriji, HPOI
- Delavnica AS-IT-IC
- Soočanje z demografskimi izzivi
- Sodelovanje, programska oprema in storitve v informacijski družbi
- Delavnica za elektronsko in mobilno zdravje ter pametna mesta
- Vzgoja in izobraževanje v informacijski družbi
- 5. študentska računalniška konferenca
- Mednarodna konferenca o prenosu tehnologij (ITTC)

Soorganizatorji in podporniki konference so različne raziskovalne institucije in združenja, med njimi tudi ACM Slovenija, Slovensko društvo za umetno inteligenco (SLAIS), Slovensko društvo za kognitivne znanosti (DKZ) in druga slovenska nacionalna akademija, Inženirska akademija Slovenije (IAS). V imenu organizatorjev konference se zahvaljujemo združenjem in institucijam, še posebej pa udeležencem za njihove dragocene prispevke in priložnost, da z nami delijo svoje izkušnje o informacijski družbi. Zahvaljujemo se tudi recenzentom za njihovo pomoč pri recenziranju.

V letu 2018 bomo šestič podelili nagrado za življenjske dosežke v čast Donalda Michieja in Alana Turinga. Nagrado Michie-Turing za izjemen življenjski prispevek k razvoju in promociji informacijske družbe bo prejel prof. dr. Saša Divjak. Priznanje za dosežek leta bo pripadlo doc. dr. Marinki Žitnik. Že sedmič podeljujemo nagradi »informacijska limona« in »informacijska jagoda« za najbolj (ne)uspešne poteze v zvezi z informacijsko družbo. Limono letos prejme padanje državnih sredstev za raziskovalno dejavnost, jagodo pa Yaskawina tovarna robotov v Kočevju. Čestitke nagrajencem!

Mojca Ciglarič, predsednik programskega odbora

Matjaž Gams, predsednik organizacijskega odbora

FOREWORD - INFORMATION SOCIETY 2018

In its 21st year, the Information Society Multiconference (http://is.ijs.si) remains one of the leading conferences in Central Europe devoted to information society, computer science and informatics. In 2018, it is organized at various locations, with the main events taking place at the Jožef Stefan Institute.

Information society, knowledge and artificial intelligence continue to represent the central pillars of human civilization. Will the pace of progress of information society, knowledge and artificial intelligence continue, thus enabling unseen progress of human civilization, or will the progress stall and even stagnate? Will ICT and AI continue to foster human progress, or will the growth of human, demographic, social and environmental problems stall global progress? Both extremes seem to be playing out to a certain degree – we seem to be transitioning into the next civilization period, while the internal and external conflicts of the contemporary society seem to be on the rise.

The Multiconference runs in parallel sessions with 215 presentations of scientific papers at eleven conferences, many round tables, workshops and award ceremonies. Selected papers will be published in the Informatica journal, which boasts of its 42-year tradition of excellent research publishing.

The Information Society 2018 Multiconference consists of the following conferences:

- Slovenian Conference on Artificial Intelligence
- Cognitive Science
- Data Mining and Data Warehouses - SiKDD
- International Conference on High-Performance Optimization in Industry, HPOI
- AS-IT-IC Workshop
- Facing demographic challenges
- Collaboration, Software and Services in Information Society
- Workshop Electronic and Mobile Health and Smart Cities
- Education in Information Society
- 5th Student Computer Science Research Conference
- International Technology Transfer Conference (ITTC)

The Multiconference is co-organized and supported by several major research institutions and societies, among them ACM Slovenia, i.e. the Slovenian chapter of the ACM, Slovenian Artificial Intelligence Society (SLAIS), Slovenian Society for Cognitive Sciences (DKZ) and the second national engineering academy, the Slovenian Engineering Academy (IAS). On behalf of the conference organizers, we thank all the societies and institutions, and particularly all the participants for their valuable contribution and their interest in this event, and the reviewers for their thorough reviews.

For the sixth year, the award for life-long outstanding contributions will be presented in memory of Donald Michie and Alan Turing. The Michie-Turing award will be given to Prof. Saša Divjak for his life-long outstanding contribution to the development and promotion of information society in our country. In addition, an award for current achievements will be given to Assist. Prof. Marinka Žitnik. The information lemon goes to decreased national funding of research. The information strawberry is awarded to the Yaskawa robot factory in Kočevje. Congratulations!

Mojca Ciglarič, Programme Committee Chair

Matjaž Gams, Organizing Committee Chair

KONFERENČNI ODBORI
CONFERENCE COMMITTEES

International Programme Committee

Vladimir Bajic, South Africa
Heiner Benking, Germany
Se Woo Cheon, South Korea
Howie Firth, UK
Olga Fomichova, Russia
Vladimir Fomichov, Russia
Vesna Hljuz Dobric, Croatia
Alfred Inselberg, Israel
Jay Liebowitz, USA
Huan Liu, Singapore
Henz Martin, Germany
Marcin Paprzycki, USA
Karl Pribram, USA
Claude Sammut, Australia
Jiri Wiedermann, Czech Republic
Xindong Wu, USA
Yiming Ye, USA
Ning Zhong, USA
Wray Buntine, Australia
Bezalel Gavish, USA
Gal A. Kaminka, Israel
Mike Bain, Australia
Michela Milano, Italy
Derong Liu, USA
Toby Walsh, Australia

Organizing Committee

Matjaž Gams, chair
Mitja Luštrek
Lana Zemljak
Vesna Koricki
Mitja Lasič
Blaž Mahnič
Jani Bizjak
Tine Kolenik

Programme Committee

Franc Solina, co-chair	Matjaž Gams	Vladislav Rajkovič
Viljan Mahnič, co-chair	Marko Grobelnik	Grega Repovš
Cene Bavec, co-chair	Nikola Guid	Ivan Rozman
Tomaž Kalin, co-chair	Marjan Heričko	Niko Schlamberger
Jozsef Györkös, co-chair	Borka Jerman Blažič Džonova	Stanko Strmčnik
Tadej Bajd	Gorazd Kandus	Jurij Šilc
Jaroslav Berce	Urban Kordeš	Jurij Tasič
Mojca Bernik	Marjan Krisper	Denis Trček
Marko Bohanec	Andrej Kuščer	Andrej Ule
Ivan Bratko	Jadran Lenarčič	Tanja Urbančič
Andrej Brodnik	Borut Likar	Boštjan Vilfan
Dušan Caf	Mitja Luštrek	Baldomir Zajc
Saša Divjak	Janez Malačič	Blaž Zupan
Tomaž Erjavec	Olga Markič	Boris Žemva
Bogdan Filipič	Dunja Mladenič	Leon Žlajpah
Andrej Gams	Franc Novak	

iii

KAZALO / TABLE OF CONTENTS

Zbornik 21. mednarodne multikonference

INFORMACIJSKA DRUŽBA – IS 2018

Zvezek B

Proceedings of the 21st International Multiconference

INFORMATION SOCIETY – IS 2018

Volume B

Kognitivna znanost
Cognitive Science

Uredila / Edited by

Toma Strle, Olga Markič

http://is.ijs.si

11. oktober 2018 / 11 October 2018
Ljubljana, Slovenia

PREDGOVOR / FOREWORD

Na letošnji konferenci Kognitivna znanost sodelujejo avtorji/ice z različnih disciplinarnih področij, ki predstavljajo tako empirične rezultate svojih raziskav kot tudi teoretska raziskovanja. Ena izmed osrednjih tem letošnje konference je "Družbene, filozofske in etične implikacije raziskovanja v kognitivni znanosti", avtorji pa se dotikajo tudi drugih področji kognitivne znanosti.

Upamo, da bo letošnja disciplinarno in metodološko bogata kognitivna konferenca odprla prostor za izmenjavo zanimivih misli in idej ter povezala znanstvenike/ice različnih disciplin, ki se ukvarjajo z vprašanji kognitivnih procesov.

Olga Markič
Toma Strle

*

Authors at this year's Cognitive Science conference come from numerous disciplines and present their empirical as well as theoretical work. One of this year's main topics is "Social, philosophical and ethical implications of cognitive science" but authors also present research form other areas in cognitive science.

We hope that this year's cognitive conference, being extremely diverse in disciplines and methodologies, will become a welcoming space for exchanging intriguing ideas and thoughts as well as for bringing together scientists from all the different areas exploring the questions of cognitive processes.

Olga Markič
Toma Strle

PROGRAMSKI ODBOR / PROGRAMME COMMITTEE

Olga Markič, Univerza v Ljubljani, Filozofska fakulteta

Toma Strle, Univerza v Ljubljani, Pedagoška fakulteta, Center za kognitivno znanost

Urban Kordeš, Univerza v Ljubljani, Pedagoška fakulteta, Center za kognitivno znanost

Matjaž Gams, Odsek za inteligentne sisteme, Institut »Jožef Stefan«, Ljubljana

Razumevanje odraslih partnerskih odnosov v luči spolne zlorabe v otroštvu

asist. dr. Sara Jerebic, spec. ZDT
Teološka fakulteta
Katedra za zakonsko in družinsko terapijo ter psihologijo in sociologijo religije
Poljanska cesta 4, 1000 Ljubljana, Slovenija
Tel: +386 1 4345810; fax: +386 1 4345854
Družinski inštitut Bližina
Vodnikova ulica 11, 3000 Celje, Slovenija
Tel: +386 3 4925580
e-mail: sara.jerebic@teof.uni-lj.si; sara.jerebic@blizina.si

POVZETEK[1]

Spolna zloraba v otroštvu je travmatična izkušnja s številnimi daljnosežnimi posledicami. Dokler travma ni razrešena, se pojavlja v mislih, čustvih in vedenju. Pojavi se takrat, ko je prebujen spomin iz preteklosti, kar se pogosto zgodi v odraslih partnerskih odnosih. Raziskave kažejo, da so težave v partnerskih odnosih bolj pogoste pri tistih, ki so doživeli spolno zlorabo, kot pri tistih, ki spolne zlorabe niso doživeli. Pari med drugim poročajo o oslabljenem zaupanju, oteženem čustvenem izražanju in težavah na področju spolnosti. Partnerski odnos se tako znajde na preizkušnji, oba v paru pa se spoprijemata z različnimi stiskami, katerih pogosto ne pripisujeta pretekli travmi. Prepoznavanje, razumevanje z elementi dvojnega zavedanja in predelava posledic, so bistvenega pomena za kakovostni partnerskih odnos. S prispevkom želimo poglobiti razumevanje delovanja odraslih partnerskih odnosov kot posledice spolne zlorabe v otroštvu in na podlagi klinične prakse opozoriti na možne oblike strokovne terapevtske pomoči pri predelavi posledic spolne zlorabe v otroštvu v partnerskih odnosih.

Ključne besede

Spolna zloraba v otroštvu, intimni partnerski odnos, zakonska terapija.

1. SPOLNA ZLORABA V OTROŠTVU

Spolna zloraba v otroštvu je travmatični dogodek, kjer je otrok uporabljen kot spolni objekt za zadovoljevanje potreb ali želja odrasle osebe, mladostnika ali starejšega otroka oz. storilca, ki namerno išče in spodbuja zlorabo z grožnjo, silo, ustrahovanjem ali manipulacijo [1]. V navedeno definicijo sodijo vrste vedenj s telesnim stikom ali brez telesnega stika, vključno z navezovanjem stikov za zlorabo, spolno sugestivnim jezikom, pornografijo, vojerizmom, ekshibicionizmom, ljubkovanjem, masturbacijo in penetracijo. Otrok je lahko zlorabljen s strani nekoga v družini ali izven nje, ki zlorabi svoj položaj moči in zaupanja [2], kar pomeni, da se zloraba zgodi znotraj odnosa. V literaturi je navedena kot relacijska travma, ki je bolj psihopatogena kot kateri koli drugi socialni ali fizični stres [3]. Številne posledice so lahko vidne takoj ali pa se pojavijo tekom odraščanja, pogosto pa se manifestirajo v intimnih odraslih partnerskih odnosih.

2. POSLEDICE SPOLNE ZLORABE V PARTNERSKIH ODNOSIH

Raziskave kažejo, da so psihološke posledice in težave v medosebnih odnosih bolj pogoste pri tistih, ki so doživeli spolno zlorabo, kot pa pri tistih, ki zlorabe niso doživeli [4]. Moški in ženske, ki so doživeli spolno zlorabo v otroštvu, v primerjavi s tistimi, ki je niso doživeli, poročajo o več medosebnih težavah na področju zaupanja, kontrole, odgovornosti, asertivnosti in občutku odtujenosti od svojega partnerja [5], kar se kaže kot nezadovoljstvo v zakonskem odnosu [6]. Številne raziskave so pokazale, da travma prizadene intimnost s partnerji [7, 8], npr. oslabi sposobnost zaupanja in oteži čustveno izražanja, izraža pa se tudi v obliki težav na področju spolnosti [9, 10, 11]. Najpogosteje so težave v tuji literaturi opisane kot zasvojenost s spolnostjo, spolne disfunkcije, ki posameznike zaznamujejo na področju spolne želje, vzburjenja in orgazma [12, 13], pogoste pa so tudi bolečinske motnje [13]. Partnerski odnos se zaradi posledic znajde na preizkušnji, oba v paru pa se spoprijemata z različnimi stiskami, katerih pogosto ne pripisujeta pretekli spolni zlorabi. Da bi lahko razmejila preteklost od sedanjosti [14] in kakovostno zaživela v partnerskem odnosu, jima je lahko v pomoč relacijska zakonska terapija.

3. RELACIJSKA ZAKONSKA TERAPIJA

Relacijska zakonska terapija [15, 16, 17, 18] temelji na premisi, da zakonca drug v drugem lahko prebudita najhujše nočne more, ki sta jih že kot otroka doživljala ob pomembnih drugih, in sicer z nezavednim namenom razrešitve [16]. V paru tako prisilno, ponavljata sistemske, medosebne in notranjepsihične zaznave, kjer ponovno nezavedno ustvarita konfliktne situacije, travme, vedenje, mišljenje in čutenje [15]. Tako se spolna zloraba z vso svojo razdiralnostjo ponavlja v intimnih partnerskih odnosih in išče razrešitev. Za boljše razumevanje v nadaljevanju predstavljamo vpogled v *del* kvalitativne raziskave, v kateri smo ugotavljali značilnosti intimnega partnerskega odnosa pri parih, ki so doživeli travmo spolne zlorabe v otroštvu in so bili vključeni v proces relacijske zakonske terapije [19].

4. METODOLOGIJA

4.1 Metoda in udeleženci

V raziskavo je bilo vključenih pet parov oz. deset udeležencev, od katerih je vsaj eden izmed partnerjev doživel travmo spolne zlorabe v otroštvu, obenem pa je ob vključitvi v partnersko terapijo njun odnos trajal vsaj leto dni. Raziskavo smo izvedli z metodo osnovane teorije, ki je omogočila sistematično zbiranje, analizo in razvoj teorije na podlagi zbranih in analiziranih podatkov [20].

4.2 Potek raziskave

Zbiranje podatkov je potekalo skozi terapevtski proces relacijske zakonske terapije, ki je trajala 6 mesecev (2 terapevtska cikla po 12 srečanj) oz. 24 srečanj, in sicer enkrat tedensko po 50 minut. Pred vključitvijo v terapijo smo pridobili soglasje za udeležbo v raziskavi z dovoljenjem za snemanje. Posnete terapije smo transkribirali in analizirali, vključili pa smo tudi svoja lastna opažanja, odzive in zapise supervizijskih srečanj. Raziskava je potekala na Družinskem inštitutu Bližina. V nadaljevanju predstavljamo rezultate dela raziskave, ki se nanašajo na področje težav v spolnosti in nekatere posege terapevtke v procesu relacijske zakonske terapije, ki poleg varnega terapevtskega odnosa omogočajo spremembe [19].

5. REZULTATI

5.1 Medosebne težave na področju spolnosti

V obdelavi zbranih podatkov je bilo ugotovljenih šest kategorij, ki predstavljajo težave na področju spolnosti: umanjkanje spolnih odnosov, odsotnost želje po spolnosti, spolnost kot dolžnost, doživljanje telesa v spolnosti, nezadovoljstvo v spolnosti, doživljanje občutij v spolnosti [19].

5.1.1 Umanjkanje spolnih odnosov

M[2]: *»Je bilo že bolš, sedaj pa je »slaba volja« med nama ...«* Ž[3]: *»Spolnih odnosov z možem zaenkrat še nimava, kar me straši, saj to traja že pol leta.«*

Vsi pari so na neki točki v terapiji poročali o umanjkanju spolnih odnosov. Nekateri so poročali o umanjkanju spolnih odnosov že pred pričetkom terapevtskega procesa, pri drugih je do umanjkanja spolnih odnosov prišlo že v prvem delu terapevtskega cikla. Za pare je umanjkanje spolnih odnosov predstavljalo težavo. Spraševali so se, kaj je narobe z njimi, da nimajo spolnih odnosov. Namreč, ko se spolna zloraba prebudi v partnerskem odnosu, pomeni, da je oseba pričela doživljati potlačena čutenja, ki so vezana na samo spolno zlorabo v otroštvu. Preko različnih sprožilcev, kot je lahko pri spolnih odnosih dotik, vonj ali okus [14], so se udeleženci z umanjkanjem spolnosti nezavedno zaščitili pred gnusom, strahom, sramom, jezo in krivdo, ki je povezana s samo zlorabo.

5.1.2 Odsotnost želje po spolnosti

Ž: *»Jaz nimam želje po spolnih odnosih, <u>on pa razume, kot da ga ne maram.</u>«*

O odsotnosti želje po spolnosti so poročale samo ženske, vendar pa je odsotnost želje pri enem v paru povzročila neprijetna občutja pri obeh, kar je povedala ena izmed udeleženk (podčrtano) in kar se je za pomembno izkazalo tudi med samim procesom terapije. Ena izmed udeleženk je npr. povedala, da ji odleže, če gre z otrokom spat, saj nima nobene želje po spolnosti.

[2] moški

[3] ženska

5.1.3 Spolnost kot dolžnost

Ž:*»Spolnost sem jemala, ali pa jo še kot dolžnost, ki jo je potrebno opravljati v zakonu, v smislu brez spolnosti ni zakona.«* Ž:*»Tako sem mislila, nisem vedela, da imam pravico reči ne, in je bilo veliko spolnosti, ki si je nisem želela.«*

Spolnost kot dolžnost je doživljalo vseh pet žensk v paru in dva moška. Nekateri so o doživljanju spolnosti kot dolžnosti poročali retrospektivno, tj. pred prihodom na terapijo, drugi pa so spolnost kot dolžnost doživljali med terapevtskim procesom. Navedeno doživljanje lahko povežemo s samo spolno zlorabo. Storilec je otroka izkoristil za zadovoljevanje svojih potreb in otrok je moral biti storilcu na razpolago za zadovoljevanje njegovih potreb. Storilec je svoje hotenje (moraš, dolžan si mi) prenesel na otroka, ki se je čutil dolžnega. V partnerskih odnosih se občutek dolžnosti prebudi ob partnerju, ki si spolnosti želi, posameznik pa doživlja, da mu mora biti na razpolago in da je dolžan z njim imeti spolni odnos in tako zadovoljiti potrebe partnerja.

5.1.4 Doživljanje telesa v spolnosti

Ž: *»... še vedno me peče, šla sem po antibiotike ... nič, nožnica me še vedno peče.«*

M: *»... vsakič, ko ji samo povem, da čutim vzburjenje, ona pove, da takoj čuti bolečino, ta pekoč občutek, samo da omenim ali najmanjši dotik, ki bi lahko imel to konotacijo, da bo šlo za neko približevanje ali predigra ali spolnost ...«*

Pari so telo doživljali različno, od prisotne fizične bolečine, ki se je pri eni udeleženki konstantno pojavljala, kadar je njen mož čutil vzburjenost. Zaradi pekoče bolečine je pri ginekologinji večkrat iskala pomoč, vendar pa fizičnega vzroka niso odkrili, prav tako ji tudi antibiotična zdravila niso pomagala odpraviti pekoče bolečine. Tudi druga udeleženka je poročala o bolečini med samim spolnim odnosom. Dve udeleženki sta svoje telo doživljali kot breme, ki bi ga bilo potrebno »odstraniti«, ena udeleženka in eden udeleženec med spolnim odnosom telesa nista čutila oz. je prišlo do disociacije le-tega. Trije udeleženci nad svojim telesom niso imeli kontrole. Pri enem ni bilo mogoče kontrolirati odziva telesa, ko je prihajalo do neželenih izlivov, pri dveh pa je bilo izrazito samozadovoljevanje. Samozadovoljevanje se je kazalo kot odraz tesnobe in napetosti, kar je vodilo v občutja krivde.

5.1.5 Nezadovoljstvo v spolnosti

M: *»Odnos je žal zelo neusklajen, saj je tako fizičnega kot spolnega kontakta premalo ... in tudi takrat ko je, je to po njeni odločitvi.«*

Ž: *»Spolnost je v bedu.«*

Vsi moški udeleženci so doživljali nezadovoljstvo, ki se je kazalo v obliki prepirov, razočaranja, nesproščenosti, nepredanosti in nezmožnosti vplivanja na to, kdaj bosta imela spolni odnos. Ženske udeleženke so poročale o prepirih, nesproščenosti in razočaranju. Glede na to, da so pari doživljali spominske preblike, disociacije, bolečinske motnje, strahove, tesnobna razpoloženja, odpor in nezaupanje, je nezadovoljstvo v spolnosti normalni odgovor na nenormalno izkušnjo v otroštvu oz. normalna posledica prebujene zlorabe v partnerskem odnosu.

5.1.6 Doživljanje občutij v spolnosti

Ž: *»Me pa velikokrat prešine misel, da tega ne smem početi, da ne smem uživati v spolnem odnosu, ker bom kaznovana.«* M: *»... Ma me jezi, ja! Najprej ja, potem ne ... daj, odloči se že enkrat!«* V raziskavi se je izkazalo, da so pari doživljali različna občutja, od negotovosti do krivde, strahu, sramu, jeze, žalosti, gnusa, odpora, stiske in zmedenosti. Občutja so bila močno prisotna tudi v terapevtskem procesu, pri katerem je bila osnovna terapevtska intervencija regulacija le-teh. Ženske udeleženke so doživljale

močna občutja krivde. Pri treh udeleženkah se je občutje prebudilo po spolnem odnosu, v enem primeru je bilo povezano s kaznijo, pri dveh udeleženkah pa s tem, ko zavrneta spolni odnos. Eden izmed udeležencev moškega spola je prav tako vedno po spolnem odnosu doživljal občutke krivde. Pri krivdi gre za samoobtoževanje, ki je posledica spolne zlorabe. Krivijo se, ker niso preprečili ali ustavili zlorabe, pogosto sebe krivijo, da je do spolne zlorabe sploh prišlo. To se kaže tudi v izjavi ene izmed udeleženk, ki sta jo zlorabljala dva storilca: *»... Ne vem ... nekaj je že moglo biti na meni ... mogoče moje oči ...«* Moški udeleženci so doživljali žalost, jezo in tudi gnus. Eden izmed udeležencev je pojasnil, da zelo pogosto med samim spolnim odnosom doživi gnus, ki ga popolnoma zablokira. Navedeno lahko pojasnimo s sistemsko razsežnostjo travme, ko se na moža prenesejo občutja, ki jih žena nosi globoko v sebi. Ob spolni zlorabi se kot otrok ni smela oz. mogla jeziti, zaradi izdajstva je bila žalostna, doživljala je tudi občutja gnusa. Občutja, ki jih je povzročil storilec in so ostala »zaklenjena v telesu«, so se preko mehanizma projekcijske identifikacije prenesla na moža oz. partnerja, ki prične čutiti in doživljati vsa občutja na svojem telesu. Navedene posledice predstavljajo stiske, ki jih doživljajo pari v partnerskih odnosih. Klinična praksa kaže, da posledic pogosto ne povezujejo s preteklo zlorabo, neredki pa se tudi ne zavedajo, da so določene izkušnje iz njihovega otroštva po svoji naravi pravzaprav bile spolna zloraba. Da bi lahko resnično zaživeli kakovostni partnerski odnos so potrebni prepoznavanje, razumevanje z elementi dvojnega zavedanja in predelava posledic, kar omogoča zakonska oz. partnerska terapija.

5.2 Proces relacijske zakonske terapije

V samem procesu relacijske zakonske terapije smo s posegi terapevtke zasledovali spremembe, ki so tudi cilj terapevtskega procesa in sicer spremembe na vedenjski, kognitivni in čutenjski ravni kot tudi spremembe na intrapsihičnem, medosebnem in sistemskem delovanju.

5.2.1 Zagotavljanje varnosti

Na začetku terapevtskega procesa je bila prva naloga terapevtke zagotavljanje varnosti. Tedenska struktura je parom zagotavljala predvidljivost, ki je v svojem domačem okolju niso bili vajeni, in jim je s tem omogočala spremembo. Obenem smo v terapevtskem procesu naslavljali njihova doživljanja glede vključevanja v raziskavo in načinom, da se srečanja snemajo. Ob tem smo parom obrazložili, da so v svoji odločitvi popolnoma svobodni in v kolikor bi se odločili, da ne želijo več sodelovati v raziskavi, njihova odločitev ne bo vplivala na kakovost terapevtske obravnave ter da že podpisano soglasje lahko tudi med terapevtskim procesom kadarkoli iz kakršnihkoli razlogov prekličejo. S tem smo parom ob zagotavljanju varnosti omogočili tudi možnost kontrole. Namreč ob samem dogodku zlorabe je bil otrok popolnoma brez kontrole, saj je bil odrasli tisti, ki je vse nadziral, so med terapevtskim procesom dobili možnost, da le-tega lahko kontrolirajo, tudi s temami, ki so jih na srečanja prinašali. Terapevtka jim je vedno zagotavljala, da v kolikor o nečem ne želijo govoriti, jim to ni potrebno oz. lahko povedo samo toliko, kolikor se jim zdi varno povedati, obenem pa vedno lahko ustavijo, vprašajo oz. izrazijo svoje želje glede samega procesa terapije. Na koncu vsakega srečanja je bila narejena refleksija v smislu, kako bo oditi po posamičnem srečanju. Pari so imeli tako možnost spregovoriti o svojem čustvenem doživljanju in izraziti svoje misli v samem terapevtskem procesu. Glede same varnosti smo v terapevtskem procesu tudi preverjali, kako je s trenutno varnostjo izven terapevtskega procesa (doma, v širši družini, v službenem okolju). Da bi lahko v polnosti pridobili

občutek varnosti in kontrolo nad tem, kar se jim dogaja, so najprej potrebovali vzpostaviti kontrolo nad telesom in čutenji, s katerimi so postopoma prihajali v stik. Terapevtka je tudi kasneje, ne samo na začetnih srečanjih, zagotavljala varnost s terapevtskim odnosom, v katerem je parom ovrednotila njihovo doživljanje in s sočutnimi odzivi nudila občutek sprejetosti.

5.2.2 Priznanje, da se je zloraba res zgodila

Pomemben element v terapiji je bilo razkritje zlorabe. Osebe, ki so doživele spolno zlorabo, težko zaupajo, ne samo partnerju, tudi terapevtu. Najbolj jih ovira sram, ki ohranja negativno podobo o sebi. Ko govorimo o razkritju spolne zlorabe, to ne pomeni, da mora oseba grafično opisovati spolno zlorabo, temveč da spregovori o svojem doživljanju. Prvi korak je bil, da je lahko naglas priznala, da se je zloraba res zgodila. Ne gre samo za besede oz. kognitivno komponento, temveč tudi za afektivni vidik. Ko spregovorijo o čustvenem doživljanju, jim ni več potrebno skrivati sramu, kar jih razbremeni. Navedeno tudi pripomore, da »težkih vsebin« ne odrežejo, temveč se pričino soočati tudi s posledicami, kot so npr. spominski prebliski. Do navedenega lahko pride, kadar je v odnosu dovolj varnosti, ki pa jo mora najprej zagotoviti terapevtka. Takrat potem lahko oba v paru sprocesirata žalost, jezo, sram, strah in gnus.

5.2.3 Psihoedukacija

Del terapevtskega procesa je bila psihoedukacija, v kateri je terapevtka pojasnila samo naravo travme spolne zlorabe, vpliv, ki ga je imela na osebo, ki jo je doživela, in posledice, ki jih doživlja spolno zlorabljena oseba ter tudi njen partner. Pari, ki so doživeli spolno zlorabo v otroštvu so doživljali veliko simptomov, kot so nočne more, strahovi, hude stiske in so se spraševali, če so sploh še normalni. Terapevtka je normalizirala njihovo doživljanje v povezavi s preteklo spolno zlorabo. Pomembno je bilo, da oba v paru razumeta, kar se je dogajalo, saj je partnerja, ki je doživel zlorabo sram, da tako čuti in doživlja. Terapevtka je parom omogočala občutek kontrole, da so lahko sami izbirali teme in se pogovarjali o tem, kar so želeli. Če je npr. klientka dejala, da o spolnosti še ne želi govoriti, jo terapevtka nikoli ni silila, temveč jo je zanimalo, kaj bi čutila, če bi govorila. Pogosto je bil sram tisti, ki ji je onemogočal spregovoriti o stiski. Pomembno je, da razumeta oba, da v primeru raznovrstnih težkih simptomov ne gre za patologijo osebe, ki je bila spolno zlorabljena v otroštvu, temveč za normalne odzive na nenormalne, travmatične dogodke. Gre za del pomoči vzpostavljanja kontrole nad situacijo, ko vidi, kaj se dogaja, in ve, da gre samo za simptom, da ni z njo nekaj narobe, da razume najprej na kognitivnem nivoju, kaj se dogaja.

5.2.4 Postavljanje meja

Ena izmed pomembnih tem je bila spoštovanje do lastnega telesa. Kadar je terapevtka čutila, da si ženska ne želi spolnih odnosov, kljub temu pa se vanje spušča, ji je dala dovoljenje, da ji tega ni več potrebno početi. Ženske so v terapiji povedale, da jim je slednje zelo pomagalo, da so lahko pričele poslušati svoje telo, obenem pa je navedeno razbremenilo tudi njihove partnerje.

5.2.5 Razmejevanje preteklosti od sedanjosti

Pomemben poseg terapevtke je bilo tudi razmejevanje preteklosti od sedanjosti oz. dvojno zavedanje. Oba zakonca oz. partnerja v paru sta pogosto doživljala hude čustvene stiske, ki sta jih pripisovala drug drugemu. Preteklost se je preko čustev in telesa prebujala v sedanjosti, drug ob drugem. Navedena dinamika je bila pogosta v spolnih odnosih, pri katerih je žena svojega moža pričela doživljati kot storilca. Ko je žena zavračala spolne odnose, se je mož počutil zavrnjenega. Terapevtka jima je pomagala razumeti, da žena ne zavrača moža, temveč čutenja, kot so sram,

gnus v povezavi s poniženjem in razvrednotenjem, kar se prenaša preko relacijskega mehanizma projekcijske identifikacije iz preteklosti v sedanjost in se odigrava v njunem partnerskem odnosu. V tem delu je bilo pomembno razumevanje, da se zloraba ne dogaja več, so pa čutenja tista, ki se prebujajo v sedanjosti.

5.2.6 Povezovanje v paru

Bistveno v relacijski zakonski terapiji je povezovanja para na nov način. Namreč glavnina procesa poteka na njunem partnerskem odnosu v sedanjosti, ne na pretekli travmi. V kolikor sta se zmožna čustveno povezati na drugačen način, potem partnerski odnos nudi toliko varnosti, da jima ni potrebno več ohranjati obrambnih drž, kot so obtoževanje in umikanje, temveč lahko delita najbolj ranljiva čutenja. Šele, ko sta v polnosti ranljiva drug z drugim, sta sposobna razviti sočutje drug do drugega, do male deklice in malega dečka, ki sta še vedno globoko v njima, povezana s težkimi občutji in moledujeta za pozornost, sprejetost in ljubezen.

6. ZAKLJUČEK

Pari z izkušnjo spolne zlorabe v otroštvu se srečujejo s težavami v medosebnih intimnih odnosih in potrebujejo strokovno pomoč, ki jim pomaga razumeti njihove trenutne težave v luči pretekle spolne zlorabe. Ključni pri pomoči so lahko zakonski terapevti, ki s pomočjo relacijske zakonske terapije, svojim sočutnim odnosom in ustreznim znanjem pripomorejo k zdravljenu preteklih travm in skupni rasti v partnerskem odnosu. Glede prizadevanja pomoči parom bi bilo potrebno več pozornosti nameniti preventivnim programom, ki bi s svojo izobraževalno naravo pripomogli k lajšanju marsikatere stiske, ki jo doživljajo pari.

7. REFERENCE

[1] Finkelhor, D. in Korbin, J. 1988. Child abuse as an international issue. *Child Abuse & Neglect* 12, 1, 3-23.

[2] Sanderson, C. 2006. *Counselling adult survivors of child sexual abuse*. London, Philadelphia: Jessica Kingsley.

[3] Schore, A. N. 2003. Affect regulation and the repair of the self (norton series on interpersonal neurobiology). New York: WW Norton & Company.

[4] Briere, J. N. in Elliott, D. M. 1994. Immediate and Long-Term Impacts of Child Sexual Abuse. *The future of children* 4, 2, 54-69.

[5] Whiffen, V. E., Thompson, J. M. in Aube, J. A. 2000. Mediators of the link between childhood sexual abuse and adult depressive symptoms. *Journal of Interpersonal Violence* 15, 10, 1100-1120.

[6] Liang, B., Williams, L. M. in Siegel, J. A. 2006. Relational outcomes of childhood sexual trauma in female survivors a longitudinal study. *Journal of Interpersonal Violence* 21, 1, 42-57.

[7] Davis, J. L. in Petretic-Jackson, P. A. 2000. The impact of child sexual abuse on adult interpersonal functioning: A review and synthesis of the empirical literature. *Aggression and violent behavior* 5, 3, 291-328.

[8] Rumstein-McKean, O. in Hunsley, J. 2001. Interpersonal and family functioning of female survivors of childhood sexual abuse. *Clinical Psychology Review* 21, 3, 471-490.

[9] Colman, R. A. in Spatz Widom, C. 2004. Childhood abuse and neglect and adult intimate relationships: a prospective study. *Child Abuse & Neglect* 28, 11, 1133-1151.

[10] Leonard, L. M. in Follette, V. M. 2002. Sexual functioning in women reporting a history of child sexual abuse: review of the empirical literature and clinical implications. *Annual Review of Sex Research* 13, 1, 346-388.

[11] Najman, J. M., Dunne, M. P., Purdie, D. M., Boyle, F. M. in Coxeter, P. D. 2005. Sexual abuse in childhood and sexual dysfunction in adulthood: An Australian population-based study. *Archives of sexual behavior* 34, 5, 517-526.

[12] Carnes, P. in Moriarity, J. 1997. *Sexual anorexia: overcoming sexual self-hatred*. Center City, Minn.: Hazelden.

[13] Maltz, W. in Holman, B. 1987. *Incest and sexuality: a guide to understanding and healing*. Lexington: Heath and Com.

[14] Rothschild, B. 2000. *The body remembers: the psychophysiology of trauma and trauma treatment*. New York: Norton.

[15] Gostečnik, C. 2004. *Relacijska družinska terapija*. Ljubljana: Brat Frančišek in Frančiškanski družinski inštitut.

[16] Gostečnik, C. 2007. *Relacijska zakonska terapija. Vol. 9, Monografije FDI*. Ljubljana: Brat Frančišek; Frančiškanski družinski inštitut.

[17] Gostečnik, C. 2011. *Inovativna relacijska družinska terapija: inovativni psiho-biološki model*. Ljubljana: Brat Frančišek, Teološka fakulteta in Frančiškanski družinski inštitut.

[18] Gostečnik, C. 2015. *Inovativna relacijska zakonska terapija*. Ljubljana: Brat Frančišek, Teološka fakulteta in Frančiškanski družinski inštitut.

[19] Jerebic, S. 2016. *Travma spolne zlorabe v otroštvu in intimni partnerski odnos*. Doktorska disertacija. Univerza v Ljubljani, Teološka fakulteta.

[20] Strauss, A in Corbin, J. 1990. Basics of Qualitative Research. Newbury Park, CA: Sage

Do healthy participants make advantageous decisions? Evidence from the Iowa Gambling Task

Maša Vukčević Marković

Psychosocial Innovation Network, Belgrade and Belgrade University, Faculty of Philosophy, Laboratory for
Research of Individual Differences
Gospodar Jevremova 48
11000 Belgrade, Serbia
+381637234658
masa@psychosocialinnovation.net

ABSTRACT

Somatic Marker Hypothesis (SMH) suggests that decision making in uncertainty relays on somatic markers – emotional reactions reflected in bodily states which lead person towards advantageous decision making. Authors of SMH created a task aiming to access decision making in uncertainty – Iowa Gambling Task (IGT), suggesting that since healthy participants tend to develop advantageous decision-making strategy, they will have good IGT performance. Recent studies however, question good IGT performance among healthy participants thus imposing question on their advantageous decision-making strategies. The aim of this study was to explore IGT performance among healthy participants. Participants (N=268) filled standard version of IGT, consisting of 100 trials in which participants select cards from four decks: A and B, that bring higher rewards and penalties (good decks) and C and D that bring small rewards and penalties (bad decks). At the beginning of the IGT, participants get 2000$ of play money, with instruction to try to gain as much money as possible. Results show that participants tend to lose rather then gain money, with almost half of participants showing impaired IGT performance. In addition, results indicate that frequency of punishment, rather than overall goodness or decks is factor relevant for specific deck selection. Study results question wheatear healthy participants apply advantageous decision-making strategies when performing IGT, indicating that future studies are needed in order to explore factors contributing to advantageous decision making in uncertainty.

Keywords

Iowa Gambling Task, IGT, Somatic Marker Hypothesis, Decision Making.

1. INTRODUCTION

Numerous studies exploring decision making in uncertainty rely on Somatic Marker Hypothesis (SMH) [3, 5, 9]. This hypothesis suggests that normal decision-making in humans is led by somatic markers - bodily states which represent emotional reactions to different stimuli or to possible courses of action. Authors of SMH suggest that somatic markers, activated in situation of uncertainty can effectively reflect the goodness or badness of the possible outcome, leading a person towards advantageous decision making. Importantly, authors suggest that somatic markers represent a "gut feeling", which can operate both consciously and unconsciously. SMH, thus suggests that somatic markers can lead people to advantageous decision making even though they do not necessarily need to be aware of which decisions are good and which ones are bad [5, 9, 10].

1.1 IOWA GAMBLING TASK

Authors of Somatic Marker Hypothesis created a task aiming to access advantageous versus disadvantageous decision making – Iowa Gambling Task (IGT) [4]. In the IGT, participants are presented with four deck of cards. After selecting a card, participants either win or lose different amounts of play money. Unknown to the participants, selecting cards from two of the decks – C and D will result in eventual gain ('advantageous', or 'good' decks), whereas selecting cards from the other two decks - A and B will result in eventual loss ('disadvantageous', or 'bed' decks). The task has 100 trials, so that, according to authors' idea, participants still could not understand reword-punishment scheme. Authors suggested and demonstrated in several studies that healthy individuals will select more cards from the good than the bad decks, thus showing tendency towards advantageous decision making [4, 17]. Task validation was demonstrated in numerous studies which indicated that impaired IGT performance have persons with problems in advantageous decision making in everyday life [7, 12, 14, 15, 18], as well as patients with lesions in the brain region relevant for the judgment and decision-making in real-life settings [4, 6, 9].

Recent studies, however, question whether healthy persons do prefer good over a bed decks and indicate that healthy controls tend to select bed decks equally or even more frequently than good ones [2, 8, 11, 16]. Moreover, when selection of each deck specifically was analyzed, it was shown that bed deck B is equally or even more frequently selected then good decks C and D [8, 11], while bed deck A represent the least selected deck [11, 13, 19]. Possible explanation can be found in differences among frequency of punishment within good and bed decks, resulting in decks A and C being frequent punishment decks, while decks B and D represent non-frequent punishment decks, which was not taken into consideration in the early IGT studies. Studies that included both criteria into analysis of IGT performance, indicated that frequency of punishments, rather than overall goodness of deck seems to be leading participants towards selection of specific decks [8, 11, 13, 19]. Thus, recent studies indicated that it is not certain whether healthy controls show homogeneous tendency towards advantageous decision making and whether overall goodness of decks, or rather frequency of its punishments will be relevant factor for deck selections.

1.2 PRESENT STUDY

In this study 268 (81.7% females, average age 20 years) participants, completed standard version of Iowa Gambling Task (IGT).

2. RESULTS

On average, participants scored 1475$ on IGT, ranging from -4525 to 6075, indicating that participants tend to rather lose then gain money on IGT, i.e. that participants, on average, do not show advantageous decision-making strategy. Table 1 presents descriptive measures of each deck selection separately.

Table 1. Descriptive statistic of each deck selection, in proportion

Deck	*M*	*SD*	*Min*	*Max*	*zSk*	*zKu*
A	.17	.098	0	.75	8.18**	16.66*
B	.29	.142	0	.72	1.20	.60
C	.23	.156	0	.1	15.21*	27
D	.31	.175	0	.1	9.73**	9.51

M – mean; *SD* – standard deviation; *Min* - minimum; *Max* – maximum; *zSk* – standardized Skewness; *zKu* – standardized Kurtosis; ns - p > .05; * - p <.05; ** - p <.01.

Results show that participants the most frequently selected cards from deck D and deck B subsequently, somewhat less frequently they selected cards from deck C, while they selected cards from deck A the least frequently among all decks. Table 2. presents percentage of participants who selected cards from good versus bed decks at more than 50% of trials, as well as percentage of participants who selected cards from non-frequent punishments decks B, D versus frequent punishments decks A and C at more than 50% of trials. Results indicate that somewhat less than half of participants have impaired performance [4], i.e. they selected more cards from bed then good decks, while more than three quarters of participants selected more cards form non-frequent punishments decks B and D, in comparation to frequent-punishments decks A and C. There is no significant difference in number of participants who selected more cards from good then bed decks [$\chi^2(1) = 2.149$; *p* = .143], while significant difference was found in the number of participants who selected more cards from non-frequent punishment decks B and D in comparison to selection of cards from frequent punishment decks A and C [$\chi2(1) = 75.239$; *p* < .001].

Table 2. Percentage of participants who selected more cards from different decks

		% of participants with more cards from good decks	% of participants with more cards from bed decks
		54.5%	45.5%
% of participants with more cards from non-frequent punishments decks	76.5%	39.9%	36.6%
% of participants with more cards from frequent punishments decks	23.5%	14.6%	9%

3. DISCUSSION

Study results indicate that, contrary to IGT authors' expectations, participants tend to lose rather than win money on IGT and that almost half of participants have impaired performance, i.e. more choices from bed then from good decks. This result is in line with studies questioning whether healthy controls prefer good over a bad deck and whether they show overall advantageous decision-making strategy [2, 8, 11, 16, 17]. Results also show that majority of participants prefer non-frequent punishment decks B and D in comparation to frequent punishment decks A and C. These results indicated importance of frequency of punishment in deck selection. They are also in line with results of a previous studies that considered both frequency of punishments and overall goodness of decks, indicating that participants tend to the most frequently select cards from deck D (good, non-frequent punishments deck), then B (bed, non-frequent punishments deck), then deck C (good, frequent punishments deck), while they the least frequently select cards from deck A (bed, frequent punishments deck) [1, 8, 11].

Results of this study indicated that healthy controls do not show homogenous IGT performance, nor advantageous decision-making strategy. In addition, results indicated than frequency of punishments rather than overall goodness of deck is relevant for deck selection, showing that immediate reword/punishment seems to be more important than long term consequences in decision making. These results do not question Somatic Marker Hypothesis, just call for additional studies that will improve understanding of factors that contribute to good IGT performance, and overall advantageous decision making in situations of uncertainty.

4. REFERENCES

[1] Alarcón, D., Amián, J. G., and Sánchez-medina, J. A. 2015. Enhancing emotion-based learning in decision-making under uncertainty. *Psicothema*, *27*(4), 368–373. https://doi.org/10.7334/psicothema2015.45

[2] Bark, R., Dieckmann, S., Bogerts, B., and Northoff, G. 2005. Deficit in decision making in catatonic schizophrenia : An exploratory study, *Psychiat res 134*, 131–141. https://doi.org/10.1016/j.psychres.2004.04.013

[3] Bechara, A., and Damasio, A. 2005. The somatic marker hypothesis: A neural theory of economic decision. *Game Econ Behav*, *52*(2), 336–372. https://doi.org/10.1016/j.geb.2004.06.010

[4] Bechara, A., Damasio, A., Damasio, H., and Anderson, A. 1994. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7–15.

[5] Bechara, A., Damasio, H., Tranel, D., and Damasio, A. 2005. The Iowa Gambling Task and the somatic marker hypothesis: some questions and answers. *Trends Cogn Sci*, *9*(4), 159-62; discussion 162–4. https://doi.org/10.1016/j.tics.2005.02.002

[6] Bechara, A., Tranel, D., and Damasio, H. 2000. Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. *J Neur*, *123 (Pt 1*, 2189–202. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11050020

[7] Brevers, D., Cleeremans, A., Goudriaan, A., Bechara, A., Kornreich, C., Verbanck, P., and Noël, X. 2012. Decision making under ambiguity but not under risk is related to problem gambling severity. *Psychiat Res*, *200*(2–3), 568–74. https://doi.org/10.1016/j.psychres.2012.03.053

[8] Caroselli, J. S., Hiscock, M., Scheibel, R. S., and Ingram, F. 2006. The simulated gambling paradigm applied to young adults: an examination of university students' performance. *Appl Neuropsychol*, *13*(4), 203–12. https://doi.org/10.1207/s15324826an1304_1

[9] Damasio, A. 1994. Descartes' error: Emotion, reason, and the human brain. New York, NY: Putnam.

[10] Damasio, A. 1999. The Feeling of What Happens.

[11] Fernie, G., and Tunney, R. J. 2006. Some decks are better than others: the effect of reinforcer type and task instructions on learning in the Iowa Gambling Task. *Brain Cognition*, *60*(1), 94–102. https://doi.org/10.1016/j.bandc.2005.09.011

[12] Hanson, K., Luciana, M., and Sullwold, K. 2008. Reward-related decision-making deficits and elevated impulsivity among MDMA and other drug users. *Drug Alcohol Depen*, *96*(1–2), 99–110. https://doi.org/10.1016/j.drugalcdep.2008.02.003

[13] Kjome, K. L., Lane, S. D., Schmitz, J. M., Green, C., Ma, L., Prasla, I., … Moeller, F. G. 2010. Relationship between impulsivity and decision making in cocaine dependence. *Psychiatry Res*, *178*(2), 299–304. https://doi.org/10.1016/j.psychres.2009.11.024

[14] Lev, D., Hershkovitz, E., and Yechiam, E. 2008. Decision making and personality in traffic offenders: a study of Israeli drivers. *Accident Anal Prev*, *40*(1), 223–30. https://doi.org/10.1016/j.aap.2007.05.009

[15] Noël, X., Bechara, A., Dan, B., Hanak, C., and Verbanck, P. 2007. Response inhibition deficit is involved in poor decision making under risk in nonamnesic individuals with alcoholism. *Neuropsychology*, *21*(6), 778–86. https://doi.org/10.1037/0894-4105.21.6.778

[16] Sevy, S., Burdick, K., Visweswaraiah, H., Abdelmessih, S., Lukin, M., Yechiam, E., and Bechara, A. 2007. Iowa Gambling Task in schizophrenia : A review and new data in patients with schizophrenia and co-occurring cannabis use disorders. *Schizophr Res*, *92*, 74–84. https://doi.org/10.1016/j.schres.2007.01.005

[17] Steingroever, H., Wetzels, R., Horstmann, A., Neumann, J., and Wagenmakers, E.-J. 2013. Performance of healthy participants on the Iowa Gambling Task. *Psychol Assessment*, *25*(1), 180–93. https://doi.org/10.1037/a0029929

[18] Struglia, F., Tomassini, a., Stratta, P., Spaziani, D., & Rossi, A. (2009). P01-75 Cognitive functions, impulsivity and personality traits in alcohol dependent subjects. *Eur Psychiat*, *24*, S463. https://doi.org/10.1016/S0924-9338(09)70696-0

[19] Wood, S., Busemeyer, J. R., Koling, A., Cox, C. R., & Davis, H. (2005). Older adults as adaptive decision makers: evidence from the Iowa Gambling Task. *Psychol Aging*, *20*(2), 220–5. https://doi.org/10.1037/0882-7974.20.2.220

Palaeolithic toolmaking and the evolution of cognition and language

Petar Gabrić
Department of Linguistics
Faculty of Humanities and Social Sciences
University of Zagreb
pgabric@ffzg.hr

Marko Banda
Department of Archaeology
Faculty of Humanities and Social Sciences
University of Zagreb
mabanda@ffzg.hr

Ivor Karavanić
Department of Archaeology
Faculty of Humanities and Social Sciences
University of Zagreb
ikaravan@ffzg.hr

ABSTRACT

This paper reviews in short the current research on the hypothesis of coevolution between Palaeolithic stone tool manufacture on one side, and cognition and specifically language on the other. Of particular interest are behavioral and neuroimaging studies.

Keywords

language evolution, cognitive evolution, Palaeolithic, stone tools, Oldowan, Acheulean, Levallois, cultural transmission

1. INTRODUCTION[1]

In the past two decades great interest has emerged for interdisciplinary discussion on the evolution of cognition and specifically language (e.g. Janković and Šojer 2015). Recent work strongly emphasizes the role of cumulative culture, i.e. products, skills and knowledge created over generations of human lifetimes, in the evolution of specific hominin cognitive abilities, including language (e.g. Lotem et al. 2017).

Moreover, it remains a controversial issue whether evolutionary changes in hominin cognition should be viewed as structural and/or functional novelties, or as exaptations of preexisting primate structures and/or functions (e.g. Stout and Chaminade 2007: 1092). Nevertheless, ample literature now acknowledges that exaptation played a major role in hominin cognitive evolution (e.g. Kolodny and Edelman 2018), but also e.g. in the evolution of hominin musculature underlying bipedalism, tool use and speech (Diogo 2018). For language specifically, it has been e.g. argued that both child first language and adult second language acquisition are served by general-purpose learning systems of declarative and procedural memory (Hamrick et al. 2018), and that linguistic semantic processing is highly dependent on modality-specific processing mechanisms of the sensorimotor system and not only on amodal or "abstract" semantic operations (e.g. Pulvermüller 2013).

Still, empirical research in the evolution of cognition and language remains relatively scarce, mostly due to serious methodological limitations (see §3). One approach has been to find behavioral correlates of specific cognitive functions in the Palaeolithic record, the archaeological period beginning with the appearance of first known hominin stone tools ~3.3 mya (Harmand et al. 2015) and finishing with the end of the last glacial period (Karavanić and Janković 2009).

---

[1] *Abbreviations*: ESA=Early Stone Age (~Oldowan and Acheulean); IPS=intraparietal sulcus; SMA=supplementary motor area; SMG=supramarginal gyrus; vPrCG=ventral precentral gyrus; vPMC=ventral premotor cortex

Research in this area has been in large part focused on inferring "modern" and "symbolic behavior" from archaeological data, such as ochre and ornament use, figurative art production, subsistence strategies etc. (e.g. d'Errico et al. 2003). The research was led by the premise that such behaviors indicate the existence of a package of "higher" cognitive abilities. However, these discussions have been heavily criticized from a methodological point of view (e.g. Botha 2009, 2012), and they currently struggle to find support in cognitive science (e.g. Garofoli 2014).

2. PALAEOLITHIC STONE TOOLS

Recently, vast literature has appeared supporting the coevolution of cognition and language on one side, and the manufacture of Palaeolithic stone tools on the other (see e.g. Stout and Hecht 2015 for a review). Archaeology has recognized several phylogenetic phases of stone tool manufacture, and research in this topic has tried to correlate these phases with evolutionary developments in cognition, most notably visuospatial processing, executive functioning, social cognition and language.

2.1 Lomekwian

The earliest known hominin stone tools dated to ~3.3 mya come from Lomekwi, Kenya, and consist of various rocks used for pounding, stone anvils and cores from which flakes were struck using a hammerstone (Harmand et al. 2015). The site predates the earliest known *Homo* specimens dated to ~2.8 mya (Villmoare et al. 2015). Cognitive implications of the Lomekwi findings, regarding among others bimanual coordination and lateralization more generally, are discussed in Lewis and Harmand (2016).

2.2 Oldowan

Oldowan represents the next stage of hominin stone technology, ranging from ~2.6 (Semaw 2000) to ~1.42 mya (Toth and Schick 2018: 7). It is generally considered to be characterized by flaked pebble tools, namely choppers and chopping tools, used for pounding or bone splitting. The knapping of these pebbles produced smaller flakes which could have been utilized for meat butchering (Toth 1987). Oldowan is most commonly associated with *H. habilis*, but it is not excluded that australopithecines engaged with Oldowan tools (Karavanić and Janković 2009: 107). Compared to Lomekwian Oldowan is described as entailing greater abundance of flakes and smaller tool dimensions (Hovers 2015). Toth and Schick (2018) argue that Oldowan findings are suggestive of the incorporation of stone tools as a critical adaptive component which presumably led to more complex subsistence strategies, social behavior and communication.

2.3 Acheulean

The oldest Acheulean finds are dated to ~1.7 mya (Diez-Martín et al. 2015) and lasted in some parts of the world up to ~100 kya. Acheulean is most commonly associated with *H. erectus* and Middle Pleistocene hominins, such as *H. heidelbergensis* (Torre 2016). The most commonly recognized Acheulean tool is the handaxe and its iterations, a bifacially flaked stone tool usually of amygdaloidal form (Torre 2016). Furthermore, Acheulean assemblages also contain various flake tools. *H. ergaster/erectus* was the first hominin species to expand out of Africa settling vast areas of Asia and Europe. However, this expansion did not necessarily entail the spread of Acheulean, because the earliest known hominin sites in Europe exhibit a stone industry with an Oldowan tradition (Karavanić and Janković 2009: 120). ~500 kya, many European stone industries start to resemble the later Acheulean with its well-formed handaxes (Karavanić and Janković 2009: 124). It has been argued that Acheulean tool production is more cognitively demanding relative to Oldowan due to its supposed higher hierarchical and sequential complexity (e.g. Torre 2016: 8ff.).

Due to space limitations and the fact that empirical studies almost exclusively focused on Oldowan and Acheulean, later hominin technologies will not be discussed here (but see §3). Additionally, further insights into the cognitive implications of Palaeolithic tools have come from the recent discoveries of primate tool-related behaviors (see Haslam et al. 2017 for a review).

3. EMPIRICAL STUDIES

Empirical studies concerned with the coevolution of cognition and stone tool manufacture include (1) behavioral studies investigating the effects of different cultural transmission conditions on the acquisition of stone tool manufacture, and (2) neuroimaging studies of stone toolmaking or its observation. More detailed and critical analyses of the studies in question are, unfortunately, beyond the limitations of this paper.

Some methodological constraints include (1) difficulties in neuroimaging of stone tool manufacture, e.g. due to the static nature of the conventional methods such as fMRI, or the possibly more harmful effects of FDG-PET on subjects due to longer and more demanding activities (Stout and Chaminade 2007: 1096), (2) small numbers of subjects and/or shorter learning periods due to the need of collecting large amounts of raw material for tool production and the known infrastructural limitations of neuroscientific research on the number or the mere existence of subjects, and (3) recruiting modern humans to study cognitive abilities of extinct hominin species (e.g. Putt et al. 2017: 1).

3.1 Cultural transmission studies

To our knowledge only five such studies have been conducted while interpreting the results has been challenging. All have included subjects in the earliest learning stages. Regarding Oldowan, Morgan et al. (2015) compared transmission efficiency during flaking acquisition across five communication conditions. They found i.a. that the transmission improved with gestural and verbal teaching relative to imitation/eumlation as seen e.g. in flake quality and the number of viable flakes produced. Furthermore, they found that verbal teaching improves performance relative to gestural teaching. Morgan and colleagues argued that ESA technologies possibly generated selection for increasingly complex transmission modes. These results were in some measure replicated by Lombao et al. (2017) who compared the efficiency of imitation/emulation, and gestural and verbal teaching in the acquisition of the alternating flaking method. They reported that

both teaching conditions improved performance compared to imitation/emulation and that the effects were most pronounced in the verbal teaching condition. Additionally, Cataldo et al. (2018) found in their study of flaking that subjects instructed with gesture-free verbal teaching underperform relative to subjects instructed by gestural or "full-language" teaching. Cataldo and colleagues thus conclude that while gestural communication was likely under selective pressures in the Oldowan populations, this is not necessarily the case for (spoken) language.

As to post-Oldowan techniques, Putt et al. (2014) compared the efficiency of imitation/emulation and verbal teaching in the acquisition of Acheulean handaxe manufacture. They reported no significant differences e.g. in shape and symmetry of the tools between the groups, but the non-verbal group produced more efficient flakes as seen in higher ratios of platform width to platform thickness and size to mass compared to the verbal group. Putt and colleagues concluded that (spoken) language wasn't necessary for the transmission of handaxe manufacture and that its implementation in the earliest learning stages might hinder progress. Ohnuma et al. (1997) compared the efficiency of "verbal" and "non-verbal demonstration" in the acquisition of Levallois flaking. There were no significant differences between the two conditions in the acquisition rates and mean times, and flaking success from which Ohnuma and colleagues concluded that (spoken) language was not necessary for Levallois flaking. Nonetheless, Levallois has been linked to increased demands in sequential and hierarchical planning as well as self-monitoring (e.g. Schlanger 1996: 246ff.).

It should, however, be noted that there exist considerable methodological variations in the observed studies as in the neuroimaging studies, e.g. in the selection, homogeneity and preparatory modification of raw material, learning duration and rates, presence of an experienced teacher, number of subjects etc.

3.2 Neuroimaging studies

Neuroimaging studies have been mainly trying to determine the cortical structures which would have been under selective pressures, and then tying these structures to specific cognitive functions. Hecht et al. (2014) conducted a longitudinal DTI study with participants who in a two-year program acquired basic Oldowan flaking, Acheulean handaxe manufacture and Levallois flaking. They recorded fractional anisotropy changes in branches of the superior longitudinal fasciculus leading into the left SMG and vPrCG, and right pars triangularis (part of Broca's area). They thus concluded that the acquisition of Palaeolithic stone tool manufacture entails structural remodeling of inferior frontoparietal areas. Stout et al. (2008) further report from their PET study of Oldowan flaking and Acheulean handaxe manufacture by expert subjects that Oldowan and Acheulean manufacture shared increased activation compared to the control condition in specific bilateral parietal clusters in the superior parietal lobule, IPS and SMG, and additionally in BA 17 and 18 in the occipital lobe. Stout et al. (2011) conducted an fMRI study in which subjects with different degrees of experience in Palaeolithic toolmaking viewed 20-second clips of Oldowan flaking and Acheulean handaxe making. Compared to the control condition increased activation during observation of ESA toolmaking was yet again documented in the occipital and inferior parietal areas, and the precentral gyri, and was further found in the inferior temporal cortices and the right Broca's area (BA 44 and 45). These results suggest that early Palaeolithic toolmaking relies largely on motor and visuospatial processing, but two studies have

also recorded increased activity in the prefrontal cortex suggesting the involvement of hierarchical and sequential action processing.

Oldowan toolmaking has been described as involving mainly the frontoparietal sensorimotor areas, most notably the vPrCG, SMA and IPS, and the cerebellum while it is not associated with prefrontal activity (Stout 2000, 2007). Studies have concluded that Oldowan toolmaking relies mostly on motor and visuospatial processing with no apparent role of e.g. executive functioning, suggestive of more "ape-like" cognitive abilities (Putt et al. 2017: 4). It is here notable that the parietal cortex has witnessed a significant enlargement during hominin evolution and it is suggested that the most pronounced changes occurred around the IPS as evidenced by an inferior displacement of the lower parietal areas (Bruner 2010). Furthermore, changes in cerebellum size have been noted as well in a computational study comparing Neanderthal and modern human brains (Kochiyama et al. 2018).

Stout et al. (2008) found higher activation during Acheulean relative to Oldowan toolmaking bilaterally in the vPMC, inferior parietal areas and the right Broca's area. Furthermore, Uomini and Meyer (2013) conducted an fTCD study with expert subjects engaged in the Acheulean handaxe production and silent phonemic fluency. They noted high correlations between the hemodynamic lateralization patterns during the two tasks. It has been argued that there is a strong evolutionary connection between functional lateralization, as seen e.g. in bimanual coordination, and language (e.g. Uomini 2015). However, Putt et al. (2017) reported from their fNIRS study that subjects acquiring Acheulean handaxe manufacture in a verbal teaching condition had increased activation in the right pars triangularis compared to the non-verbal condition. These results cast potential doubt on the validity of previous results regarding Broca's area. However, comparisons between the toolmaking and control conditions were not reported. Furthermore, the study included only subjects in their early learning phases. Still and all, it has been proposed that Broca's area poses a possible connection between the evolution of toolmaking and language because of its prominent role in schematic body representation as well as sequential and hierarchical goal-directed action processing (e.g. Ruck 2014). Additionally, Kemmerer (2012) suggested that the cross-linguistically most prevalent word orders (SOV and SVO) reflect the ways Broca's area processes actions and/or events. Moreover, Putt et al. (2017) found bilateral temporal Acheulean-related activity which they associated with auditory working memory. Finally, Putt and Wijeakumar (2018) conclude based on their earlier study that Acheulean handaxe manufacture and modern language rely on different components of working memory, with Acheulean recruiting visual and auditory working memory components not typically related with modern human linguistic processing. However, this auditory component might represent a precursor to modern verbal working memory. A further point of convergence might have been the vPMC (Putt and Wijeakumar 2018: 282).

Miura et al. (2014) conducted an fMRI study comparing observation of videos of Mousterian toolmaking and a man pronouncing Uzbek words, a language unfamiliar to the subjects. Among others, increased activation during Mousterian observation relative to the perception of Uzbek words was found in the right pars opercularis (part of Broca's area) and bilaterally in BA 6 (entailing the PMC and SMA).

4. CONCLUSION

Behavioral studies tentatively demonstrate that language was not a prerequisite for early stone tool transmission. However, they suggest that non-linguistic gestures might have been under selective pressure as early as Oldowan. Neuroimaging studies show that Acheulean compared to Oldowan activates prefrontal areas suggesting that Acheulean possibly requires significantly more executive control. While the prefrontal activity is mostly limited to BA 44 and 45 (Broca's area), the nature of the coevolution of toolmaking and language remains largely unsettled. Future research should include later hominin technologies and seek to determine the particular cognitive functions associated with Acheulean- and Mousterian-related prefrontal activity. Further accumulation of research will hopefully reveal new methodological possibilities in language evolution research.

5. REFERENCES

[1] Botha, R. (2009). "Theoretical underpinnings of inferences about language evolution: the syntax used at Blombos Cave". In: Botha, R., Knight, C. (eds.). *The Cradle of Language*. Oxford [etc.]: Oxford University Press, 93–111.

[2] Botha, R. (2012). "Inferring modern language from ancient objects". In: Tallerman, M., Gibson, K. R. (eds.). *The Oxford Handbook of Language Evolution*. Oxford [etc.]: Oxford University Press, 303–12.

[3] Bruner, E. (2010). "Morphological differences in the parietal lobes within the human genus". *Curr Anthropol*, 51, S77–88.

[4] Cataldo, D.M., Migliano, A.B., Vinicius, L. (2018). "Speech, stone tool-making and the evolution of language". *PLoS ONE*, 13(1), e0191071.

[5] d'Errico, F., Henshilwood, C., Lawson, G., Vanhaeren, M., Tillier, A.-M., Soressi, M., Bresson, F., Maureille, B., Nowell, A., Lakarra, J., Backwell, L., Julien, M. (2003). "Archaeological evidence for the emergence of language, symbolism, and music – an alternative multidisciplinary perspective". *J World Prehist*, 17(1), 1–70.

[6] Diez-Martín, F., Sánchez Yustos, P., Uribelarrea, D., Baquedano, E., Mark, D. F., Mabulla, A., Fraile, C., Duque, J., Díaz, I., Pérez-Gonzáles, A., Yravedra, J., Egeland, C. P., Organista, E., Domínguez-Rodrigo, M. (2015). "The origin of the Acheulean: the 1.7 Million-year-old site of FLK West, Olduvai Gorge (Tanzania)". *Sci Rep*, 5, 17839.

[7] Diogo, R. (2018). "First detailed anatomical study of bonobos reveals intra-specific variations and exposes just-so stories of human evolution, bipedalism, and tool use". *Front Ecol Evol*, 6, 53.

[8] Garofoli, D. (2014). "Do early body ornaments prove cognitive modernity? A critical analysis from situated cognition". *Phenomenol Cogn Sci*, 14(4), 803–25.

[9] Hamrick, P., Lum, J. A.G., Ullman, M. T. (2018). "Child first language and adult second language are both tied to general-purpose learning systems". *Proc Natl Acad Sci U S A*, 115, 1487–92.

[10] Harmand, S., Lewis, J.E., Feibel, C.S., Lepre, C.J., Prat, S., Lenoble, A., Boës, X., Quinn, R.L., Brenet, M., Arroyo, A., Taylor, N., Clément, S., Daver, G., Brugal, J.-P., Leakey, L., Mortlock, R.A., Wright, J.D., Lokorodi, S., Kirwa, C., Kent, D.V., Roche, H. (2015). "3.3-million-year-old stone tools

from Lomekwi 3, West Turkana, Kenya". *Nature*, 521, 310–5.

[11] Haslam, M., Hernandez-Aguilar, R.A., Proffitt, T., Arroyo, A., Falótico, T., Fragaszy, D., Gumert, M., Harris, J.W.K., Huffman, M.A., Kalan, A.K., Malaivijitnond, S., Matsuzawa, T., McGrew, W., Ottoni, E.B., Pascual-Garrido, A., Piel, A., Pruetz, J., Schuppli, C., Stewart, F., Tan, A., Visalberghi, E., Luncz, L.V. (2017). "Primate archaeology evolves". *Nat Ecol Evol*, 1, 1431–7.

[12] Hecht, E.E., Gutman, D.A., Khreisheh, N., Taylor, S.V., Kilner, J., Faisal, A.A., Bradley, B.A., Chaminade, T., Stout, D. (2014). "Acquisition of Paleolithic toolmaking abilities involves structural remodeling to inferior frontoparietal regions". *Brain Struct Funct*, 220(4), 2315–31.

[13] Hovers, E. (2015). "Archaeology: tools go back in time". *Nature*, 521, 294–5.

[14] Janković, I., Šojer, T. (2015). "The evolution of speech and language." *Opvscvla archaeologica*, 37/38, 11–48.

[15] Karavanić, I., Janković, I. (2009). *Osvit čovječanstva: početci našega biološkog i kulturnog razvoja*. Zagreb: Školska knjiga.

[16] Kemmerer, D. (2012). "The cross-linguistic prevalence of SOV and SVO word order reflects the sequential and hierarchical representation of action in Broca's area". *Lang Linguist Compass*, 6(1), 50–66.

[17] Kochiyama, T., Ogihara, N., Tanabe, H.C., Kondo, O., Amano, H., Hasegawa, K., Suzuki, H., León, de M.S.P., Zollikofer, C.P.E., Bastir, M., Stringer, C., Sadato, N., Akazawa, T. (2018). "Reconstructing the Neanderthal brain using computational anatomy". *Sci Rep*, 8, 6296.

[18] Kolodny, O., Edelman, S. (2018). "The evolution of the capacity for language: the ecological context and adaptive value of a process of cognitive hijacking". *Phil. Trans. R. Soc. B*, 373, 20170052.

[19] Lewis, J.E., Harmand, S. (2016). "An earlier origin for stone tool making: implications for cognitive evolution and the transition to *Homo*". *Phil. Trans. R. Soc. B*, 371: 20150233.

[20] Lombao, D., Guardiola, M., Mosquera, M. (2017). "Teaching to make stone tools: new experimental evidence supporting a technological hypothesis for the origins of language". *Sci Rep*, 7, 14394.

[21] Lotem, A., Halpern, J.Y., Edelman, S., Kolodny, O. (2017). "The evolution of cognitive mechanisms in response to cultural innovations". *Proc Natl Acad Sci U S A*, 114(30), 7915–22.

[22] Miura, N., Nagai, K., Yamazaki, M., Yoshida, Y., Tanabe, H.C., Akazawa, T., Sadato, N. (2014). "Brain activation related to the imitative learning of bodily actions observed during the construction of a Mousterian stone tool: a functional magnetic resonance imaging study". In: Akazawa, T. et al. (eds.). *Dynamics of Learning in Neanderthals and Modern Humans. Volume 2: Cognitive and Physical Perspectives*. Tokyo [etc.]: Springer, 221–32.

[23] Morgan, T.J.H., Uomini, N.T., Rendell, L.E., Chouinard-Thuly, L., Street, S.E., Lewis, H.M., Cross, C.P., Evans, C., Kearney, R., Torre, I. de la, Whiten, A., Laland, K.N. (2015). "Experimental evidence for the co-evolution of hominin tool-making teaching and language". *Nat Commun*, 6, 6029.

[24] Ohnuma, K., Aoki, K., Akazawa, T. (1997). "Transmission of tool-making through verbal and non-verbal communication: preliminary experiments in Levallois flake production". *Anthropol. Sci*, 105(3), 159–68.

[25] Pulvermüller, F. (2013). "How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics". *Trends Cogn Sci*, 17(9), 458–70.

[26] Putt, S.S., Wijeakumar, S. (2018). "Tracing the evolutionary trajectory of verbal working memory with neuro-archaeology". *Interact Stud*, 19(1–2), 272–88.

[27] Putt, S.S., Wijeakumar, S., Franciscus, R.G., Spencer, J.P. (2017). "The functional brain networks that underlie Early Stone Age tool manufacture". *Nat Hum Behav*, 1, 0102.

[28] Putt, S.S., Woods, A.D., Franciscus, R.G. (2014). "The role of verbal interaction during experimental bifacial stone tool manufacture". *Lithic Technology*, 39(2), 96–112.

[29] Ruck, L. (2014). "Manual praxis in stone tool manufacture: implications for language evolution". *Brain Lang*, 139, 68–83.

[30] Schlanger, N. (1996). "Understanding Levallois: lithic technology and cognitive archaeology". *Cambridge Archaeological Journal*, 6(2), 231–54.

[31] Semaw, S. (2000). "The world's oldest stone artefacts from Gona, Ethiopia: their implications for understanding stone technology and patterns of human evolution between 2·6–1·5 million years ago". *J Archaeol Sci*, 27, 1197–214.

[32] Stout, D., Chaminade, T. (2007). "The evolutionary neuroscience of tool making". *Neuropsychologia*, 45, 1091–100.

[33] Stout, D., Hecht, E. (2015). "Neuroarchaeology". In: Bruner, E. (ed.). *Human Paleoneurology*. Cham: Springer, 145–75.

[34] Stout, D., Passingham, R., Frith, C., Apel, J., Chaminade, T. (2011). "Technology, expertise and social cognition in human evolution". *Eur J Neurosci*, 33, 1328–38.

[35] Stout, D., Toth, N., Schick, K., Chaminade, T. (2008). "Neural correlates of Early Stone Age toolmaking: technology, language and cognition in human evolution". *Phil. Trans. R. Soc. B*, 363, 1939–49.

[36] Stout, D., Toth, N., Schick, K., Stout, J., Hutchins, G. (2000). "Stone tool-making and brain activation: position emission tomography (PET) studies". *J Archaeol Sci*, 27, 1215–23.

[37] Torre, de la I. (2016). "The origins of the Acheulean: past and present perspectives on a major transition in human evolution". *Phil. Trans. R. Soc. B*, 371, 20150245.

[38] Toth, N. (1987). "The first technology". *Scientific American*, 256(4), 104–13.

[39] Toth, N., Schick, K. (2018). "An overview of the cognitive implications of the Oldowan Industrial Complex". *Azania*, 53(1), 3–39.

[40] Uomini, N.T. (2015). "Paleoneurology and behaviour". In: Bruner, E. (ed.). *Human Paleoneurology*. Cham: Springer, 121–43.

[41] Uomini, N.T., Meyer, G.F. (2013). "Shared brain lateralization patterns in language and Acheulean stone tool production: a functional transcranial doppler ultrasound study". *PLoS ONE*, 8(8), e72693.

[42] Villmoare, B., Kimbel, W.H., Seyoum, C., Campisano, C.J., DiMaggio, E.N., Rowan, J., Braun, D.R., Arrowsmith, J.R., Reed, K.E. (2015). "Early Homo at 2.8 Ma from Ledi-Geraru, Afar, Ethiopia". *Science*, 347(6228), 1352–5.

Psihonevrobiološka razvojna perspektiva razumevanja nasilja

izr. prof. dr. Barbara Simonič

Teološka fakulteta

Katedra za zakonsko in družinsko terapijo ter psihologijo in sociologijo religije

Poljanska c. 4, 1000 Ljubljana, Slovenija

Tel: +386 1 4345810; fax: +386 1 4345854

e-mail: barbara.simonic@teof.uni-lj.si

POVZETEK[1]

Kompleksnost nasilja se kaže v različnih pojavnih oblikah, vzročnosti, vzorcih, posledicah in kontekstih. Zaradi te kompleksnosti se pojavljajo tudi težave pri preventivi in celoviti obravnavi nasilja. V prispevku so prikazane nekatere dileme pri definiciji in razumevanju nasilja, s poudarkom na psihonevrobiološki perspektivi razvoja pa je prikazana ena izmed možnosti razumevanja agresivnega vedenja oz. nasilja, ki lahko služi kot izhodišče za dolgoročno preprečevanje nasilja.

Ključne besede

agresivnost, nasilje, splošni model agresivnosti, psihonevrobiologija, razvoj, regulacija afekta

1. UVOD

Agresivnost (nasilje)[2] je kompleksen pojav, ki ga srečamo v različnih pojavnih oblikah in v različnih dimenzijah družbene in osebne stvarnosti. Eden od razlogov, zakaj v določenih primerih ostaja ta pojav prezrt, je pomanjkanje jasne definicije, ki ni strogo znanstvena, ampak je velikokrat tudi stvar presoje, ki je pogojena tudi s kulturo, vrednotami in družbenimi normami trenutnega obstoja [1], vse to pa predstavlja tudi težavo pri načrtovanju ukrepov preventive nasilja. Najbolj osnovne definicije nasilja opredeljujejo kot vedenje, ki je namerno usmerjeno k povzročitvi škode in ki dejansko povzroči fizično ali psihično škodo. Svetovna zdravstvena organizacija [2] opredeljuje nasilje kot namerno uporabo fizične sile ali moči, ki obstaja v obliki grožnje ali pa v dejanjih zoper samega sebe, drugo osebo ali proti skupini ali skupnosti, za posledico pa ima visoko verjetnost, da povzroči telesno poškodbo, smrt, psihično škodo, zastoj v rasti (razvoju) ali prikrajšanje (odvzem za razvoj potrebnih stvari). Zaradi škode, ki jo povzroča nasilje, je seveda primarnega pomena razmišljati o preventivi nasilja, pri čemer pa je primarnega pomena razumevanje vzročnosti agresivnega vedenja, česar pa definicije večinoma ne opredeljujejo. Pri tem so lahko v pomoč številni pogledi, teorije in modeli, ki na različne, a sorodne načine razlagajo etiologijo nasilja [3].

2. PROBLEMI DEFINICIJ AGRESIVNOSTI IN NASILJA

Nasilje oz. agresivno vedenje se običajno smatra za nezaželeno, čeprav so pri tem nekatere izjeme glede na to, kakšno vedenje se v določeni situaciji pojavlja. V določenih okvirih se namreč agresivno vedenje lahko smatra kot zaželeno (npr. samoobramba). Zaradi teh različnih okoliščin in vzrokov (motivacija, učinki, dejanja, psihološki, socialni in politični pomen), v katerih se agresivno vedenje pojavlja, imamo o nasilju različne definicije. Ta različnost ustvarja tudi neenotnost pri razumevanju tega, kako se nasilje razume, kateri vzorci vedenja so prepoznani kot nasilni, kako so povezani faktorji tveganja ter tudi katere intervencije in politike so primerne pri preprečevanju nasilja. Pomanjkanje soglasja je ovira za usklajevanje raziskovanja, načrtovanja in oblikovanja učinkovitih ukrepov za preprečevanje in obravnavo različnih oblik nasilja, posledice tega pa se odražajo v splošnem blagostanju družbe. Skoraj vsakdo lahko prepozna, ali je v določeni situaciji neko dejanje ali situacija nasilna. Težje pa je v splošnem določiti, kaj je nasilno in kaj ni [4]. S tega stališča je kljub občutku jasnosti, kaj nasilje je, to vseeno dvoumen pojav, kar se pokaže pri poskusu splošnega označevanja in kategorizacije njegovih pojavov, logike pojavljanja in stopnjevanja, iskanju vzročnih razlag in njegovega vrednotenja, pri tem pa lahko nevarno zaidemo na področje zagovarjanja pravičnosti nasilja [5].

Kljub temu, da obstaja splošno soglasje o tem, da nasilje deluje uničujoče in da za sabo vedno pušča žrtve, pa vseeno ostajajo vrzeli nedoločenosti, saj splošne definicije ne povedo, kdo ali kaj je bilo poškodovano in kako resna je ta poškodba (npr. ali je lahko neko agresivno vedenje zaradi težke psihiatrične diagnoze smatrano kot nasilje?). Učinkov nasilja, glede na kontekst, ni mogoče vedno jasno označiti kot pozitivne ali negativne [6]. Zdi se, da vsak poskus dokončnega določanja, kaj je nasilje, vedno nekomu dela krivico (ali osebi, ki je agresivna, ali pa žrtvi). Kar pa lahko z gotovostjo rečemo pri vseh poskusih definicije nasilja pa je, da se pri vseh pojavlja osnovno načelo prekoračitve meja – določiti, kaj pa so meje, pa se v dobi, ko se pojavlja opuščanje in relativizacija moralnih, spolnih, izobraževalnih, pravnih standardov, zdi spet zelo težka naloga [5].

[2] V pričujočem prispevku so izmenično uporabljeni različni izrazi, ki so velikokrat (po navedbah različnih avtorjev), razumljeni kot sinonimi: 'nasilje', 'agresivnost', 'agresivno vedenje'. V definicijah se pojma 'nasilje' in 'agresivnost' praviloma istočasno pojavljata in nista vedno razmejena ali pojasnjena. V prispevku tako uporabljamo izraza 'nasilje' in 'agresivno vedenje' kot sinonima, ki sta odraz zunanje manifestacije 'agresivnosti', ki pa označuje bolj notranje stanje oz. notranjo napetost v posamezniku in se nujno ne manifestira v zunanjem 'agresivnem vedenju' oz. 'nasilju'.

3. MODELI AGRESIVNOSTI IN AGRESIVNEGA VEDENJA

Zaradi omenjene problematike se kot izhodišče razumevanja in načrtovanja ukrepov za preprečevanja nasilja zdi bolj smiselno izhajati iz celostnih perspektiv razumevanja, kakršne predstavljajo različni modeli in ki vključujejo različne vzročne dejavnike tveganja za nasilje. Svetovna zdravstvena organizacija je npr. za razumevanje kompleksnosti narave nasilja izoblikovala t.i. *ekološki model* (angl. ecological) model, ki je dinamičen in odprt nenehnim dopolnitvam, saj upošteva interakcijo in spreminjanje različnih dejavnikov, ki vplivajo na povečanje tveganja za izvajanje nasilja, in dejavnikov, ki rizično vplivajo na to, da je nekdo žrtev nasilja. Ti dejavniki so razvrščeni na štiri ravni. Prva raven obravnava zgodovinske biološke in osebne dejavnike, ki vplivajo na ravnanje oseb in povečujejo možnost, da oseba postane bodisi nasilnež bodisi žrtev nasilja. Druga raven se omejuje na obravnavo bližnjih odnosov (npr. odnosi v družini, s prijatelji, intimnimi partnerji in ljudmi istega stanu) in raziskuje, kako ti odnosi povečujejo možnost, da oseba postane bodisi nasilnež bodisi žrtev nasilja. Tretja raven se osredotoča na organizirane skupnosti, v katerih potekajo družbene interakcije (npr. šole, delovna mesta, soseske) in išče značilnosti, ki povečujejo tveganje za izbruh nasilja. Četrta raven pa proučuje širše družbene dejavnike, ki oblikujejo klimo, ki nasilje bodisi spodbuja bodisi ga zavira [2].

Ekološki model razjasnjuje vzroke nasilja kakor tudi njihovo kompleksno vzajemno delovanje, zato predlaga, kaj je treba na različnih ravneh simultano storiti, da bi se nasilje preprečilo. Temu podoben je *splošni model agresivnosti* (angl. General Aggression Model – GAM), ki predstavlja celovit in integrativen model za razumevanje agresivnosti in agresivnega vedenja, nudi pa tudi razmislek o ključnih vidikih intervencij, ki bi lahko vodile k preprečevanju nasilnih in uničujočih odzivov [7].

Splošni model agresivnosti predpostavlja, da na človekovo agresivnost močno vplivajo zaznavne strukture, ki zaznamujejo veliko socialno-kognitivnih fenomenov, kot so percepcija, odločanje in vedenje. Nekatere teh najpomembnejših struktur vključujejo tudi prepričanja, sheme dojemanja, sheme pričakovanj in vedenjske skripte. Gre za strukture, ki se razvijajo preko izkušenj in lahko vplivajo na dojemanja na različnih nivojih, od preprostega dojemanja predmetov do kompleksnega dojemanja socialnih dogodkov [8].

Splošni model agresivnosti (slika 1) v razlaganju agresivnosti združuje dve skupini procesov: 1. *bližnji procesi* (angl. proximate processes) so tisti, ki so neposredno povezani z agresivnim vedenjem v danem trenutku; 2. *oddaljeni procesi* (angl. distal processes) pa so tisti, ki vplivajo na bližnje procese preko globljih tendenc in struktur (gre za aspekt, kako biološki in okoljski dejavniki sodelujejo in vplivajo na osebnost osebe) [9]. V teh procesih pa splošni model agresivnosti izpostavlja tri glavne faze oz. elemente interakcij med osebo in situacijo, ki se pojavljajo kot 1. *vnosi* (angl. inputs), 2. *poti* (angl. routes) in *rezultati* (angl. outcomes) [8]. Prva faza so vnosi, kjer gre za dejavnike dveh vrst. Prvi so povezani z osebo (*osebni dejavniki*, kot so spol, prepričanja, osebne lastnosti, vrednote, dolgoročni cilji, razpoloženje...) in predstavljajo stanje, v katerem se poveča možnost za agresivnost. Drugi dejavniki so povezani s situacijo (*situacijski dejavniki*, kot so provokacija, socialni stres, socialno zavračanje, frustracija, alkoholiziranost, nasilnost medijev, bolečina ali neudobje, visoka temperatura, hrup, prisotnost orožja, ogrožajoči ali strah vzbujajoči dražljaji...), v kateri se poveča možnost za agresivne reakcije [9]. Posamezni dejavniki ali kombinacija osebnih in situacijskih dejavnikov predstavlja izhodišče za povečevanje ali

znileževanje možnosti agresivnosti in agresivnega vedenja, saj vplivajo na kognicije, afekte in stopnjo telesnega vznemirjenja. V splošnem je tako, da več kot je dejavnikov tveganja za agresivnost, večja je možnost agresivnosti in agresivnega vedenja [10].

V drugi fazi se torej odvija vpliv vnosov (osebnih in situacijskih dejavnikov) na poti, ki vodijo v procese ocenjevanja in odločanja (agresivne ali neagresivne rezultate). Osebni in situacijski dejavniki lahko spremenijo *afektivno stanje osebe*, *kognicije* in stopnjo *telesnega vzburjenja*. Vsi ti trije dejavniki oblikujejo trenutno notranje stanje osebe in so med seboj prepleteni. Npr. afektivno stanje lahko vpliva na spoznavne procese in stopnjo fizične vznemirjenosti [11]. Jeza npr. lahko spodbudi nasilno razmišljanje in poveča telesno vznemirjenost. Prav tako lahko spoznavni procesi in vznemirjenje vplivajo na afektivno stanje (npr. če interpretiramo situacijo kot nasilno, se lahko poveča jeza, kar vodi v višjo stopnjo telesnega vznemirjenja) [8].
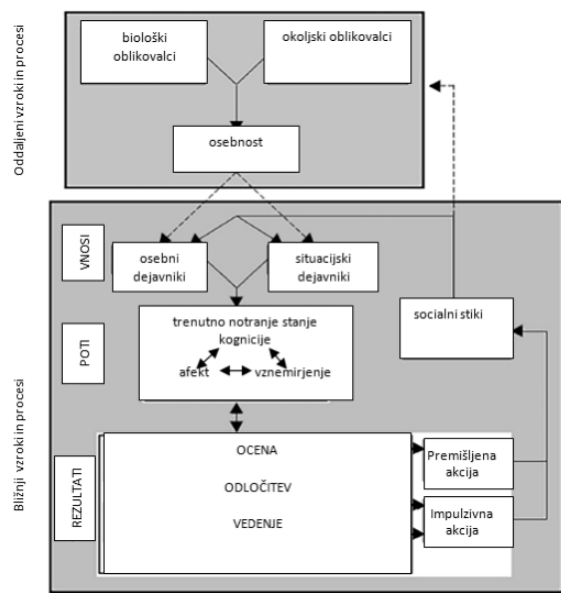
Slika 1: Splošni model agresivnosti: bližnji in oddaljeni vzroki in procesi [12].

Skupaj te poti ustvarijo notranje stanje, ki poveča verjetnost izida agresivnega vedenja [13]. To je tretja faza, kjer govorimo o *rezultatih* (procesi vrednotenja oz. ocene, ki so lahko tudi avtomatski, procesi odločitve in agresivno ali neagresivno vedenje). V tej fazi torej oseba ovrednoti situacijo in se odloči, kako se bo odzvala. Aktivnost, za katero se odloči, vpliva na socialne stike, ki povratno vplivajo na osebne in situacijske dejavnike, kar vodi v nov začetek cikla oz. epizode agresivnosti. Vsako ponavljanje epizode agresivnosti (ali neagresivnosti) služi kot učni proces, ki lahko vpliva na razvoj agresivnih zaznavnih struktur (in tako osebnosti) skozi čas [8].

4. PSIHONEVROBIOLOGIJA RAZVOJA KOT DEJAVNIK AGRESIVNOSTI

V zgoraj omenjenih dimenzijah agresivnosti in nasilja, ki jih opredeljujejo različni modeli in definicije, se zdi, da so nekateri dejavniki tveganja za nasilno vedenje bolj dostopni posameznikovi kontroli, drugi pa manj. Predvsem zaskrbljujoči so slednji, saj le na te posameznik težko neposredno vpliva, bodisi ker so vkoreninjeni v

splošne sheme prepričanja na družbeni ravni, ali pa ker predstavljajo nezavedne sfere njegove osebnosti. Glede na splošni model agresivnosti gre pri tem za oddaljene procese, kamor se uvrščajo biološki dejavniki (npr. ADHD, hormonsko neravnovesje, hipersenzitivnost itd.) in okoljski dejavniki (npr. kulturne norme, ki podpirajo nasilje, nefunkcionalne družine in starševstvo, težki življenjski pogoji, nasilne soseske in vrstniške skupine, nasilni mediji itd.), ki vplivajo na osebnost posameznika [14]. Predvsem pri okoljskih dejavnikih, kjer so vključeni odnosi z bližnjimi in predstavljajo okolje, v katerem se posameznik razvija, se zdi smiselno iskati tudi možnosti interveniranja, ki bi dolgoročno pripomoglo k preprečevanju kasnejšega agresivnega vedenja.

Zgodnje izkušnje s starši vplivajo na razvoj organske podlage naše osebnosti, in sicer se zapišejo v možgane, ki so dogovorni za procesiranje zavednih in nezavednih informacij skozi življenje. Med nosečnostjo in v prvih dveh letih življenja so vzpostavljeni temelji čustvenega delovanja in čustvenih virov odraslega človeka. Takrat se oblikujejo »socialni možgani«, del možganov, ki se uči, kako obvladovati občutke v odnosu z drugimi ljudmi in v različnih situacijah. Primarni odnos otroka, zlasti z materjo (ali drugim primarnim skrbnikom), tako zagotavlja osnovo za to, kako se bodo vzpostavili procesi regulacije afekta v otrokovih možganih. To bo tudi določilo posameznikovo večjo ali manjšo sposobnost za vstop v čustvena in socialna razmerja in situacije pozneje v življenju. Odnos med otrokom in odraslim tako zagotavlja prostor za razvoj otroka [15].

Ob rojstvu je otrok še »nedokončan«. Čeprav ima vse anatomske dele, pa je kot »surov material« z genetskim načrtom in edinstveno paleto možnosti, ki se morajo na določen način še razvijati skozi izkušnje, zlasti socialne [16]. Izkušnje v zgodnjem življenju so lahko še posebej pomembne pri organiziranju načina razvijanja osnovnih regulativnih struktur v možganih. Vsakodnevne izkušnje tako oblikujejo strukturo možganov, ker aktivirajo določene poti v možganih, krepijo obstoječe povezave in ustvarjajo nove. Psihonevrobiološka perspektiva razvoja poudarja, da je zgodnje otroštvo obdobje, ko je določena osnovna »arhitektura« v možganih [17].

Zgodnje socialno okolje (odnosi s primarnimi skrbniki) tako neposredno vplivajo na končno vzpostavljenost krogotokov v možganih otroka, ki so odgovorni za prihodnje družbeno in čustveno življenje. Odnosi navezanosti neposredno oblikujejo zorenje otrokove desne hemisfere, posebej sisteme prefrontalnega dela desne hemisfere, ki so odgovorni za regulacijo čustev, vključno s pozitivnimi čustvi (npr. veselje in zanimanje), pa tudi negativnih čustev, kot sta strah in agresija [18]. Vzorci varne navezanosti vodijo tako v organiziranost primernega regulatornega sistema v prefrontalnih območjih desne hemisfere, učinkovito delovanje v tem delu možganov pa je osrednjega pomena za otrokovo naraščajočo sposobnost samoregulacije, sposobnosti fleksibilnega uravnavanja stresnih čustvenih stanj zaradi interakcij z drugimi ljudmi in samoregulacije v avtonomnih kontekstih. Gre za način, kako varni odnosi navezanosti (epizode nežnosti, uglašenosti, sinhronije) spodbujajo razvoj desne možganske hemisfere, pospešujejo učinkovito regulacijo afekta in gradijo temelje otrokovega psihičnega zdravja. Skrbnikova senzitivnost je tako zunanji organizator otrokove biološke podlage vedenjske regulacije. Na podlagi tega se v določenem trenutku razvije tudi samoregulacija otroka, ko se v relacijskem okolju dogradijo notranji regulatorni mehanizmi [18].

V primeru slabih relacijskih okvirov razvoja lahko govorimo o relacijski travmi, ki v nevrobiološkem delovanju otroka pušča povsem drugačne vzorce, saj negativno vpliva na od izkušenj odvisne razvojne procese v delovanju možganov. To pušča dolgoročen negativen vpliv

na sam razvojni proces in tudi strukturo osebnosti posameznika [19]. Travmatični odnosi in čustveno zanemarjanje negativno vplivajo na strukturo in funkcijo desne hemisfere, subkortične strukture in frontalni korteks [18]. Desna možganska hemisfera in subkortične strukture so temeljnega pomena v čustvenem procesiranju, prefrontalni korteks pa igra pomembno vlogo pri regulaciji afekta [20]. Tako ni čudno, da ti ljudje tudi kot odrasli trpijo posledice, ki uničujoče delujejo v njihovem življenju in v življenjih njihovih bližnjih. V poznejših življenjskih obdobjih so otroci z varno navezanostjo v večji meri sposobni razviti in vzdrževati dobre bližnje odnose, zlasti s starši, vrstniki in partnerji kot ne-varno navezani otroci. Pri varno navezanih se v otroštvu, najstništvu in v odraslosti razvijejo različne bolj zaželene osebnostne lastnosti, pri njih je večja verjetnost prisotnosti konstruktivnih oblik čustvovanja in čustvene samoregulacije, izražajo bolj pozitivno samospoštovanje [21]. Različne motnje v odraslosti imajo svoje korenine v otroštvu, npr. motnje hranjenja, zloraba substanc, preveč impulzivno vedenje, pomanjkanje empatije, depresivnost, tesnobnost, agresivnost itd., kar je vse posledica tega, da zaradi slabe razvitosti čustvene regulacije posamezniki niso sposobni primernega upravljanja s svojimi občutki [16]. V povezavi s tem je pomembno poudariti, da so posamezniki, ki niso sposobni primerne regulacije afektov, kot je jeza in druga afektivna stanja, bolj nagnjena k agresivnemu vedenju, da bi s tem spremenili, odstranili ali se izognili neprijetnim čustvenim stanjem (vzniku agresivnega občutja) [22]. Gre za stanje t.i. afektivne disregulacije, o kateri govorimo, ko oseba ni sposobna primerno občutiti čustev, ko jo čustva prevzamejo ali ko ostanejo nerazrešena [23].

Z vidika psihonevrobiološke perspektive razvoja kakovost odnosov s starši v zgodnjem otroštvu vpliva na razvoj možganskih področij, ki so odgovorna za različne vrste regulativnih strategij. Šibko razvit prefrontalni korteks je značilen pri številnih psihopatoloških in negativnih vedenjskih strategijah. Brez močnega prefrontalnega korteksa mehanizmi samokontrole, pomirjanja sebe in občutka povezanosti z drugimi, ne dozorijo [16]. V nezmožnosti uravnavanja močnih čustev zaradi nerazvitih procesov regulacije je tako mogoče prepoznati vir nasilja [24]. Do tega pride zato, ker nedozoreli orbitofrontalni sistem ni sposoben regulirati impulzov (tudi agresivnih) iz nižjih limbičnih struktur, kar povečuje možnost agresivnega vedenja, ki predstavlja neprimerno regulacijo afektivnih stanj, ki vzniknejo v določeni situaciji (jeza, sram, strah...) [18].

5. SKLEP

Sposobnost regulacije čustev (tudi regulacija agresivnih impulzov) od otroštva naprej služi kot zaščitni faktor pred agresivnim vedenjem posameznika. Čeprav se ta sposobnost lahko razvija in dopolnjuje še kasneje, se v prvih dveh letih življenja oblikujejo osnovni vzorci za soočanje z negativnimi in težavnimi čustvenimi situacijami. Človeška bitja so odprti sistemi, ki jih oblikujejo drugi ljudje. Človeški psihološki in duševni sistemi se razvijajo v odnosu z drugimi ljudmi - in to se najbolj intenzivno dogaja v otroštvu. Lahko rečemo, da je človeški otrok najbolj družbeno vplivano bitje (ali biološki sistem) na zemlji [16]. Biološki in socialni dejavniki so v človekovem razvoju zelo prepleteni in povezani, še posebej v obdobju zgodnjega razvoja. Seveda imajo genetika in okoljski dejavniki vpliv na dojenčke, vendar so socialni dejavniki, zlasti senzitivna nega in skrb skrbnika tisti, ki vplivajo in dejansko spremenijo biološke elemente (kot sta kemija možganov in možgansko tkivo). Zelo pomembno je, da ne zanemarjamo razmerja med biologijo, izkušnjami in vedenjem, saj je to lahko škodljivo in uničujoče za našo celotno družbo [24]. Zato je v povezavi s tem bistveno razmisliti in razviti ustrezne programe za zgodnje ukrepanje, ki bi lahko delovali tudi kot zaščitni dejavnik proti

tveganju agresivnega vedenja in nasilja. Tovrstno ukrepanje, ki bi spodbujalo zdrave odnose med primarnimi skrbniki in otrokom (preprečevanje relacijske travme in spodbujanje učinkovitega in senzitivnega starševstva), se mora začeti še pred otrokovim rojstvom in ga je treba podaljšati skozi perinatalno in postnatalno obdobje. Programi za učinkovito starševstvo (na katerem koli področju - pediatrija, psihiatrija, psihologija, socialno delo, izobraževanje, pastoralna oskrba ...) lahko ustvarijo razvojni kontekst za preoblikovanje ne-varne navezanosti v varno in na ta način spodbudijo od izkušenj odvisno nevrobiološko zorenje desne hemisfere, ki je ključnega pomena pri primerni regulaciji različnih motivacijskih (afektivnih) stanj, vključno z agresivnimi [18]. Dejstva, ki jih je razkrila (nevro)znanost v preteklih desetletjih, so jasna in jih ni mogoče zanikati: obstaja velika priložnost za nekaj, kar bi posamezniku omogočilo optimalen začetek, da bo bil kar najbolje čustveno opremljen za soočanje z izzivi življenja.

6. REFERENCE

[1] Cerar, J. (2017). *Nasilje žensk v partnerskem odnosu skozi prizmo relacijske družinske paradigme.* Doktorska disertacija. Teološka fakulteta Univerze v Ljubljani.

[2] Svetovna zdravstvena organizacija. (2002). *World report on violence and health: summary.* Geneva: World Health Organization.

[3] Hampton, R. L. (1999). *Family Violence.* London: Sage.

[4] Tolan, P. H. (2007). Understanding Violence. V: D. J. Flannery, A. T. Vazsonyi in I. D. Waldman (ur.), *The Cambridge Handbook of Violent Behavior and Aggression* (str. 5-18). Cambridge: Cambridge University Press.

[5] Heitmeyer, W. in Hagan, J. (2003). Violence: The Difficulties of a Systematic International Review. V: W. Heitmeyer in J. Hagan (ur.), *International Handbook of violence research* (str. 3-11). New York: Kluwer Academic Publishers.

[6] Imbusch, P. (2003). The Concept of Violence. V: W. Heitmeyer in J. Hagan (ur.), *International Handbook of violence research* (str. 13-39). New York: Kluwer Academic Publishers.

[7] Baron, R. A. in Byrne, D. (1995). *Social psychology: Understanding Human Interaction.* New Delhi: Prentice Hall.

[8] Allen, J. J, Anderson, C. A. in Bushman, B. J. (2018). The General Aggression Model. *Current Opinion in Psychology, 19,* 75-80.

[9] Anderson, C. A. in Groves, C. (2013). General aggression model. V: M. S. Eastin (ur.), *Encyclopedia of Media Violence* (str. 182-187). Los Angeles: Sage.

[10] Gentile, D. A. in Bushman, B. J. (2012). Reassessing media violence effects using a risk and resilience approach to understanding aggression. *Psychology of Popular Media Culture, 1*(3), 138-151.

[11] Bower, G. H. (1981). Mood and memory. *American Psychologist, 36*(2), 129-148.

[12] Allen, J. J. in Anderson, C. A. (2017). General Aggression Model. V: P. Roessler, C. A. Hoffner in L. van Zoonen (ur.), *International Encyclopedia of Media Effects.* Hoboken: Wiley-Blackwell.

[13] Bishmi, P. D. (2017). A comparative study on level and direction of aggression in high and low rejection sensitive rejection. *Indian Journal of Health and Wellbeing, 8*(7), 729-733.

[14] Anderson, C. A. in Carnagey, N. L. (2004). Violent evil and the General Aggression Model. V: A. G. Miller (ur.), *The Social Psychology of Good and Evil* (str. 168-192). New York: Guilford Press.

[15] Simonič, B. in Poljanec, A. (2014). Building Motherhood in the Young Mothers' Group. *Child Care in Practice, 20*(3), 270-285.

[16] Gerhard, S. (2004). *Why Love matters.* London: Routledge.

[17] Siegel, D. J. (2012). *The Developing Mind.* New York: Norton.

[18] Schore, A. N. (2003). Early relational trauma, disorganized attachment, and the development of a predisposition to violence. V: M. F. Solomon in D. J. Siegel (ur.), *Healing trauma* (str. 112-113). New York: W. W. Norton & Company.

[19] Repič Slavič, T. in Gostečnik, C. (2017). Relational Family Therapy as an Aid Toward Resolving the Trauma of Sexual Abuse in Childhood in the Process of Separation in the Couple Relationship. *Journal of Marital and Family Therapy, 43*(3), 422-434.

[20] Fosha, D. (2003). Dyadic regulation and experiental work with emotion and relatedness in trauma and disorganized attachment. V: M. F. Solomon in D. J. Siegel (ur.), *Healing trauma* (str. 221-281). New York: W. W. Norton & Company.

[21] Thompson, R. A. (2008). Early Attachment and Later Development Familiar Questions, New Answers. V: J. Cassidy in P. R. Shaver (ur.), *Handbook of attachment* (str. 348-365). New York: Guilford.

[22] Roberton, T., Daffern, M. in Bucks, R. S. (2012). Emotion regulation and aggression. *Aggression and Violent Behavior, 17*(1), 72-82.

[23] Heller, L. in LaPierre, A. (2012). *Healing devlopmental trauma.* Berkley: North Atlantic Books.

[24] Karr-Morse, R. in Wiley, M. S. (2013). *Ghosts from the Nursery.* New York: Atlantic Monthly Press.

Emergence of Visual Consciousness in ADHD Children

Katja Zupanič
University of Ljubljana
Maribor
katja.zupanich@gmail.com

Rawan Y. Fakhreddine
University of Texas at Austin
Austin, Texas
rawan-f10@utexas.edu

Zoltán Nádasdy
Eötvös Loránd University
Budapest
zoltan@utexas.edu

ABSTRACT

The nature of consciousness has excited the imagination of scientists and researches for years. In the past, some were of the opinion that a single part of the brain is important fort he emergence of consciousness, while some disagreed, one of the first being the psychologist William James, who claimed that what is important for its emergence is the function of the brain as a whole and that parallel function of different parts is key. In our study we focus mainly on the research of visual consciousness. We want to find out whether there is a difference in the emergence of visual consciousness between healthy participant (children ages 6-16) and children diagnosed with ADHD. We attempt to measure the emergence of visual consciousness using an experimental paradigm designed by dr. Zoltán Nádasdy and his team.

Keywords

visual consciousness, ADHD, visual attention, visual integration

1. INTRODUCTION

The nature of consciousness has puzzled mankind since the beginning of time. What is consciousness, how can our physical brain interpret subjective experience, what are qualia? The question also presents itself as in where in the brain does consciousness lie. All these are questions that scientists have been posing for years and still today.

What is consciousness, where is consciousness formed, which neural correlates are responsible for consciousness, all these are questions that scientists and philosophers are occupied with today and were occupied with in the past. The search for rigorously defined neural correlates has been going on for decades.Some recent studies these days identify the thalamus as a central structure implicated in conscious awareness [4]. The thalamus works as a relay station or a passageway of sensory information into certain parts of the cerebral cortex [4], which means that all information we are consciously aware of (except olfactory) meets the thalamus at some point.

The study we present in the paper, tests visual awareness using an experimental paradigm developed by dr. Nádasdy and colleagues. The task comprised visual integration of image fragments presented on a computer screen. The project connects the neuroscientific and psychological aspects of visual consciousness.

We assume a critical point in time when visual consciousness (of recognition of an image as a whole or non-whole) emerges and we posit that this time of emergence is delayed with certain developmental or neurological (neuropsychological) disorders, namely in ADHD children.

2. THEORETICAL BASIS

As the body of neurological knowledge has dramatically grown in the past few decades, so too has grown the interest in and search for neural correlates of consciousness. While some were occupied with the search of a common center of the brain responsible for conscious awareness, others search for multiple regions or systems that play an important role in the production of consciousness. William James was one of the first advocates of this position. He claimed that consciousness is a result of the common effort of the whole brain (Principles of Psychology, 1890/1950, as cited in [3]). However, an increasing quantity of empirical evidence shows that some parts of the brain are much more important than others when it comes to the production and expression of conscious awareness. These are mainly the thalamocortical system and the reticular system of the brain stem [3].

In the search for brain regions involved in the creation of consciousness there is a structure that has gathered a lot of attention – the thalamus. The thalamus is a structure in the midbrain found at the top of the brain stem. In the mid-1800 the idea already appeared that the thalamus is the sensory center of the brain. Neuroanatomical examinations of patients showed that thalamic regions in an otherwise unaffected brain caused major sensory dysfunction. These neuroanatomical findings showed that the thalamus plays a key role in the processing of sensory information [4]. Later neuroanatomical studies of the thalamus and its connections to other parts of the brain highlighted its function in sensory processing and integration. It is important to note that it was discovered that all sensory information travels through the thalamus before reaching the corresponding parts of the cortex [3].

Francis Crick [1] contributed much to the literature on consciousness by suggesting that the TRN (thalamic reticular nucleus) and the dorsal thalamus play a key role in consciousness by means of controlling and expressing an internal attentional searchlight as first suggested by Treisman and colleagues (1991, as cited in [1]). Crick suggested the TRN produces rapid firing in thalamic relay neurons to control the brain's attentional searchlight, acting on Malsburg synapses to create conjunctions of neurons, expressed by cell assemblies, which Crick believes to be the neural expression of the attentional searchlight.

By researching attentional capacity of the brain at tasks comprising different types of search, researchers noted that the brain searches consecutively, focusing on one object at a time, before moving on to the next [1].

In his attempt to explain how the thalamocortical system controls attention, Crick suggested the following mechanism: when sensory signals pass through the TRN on their way to the cortex, this causes excitation. Crick suggested that the TRN and dorsal thalamus work together to increase activity at the most active thalamocortical pathways and at the same time reduce activity in less active parts of the thalamocortical network. He described this mechanism as "… the heating up of the warmer parts of the thalamus and the cooling down of the colder parts" [1] (page 4587).

3. HYPOTHESIS

Our hypothesis is that there is a certain amount of information needed to be able to put together image fragments with confidence (confidence here not referring to the reported level of confidence by subjects). Hence, the decision point to reach a certain level of confidence for a given image is consistent across healthy subjects. This point can be standardized for a selected set of pictures and the test can be validated. Having the image set standardized, the test should be sensitive to disorders implicated in consciousness, attention, mental focus, and visual integration processes such as functions compromised in ADHD. We also wish to find out whether our test is appropriate for the detection of such disorders.

The main research goal is to test the hypothesis and based on this construct a theoretical framework describing the differences at the emergence of visual consciousness in healthy children and children with ADHD.

4. THE EXPERIMENT

4.1 Method

Participants sit down in front of a computer screen on which four very similar pictures appear, divided by black lines. The participant decides whether the images compose one joint picture or whether they are different. They record their decision by pressing a key on the computer keyboard. At every keystroke the lines between the images thin until the entire picture is revealed and it becomes clear whether the images form a whole. The pictures include themes from everyday life (houses, cars, landscape, furniture etc.), various kinds of animals and abstract pictures, like pictures of the universe, microscopic cells etc.

In our study, visual consciousness is experimentally tested using a between-subjects design. The independent variable is the manipulation of trials. During solving time, the task has seven levels equivalent to every phase of testing. The dependent variable is the judgment of participants about the continuity of the images presented in every phase (by pressing a computer key).

4.2 Progress

So far 48 children have taken part in the study. The control group consists of 20 students of the Kidričevo elementary school, (10 female, 10 male), aged between 8 and 9 years. The participants were primary school children that took part in the research study voluntarily with parental or legal guardian consents. There are 28 children (27 male, 1 female) in the target group, 8 children age 6-8, 15 age 11-14 and 5 age 10-12.

The criterion for inclusion was an age of above 5 years and under 15 years. The final sample of the group of all participants was planned to encompass at least 50 children (25 tested children in the control group and 25 tested children with a psychiatric ADHD diagnosis).

4.3 Procedure

The test is composed of 100 trials, presented in series on a computer screen. In every trial, 4 similar images were shown on the screen, divided by black borders slightly covering the images (in the shape of a cross). Every trial included eight phases where the borders between images diminish in 7 steps from thick to very narrow. Lastly, in the 8th phase, it is revealed whether the four fragments join into a whole or remain different.

By pressing the "S" key it is meant that the subject believed that the presented fragments are part of a continuous picture, pressing "R", the person believed that the images are part of four different pictures and pressing "N", the subject was unsure about the continuity of the presented image fragments.

An exit command was also added to the program in case any participant wanted to stop before finishing the task. In this case, the participant had the option of pressing the "escape" key on the keyboard, which would save their data and immediately leave the program.

4.4 Expected Results

We call the moment during the task when subjects suddenly see and decide that the picture fragments are merging an "aha" experience. We define this moment as the emergence of visual consciousness. The moment when different subjects reach this level of visual consciousness, we predict, should depend on how effectively they process and analyze image fragments. Children diagnosed with ADHD may not be able to integrate the fragments of visual information as effectively as children with normal cognitive control do. Hence, with this task we aim to quantify the execution of visual consciousness.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] F. Crick. Function of the thalamic reticular complex: the searchlight hypothesis. *Proceedings of the National Academy of Sciences*, 81(14):4586-4590, 1984.

[2] F. Crick, C. Koch. Towards a neurobiological theory of consciousness. *Seminar in the Neurosciences*, 2: 263-275, 1990.

[3] G. Edelman. Consciousness: The remembered present. *Annals of the New York Academy of Sciences*, 929(1): 111-122, 2006.

[4] E. G. Jones. *The Thalamus*. Plenum Press. New York, NY 1985

European Legal Regulation of Self-learning Systems

Tjaša Zapušek
University of Copenhagen
Faculty of Law - LL.M.
Njalsgade 76, København, Denmark
tjasa.zapusek@gmail.com

ABSTRACT

Modern technology often gives us an impression of a better life, yet simultaneously raises new risks that have not been present ever before. This paper will present differences between three different 'decision-making processes' that occur as a consequence of three different types of systems, namely, AI can be equipped with a deterministic algorithm, a supervised, or an unsupervised learning algorithm. Through this systematization, concerns regarding foreseeability and certainty of the system's actions will be examined. The scope of the thesis is limited to foreseeability regarding harm caused by physical machines or products, equipped with advanced algorithms. This paradox raises questions regarding the role of scientists in situations when they are facing uncertain risks, and legal scholars when regulating advanced technology.

Keywords

self-learning systems, algorithms, strict liability, development risk defence, product liability, product safety, machine learning

1. INTRODUCTION

Presence of artificial intelligence (AI) in our everyday life will call into question the efficiency of existing legal rules. Yet before we focus on regulating liability concerning the harm caused by a machine equipped with advanced algorithms we should focus on the question 'is risk assessment is even possible, can we even predict (with existing scientific knowledge) potential consequences of these machines?'

When policy makers and judicial authorities are confronted with an uncertain situation they turn to experts for definite answers and conclusive evidences, even though uncertainty and absence of foreseeability clearly precludes definitiveness and conclusiveness.

The following research questions were developed in order to define the research problem and further, to address the gaps in existing legislation.

(i) What is the nature of algorithmic uncertainty?

(ii) How does the European legislation address the concept of uncertainty and un-foreseeability?

2. ALGORITHMS

The algorithmic breakthrough is often connected with Alan Turing, whose one of the most ground-breaking achievements was breaking the German Enigma code during the Second World War. [Roberts 2016] In 1952, he published a set of mathematical equations through which he wanted to explain the pattern we see in nature, such as zebra stripes and leopard spots. [Serna 2017] These algorithms are still in use when describing patterns that emerge in

nature. [Deangelis 2014] On 20 February 1947, Alan Turing had a lecture at the London Mathematical Society, where he enlightened the academic public with the statement that "what we want is a machine that learns from experience".[Press 2017, 1] Alan Turing concluded his lecture with these words: "Putting the same point differently, the machine must be allowed to have contact with human beings in order that it may adapt itself to their standards." [Turing 1947, 14]

In these few decades, since Alan Turing first introduced machine learning and algorithms, the latter have become a big and a ubiquitous part of our lives. They are used for data processing, calculations and automated reasoning. [Deangelis 2014]

"To make a computer do anything, you have to write a computer program. To write a computer program, you have to tell the computer step by step, exactly what you want it to do. The computer then 'executes' the program, following each step mechanically, to accomplish the end goal. When you are telling the computer what to do, you also get to choose how it's going to do it. That's where computer algorithms come in. The algorithm is the basic technique used to get the job done". [Deangelis 2014, 1] Pursuant to Deangelis, the point where this explanation could be slightly adjusted is where the text indicates that 'you have to tell the computer step by step what you want it to do'. [Deangelis 2014] Instead of following pre-defined and programmed instructions, some computer algorithms are designed in a way that enable a computer to learn on their own, in fact they facilitate machine learning which may result in a creation of unique corrections between obtained data that sometimes 'produce' unforeseeable outcomes [Deangelis 2014] Legal concerns that have been emphasized in this thesis can be summarized with the statement by Dr. Panos Parpas, who is a lecturer at Imperial College London, that reservations of academics and people around the world are not directed towards the algorithms per se, but towards the structure of a society and its ability/inability to cope with algorithms (data use). [Hickman 2013] At the moment, we are witnessing an awkward marriage between algorithms and data. [Deangelis 2014] Together with technological progress there will be mistakes, but it should be taken into account that machines are just a tool and tools should not be blamed. [Deangelis 2014] Learning as such can present either acquiring or enhancing existing knowledge. As stated by Herbert Simon, machine learning refers to adaptive changes in a particular system that enable the system to perform the same task (or tasks which have been drawn from the same population) more precisely and effectively next time. [Sathya 2013] The text below will introduce different types of algorithms and their main features. Through this systematization, concerns regarding certainty and foreseeability of their actions and consequences will be presented.

2.1. Deterministic algorithms

This algorithm is constructed to prevent a machine from making autonomous decisions, creating new patterns, and consequently creating unpredictable outcomes. [Zapušek 2017] Robots equipped with deterministic algorithms follow predefined paths. [Hildebrandt 2016] The major significance that distinguishes the first group of algorithms from the others is that the final action, machine's response or steps have already been predicted by a human being. [Hildebrandt 2016] In the case of a deterministic algorithm, we will always find someone liable for a robot's actions – as it will be presented in the text bellow, existing legal rules concerning the liability of the producer will suffice.

2.2. Supervised learning algorithms

More advanced forms of algorithms can enable machine learning and supervised learning algorithm is one of them. [Hildebrandt 2016] In the case of supervised learning algorithms, a machine is fed with a so-called training set by human supervisors, which consists of particular data and predefined patterns that provide some sort of a background on what counts as a desired and satisfactory solution or outcome. [Ploj 2013] Hence, in the case of supervised learning, the machine already knows the output of the algorithm, before it even starts learning. [Van Loon 2018] Since the outcome is already known, all that a system needs to do is to execute steps and processes that are needed to get from input to the desired output. In a situation, when algorithms produce completely different and unexpected results, training data serves as a guidance to steer the algorithm back towards the right path. [Van Loon 2018] For this reason, a supervised model can also be called error back-propagation algorithm. [Van Loon 2018] Such error correction-learning algorithms have been thought to train the network based on the input and output samples and to find error signals, which denote the difference of the desired output and output calculated. [Van Loon 2018]

2.3. Unsupervised learning algorithms

This one is, compared to a supervised learning algorithm, not (yet) as widespread and frequently. [Brummette 2017] Although the unsupervised learning algorithm has not been used on a wider scale yet, its capabilities and methodology represent the future of machine learning. [Van Loon 2018] A good illustration of this type of machine learning would be imagining a child who has just started discovering the world. [Brummette 2017] After his first interaction with a four-legged animal, he might hear someone call out the word "dog". [Brummette 2017] After a while, when the child sees other four-legged animals, such as a cow, a cat, or a horse, he thinks they are all dogs. [Brummette 2017] "This is because the natural classification methods installed in a human brain informed him that the trait 'four legs' is associated with a specific animal type. As the child grows and sees more four-legged animals, additional detailed classifications emerge. Dogs, cows, and horses are all discovered to have distinct traits and become a subset of four-legged animals in the child's mind." [Brummette 2017, 1] The idea of comparing machine learning process with a child's brain was also presented by Alan Turing in his paper 'Computing machinery and intelligence'. [Turing 1950] He compared a child's brain with a notebook. According to him there are "rather little mechanisms, and lots of blank sheets". [Turing 1950, 19] Hence, heading back to scientific conclusions about the unsupervised machine learning technique, which makes correlations between obtained data without previously hypothesizing them. [Zapušek 2017] [Hildebrandt 2016] Without the provision of a training set, consisting of data and predefined patterns, the machine cannot familiarize itself with information and what counts as a desirable solution. [Zapušek

2017] A machine that is in the process of unsupervised learning finds input data and classifies it (according to its own criteria) in different categories. [Ploj 2013] We use unsupervised learning algorithms to identify hidden patterns and unlabelled input data, [Sathya 2013] since they are capable of mining new data and creating novel, unexpected patterns. [Ploj 2013] The system is blinded when it goes into an operation. [Van Loon 2018] It carries out its own immense and faultless logical operations that serve as guidance, but still due to the lack of exact input and output data (the system has no reference data at all) the transparency of the process and steps is not clear. [Van Loon 2018] Despite the lack of transparency this algorithm has the powerful ability to interpret limitless amount of data, through its binary logic. [Van Loon 2018]

3. UNCERTAINTY

Uncertainty, as a concept, has been around for a long time, Its roots stretch back to Plato and Socrates, who exposed their doubt whether scientific knowledge, no matter how elaborated and comprehensive, reflects reality. [Tannert 2007] Kant and Prauss extended the idea about uncertainties with the statement that the more detail we have about the mysteries of nature, the more we become aware of limitations of our knowledge about what things as such are, how do they work, etc. [Kant 2013] Limitations to our understanding are the main reason for our inability to foresee future events and their effects. [Tannert 2007] When policy makers and judicial authorities are confronted with an uncertain situation they turn to experts for defined answers and conclusive evidences, even though uncertainty clearly precludes definitiveness and conclusiveness. [Fox 2009] The paradox causes questions regarding the role of scientist and scholars in situations when they are facing uncertain risks when regulating advanced/modern technology. In the analysis of cases of EU risk regulations it has been observed that the uncertainty paradox leads to non-effective and unintelligible policy-making processes. [Fox 2007] As emphasized by Mourik [Mourik 2004], without sufficiently designed infrastructure, not only will questions concerning responsibility eventually emerge, the non-sufficient regulation may also lead to a state called 'organized irresponsibility'. [Fox 2009] Ulrich Beck uses this term when society is ill-prepared and is consequently unable to effectively respond to the "inevitable surprises, negative consequences and/or long-term impacts associated with uncertain risks". [Fox 2009, 1] The idea behind 'risk society' suggests that a world has become more hazardous, yet this is not entirely true. [Giddens 1999] [Beck 2009] The society has become increasingly preoccupied with the future and safety that generate the notion of risk. [Giddens 1999] With the term 'risk' we are describing potential events with particular consequences which are evaluated as negative. [Fox 2009] The economist Knight emphasizes that risk and uncertainty are not synonyms per se. He views a risk as calculable, hence controllable island in the ocean of uncertainty. [Van Asselt 2006] Risk will usually be applied to the situations where all potential future outcomes can be specified, uncertainty as a bare term can be applied in the situations where all potential future outcomes cannot be defined or specified before their occurrence. [Faber 1992] It is natural that whenever we make a decision that contains unpredictable factors, we need to weight not only potential outcomes, but also their benefits and risks. The scale of potential consequences of decisions that are carried out on a national, European or even international level have a much greater influence on people in comparison with individual decisions. Uncertainty as such has no ethical quality, it is inherent to an individual situation. [Tannert 2007] But, if there are concerns regarding dangerousness of a situation, uncertainty itself may trigger ethically adjusted behavior, the main purpose of which is

the avoidance of danger and diminution of existing risks – which might be the case when the society is confronted with products, equipped with self-learning algorithms that lead product's actions. [Tannert 2007] "When it comes to decisions that affect people's lives and health [...] carrying out research to diminish uncertainty and, consequentially, risks can become an ethical duty." [Tannert 2007, 892]

4. FORESEEABILITY IN TORT LAW

Cosmologists suggest that the Universe is comprised largely of 'dark matter', for ninety per cent of matter in the Universe does not glow, but is dark. [Baker 2010] Even though we cannot directly see dark matter, we can detect its mass through its gravitational pull on other astronomical objects in the Universe. [Baker 2010] This invisible stuff is powerful enough to bind all things together. According to professor David Owen, foreseeability is the "dark matter" of tort which connects its components, and "gives moral content to the law of negligence, controlling how each element fits together and, ultimately, whether one person is bound to pay another for harm" [Owen 2009, 1277]. From a legal perspective the test of foreseeability can be described with a sentence "whether one can see a systematic relationship between the type of accident that the plaintiff suffered and...the defendant's (by someone/something defendant is responsible for) wrongdoing." [De Villiers 2015, 344] Considering the aspect of foreseeability a defendant may escape liability if scientists could not predict (ex ante) a systematic relationship between wrongdoing and harm. [De Villiers 2015] For many years, in civil law systems of most European countries, tort liability was based on a broadly defined general clause of fault liability. [Werro 2004] However, in the 19th century, Europe was confronted with an increase in technical and industrial risks. [Werro 2004] For this reason, the majority of European legal systems established liability rules which provide some form of strict liability. [Werro 2004] A distinction between fault liability and strict liability lies in their conceptual levels. "The distinctive factor is whether or not liability rests on the judgment that the defendant should have behaved otherwise than he did". [Werro 2004, 7] In order to decide whether a defendant acted with fault, we need to determine all relevant standards of conduct and then establish that a defendant did not meet these standards. [Werro 2004] The determination regarding the standard of care is based on objective criteria. [Werro 2004] Legal systems of other European countries on the other hand focus primarily on the so-called 'what could reasonably be expected of the defendant' standard. [Werro 2004] Therefore, we may conclude that the limitation of negligence liability to so-called foreseeable harm prevents us to hold a defendant liable for not taking precautions with respect to harm which was not reasonably foreseeable. Fault liability can be explained in one sentence as a liability for reasonably foreseeable and avoidable harm. [Werro 2004] Compared to fault liability, strict liability is "concerned with the precautionary consequences of harm not attributable to a lack of precaution on the part of the defendant, or with what may be referred to as 'accidental harm'." [Werro 2004, 9] Besides liability for reasonably foreseeable and avoidable harm, it can also extend to unforeseeable and unavoidable harm. [Werro 2004] [Viney 1998] Even though the concept of reasonably foreseeable and avoidable harm is primarily used in negligence cases, it would be incorrect if we claimed that probability and foreseeability issues are irrelevant to strict liability.

5. DEVELOPMENT RISK DEFENCE

Existing legislation of the majority of European member states regulates strict liability in the way that excludes liability for harm if the latter lies beyond certain limits of foreseeability. Legal

regulation on the European level follows the same pattern. The EC Directive on Liability for Defective Products (hereinafter: Liability Directive) [Council Defective Products], adopted in 1985 ensures that producers take responsibility for their products vis-a-vis consumers. It was one of the first pieces of European legislation that explicitly focused on the consumers' protection and which introduced the concept of strict liability, where producers are held liable regardless of whether the defect is their fault. In its article 7(e) the producer can escape liability for harm caused by the defective product by showing that the state of scientific and technical knowledge at the time when they put the product into circulation was not such as to enable the discovery of the existence of the defect. [Council Defective Products] The Liability Directive introduced a new common scheme of strict liability that in comparison with the law of negligence does not require a producer's negligent behaviour. [Alaimo 2014] In negligent cases the claimant has to prove that the defendant knew/ought to have known about the present risk. [Van Dam 2013] However, the new scheme presented by the Liability Directive sets the requirement in its Article 4 regarding proof of the existence of a product's defect, harm, and a causal link between these two. [Council Defective Products] The creators of the Liability Directive wrote the provisions that allow national legislations to limit and dilute strict liability. [Alaimo 2014] One of these provisions is also the "development risk defence" which is still considered one of the most controversial parts of this Directive. [Elliott 2011] The latter is very similar to the 'state of art' defence that has been developed in negligence cases. [Arbour 2014] This so-called development risk defence is enshrined in previously mentioned Article 7(e) of the Liability Directive under which the producer can escape liability for harm caused due to lack of knowledge. According to Stapleton, the above mentioned Article does not refer to risks of a product, rather it is directed at the discoverability of the defect itself. [Stapleton 1994] Consumer associations' opinions believed that it is crucial to protect consumers from unknown and unforeseeable risks, and that the adoption of the mentioned defence (and consequently the exclusion of liability for such risks) would create a gap [Petitpierre 1974] in a general protection of consumers. [Alaimo 2014] The majority of these associations claimed that the development risk defence weakens the principle of strict liability, and due to un-foreseeability it puts an unreasonable burden on consumers. [Alaimo 2014] Producer associations' hold the opposite position and argue that the exclusion of such a defence would significantly discourage scientific and technical research and prevent marketing of new high tech products. [Alaimo 2014] The idea behind the protection of producers against development risks is that despite taking into account all available scientific and technical knowledge it is impossible for the producer of the product to foresee the risk in question. [Van Dam 2013] To sum up, producers cannot be considered negligent if they have (within reasonable limits) analyzed and gathered all available scientific and technical knowledge. [Alaimo 2014] Making a final decision concerning liability and compensation after the occurrence of harm is important, yet not enough. In order to provide comprehensive regulation of modern technologies, filling the liability gaps will simply not be enough. Deciding on liability questions must be carried out simultaneously with controlling such risks before they cause harm. However, the prohibition of the development risk defence, meaning the extension of a producer's liability to development risks would almost certainly cause more harm than benefit [Alaimo 2014], because: (1) complete liability regarding an unforeseeable event would most probably lead producers to irrational decisions in research and development, as well as innovation. [Alaimo 2014] The adoption of the development risk

defence suggests that too much liability might chill innovation; [Arbour 2014] (2) holding producers completely liable could become an incentive for excessive litigation because plaintiffs will get compensation as soon as they prove the existence of a defect and a causal link between the cause and harm; [Alaimo 2014] (3) it will cause financial problems, since it will be difficult to provide enough finances to compensate victims.[Alaimo 2014]

The Liability Directive complements European product safety legislation and this is known as the 'New Approach' to product safety. The latter focuses on the prevention of accidents by setting common safety rules which allow single market for goods to reduce administrative burden and to function as smoothly as possible. [Report 2018] In theory, as stated in Section 36 of the preamble of the Safety Directive, [Council Product Safety] "this Directive (the Safety Directive) should not affect victims' rights within the meaning on Council Directive 85/374/EEC of 25 July 1985... concerning liability for defective products (the Liability Directive)." [Council Defective Products] However, in legal practice the provisions and the scope of matters regulated by the Safety Directive affect consumer's liability. A requirement for deeper research results in greater awareness, not only of producers and programmers, but also consumers. It can be claimed that removal of a development risk defence could lead producers towards stricter internal policies on safety. [Alaimo 2014] However, keeping the development risk defence as it is – as a safe legal bubble that allows producers to escape liability for unforeseeable and unpredictable consequences is a risky option as well. In Rosselli's recommendations for the European Commission it has been emphasised that researchers need to focus on the analysis of mutual impact between the development risk defence and general product safety legislation, which has not yet been carried out. [Alaimo 2014] In Europe, we must achieve greater coherence between the Liability Directive and the Safety Directive (and their transpositions in national legislation). In order to provide a comprehensive and an efficient legal response to modern technologies it is necessary to place involve independent scientists and researchers to evaluate product safety. [Alaimo 2014] The latter would help producers and the society avoid the so-called 'junk-science' getting a place in the courtroom. [Arbour 2014] Even though the boundaries of the development risk defence have been set by the judiciary, the courtroom is not an appropriate place for scientific guesswork, Justice Posner concluded in the case Rosen vs Ciba-Geigy Corp. [Rosen 1996]

Requirements in the Safety Directive regarding detailed and thorough evaluation referring to potential consequences (evaluators must foresee the unforeseeable and pass this knowledge to the producers or manufacturers) are very important for further evaluation regarding the liability for harm caused by a machine or a product. Pursuant to the Report of European Commission "2018 is not 1985 (when The Liability Directive was adopted). The EU and its rules on product safety have evolved, as have the economy and technologies. Many products available today have characteristics that were considered science fiction in the 1980s. The challenges we are facing now and even more acutely in the future — to name but a few — relate to digitalization, the Internet of Things, artificial intelligence and cybersecurity." [Report 2018, 1] Absence of detailed safety requirements will leave a lot of potential knowledge about the operation of products undiscovered, and in such situations Article 7(e) of the Liability Directive will protect producers and allow them to exculpate its liability if state of scientific and technical knowledge at the time when he put the product on the market was not such as to enable the existence of the defect to be discovered. We want to avoid a situation when no one

will be held responsible for the harm caused by modern technology, especially if we have the capacity (knowledge) to identify the risks and through thorough research predict potential consequences of their operation. According to Arbour "(...) it may not be in a producer's best interest to know too much (...) as too much knowledge could lead to the exclusion of the DRD (development risk defence), and, consequently, expose producers to liability." [Arbour 2014, 932] However, the pile of iron and cables will (in the majority of cases – except 'vis major') not be the main reason for unforeseeable consequences, the major issue is the operation of the machine, its program and its self-learning capabilities. This is why the provisions and the regulatory scope of not only the Safety Directive, but also the Liability Directive needs to pay attention to software and algorithms – How does it work, which factors are relevant for the final outcome, and is the power to monitor these factors in our hands?

6. EU RULES ON PRODUCT SAFETY

For a long time, product safety as a sector has been regulated at a national level as a part of public law (in Germany 'Gerätesicherheitsgesetz' [GESETZ]). [Cafaggi 2009] "These norms provided a set of rules for the producers regarding product safety, supplemented by technical standards for specific products set by hybrid state-private bodies". [Cafaggi 2009, 245] However, the European Commission quickly realized that differences between technical regulations and standards among member states of the European Union presented an obstacle for trade. [Cafaggi 2009] For this reason, in order to improve free movement of goods, the European Community started to follow the concept of 'full harmonization' of technical requirements concerning a specific type of product. [Cafaggi 2009] Despite the effort, the European Commission soon realized that this approach will not be successful due to the inflexibility of the European legislation. The latter was not able to keep pace with fast technological progress and this led to the abolishment of the idea about full harmonization. [Cafaggi 2009] However, the European Commission found a solution through the judgment of the European Court of Justice in the 'Cassis the Dijon' case where the Commission developed a well-known principle of mutual recognition that resulted in a 'New Approach' [ECJ 120/78], which should soon become the main concept for product safety regulations. [Cafaggi 2009] The latter is based on the so-called hybrid regulation, meaning that main/essential requirements regarding product safety, which are in the public interest, are harmonized in EU directives, yet special/unique requirements on the other hand are listed in technical standards (harmonized standards), developed by private European standards organizations (such as CEN, CENELEC, ETSI) and are officially mandated by the Commission. [39] The benefit of keeping manufactured products in conformity with harmonized standards is the presumption that products are in line with the corresponding essential requirements. In other words, conformity with technical standards will relieve the producer of the burden of proof that products meet all the requirements concerning product safety. [Cafaggi 2009] On this point it is important to mention that only the essential safety requirements are mandatory and consequently legally binding in comparison with harmonized standards, the application of which remains optional. [Cafaggi 2009] Compared to harmonized standards, which are set by private organizations, most national non-harmonized regulations (national sector) rely on technical standards set by a mixture of industry (private body) and state representatives. [Cafaggi 2009] Even though a process of standardization is left to private bodies states want to have influence in the standardization process. [Cafaggi 2009] Despite the fact that the process of standardization is

organized in private forms they present a guideline for administrative decisions, take into account interests of the general public, want to satisfy common safety expectations and concretize due diligence in civil legislation. [Cafaggi 2009] However, governmental influence and surveillance regarding product standardization is very important from European point of view. The need derives from 'the obligation to reduce restrictions or measures having equivalent effect on free movement of goods. [Cafaggi 2009] Governments are sending their state representatives to standardization boards in order to provide further suggestions concerning alternative approaches to certain product areas for standardization. [Cafaggi 2009] Private standards (product safety) de facto observe double rules; first, they denote the required level of safety, and second they denote the required level of care. [Cafaggi 2009] The interaction between product safety (outlined in standards) and product liability can be described in the following way: (1) Public law sets a minimum level of safety and formulates basic requirements. Moreover, the administrative procedures are limited and formulated in ways that provide protection to the essential public and private goods like health and life; (2) Technical standards provided by private agencies specify a minimum level of safety that is provided by the public law; (3) Tort (civil) law focuses on individual circumstances and tries to define and attach the obligation on the producer. In addition, it fills the gaps left by technical standards and public law. [Cafaggi 2009] At first sight, it seems that this briefly described structure regarding product safety and liability provides the protection against all types of dangers. However, it fails to cover some crucial products/parts of the product, such as software, especially software equipped with advanced machine learning algorithms; devices connected with the internet, etc. [Cafaggi 2009] Here, product liability rules fail to cover gaps that were left by public law. This is the repercussion of our mistaken perception under which we understand product safety in the traditional sense. If we look at European regulation we will not find any specific legal provisions, even though such embedded systems (ES) are put into robotic machines (such as medicinal products). [Cafaggi 2009] Not only at the European level, the situation at the national level is very similar. There are hardly any safety regulations concerning ES. [Cafaggi 2009] The directive 2001/95/EC on general product safety does not encompass ES and consequently producers/programmers of ES are not obliged by specific product safety regulations, since they do not exist. [Cafaggi 2009] As mentioned before, the gap left by public law (such as lack of safety regulation about ES) cannot be filled exclusively by reorganization or updating of civil liability rules.

7. CONCLUSION

When policy makers and judicial authorities are confronted with an uncertain situation they turn to experts for definite answers and conclusive evidences, even though uncertainty clearly precludes definitiveness and conclusiveness. The uncertainty paradox leads to an ineffective policy-making processes. If authorities want to provide a comprehensive and an efficient legal response to modern technologies and liability for their actions it is necessary to involve sound, independent scientists and researchers to evaluate product safety. Authorities' decisions about liability questions must be carried out simultaneously with the control of such risks before they cause harm. Such a defence must exist alongside measures (the majority of which must focus on safety regulation) that may help to achieve the following four objectives: (1) "provide industry with unambiguous rules in order to prevent situations of jeopardization of products' safety due to insufficient knowledge. Industry and sectors will have to accept that such rules will result in a slower pace of innovation; (2) promote continuous transfer of knowledge

regarding safety in industry; (3) provide a system to restrict the potentially harmful impact of development risks; and (4) provide sufficient compensation for victims." [Cafaggi 2009] What seems to be the main problem regarding machines, equipped with programs that due to their algorithms enable self-learning, risk assessment of such programs is not required by the Safety Directive, even though a machine (equipped with such a program) may cause harmful consequences as a result of to the program. Without sufficient and regular evaluation of product's safety more consequences that will occur in the future will, from a legal perspective, be recognized as unforeseeable and this will prevent the judicial branch to provide satisfaction for harm suffered by victim.

8. REFERENCES

1. ALAIMO, A., et al. 2014. Analysis of the Economic Impact of the Development Risk Clause as provided by Directive 85/374/EEC on Liability for Defective Products. *Fondazione Rosselli, appointed by the Directorate for the Internal Market of the European Commission (Contract No. ETD/2002/85)*, 19-22, 59-66.

2. *ARBOUR, M.E. 2014. Portrait of Development Risk as a Young Defence. Technological Innovation and Civil Responsibility, 59(4), 915-917, 932.*

3. BAKER, J. 2010. *50 ideas you really need to know Universe.* Quercus Publishing Plc, 72.

4. BECK, U. 2009. Critical Theory of World Risk Society: A Cosmopolitan Vision. *Constellations 16(1)*, 7.

5. BRUMMETTE, W. 2017. Supervised vs Unsupervised Machine Learning, *ClickUp*. https://clickup.com/blog/supervised-vs-unsupervised-machine-learning/.

6. CAFAGGI, F., MUIR WATT, H. 2009. *The Regulatory Function of European Private Law.* Edward Elgar Publishing Limited, 245-253.

7. DEANGELIS, S. F. 2014. Artificial Intelligence: How Algorithms Make Systems Smart, *Wired.* https://www.wired.com/insights/2014/09/artificial-intelligence-algorithms-2/.

8. DE VILLIERS, M. 2015. Foreseeability Decoded. *Minnesota Journal of Law, Science & Technology 16(1)*, 344-345.

9. ELLIOTT, C., QUINN, F. 2011. *Tort Law.* Longman, Pearson, 191.

10. FABER, M., MANSTETTEN, R., PROOPS, J.L.R. 1992. Humankind and the Environment: An Anatomy of Surprise and Ignorance. *Environmental Values, Cambridge, 1(3)*, 224.

11. FOX, T. 2009. The role of uncertainty intolerance in European Risk Governance. *King's College Risk Research Symposium*, 1-3.

12. GESETZ UBER TECHNISCHE ARBEITSMITTEL OF 24 JUNE. 1968. Bundesgesetzblatt I, 717.

13. GIDDENS, A. 1999. Risk and Responsibility. *The Modern Law Review 62(1)*, 3.

14. HICKMAN, L. 2013. How algorithms rule the world, *The Guardian, Section: Mathematics.* https://www.theguardian.com/science/2013/jul/01/how-algorithms-rule-world-nsa

15. HILDEBRANDT, M. 2016. *Smart Technologies and the End(s) of Law*, Edward Elgar Publishing Limited, 23-25.

16. KANT, I. 2013. *Kritik der reinen Vernunft*. Felix Meiner Verlag, Hamburg, 49.

17. MOURIK, R. M. 2004. Did Water Kill the Cows. *Pallas Publications.*

18. OWEN, D. G. 2009. Figuring Foreseeability. *Wake Forest Law Review 44*, 1277.

19. PETITPIERRE, G. 1974. *La responsabilité du fait des produits : les bases d'une responsabilité spéciale en droit suisse, à la lumière de l'expérience des Etats-Unis.* Genève, Georg, 165.

20. PLOJ, B. 2013. Metoda mejnih parov za učenje umetnih nevronskih mrež, *Doktorska dizertacija, Fakulteta za elektrotehniko, računalništvo in informatiko Maribor*, 7.

21. PRESS, G. 2017. Alan Turing Predicts Machine Learning And The Impact Of Artificial Intelligence On Jobs, *Forbes.* https://www.forbes.com/sites/gilpress/2017/02/19/alan-turing-predicts-machine-learning-and- the-impact-of-artificial-intelligence-on-jobs/#651077b01c2b.

22. ROBERTS, E. 2016. The Enigma Machine, *Stanford University class* (1). http://stanford.edu/class/ archive/cs/cs106a/cs106a.1164/handouts/29 TheEnigmaMachine.pdf.

23. Rosen v Ciba-Geigy Corp, 78 F (3d) 316 at 319, 64 USLW 2612 (7th Cir 1996).

24. SATHYA, R. 2013. Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification. *(IJARAI) International Journal of Advanced Research in Artificial Intelligence, 2 (2)*, 34-35.

25. SERNA, H., BARRAGÁN, D. 2017. Patterns in nature: more than an inspiring design. Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales, Nat. 41, 352.

26. STAPLETON, J. 1994. *Product Liability.* Butterworths, London, 236.

27. TANNERT, C., ELVERIS, H.D., JANDRIG, B. 2007. The ethics of uncertainty. *EMBO reports: Science and Society, 8(10)*, 892-893.

28. TURING, A. M. 1950. Computing Machinery and Intelligence. *Research paper,* 19.

29. TURING, A. M. 1947. Lecture to the London Mathematical Society on 20 February 1947, *Seminar text*, 13.

30. VAN ASSELT, M.B.A., VOS, E. 2006. The Precautionary Principle and the Uncertainty Paradox. *Journal of Risk Research 9(4)*, 2.

31. VAN DAM, C. 2013. *European Tort Law: Second edition.* Oxford University Press, 434-435.

32. VAN LOON, R. 2018. Machine learning explained: Understanding supervised, unsupervised and reinforcement learning, *Big Data made simple*. http://bigdata-madesimple.com/machine-learning-explained-understanding-supervised-unsupervised-and-reinforcement-learning/.

33. VINEY, G., JOURDAIN, P. 1998. Les conditions de la responsabilité. *L.G.D.J.,* 163.

34. ZAPUŠEK, T. 2017. Artificial intelligence in medicine, *The Asia Pacific Journal of Health Law & Ethics, 11*, 117.

35. WERRO, F., PALMER, V.V. 2004. *The Boundaries of Strict Liability in European Tort Law.* Carolina Academic Press, Stämpfli Publishers Ltd, Bruylant, 3-10.

36. Art. 30 of the European Economic Community - EEC Treaty, now Art. 28 EC Treaty

37. Council Directive 2001/95/EC of the European Parliament and of the Council of 3 December 2001 on general product safety [2002] OJ L 11, Preamble.

38. Council Directive 2001/29/EC on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products [1985] OJ L 210.

39. ECJ, Case 120/78, Rewe-Zentral AG *vs*. Bundesmonopolverwaltung für Branntwein (1979), p. 649.

40. REPORT FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL AND THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE on the Application of the Council Directive on the approximation of the laws, regulations, and administrative provisions of the Member States concerning liability for defective products (85/374/EEC), 7.5.2018.

41. Rosen v Ciba-Geigy Corp, 78 F (3d) 316 at 319, 64 USLW 2612 (7th Cir 1996).

42. The European Consumer Association (BEUC). 1999, Response to the Commission Green Paper on Liability for defective products, *COM (99) 396 final, 28 July 1999*, 8.

Legal Practitioners' Views on Neuroscientific Evidence in Court Proceedings

Cristina Siserman-Gray
Faculty of Law, University of Vienna
Juridicum, Schottenbastei 10-16
1010, Vienna, Austria
+436802300653
cristinasiserman@gmail.com

ABSTRACT

This paper presents the results of an empirical study conducted on a number of legal practitioners in Austria, Romania and Slovenia to measure the relevance of the neuroscientific evidence when determining the legal responsibility of a defendant during court proceedings. The legal practitioners were presented five case scenarios in which they were asked to analyze the admissibility of neuroscientific evidence. The results revealed certain important challenges when interpreting neuroscientific evidence for use in legal proceedings. This study summarizes those challenges and further concludes that the application of neuroscientific data is not a straightforward matter. As such, it also argues that neuroscientific experts must play a greater role in court proceedings in order to provide better guidance to judges and other legal practitioners.

Keywords

Legal practitioners, neuroscience, law, courtroom, criminal responsibility etc.

1. INTRODUCTION

Neuroimaging scans have been used in a variety of cases in order to establish a defendant's mental state. In civil cases, neuroimaging has been requested in personal injury, disability belief and contract cases, while in criminal cases it has been mainly invoked in order to negate the *mens rea* of a crime and to avoid convictions. For instance, statistics show that in some countries such as the USA over 1500 judicial opinions issued during 2005-2012 discuss the use of neuroscience by criminal defendants [1]. However, despite the wide use of scans in courts, there is concern among scientists and legal practitioners regarding the practical consequences of neuroscientific developments. The aim of this empirical study seeks to understand how European legal practitioners relate to "neurolaw cases" and the degree to which they admit neuroscientific evidence in court in order to determine the moral and legal responsibility of an individual.

2. PREVIOUS RESEARCH

To our knowledge this is one of the first studies of its kind to assess the views of legal practitioners regarding neuroscientific evidence and the impact of those views on their decisions when assessing the responsibility of an individual. While a number of studies have concluded that neuroscientific evidence affects legal practitioners, few have attempted to qualify that effect. Therefore, in order to test the degree to which legal practitioners use neuroscientific evidence in the assessment of responsibility, we based our study on a hypothesis tested by Aspinwall et al. in 2012 who argued that there is a "correlation between the inclusion of the biological explanation of neurological disorder with significantly reduced sentence length and increased number of mitigating factors listed" [2]. Though this correlation has seen corroborated by the US president's Council on Bioethics, which claimed that "judges and juries have recognized the persuasive allure of brain scans" [3] and Munro's 2014 study, which concluded that the public is more likely to trust diagnosis when brain imaging evidence is presented [4], our study will step beyond these conclusions to understand *how* neuroscientific evidence affects legal practitioners in legal proceedings.

3. DESIGN OF THE EXPERIMENT

3.1 Participants

The participants selected for this study were legal practitioners (judges, lawyers and attorneys) or graduates of legal studies who were already working in the legal profession. The participants' views were assessed via an online questionnaire that was completed by 91 participants (N=91).

The participants were categorized in three main groups, a group from Slovenia (=Slovenian group comprising 41 practitioners with an average of 28.2 years old); a second group from Romania (=Romanian group comprising 22 practitioners with an average age of 34.1 years old); and a third group (=Mixed group comprising 28 legal practitioners with an average age of 28.6 years old) representing legal practitioners residing in Austria, but coming from other countries in Europe and North America.

3.2 Instruments

The legal practitioners were asked to analyze five case scenarios that were designed in such a way as to determine a) the types of neuroscientific evidence that would be acceptable in courts and b) how the legal practitioners' assessment of the individual's responsibility (i.e. guiltiness, type and length of punishment) would be influenced by this type of evidence. The case scenarios were also designed to assess how legal practitioners view "new science", such as advancements in neuroscience.

Before briefly introducing each case scenario and the results, a few remarks are required to fully understand why each case was formulated as it was. Sloboghin identified five categories of neuroscientific evidence based on the cases presented in American courts [5], which we adopted as the basis for our scenarios. The 5 types of evidence are: (a) evidence of abnormality; (b) cause-of-an-effect evidence; (c) effect-of-a-cause evidence; (d) individualized neuro-psychological findings compared against known performance baselines; (e) individualized neuroscience findings compared against known legal baselines.

4. RESULTS AND DISCUSSION

In the following we will discuss the questionnaire results of each scenario, as well as the main problems that were raised by the respondents. Each case scenario presents an offence and the type of neuroscientific evidence that the defendant invoked in order to either excuse his/her behavior or diminish his/her legal responsibility. The legal practitioners were asked to perform three tasks in order to determine: (i) whether the individual in the scenario should be held responsible for the act; (ii) the kind of punishment or treatment that should be applied to an individual; and (iii) the length of punishment (if any). We will additionally highlight some of the challenges identified by the legal practitioners when analyzing the type of neuroscientific evidence presented in each case scenarios.

4.1 Difficulties demonstrating causality

As reported by legal practitioners, demonstrating causality is among the most important problems that arise in practice when using neuroscientific evidence. The problem of causality was specifically included in one of this first case scenarios, in which the legal practitioners were presented with a "cause-of-an-effect" evidence, and asked to determine whether the evidence showing a link between the damage to the frontal lobe and abnormal sexual behavior is acceptable. They were presented the following scenario:

> *"Mr. Jones, aged 35, was arrested for having raped three young women. In the pre-trial investigation, a scan of Mr. Jones' brain using PET (positron emission tomography) revealed serious damage to his frontal lobe, apparently as a result of a stroke. The medical expert showed that during the stroke, the frontal lobe (which is involved in judgment, impulse control and sexual behavior) has been irreversibly damaged and this explains Mr. Jones' abnormal sexual behavior".*

In principle, a "cause-of-an-effect" evidence could be highly exculpatory in order to decide on a volitional impairment. However, the results indicated that our participants did not generally agree in this case. In fact, some Slovenian legal practitioners reported that this type of evidence is not very useful for them as a basis to decide the legal responsibility of an individual. Though this argument was expressed by only some Slovenian respondents, it could generally explain the answers of the Slovenian Group, which agreed by a large majority (80.2%) that the accused in this case should be held responsible for rape. The same view was expressed in a proportion of 64.3% by the Mixed Group.

Interestingly, the Romanian judges were of a different opinion, with only 41% of them agreeing that the accused should be held responsible. The Romanian legal practitioners believed that the evidence was acceptable to diminish the responsibility of the defendant, but not sufficient to remove it completely. The Romanian legal practitioners did not challenge the evidence *per se* in as much detail as the Slovenians did, but the Romanian legal practitioners that did agree that the accused should be held responsible also reported that they were not convinced beyond any reasonable doubt about the causal link between the stroke and the abnormal sexual behavior.

Additionally, some Slovenian legal practitioners noted disparities in the way scientific conclusions could be interpreted. In their view, the conclusion presented in this case, which argues for a causal link between a stroke and abnormal behavior, could also be interpreted as accepting that a high proportion of people that commit rape had a frontal lobe stroke at some point,

which, of course, is not necessarily true. In their view, the evidence submitted does not indicate the probability with which people with strokes would also commit rape, which is the central question that a judge would want to have answered and which was not done in a satisfactory manner here.

Consequently, legal practitioners noted that neuroscientific studies presented in courts can rarely be presumed to be fully conclusive, despite the fact that some researchers often invoke assumptions of cause and effect. In their view, the notion of "cause" presents differences in meaning between the social sciences and the criminal law. It might be that because of this reason the two fields frequently clash. According to many legal practitioners, among which Denno, as well, scientists do not often use terms such as "cause-and-effect", but instead use concepts such as "laws of change", "paradigms", "models and theories", "hypothesis testing" and "falsification" [6]. According to them, these various "jargons" make their work more complicated.

4.2 The effect of methodological weaknesses

There is an awareness among legal practitioners that many scientific studies have methodological weaknesses, which have an unknown degree of impact on the overall results. Because of these methodological weaknesses, many Courts have showed certain resistance in considering certain statistical data in their decision-making. Therefore, when deciding to accept probabilistic results as evidence, legal practitioners generally discard causal links that are assessed as not being fully convincible, regardless of the type of evidence, neuroscientific or otherwise.

This was the solution reached by a majority of legal practitioners when presented the following scenario which contained a type of evidence that can be categorized as an individualized neuro-psychological finding:

> *Mr. Johnson, aged 27, was brought in front of the court for having sold methamphetamines. The defense lawyers informed the judges that the defendant found out 8 years before that he had an unusually large pituitary tumor, which caused irreversible brain tumors. The medical expert showed that pituitary tumors may affect thyroid production, causing mood disorder and damage to the frontal, temporal and thalamic regions, which may cause problems in decision-making, mental flexibility and overall intellectual capacity. The medical expert supported the defense who claimed that because of the tumor, the defendant was more susceptible to be influenced and manipulated by the drugs dealers and that is why he accepted to sell the drugs".*

In general, individualized neuro-psychological findings compared against known performance baselines (such as those in this scenario) are accepted in courts because they provide insight into the particular defendant's biological functioning. Despite this fact, a great majority of respondents across all three groups (average of 87.9%) rejected the evidence presented and agreed that the accused should be held responsible for selling drugs. An argument invoked by the legal practitioners for this conclusion was that it is generally very difficult for scientists to come up with convincing and compelling data relevant for all necessary demographic groups in order to make a useful comparison with the defendant. That is because findings from neurological testing can vary significantly based on a series of variables, including gender, age, and education, among others. In the view of the legal practitioners, interpreting specific results can be very difficult without this baseline information.

Another concern noted by legal practitioners was that even if these baseline measurements would be obtained, a correct determination about relative impulsivity of the defendant at one point in time does not prove anything about the impulsivity of the defendant at the time the crime occurred (in our case when the accused decided to sell the drugs). Consequently, our respondents confirmed conclusions reached in other studies, such as that of Sloboghin who noted "that science cannot currently answer the normative question of how far below the average a defendant would have to register on a particular performance task to be considered legally impaired" [5].

4.3 Limits regarding potential simultaneous effects and interrelationships

Another challenge of neuroscientific evidence noted by legal practitioners was that they were seldom presented studies that examined all the possible variables. By ignoring potential simultaneous effects and interrelationships that may exist among the numerous other biological, social, and economic factors, they concluded that very few of these studies were complete, leaving room for interpretation. Therefore, a majority of judges in our study continue to believe that most biological and environmental studies of crime remain isolated in their particular disciplines, which often makes them unacceptable as evidence in the court.

This was one of the main conclusions reached based on the results obtain for the following case scenario:

> "Mr. Green, aged 47 is brought in front of the criminal court for assault and injuries on a pedestrian. He assaulted the pedestrian on the reason that he almost generated an accident while walking on the lane destined only for bikers. Mr. Green's lawyer pleaded that the defendant is a peaceful person never having been involved in fights and that his violent behavior was the result of the defendant's level of testosterone which changed its normal level because of the ingestion of some steroids that Mr. Green took during sport training. The medical expert showed that the saliva samples collected from the defendant exhibited abnormal level of testosterone and this generated his aggressive behavior. He also added that there is no unanimity within the scientific community regarding the positive correlation between testosterone level and aggression in general, but that in his opinion, in this particular case, the correlation is evident".

This case scenario contains "effect-of-a-cause" evidence, which basically compares the prevalence of criminal behavior among individuals presenting a neurological impairment to those that do not have such an impairment. More precisely, the evidence refers to research that indicates a higher prevalence rate of violence among people with high level of testosterone. When asked to evaluate whether the defendant should be held responsible for assault and injuries on a pedestrian, the great majority (94.5%) of respondents from all three groups agreed that he should be held responsible. Compared to the previous two case scenarios, this one contained less divergent views.

As mentioned by the respondents in the comments sections, and confirmed by other studies, evidence of clear association between androgen levels and aggression in human males is currently inconclusive. Legal practitioners took note of studies that noted that behavioral "traits of dominance and aggression in the human male have been associated with higher levels of testosterone" [6]. However, they were also aware that, as Archer concluded, "direct associations between androgen levels, primarily testosterone, and criminality show somewhat conflicting results, possibly because of the different types of hormone measures used" [7]. As such, it seems that legal practitioners view research on some of these factors as statistically weak and believe that more carefully performed studies are needed before they can use this type of evidence in the courtroom.

Nonetheless, the legal practitioners also argued this type of data must be put in context. In their view, without knowing the general base rate for violence in people with normal levels of testosterone, crime prevalence in groups of people with high levels of testosterone is not of much help to legal practitioners.

4.4 Too much focus on incrimination rather than exoneration

Connected to the conclusion above, an interesting view of many legal practitioners was the acknowledgement that current neuroscience evidence is mostly focused on culpability issues (whether an individual is morally and legally responsible for their acts) rather than providing a basis for exoneration. Because of that, in deciding on the relevance of such evidence, legal practitioners and experts are aware that they have to pay attention both to the precise nature of the evidence in question and the specific legal doctrine to be addressed. Though some may argue that neuroscience evidence alone could be sufficient to remove the responsibility of an individual, in reality, in most cases, it is not. This was what the legal practitioners decided in the case of the following case scenario:

> "Ms. Black, aged 47, accountant, was filed for conducting fraudulent insurance practices for a period of 8 years. She was diagnosed 10 years before with anoxic encephalopathy caused by a myocardial infarction. A SPECT (single-photon emission computed tomography) was performed, which indicated a reduction in blood flow in the temporal and frontal lobes (which are associated with executive functioning and memory). The medical expert showed that her fraudulent behavior could be explained by her impairment".

For this scenario, the majority of the legal practitioners considered that the SPECT evidence that was included to demonstrate a reduction in blood flow in the temporal and frontal lobe should not be received as an exculpatory evidence. Their main argument was that neuro-abnormalities have minimal relevance for cause in this case, particularly when having to explain a causal link over an extended period of time (a stroke that occurred 10 years before the trial). The causal link between the stroke and the multiple fraudulent acts was impossible to establish. Therefore, the majority discarded this evidence, with 95.2% of the Slovenian, 81.9% of the Romanian respondents and 85.7% agreeing that the defendant should be held responsible for fraudulent practices.

As noted by the legal practitioners, a reason why they believed prison punishment was appropriate was due of the concern that, in this case, improperly used neuroscience evidence could be used to mitigate punishment and thereby present a greater risk of re-offense. It seems thus that judges, when confronted with the public pressure to ensure to that there is no risk of re-offense (like in the scenario presented above), may be more inclined to admit the neuroscientific evidence in order to evaluate one's risk of re-offending. The "double-edge sword" nature of neuroscience evidence is important and was not noted by the respondents prior to this case scenario. Therefore, based on these results, it is important to be aware of the possibility that neuroscientific evidence could be misused in the future to evaluate or predict an individual's potential level of dangerousness.

4.5 Cultural differences in weighing the strength of the evidence

Finally, it is important to note that we believe that some of discrepancies included in our results are attributable to cultural differences. In fact, these cultural differences seemed particularly important in relation to the severity of punishment. This is seen most clearly in the following scenario:

"Mr. Smith, aged 37, is a teacher and he is being tried for having made sexual advances on his young stepdaughter. He was found to have a tumor in the right frontal lobe of his brain. His medical results showed that when the tumor was removed, his pedophilic behavior stopped. When the tumor recurred, the behavior also resumed".

When asked whether the defendant should be held responsible for the sexual advances committed on his stepdaughter, a majority of the Slovenian and the Mixed groups, with 83% and 60.7%, respectively, agreed that the accused should be held responsible for sexual advances. However, the majority of Romanians (59%) disagreed, believing that in this case, the individual should not be held responsible for sexual advances.

We believe that the difference in the way this defendant was evaluated by the three groups has to do with the type of neuroscientific evidence presented to the participants in the study. In this case, the legal practitioners were asked to evaluate evidence of a brain abnormality. The Slovenian and Mixed Group argued that since hypothetical cases like these are very rare in practice, very few defendants can present such a straightforward connection between neurology and behavior. As such, the majority of the respondents from those two groups believed that the behavior of the defendant could not be excused because this neurological abnormality provides an insufficient basis for making the necessary link to the unlawful behavior.

That was not the case for the Romanian Group, who in a proportion of 59.1% believed that the accused should not be held responsible for sexual advances. In this situation, Romanian legal practitioners admitted the neuroscientific evidence on the presumption that the medical conclusions were correct. Compared to the Slovenian and Mixed Groups, the Romanian Group did not challenge the accuracy of the medical results and/ or of the neuroscientific evidence that was brought to be examined in the case. The reason for this discrepancy is unknown and could warrant further study.

5. CONCLUSION AND FURTHER DIRECTIONS

The results of this experiment are addressed to both scientists and legal practitioners, who, we hope, will be able to draw some practical applications for their work. Our first recommendation for the scientists is to make sure that information they release publicly is valid and clearly states the scientific limitations. Experts in the field of neuroscience should keep in mind that as much as legal practitioners would sometimes want to accept a certain type of neuroscientific evidence, specific constraints of the legal doctrine in which they work prevent them. A second message is addressed to legal practitioners, who should begin considering the best ways to integrate new scientific discoveries. Neuroscientific evidence can provide a more nuanced understanding of neurological impairments in a way designed to humanize the legal system rather than creating motives for discriminating against individuals. This more nuanced understanding can additionally provide a more graded approach

for assigning responsibility, by minimizing arbitrary incrimination or full exonerations based only on the premise that "my brain made me do it".

We are also hopeful that the conclusions obtained in this study are sufficient to provide at least a more accurate picture of how various legal practitioners witness, from their positions, the new developments in neuroscience. However, based on the results obtained in this study it is difficult to predict whether neuroscientific evidence will play a more important role in courtrooms in Europe in the future. There are many optimistic voices that see a future for neuroscientific tools, as they may provide a more objective assessment of an individual's responsibility, particularly when compared to legal rules, which are generally insufficient and represent only a translation of the social beliefs in legal terms. Of course, some concerned voices support the view that neuroscientific techniques are still not ready to provide sufficient accuracy, and hence they are not fully reliable in the court. In our view, caution is certainly warranted in regard to neuroscientific evidence. However, with further education and better understanding of neuroscientific evidence by judges and other legal practitioners, we are certain that it holds great potential for the future.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Presidential Commission for the Study of Bioethical Issues. 2015. Gray Matters: Topics at the Intersection of Neuroscience, Ethics and Society, Vol. 2, Washington DC.

[2] Aspinwall, L., Brown, T., and Tabery, J. 2012. The Double-Edge Sword: Does Biomechanism Increase or Decrease Judges' Sentencing of Psychopaths? *Science*. 337, 845-849. DOI= 10.1126/science.1219569.

[3] President's Council on Bioethics. 2004. Staff Working Paper: An Overview of the Impact of Neuroscience Evidence in Criminal Law. (September 2004). Retrieved September 25, 2018 from https://bioethicsarchive.georgetown.edu/pcbe/background/neuroscience_evidence.html.

[4] Munro, C. and Munro, C. 2000. "Soft" versus "hard" psychological evidence: Biased evaluations of scientific evidence that threatens or supports a strong held political identity. *Basic and Applied Social Psychology*. 36, 6, 533-543. DOI= https://doi.org/10.1080/01973533.2014.960080

[5] Sloboghin, C. 2017. Neuroscience Nuance: Dissecting the Relevance of Neuroscience in Adjudicating Criminal Responsibility. *Journal of Law and the Biosciences*. 4, 3, 577-593. DOI= 10.1093/jlb/lsx033.

[6] Denno, D. 2016. The Place for Neuroscience in Criminal Law, In *Philosophical Foundations of Law and Neuroscience*, New York, Oxford University Press, 69-84.

[7] Archer, J. 1991. The influence of testosterone on human aggression, *British Journal of Psychology*. 82. 1-28. DOI= https://doi.org/10.1111/j.2044-8295.1991.tb02379.x.

FROM AI TO OUR FUTURE I
AI and the Drake equation

Matjaž Gams
Jozef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
matjaz.gams@ijs.si

ABSTRACT
This paper deals with a question, can we predict the future of the human civilization based on the analysis of the current AI research, and computational models of the Drake equation. The perspective of human civilization is analyzed through two items: report from the AI mega-conference, and the Drake equation/Fermi paradox. The mega-AI conference in Stockholm represents the state of the art in the AI field, demonstrating that major world powers are fast increasing AI funding. The computational models such as recompilations of the Drake equation are gaining attention as one of the means to estimate the longevity of the human civilization.

POVZETEK
To je prvi prispevek od dveh, ki opisujeta isti koncept: ali lahko napovemo bodočnost človeštva na osnovi analize sedanjih raziskav umetne inteligence (AI), računskih modelov na osnovi Drakove enačbe in novi socialni črni scenarij, ki bi lahko zaustavil razvoj človeštva.

V tem prispevku je analizirana bodočnost človeške civilizacije skozi dve temi: poročilo z največje svetovne AI konference in analiza Drakove enačbe oz. Fermijevega paradoksa. Mega konferenca letos v Stockholmu je predstavila trenutne raziskave AI v svetu in pokazala, da svetovne velesile izredno povečujejo financiranje AI področja. Po drugi strani pa narašča število analiz in objav s področja bodočnosti človeštva. Računski modeli med drugim dajejo ocene trajanja človeške civilizacije.

Keywords / ključne besede
Progress of human civilization, Drake equation, Fermi paradox, artificial intelligence

Razvoj človeške civilizacije, Drakova enačba, Fermijev paradoks, umetna inteligenca

1. UVOD

Večina projektov v odseku E9 na Institutu »Jožef Stefan« se ukvarja z izboljšavami življenja in zdravja ljudi, recimo s skrbjo za starejše preko inteligentnega pomočnika na pametni uri, ali pa za srčne bolnike. Del raziskav pa je bolj visoko letečih. Nekaj časa že razmišljamo, kako zgraditi inteligentne asistente iz biblije, korana, slovenske ustave. Asistenti ne bi znali samo odgovarjati na vprašanja s primernim besedilom, ampak bi bili sposobni tudi tolmačiti določen del vsebine, semantike, vrednot. Bi tako ustvarili kopijo npr. krščanskega boga? Ti programi bi bili sposobni demonstrirati svoje avtonomno razmišljanje na osnovi podanih besedil, torej bi bili v nekem smislu »živi«. Za dosego tega pa morajo imeti tudi, ali predvsem kognitivne lastnosti.

Če so trenutne raziskave oz. analize na tem nivoju, ali so v svetu raziskave že blizu superinteligence [1,2,3], ki bi morebiti znala prehiteti ljudi na vseh področjih? V tem prispevku je najprej predstavljena največja svetovna konferenca s področja umetne inteligence, sledi pa še analiza Drakove enačbe z novimi računskimi pristopi. Na osnovi teh analiz skušamo oceniti, kaj nas čaka v bližnji prihodnosti, tj. skušamo napovedati prihodnji razvoj človeške civilizacije.

Letos poleti je bila v Stockholmu največja svetovna konferenca IJCAI, združena z evropsko ECAI [4]. Skupno je bilo preko 6.000 udeležencev. Približno polovica vseh prispevkov je bila kitajskih, pol manj je bilo evropejskih in ameriških. Velesile se zavedajo, da je področje umetne inteligence eno izmed ključnih, kjer se odloča, kdo bo dominiral svetu. Tako Putin kot Trump in Ši Džinping intenzivno povečujejo sredstva za umetno inteligenco, Evropa jih bo v nekaj letih nekajkrat povečala.

Dnevno naredi umetna inteligenca neverjetnih 10 bilijonov odločitev. Dosežkov umetne inteligence samo v lanskem letu je bilo toliko, da jih lahko omenimo le minimalni delež. Recimo na področju varnosti marsikje po svetu uporabljajo sistem, ki vsak dan izdela nov urnik obhodov varnostnikov po letališčih, po pristaniščih in podobno. Kjer so bili uporabljeni, so izmerili bistveno večjo učinkovitost. V skrbi za okolje so raziskovalci pod vodstvom prof. Tambeja (tam je bil tudi naš doktorand dr. Kaluža) tovrstne programe podarili šestdesetim rezervatom po svetu, da se bodo uspešneje upirali krivolovcem. Leta 2015 so programi globokih nevronskih mrež začeli premagovati ljudi pri prepoznavanju vidnih nalog in danes nas prekašajo na mnogo področjih, npr. pri prepoznavanju malignih tkiv. Pri nekaterih nalogah, recimo pri ostrenju slike (dež, megla, sneg itd.) so programi osemkrat boljši kot ljudje. Seveda se takoj pojavi strah, ampak če čakate nekaj tednov na diagnozo, ali imate raka ali ne, v Ameriki pa to naredi umetna inteligenca v nekaj minutah bolje kot katerikoli zdravnik – kaj pravite, ali bi jo uvedli tudi pri nas? Pri nekaterih sistemih kot zamenjavi organov so programi že desetletja v uporabi in so rešila na tisoče življenj. Nekateri so novejši, recimo letos so vpeljali prvi inteligentni program, ki ugotavlja diabetes iz pregleda oči, prav tako prvi program za ugotavljanje abnormalnosti prsnega koša pri slikanju. Neverjeten razvoj je najbolj znan pri avtonomni vožnji – danes imajo povprečni avtomobili kar nekaj avtonomnih inteligentnih funkcij, modernejši kot Tesla pa vozijo praktično sami in jih samo še nadziramo v nenavadnih situacijah. Nesreč je približno stokrat manj, čeprav tistih nekaj mediji toliko bolj napihnejo.

In smo pri »moralnem stroju« - vprašanju, koga naj povozi avtonomno vozilo, če npr. izbira med otrokom ali starejšim: http://moralmachine.mit.edu/. V Nemčiji je prepovedano upoštevati leta (prepovedana starostna diskriminacija), v testih pa se izkaže, da precej ljudi raje ponudi možnost preživetja otrokom. Najpomembnejše je število, torej naj bi avto zavil med dva namesto med tri ljudi. Dobršen del anketirancev pa raje izbere brezdomca kot otroka itd. Preko spleta so zbrali že okoli petdeset milijonov odločitev in izkaže se, da imajo določene države podobne sisteme preferenc, koga povoziti in koga ne.

Tako so ugotovili značilne vzorce vrednosti pri zahodnih državah, spet druge v vzhodni Evropi itd. Anglofilske države kot Anglija, Amerika, Avstralija in Nova Zelandija so svoj skupek s podobnimi vrednotami, čeprav so geografsko precej narazen. Ta grupiranja torej pokažejo, katere države imajo podoben sistem vrednot in kako se razlikujejo od drugih.

V praksi je taka situacija izredno redka, saj imajo ljudje možnost odskočiti, avto se zaleti v oviro in zaradi varnostnega pasu potniki skoraj gotovo preživijo, medtem ko pešec kakšne posebne zaščite nima. Študija, ena izmed mnogih, pa predvsem pove nekaj o nas samih – kakšni smo kot ljudje in kako cenimo življenje drugih. Sistemi umetne inteligence zato ne rešujejo samo pereča inženirska in socialna vprašanja, ampak tudi iz nas delajo boljše ljudi. Marsikdo med nami je prepričan, da je ravno umetna inteligenca tista ključna znanstvena veda, ki bo dvignila človeštvo v novo civilizacijsko ero [11].

2. IJCAI 2018 – KAKO ZGRADITI UM

Julija 2018 v Stockholmu so se združili ICML, AAMAS, ICCBR in SoCS z IJCAI and ECAI v prvo mega-konferenco. Namen dogodka je bil ponovno integrirati področje in se namesto s čedalje bolj specifičnimi problemi ukvarjati z ključnimi splošnimi. Opis v tej sekciji je namenjen skupni IJCAI-ECCAI konferenci [4].

Okoli 7000 udeležencev je 14 dni sodelovalo v mega-konferenci. IJCAI-ECCAI konferenca je doživela ponoven rekord s 3470 poslanimi prispevki: 37 % več kot v letu 2017. Leta 2017 [5] je bilo 37% vseh prispevkov kitajskih, leta 2018 pa 46%. Ameriških in evropskih prispevkov je bilo vsakih po 20%.

Spremembe niso prišle same po sebi, ampak so rezultat ključnega vlaganja kitajskega državnega vodstva v znanost in zlasti v umetno inteligenco. Odgovor Amerike in Evrope je že tu.
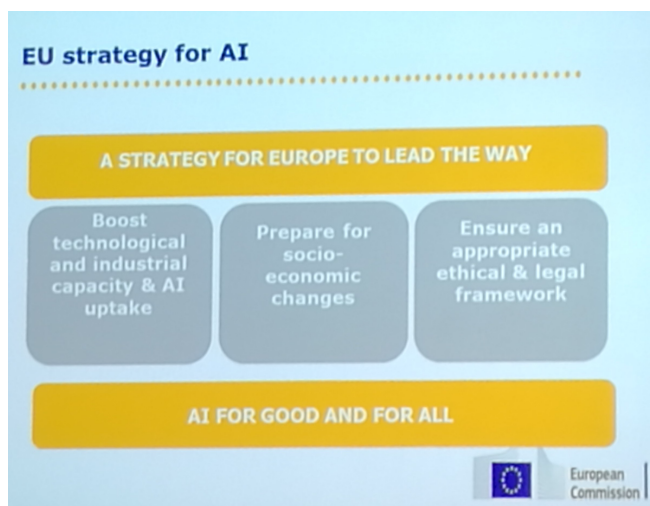
Slika 1: AI strategija EU temelji na treh elementih: znanost/tehnologija, socio-ekonomske spremembe in socialno okolje.

Na Sliki 1 je prikazana strateška usmeritev EU glede umetne inteligence. Med drugim bo EU podprla gradnjo nove odprte platforme za AI, malce podobno kot Muskov OpenAI (Slika 2). Na sliki 2 je prikazana usmeritev bodočega financiranja: najprej 70% povečanja in nato 100% in nato še 100%.

Ameriški DoD je ustanovil Joint AI Center (JAIC), ki bo financiral 600 AI projektov v vrednosti $1.7 milijarde.
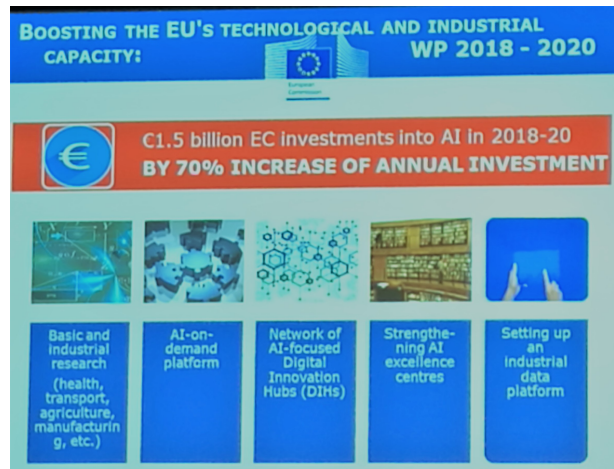
Slika 2: Tudi EU bo signifikantno povečala financiranje AI.

Najpomembnejša strateška usmeritev mega-konference IJCAI je bila, kako zgraditi um. Ker umetnega uma odraslih očitno še ne znamo zgraditi, so raziskave usmerjene predvsem v študije uma otrok od rojstva dalje. Čeprav se je vpliv AI na vsakodnevno življenje ljudi izjemno povečal, saj AI vsakodnevno naredi 100 bilijonov odločitev, so koncepti kot razumevanje, semantika, um, kognicija, čustva, »duša«, kvalia itd. ostali nedosegljivi za AI, podobno kot rešitev Turingoveg testa.

Vseeno je razvoj izredno hiter. Nekatere nove aplikacije so vizualno fascinantne – recimo konja znajo nevronske mreže prebarvati v zebro in se na zaslonu premika konj v obliki zebre. Iz seznama obrazov slavnih osebnosti znajo ti programi zgraditi nove obraze neobstoječih ljudi. Sistemi govorijo praktično kot ljudje in tudi razpoznavajo govor na tem nivoju, veliko vidnih nalog pa opravljajo precej bolje kot ljudje. Google je začel uporabljati globoke nevronske mreže za prepoznavanje vprašanj in občasno je na čudna vprašanja generiral vizije, tudi preroške vizije črne prihodnosti človeštva. Vraževerni so začeli govoriti o prebujanju prave inteligence, a so razvijalci Googla pojasnili, da so nevronske mreže prav čudna vprašanja povezala s prav čudnimi odgovori.

Obstaja *plitva*, tj. sedanja uporabna in raziskovalna AI, *globoka* AI, tj. globoke nevronske mreže, za katero velja mnenje raziskovalcev umetne inteligence, da je v resnici plitva, *prava* AI in *lažna* AI. Lažna AI je, ko vskočijo ljudje, tipično v neko komunikacijo med virtualnim asistentom in človekom, in tega ne povedo. Prava AI pa je AI, sposobna rešiti Turingov test, tista,

ki da nam ljudem vtis resnične inteligence. V Turingovem testu ljudje še vedno hitro prepoznajo računalnike, čeprav se pretvarjajo, da so ljudje, ker jim manjka prava AI.

S čedalje več raziskavami postaja tudi čedalje bolj jasno, kje se AI razlikuje od prave AI. Recimo pri učenju se otrok (Slika 3) nauči enega koncepta, npr. kako sestaviti kocke, nato pa ga praktično takoj ali z minimalnim učenjem prenese na druge domene, recimo sestavljanje figuric. Pri tem se uči univerzalno, medtem ko mora AI uporabljati specialne programe za posamezne naloge, jih ne uspe posplošiti in ne zna prilagoditi algoritmov podobnim nalogam.
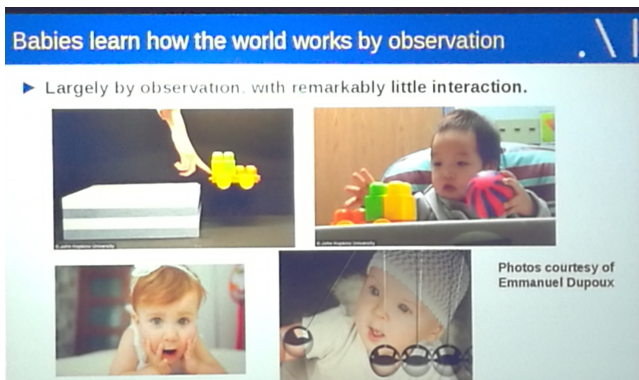

Figure 3: Otroci se učijo bistveno drugače kot sistemi AI. Zakaj se ne bi zgledovali po njih?

Drug pomemben poskus v smeri prave AI je t.i. *splošna* AI, ki skuša razvijati široko uporabne algoritme namesto ozko specializiranih. Tako npr. naj bi obstajal en algoritem za igranje vseh iger, ki bi znal izvajati vse algoritme za različne igre, in prenašal znanje med njimi. Večinsko mnenje med strokovnjaki umetne inteligence je, da bomo to vrsto inteligence razvili čez kakšnih 10 let. Splošno mnenje je, da nam ne manjka denarja ali procesorskih zmogljivosti – iščemo nove, drzne ideje, kako razviti splošno, pravo in super-inteligenco. Bolj ali manj je tudi konsenz, da bo AI pomagala reševati nakopičene socialne težave (Slika 1). Konferenca je ostala strogo znanstvena, vendar je med znanstveniki določena zaskrbljenost zaradi čedalje pogostejših napadov na znanost in prodiranje ideologij med znanstvenike.

3. FERMIJEV PARADOKS - ČRNI SCENARIJ ČLOVEŠTVA

Katere nevarnosti prežijo človeštvu? Skoraj zagotovo so največje in najbližje nevarnosti povezane s hitrim vzponom človeške civilizacije. Prav tako med najbolj perečimi ni nekaterih, ki se pogosto pojavljajo v medijih, npr. globalnega segrevanja. Čeprav se planet Zemlja segreva in čeprav je najverjetnejši krivec za to človek, je ta nevarnost obvladljiva, če se le uspemo dogovoriti. Bolj problematične so neobvladljive, nerazumljene in potuhnjene nevarnosti. V kakšnem iz naslednjih prispevkov bomo predstavili izvirno teorijo, temelječo na socialnem propadu, tu pa bodo na kratko opisane analize Drakove enačbe.

Veliki fizik Enrico Fermi je že leta 1950 med atomskimi raziskavami v Los Alamosu vprašal: »Pa kje so?«, ko je preračunal verjetnosti in ugotovil, da bi že morali vzpostaviti stik z drugimi civilizacijami. Sedaj vemo precej več: število planetov v znanem vesolju je reda velikosti 10 na 22, pregledali smo ogromno več vesolja in nismo našli nobenih znakov civilizacije! Pri tem je minilo 68 let od Fermijevega vzklika in 64 let od njegove smrti. Umrl je istega leta kot Alan Turing, računalniški Einstein in oče umetne inteligence (in se je rodil avtor tega prispevka).

Eno izmed najbolj poznanih ocen števila civilizacij je postavil Drake s svojo enačbo leta 1961 [6]. Ocena števila razvitih zunajzemeljskih civilizacij v naši Galaksiji je (x je znak za množenje) N = R* x fs x fp x ne x fl x fi x fc x L, kjer velja:
R* je razmerje zvezdnih skupin v naši Galaksiji, tj. hitrost nastajanja novih zvezd, število reda velikosti 1.
fs je delež Soncu podobnih zvezd z lastnimi planeti.
fp je delež zvezd z lastnimi planeti.
ne je povprečno število planetov, ki omogočajo življenje v razmerju z zvezdo, ki ima planete.
fl je delež planetov, kjer se domnevno lahko razvije življenje.
fi je delež planetov, kjer se dejansko razvije življenje.
fc je delež tistih, ki bi želeli ali se bili sposobni sporazumevati.
L je pričakovana življenjska doba take civilizacije.

V bistvu ta enačba temelji na oceni števila planetov; pomnoženi z verjetnostjo, da so na planetu ugodni pogoji za razvoj življenja; pomnoženi z verjetnostjo nastanka življenja; pomnoženo z verjetnostjo tehnološko napredne civilizacije; skalirano s časom trajanja civilizacije, saj nas zanima število civilizacij sedaj. V tej enačbi se trajanje naše galaksije pokrajša, ko se pojavi pri R kot produkt in pri L kot imenovalec (glej enačbo).

Ker je izračun odvisen od privzetih vrednosti in ker imajo različni strokovnjaki razne ocene, Wikipedija predlaga tri možnosti za konkretne številke:
- Ena civilizacija v naši galaksiji – komunikacija oz. zaznavanje drugih civilizacij s sedanjimi sredstvi ni mogoče.
- 100 civilizacij naši galaksiji, v povprečju oddaljenih 5000 svetlobnih let med dvema civilizacijama, ki lahko med seboj komunicirata.
- Štiri milijone civilizacij, komunikacija preprosta.

Zvezd in planetov je glede na naše znanje ogromno: galaksij je okoli 10 na 11 do 12 in podobno število je zvezd v galaksiji, torej je skupno število zvezd okoli 10 na 23 do 24. Trenutna ocena planetov je okoli 10 na 22. Če je med temi planeti le na vsakem milijardnem civilizacija, to pomeni 10 na 13 civilizacij, to je deset tisoč milijard. V naši galaksiji je okoli 10 na 11, tj. sto milijard planetov, kar po podobnem merilu prinese sto civilizacij s povprečno razdaljo med njimi pet tisoč svetlobnih let. Na taki razdalji je možno prepoznati močne energijske signale.

Verjetnost, da je življenje edino na Zemlji, je praktično nič, pa če pregledate Googla, Wikipedijo ali sami malce preračunate. Samo – sosednjih naprednih civilizacij gotovo ni, ker bi jih že opazili! Prav daljnih civilizacij trenutno najverjetneje ne moremo prepoznati, niti srednje oddaljenih, če se nočejo razkriti.

Precej vprašanj ostaja odprtih, vendar nekaj lahko vseeno sklepamo: Če so civilizacije prav redko posejane po vesolju, potem so lahko trajale dalj časa, pa se niso opazile med seboj. Če pa so gosteje posejane, tj. če se življenje seli s planeta na planet v obliki trosov ali medzvezdnih poletov, morajo propasti relativno kmalu, drugače bi jih že opazili. Kaj je to »kmalu«, ne vemo, od reda velikosti 100 do 10.000 let. V primeru redke poseljenosti imamo opravka s pojavom, da življenje nastane samo od sebe v primernih pogojih, kar smo v laboratorijih že pogosto postorili, vendar življenje ne uspe poseliti bližnjih planetov. Dovolj daleč narazen pa lahko civilizacije obstajajo dalj časa, tudi stotine milijonov let, a se ne uspejo zaznati zaradi ogromnih razdalj.

Naslednje vprašanje je, zakaj ni starejša in napredna civilizacija poselila vsega vesolja. V grobem ostajajo tri možnosti:
- Potovanje na daljše razdalje je preveč zahtevno, ker ni mogoče potovati hitreje kot s hitrostjo svetlobe
- Trajanje civilizacije je omejeno / Nekaj žre civilizacije [7]
- Napredne civilizacije pustijo razvijajoče se, kot npr. našo, in se nočejo vpletati.

Trije raziskovalci s Future of Humanity Institute (FHI) [8], Oxfordska univerza, med njimi Anders Sandberg, (s katerim je avtor prispevka intenzivno debatiral med konferenčnim izletom na kitajski zid) so objavili drugačno razlago Fermijevega paradoksa [9]. Za vsak parameter v enačbi so vstavili verjetnostni interval in izračunali končno verjetnostno distribucijo. V članku avtorji trdijo, da je verjetnost, da smo edina civilizacija v galaksiji 53 do 99,6%, da smo edini v vesolju pa 39 do 58 %. Torej drugih civilizacij ne zaznamo zato, ker so tako redke. Dokaj verjetno smo edini v galaksiji in morda tudi v vesolju.

Pomemben zaključek njihove raziskave je, da ni posebne nevarnosti, da bo naša civilizacija kmalu propadla. Izračun temelji na verjetnostni porazdelitvi in predvsem pri verjetnosti življenja da nizko oceno – kot da življenje le redko uspe, oz. le na prav posebnih planetih. Tak posebni primer naj bi bil naš planet, ki ima luno, ki je nastal s trkom s planetom velikosti Marsa itd. Tako trdi tudi Gribbin [10].

Naši izračuni še potekajo. Prve ugotovitve bodo predstavljene na konferenci Slovenian Conference On Artificial Intelligence v okviru multikonference Informacijska družba.

4. ZAKLJUČEK

Umetna inteligenca se razvija in se ukvarja z vprašanji reproduciranja osnovnih kognitivnih sposobnosti, ki se pojavljajo pri otrocih. Na ta način pričakujejo, da bodo dosegli izboljšavo sedanjih formalnih metod AI, ki nimajo praktično nobenih najbolj pomembnih človeških lastnosti. Hkrati AI čedalje bolj vpliva na vsakodnevno življenje v pozitivnem smislu – očitno je eden izmed ključnih generatorjev napredka [11].

V zadnjih desetletjih smo odkrili, da ima praktično vsaka zvezda najverjetneje nekaj planetov, torej je planetov ogromno. To je napovedoval že princip mnogoterega znanja [12], ki je tudi napovedal precejšnje število civilizacij. Hkrati pa ne opazimo nobene napredne civilizacije v delu vesolja, ki ga zaznajo naši

teleskopi. Čedalje bolj napredne analize in ocene števila civilizacij v vesolju nakazujejo, da je civilizacij malo.

Obstajata dve teoriji: po eni je civilizacij dejansko malo in smo ena redkih civilizacij v tem delu vesolja, vendar hkrati ni posebne nevarnosti, da bomo kmalu propadli. Po drugi teoriji je bilo pred nami že ogromno civilizacij relativno blizu nas, vendar so vse po vrsti propadle. Ta usoda čaka tudi nas, najverjetneje, če se ne bomo pravočasno zavedli problema in potegnili primernih potez.
Zlasti v primeru druge možnosti se je potrebno zavedati, da je čim hitrejše potovanje vsaj na bližnje lune in planete nujnost preživetja, kot večkrat poudarja Elon Musk.

LITERATURA
[1] Bostrom, N. 2014. *Superintelligence – Paths, Dangers, Strategies.* Oxford University Press, Oxford, UK.

[2] Yampolskiy, R.V. 2016. *Artificial Superintelligence*. CRC Press.

[3] http://www.tromba.si/matjaz-gams-scenariji-superinteligence-in-cloveske-civilizacije/ Superinteligenca

[4] IJCAI 2018 conference (https://www.ijcai-18.org/).

[5] Gams, M. 2017. Editor-in-chief's introduction to the special issue on "Superinteligence", AI and overview of IJCAI 2017. Informatica : an international journal of computing and informatics, ISSN 0350-5596, vol. 41, no. 4, str. 383-386

[6] https://sl.wikipedia.org/wiki/Drakova_enačba, https://en.wikipedia.org/wiki/Drake_equation, Wikipedija Drake equation

[7] http://beta.finance-on.net/files/2016-04-10/Matjaz-Gams.pdf Nekaj »žre« civilizacije

[8] Future of Life Institute, https://futureoflife.org/

[9] http://www.jodrellbank.manchester.ac.uk/media/eps/jodrell-bank-centre-for-astrophysics/news-and-events/2017/uksrn-slides/Anders-Sandberg---Dissolving-Fermi-Paradox-UKSRN.pdf Analiza Drakove enačbe

[10] Gribbin, J. 2011. *Alone in the Universe: Why Our Planet Is Unique,* ISBN-13: 978-1118147979, ISBN-10: 1118147979, John Wiley and sons.

[11] Kurzweil, R. 2006. *The Singularity Is Near: When Humans Transcend Biology,* Sep 26, Penguin Books.

[12] Gams, M. 2001. *Weak intelligence: through the principle and paradox of multiple knowledge*. Nova Science.

The Quest for Understanding: Helping People with PIMD to Communicate with their Caregivers.

Matej Cigale
Jožef Stefan Institute
Jamova cesta 39
Ljubljana, Slovenija
matej.cigale@ijs.si

Mitja Luštrek
Jožef Stefan Institute
Jamova cesta 39
Ljubljana, Slovenija
mitja.lustrek@ijs.si

Matjaž Gams
Jožef Stefan Institute
Jamova cesta 39
Ljubljana, Slovenija
matjaz.gams@ijs.si

Torsten Krämer
Pädagogische Hochschule
Heidelberg
Institut für Sonderpädagogik
Keplerstraße 87D
Heidelberg, Germany
kraemer@ph-heidelberg.de

Meike Engelhardt
Pädagogische Hochschule
Heidelberg
Institut für Sonderpädagogik
Keplerstraße 87D
Heidelberg, Germany
engelhardt@ph-heidelberg.de

Peter Zentel
Pädagogik und Didaktik im
Förderschwerpunkt
Geistige Entwicklung
Keplerstraße 87
Heidelberg, Germany
zentel@ph-heidelberg.de

ABSTRACT
People with Profound Intellectual and Multiple Disabilities (PIMD) stand for a broad and very heterogeneous spectrum of people that are characterised by some common aspects like a severe intellectual disability usually in combination with a lack of conventional and symbolic communication abilities, coupled with the need for high levels of support due to comorbidities or other possible disabilities (i.e., motor or sensorial impairments). Supporting these individuals is extremely challenging, as their communication signals are atypical and idiosyncratic. Therefore, a plethora of these behaviours are not or not easily readable for the caregivers. Without background information on a specific person with PIMD, it is hard for a caregiver, even a trained professional, to interpret the desires and mental state of the person they are interacting with, which leads to a stressful interaction for both. With advances in computer vision (CV), speech recognition technology (SRT) and artificial intelligence (AI), we are making the first steps in codifying these behaviours and attempting to mechanically extract the meaning of the communication. The INSENSION project aims to use these advancements to catalogue the actions of persons with PIMD and the environment and thus provide feedback to caregivers and enable individuals to control their surroundings. A similar system could be used to analyse the behaviour of healthy individuals so that the generalised and personalised expressions of body language could be codified and compared across cultures and individuals.

Keywords
communication, gesture recognition, PILD, AI

1. INTRODUCTION
People with PIMD experience a lot of trouble when they are attempting to communicate to the outside world. Generally, a profound intellectual disability, which complicates the learning of new skills for them, is combined with other sensory or physical impairments, which lead to an unusual communication in comparison to people without disabilities.

On this occasion, the task of their caregivers (parents, therapists, etc.) is to interpret their communication attempts and teach them to make further attempts more distinct. The problem arises from the fact that the communication attempts of the people with PIMD are indistinct, mostly unique to each individual, and hard to distinguish or interpret for people who are not close to them. The differences stem from different abilities of each person, different reception of the gestures from their caregivers and other external factors.

At present, only close caregivers are able to interpret the desires of people with PIMD in the right way, nevertheless often combined with feelings of insecurity. This makes it difficult to easily expand the circle of communication partners. The INSENSION project faces this issue by aiming to use the advances in computer vision, specifically new ways to extract posture and facial expressions from video to codify them. In the second stage of the processing, the context of the expression will be extracted, i.e. what is the state of the environment around the person. This way, the expression can be coupled with the interaction and the intent of the communication can be interpreted. This would hopefully provide improvements for both sides: for the people in contact with people with PIMD to have a window into their internal states and see their attempts and, of course, for the people with PIMD themselves. However due to their disabilities, the communication attempts of people with PIMD seem simpler and their internal working models are often assumed to be limited to temporally and spatially neighbouring desires that makes the interpretation context smaller.

This research has broader implications as a similar, but more complex, system could be used to interpret communications of individuals so that their behaviour could be objectively determined enabling a more rigid research into the communication of people and their internal state.

The rest of the paper is organised as follows: In Section 2 we look at the state of the people with PIMD and their communication attempts. In Section 3 we take a brief look at the underlying technical advancements that can facilitate the extraction of posture and facial expressions and the vocalisations of the person. In Section 4 we present the annotations that will be the input for the Machine Learning (ML) system. In Section 5 we discuss the implications of this system and present some caveats to the system. In section 6 we look at broader implications of the system as it

could be used on a more general population to systematically codify the interactions as well as further fields of inquiry that could be developed based on this system.

2.   PEOPLE WITH PIMD

People with PIMD, as the name implies, have multiple disabilities, which makes it even harder for them to participate in the large number of non-barrier-free parts of our society. Generally, PIMD means a profound intellectual disability combined with other sensory (blindness, deafness) or physical impairments (lack of mobility, problems with fine hand movements, etc.). These factors severely influence the person's ability to live without any care, support or therapy of others [1]. Individuals with PIMD have an above average risk to get additional diseases, and frequently require regular medication that also implies administrative aid. Individuals are assumed to attempt to communicate but are often not able to do so successfully, because of the inherent and external limitations [2].

A common denominator in the population we are dealing with is the limited ability to communicate coherently with their caregivers or the other way around. They usually communicate on a pre-symbolic level and their understanding of speech is severely limited. Some individuals have the ability to form joined attention with their communication partner [3] but this is not universal. While they are capable of learning, the acquisition of new skills takes significantly more time and requires frequent repetition. People with PIMD tend to exhibit not or not easily readable behaviours in order to communicate their (dis-)pleasure or to get attention. Examples of these are pushing unwanted objects away, loud vocalisations or banging to gain attention. In general, the communication attempts of people with PIMD are relying on caregivers who have been trained to understand their communications by interpreting their whole body behaviour or specific personal expressions [4].

Communication attempts are very multifaceted based on the specific individual. Some persons are capable of vocalising simple words, such as saying "Hi", but do not consistently use them in a correct manner or they grab towards toys and individuals they want to interact with. Making eye contact is possible for some individuals, which can be an orientation towards desired objects as communication attempt. Others do not have any coherent vocalisations, lack motor skills and require help holding items [3], [5].

Further complicating their behaviour or the interpretation of is stereotypy. These are actions that do not contain communication attempts but can be considered "ticks" that do not carry meaning. There is a correlation between stereotypical behaviours and low level of social interaction and stimulation. This behaviour can escalate over time to aggressive behaviour and sometimes even self-injuring. The level of these problematic behaviours seems to be correlated with communication problems [5] and would presumably point to this being an expression of frustration. These behaviours can range from hand wringing, to hitting legs, head or nearby objects, from purposeful breath holding to screaming etc.

Several attempts were made to bridge the gap of communication of people with PIMD. Some individuals have access to switches that produce specific sounds enabling an easier communication for less experienced communication partners [4], [6]. These switches can take several forms form simple push buttons to systems that attach to the individual muscles. Other systems, such as Picture Exchange Communication where the individuals with PIMD are expected to provide a picture, usually on a card, for the desired communication

or Simplified Signalling that draws the inspiration form natural gestures that are taught to the people with PIMD so that their communication is in line with general public [6]. There is also research in using Brain Computer Interface for communication [7] that uses Electroencephalography (EEG) to map the activity of the brains to interpret the desires. The system requires adaptation based on the individual and training of the individual.

3.   ANALYSING THE HUMAN

Enabling computers to interpret the desires of its users and their state of mind is a longstanding goal of computer science. The most developed systems focus on speech recognition[8], [9], but other systems are also explored. Eye tracking [10] is becoming more and more robust, moving from specialised hardware to simple web cameras [11].

More advanced systems enable facial features extraction [12], and from this psychological state of the individual can be extrapolated [13]. These systems work extremely well on typical individuals in good conditions, for example in good lighting and direct camera position [14]. In more dynamic conditions, such as unstable lighting these results are less certain but still reach acceptable levels of accuracy. Additional information, such as voice inflection or contact sensor provides additional information and greater accuracy. Research has been conducted in extracting stress of students using only the smartphone, carried by the person[15].

Another advancement in computer vision is the possibility of mapping the body parts of one or more persons from video [16]. This enables extracting of the limb and torso position, the position of fingers and facial markers. This enables researchers to qualify the position of a person and increases the robustness and ease of analysing behaviours of humans. An example of this can be seen in Figure 1.

The position of a person is only part of the puzzle of finding out what the person wants. In order to determine the context of the interaction there is another piece of the puzzle. The environment of the person must be taken into account. There are several systems that can take a video and return the objects present in the scene. One of the fastest open source solutions is YOLO [17]. Other commercial systems also exist, such as Google or Amazon Object recognition API [18].

Figure 1. An example of the image processed by OpenPose. The Fingers, arms and facial characteristics are extracted and returned as points.

4. THE EXPRESSIONS OF THE PEOPLE WITH PIMD

In our research we are working with six people with PIMD in order to provide a robust data set that will be used to train the system and extract the state of the individual and the communication they are trying to accomplish. Based on this we hope to extract some rudimentary information, such as their psychological state (pleasure, displeasure or neutral) and the mode of communication they are exhibiting such as protest, demand or comment. Based on this information and the context – the activity that was happening before, the objects that are available for interaction and the estimation of internal needs based on models such as hunger, thirst etc. – the system will propose the action the caregiver should take.

The first step for this is collecting the data. In our case this video is manually annotated in order to provide information for the system. We use several cameras, an infrared camera, a wristband to collect physiological parameters and microphones that collect sounds. The videos are collected in several interactions ranging from meal time, playing (Figure 2) and physical therapy and even some life-skills training.

Figure 1. An image of video recording of a play session with a caregiver.

The videos are annotated to indicate the position of the individuals so that the system can be trained to return the desired information such as arm position, facial expressions and actions such as rubbing parts of the body, interacting with objects or people, presence of
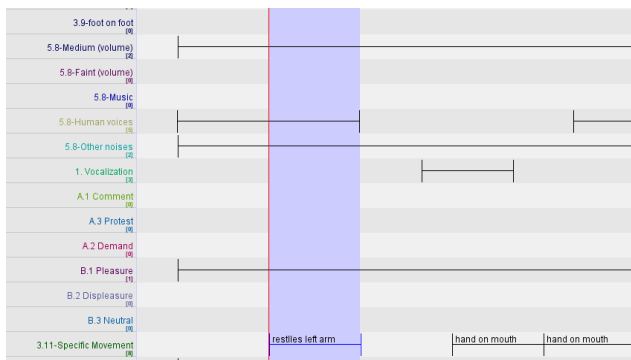

Figure 1. An image of video recording of a play session with a caregiver.

people in the scene and any external disruptions, such as loud noises etc. In Figure 3 we can see an example of the resulting annotation.

5. OVERVIEW OF THE PROPOSED SYSTEM

Our system will take the information collected from the camera, microphone and other sensors, extract the objects in the scene and the information about the communication attempts of the users and, based on this, provide a guess on the presumable mood and communicated content.

In the first round, communication classes will be extracted using unsupervised learning that will return common actions performed by the individual, as annotated or provided by the movement analysing system. This will return meaningful communication clusters. These will be simple like moving the hand from position A (on head) to position B (pointing), interacting with the desk, vocalisation or a combination of these, such as pointing and producing a distinct sound.

All these communication clusters can then be mapped onto the state and communication system. At this stage supervised learning will be used. Some actions will for instance indicate that the user is displeased - providing the information on the internal state of the individual. While others will indicate communication attempts such as demand for something. Together they can indicate to the caregivers what action they should take to provide care for their charges.

The first problem we have to solve is to remove the behaviour that does not carry any communication information. In general, it is expected that action that happen regardless of the state of the individual will be filtered out. This could become problematic as some actions can have several meanings. As such the efficacy of this is considered ongoing research.

In addition, the context, such as presence of strangers or objects in the vicinity, will be used to provide information on the desires and aversions of the individual, creating an internal database of likes and dislikes that can be taken into consideration.

6. CONCLUSIONS

As the system is not yet operational, there are several problems, that are still part of research. For instance, it is unknown which role stereotypy will play, or how accurately we can extract the information if the user is interacting with a specific object or person. The information that the user is interacting with one object may not be enough to infer the actual desire of the user.

Once this, and other problems are resolved the system can be extended to general public. Enabling researchers and others to analyse the behaviour of the individuals and extract the communication desires and psychological state of the individual to further their understanding of the motivations and desires of people, providing standardised analysis of their movements and philological states. This will in turn enable greater rigidity of inquiry providing a faster and reproducible way of analysing behaviour of people.

However, the system could be misused. Organisations and individuals could use the system to determine the state of the individual and use this to manipulate her or him for their gain. A system that can extract the information could be used to associate some products and people with this feeling, thus steering the people to the competition. Furthermore, body language could be important for determining not the information but relationships of people, creating an internal map that could result in an interaction map that

could be used to influence certain groups of people or simply to determine the connections between them and their reactions to certain events.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] H. Nakken and C. Vlaskamp, "A Need for a Taxonomy for Profound Intellectual and Multiple Disabilities," *J. Policy Pract. Intellect. Disabil.*, vol. 4, no. 2, pp. 83–87, Jun. 2007.

[2] K. Atkin and M. P. Lorch, "An ecological method for the sampling of nonverbal signalling behaviours of young children with profound and multiple learning disabilities (PMLD)," *Dev. Neurorehabil.*, vol. 19, no. 4, pp. 211–225, 2016.

[3] P. J. Yoder and S. F. Warren, "Relative treatment effects of two prelinguistic communication interventions on language development in toddlers with developmental delays vary by maternal characteristics.," *J. Speech. Lang. Hear. Res.*, vol. 44, no. 1, pp. 224–237, 2001.

[4] C. Harding, G. Lindsay, A. O'Brien, L. Dipper, and J. Wright, "Implementing AAC with children with profound and multiple learning disabilities: A study in rationale underpinning intervention," *J. Res. Spec. Educ. Needs*, vol. 11, no. 2, pp. 120–129, 2011.

[5] Me. Laura Roche, "Evaluating and Enhancing Communication Skills in Four Adolescents with Profound and Multiple Disabilities," Victoria University, 2017.

[6] J. Goldbart and S. Caton, "Communication and People with the most complex needs: what works and why this is essential," *What is communication and what are complex communication needs?*, no. July. pp. 3–4, 2010.

[7] H. Adeli, P. Rahimian, and N. Tabrizi, "Communicating with People with Profound Intellectual Disabilities Using Brain Computer Interface," *J. Technol. Pers. with Disabil.*, vol. 4, pp. 133–145, 2016.

[8] M. Benzeghiba *et al.*, "Automatic speech recognition and speech variability: A review," *Speech Commun.*, vol. 49, no. 10–11, pp. 763–786, 2007.

[9] M. Anusuya and S. Katti, "Speech recognition by machine: A review," *Int. J. Comput. Sci. Inf. Secur.*, vol. 6, no. 3, pp. 181–205, 2009.

[10] C. H. Morimoto and M. R. M. Mimica, "Eye gaze tracking techniques for interactive applications," *Comput. Vis. Image Underst.*, vol. 98, no. 1, pp. 4–24, 2005.

[11] J. San Agustin *et al.*, "Evaluation of a low-cost open-source gaze tracker," *Proc. 2010 Symp. Eye-Tracking Res. Appl. - ETRA '10*, p. 77, 2010.

[12] W. S. Chu, F. De La Torre, and J. F. Cohn, "Selective transfer machine for personalized facial expression analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 3, pp. 529–545, 2017.

[13] Z. Yin, M. Zhao, Y. Wang, J. Yang, and J. Zhang, "Recognition of emotions using multimodal physiological signals and an ensemble deep learning model," *Comput. Methods Programs Biomed.*, vol. 140, pp. 93–110, 2017.

[14] H. Gunes and H. Hung, "Is automatic facial expression recognition of emotions coming to a dead end? The rise of the new kinds on the block," *Image Vis. Comput.*, vol. 55, no. 1, pp. 6–8, 2016.

[15] M. Gjoreski, H. Gjoreski, M. Lutrek, and M. Gams, "Automatic Detection of Perceived Stress in Campus Students Using Smartphones," *2015 Int. Conf. Intell. Environ.*, pp. 132–135, 2015.

[16] S.-E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional Pose Machines," in *Conference on Computer Vision and Pattern Recognition*, 2016.

[17] J. S. D. R. G. A. F. Redmon, "(YOLO) You Only Look Once," *Cvpr*, 2016.

[18] R. Buyya, C. S. Yeo, and S. Venugopal, "Market-oriented cloud computing: Vision, hype, and reality for delivering IT services as computing utilities," *Proc. - 10th IEEE Int. Conf. High Perform. Comput. Commun. HPCC 2008*, pp. 5–13, 2008.

Kognitivna raven kot del relacijske družinske terapije pri predelovanju spolnega nasilja

Tanja Repič Slavič

Teološka fakulteta, Poljanska 4,
Ljubljana, katedra za Zakonsko in
družinsko terapijo
ter psihologijo in sociologijo religije

Frančiškanski družinski inštitut,
Prešernov trg 4, Ljubljana

E-mail: tanja.repic@teof.uni-lj.si

POVZETEK

Posamezniki, ki so v otroštvu doživeli spolno nasilje ter se v odraslosti soočajo s posledicami, potrebujejo za globinsko preoblikovanje vzorcev, odnosov in notranje drže, strokovno pomoč. Brez dela na sebi se pogosto lahko znajdejo nazaj v začaranih krogih nasilja, ali kot žrtve ali kot storilci, in doživljajo neprijetne posledice nasilja, ki jih celo nezavedno prenašajo naprej na otroke. Predno pa je posameznik sposoben iti v globino čustev in nezavednega spomina, je v relacijski družinski terapiji (RDT) s strani terapevta nujno poskrbeti za varen prostor, ki ga med drugim omogoča prav vrednotenje travme na kognitivni ravni, o čemer bomo pisali v prispevku. V tem kontekstu bomo najprej predstavili nekatera teoretična izhodišča v povezavi z RDT in spolnim nasiljem, potem pa bomo bolj natančno osvetlili stopnje procesa RDT na kognitivni ravni.

Ključne besede

Spolno nasilje, relacijska družinska teorija in terapija, kognitivna raven, razreševanje travme.

1. SPOLNO NASILJE

Spolno nasilje je vsako vedenje ali dejanje s spolno vsebino, ki si ga žrtev ne želi in pri njej vzbuja občutek, da je njeno telo le inštrument za zadovoljitev potreb [1]. Pri spolnem nasilju gre za vsa spolna dejanja povezana s spolnostjo, ki jih oseba čuti kot prisilo, ob njih doživlja sram in gnus ter da so bile prekoračene telesne meje in meje osebnega dostojanstva. Na drugi strani pa storilec to stori z namenom spolnega vzburjenja [2, 3]. Nasilje se pojavlja na različnih nivojih, še posebej boleče pa je, kadar nastopa kot relacijska travma, kar pomeni da se izvaja s strani žrtvi bližnje osebe, npr. družinskega člana. Še zlasti težki obliki nasilja, ki se izvajata v družini ali pa neučinkovito razrešujeta, sta fizično in spolno nasilje.

Žrtve nasilja so zaznamovane z dolgotrajnimi posledicami. Raziskave med najpogostejšimi posledicami omenjajo posttravmatsko stresno motnjo, neprimerno regulacijo afekta, depresijo, anksiozne motnje, zlorabo drog in alkohola, kompulzivno vedenjsko ponavljanje travmatičnih scenarijev, občutke sramu in krivde, destruktivnost do sebe in drugih, vedenjsko impulzivnost, agresivne izpade, avtoagresijo, spremenjeno podobo o sebi in podobno [4, 5].

Take izkušnje imajo velik vpliv tudi na telesno zdravje [6, 7].

Samo nasilje je za žrtev lahko tako travmatično, da če se le ta ne zdravi, lahko posledice postanejo in ostanejo doživljenjske [8]. Pomembno je poudariti, da otrok, ki je bil žrtev spolnega nasilja, v obdobju mladostnika ali odraslega posameznika predstavlja veliko tveganje, da sam postane povzročitelj spolnega nasilja nad otroci [9]. Raziskave so pokazale, da je spolna zloraba v otroštvu zelo pogosta med spolnimi prestopniki, saj ponavljajo priučeno vedenje, ki so ga bili sami deležni, torej jih je v otroštvu tako zelo travmatiziral in zaznamoval. Spolni prestopnik se kot otrok nauči pozitivnega odnosa do spolnega vedenja med otrokom in odraslim človekom [10]. Za žrtve posilstva obstaja velika verjetnost, da bodo ponovno žrtve, za deklice, ki so bile telesno ali spolno zlorabljene, pa je velika verjetnost, da bodo ponovno zlorabljene, kot odrasle ali da bodo postale prostitutke [11].

2. RELACIJSKA DRUŽINSKA TEORIJA IN SPOLNO NASILJE

RDT poudarja, da zgodnje travme, kamor sodi tudi spolno nasilje, lahko povzročijo najpogosteje nezmožnost pri reguliranju agresivnih impulzov, strahu, sramu in gnusa [12]. Ti afekti ostanejo zelo globoko ukoreninjeni v psihobiološki strukturi otroka in še v odrasli dobi lahko silovito zaznamujejo posameznika, ki je bil žrtev zlorabe, zlasti pa se prebudijo, ko pride v odraslem odnosu do intime. Zaznamuje pa tudi njegovo okolico, ki se večkrat z velikim nerazumevanjem do njegove bolečine odvrača od njega, saj ne razume, zakaj se vedno znova zapleta v odnose, kjer je spet zlorabljena [13, 2]. Gostečnik [12, 14] meni, da se med spolnim nasiljem vzpostavi izredno močna afektivna povezava med rabljem in žrtvijo. Zato lahko žrtev zlorabe ravno na osnovi tega odnosa, ki je bil kruto vsiljen, v novih situacijah nezavedno išče zlorabo oziroma v zlorabi išče podobne situacije, vzdušja in ljudi, ki jo bodo ponovno zlorabili. Pri tem posameznik podoživlja afekte zlorabe, kot bi se zloraba dogajala tukaj in sedaj, ne pa kot del preteklosti. Nezavedno čutijo, da morajo ostati zvesti tem občutjem, saj bi sicer čutili obupno praznino, močno hrepenenje, ki celo organsko boli, dokler ne pride do ponovitve začaranega kroga in pomiritve. To praznino vsekakor lahko pomirijo tudi z raznimi oblikami zasvojenosti, z avtodestruktivnimi vedenji, kot so samopoškodovanje in motnje hranjenja, ... [15]. Zloraba pomeni odnos, in odpovedati se odnosu za zlorabljenega pomeni izgubiti

vse, zato pogosto vztrajajo v še tako bolečih odnosih, v upanju na razrešitev [16].

S področja pomoči pri reševanju posledic spolnega nasilja je sicer veliko različnih tehnik, svetovalnih metod in pristopov, kar pa omogoča le razreševanje na bolj površinski ravni, saj ni dovolj možnosti za naslavljanje vseh dinamik in čustvenih stanj, ki jih pusti travma nasilja. Sodobna relacijska psihoterapevtska paradigma, kamor se uvršča tudi model relacijske družinske terapije (RDT), poudarja prav to: nerazrešene psihične vsebine, ki so zaznane tudi na nevrobiološki ravni, prihajajo na dan v različnih simptomatičnih in velikokrat nefunkcionalnih oblikah vedenja. RTD preko mehanizmov projekcijsko introjekcijske identifikacije (transferja in kontratransferja) prodira v nezavedni spomin in tako z ozaveščanjem le tega pomaga regulirati posameznikova psihobiološka stanja in afekte, ki izhajajo iz teh stanj [12, 17]. Ker je razreševanje travme zelo kompleksen proces, saj potlačene psihosomatske vsebine prihajajo postopno na dan, je nujno v terapiji zagotoviti varen prostor.

3. PROCES RELACIJSKE DRUŽINSKE TERAPIJE PRI RAZREŠEVANJU SPOLNEGA NASILJA

Relacijska družinska terapija postavlja v središče terapevtskega dela odnos, ki omogoča spremembo osnovnih relacijskih struktur. Velik pomen daje ozaveščanju potlačenih vsebin in zagotovitvi manjkajočih zgodnjih doživetij. Vendar pa ne gre zgolj za drugačna doživetja in popravljanje starih, nedokončanih procesov razvoja, temveč za razumevanje temeljnega vzorca odnosov, ki ga živi posameznik, in za spremembo teh temeljnih vzorcev odnosov. Klient zagotovo potrebuje profesionalno pomoč, zlasti pri premagovanju in soočanju s strahom, samoobtoževanjem, krivdo, z jezo in drugimi močnimi čutenji ter afekti, ki jih pusti spolna zloraba [12, 18]. S pomočjo novih izkušenj, ki se zgodijo v terapevtskem odnosu, bo prihajalo pri klientu do sprememb. Te spremembe ne bodo povezane samo z vsakdanjim življenjem, ampak tudi s spremembami v možganih [19], kar posledično vpliva tudi na hormonski sistem ter preusmeritev privlačnosti, kar v praksi pomeni, da zlorabljenemu niso nezavedno več privlačni ljudje, ob katerih bo ponovno podoživljal vzdušje in afekte originalne travme [12, 17].

Rothschildova [16, 20] meni, da mora vsako terapijo, povezano s travmatičnimi doživetji, spremljati deset temeljnih načel, ki so:

1. vzpostavitev varnosti za posameznika, tako na terapiji, kot izven nje; 2. razviti dober in učinkovit terapevtski odnos; 3. sposobnost in zmožnost ustavitve terapevtskega procesa; 4. identificirati in graditi posameznikove notranje in zunanje vire; 5. graditi mehanizme za soočenje s travmo; 6. vedno znižati pritisk; 7. prilagoditi se posameznikovim potrebam; 8. poznati širok spekter psihofizioloških teorij o travmi in PTSD; 9. upoštevati posameznikove svojske razlike in posebnosti ter 10. terapevt mora biti pripravljen opustiti določene ali vse terapevtske tehnike.

V tem kontekstu temeljnih načel dela s travmo, se prepletajo v samem procesu razreševanja različne teme in stopnje, ki pa se pri vsakem posamezniku odvijajo v drugačnem vrstnem redu. Prav tako ni značilno, da bi se neka stopnja zaključila in začela druga, ampak se lahko iste stopnje večkrat ponovijo, vendar vsakič na drugačen način, na drugem nivoju. V tem oziru proces okrevanja spominja na spiralo, kjer se ponovijo isti nivoji, a vsakič raven višje. Pomembno je vedeti, da ima vsak posameznik svojo pot okrevanja. Ni boljših in hitrejših poti, so le različne poti, ki vodijo k istemu cilju. V relacijski družinski terapiji je zagotovo cilj, da se spolno nasilje razrešuje na kognitivni ravni, na ravni čutenj in afektov ter na telesni ravni [2, 21]. V nadaljevanju bomo predstavili kognitivno raven, ki je na nek način najbolj na površini, a brez nje do ostalih dveh naprej ni mogoče prodreti.

4 KOGNITIVNA RAVEN PRI PREDELAVI SPOLNEGA NASILJA

Klinične izkušnje [2] kažejo, da zlorabljeni začnejo predelovati spolno zlorabo najprej na kognitivni ravni . Ko govorimo o kognitivni ravni, imamo v mislih racionalno dojemanje, miselne procese posameznika. V tem okviru se prepletajo različne teme in stopnje procesa, ki pa se ne odvijajo pri vseh v istem vrstnem redu.

a.) ODLOČITEV: Najprej se mora zlorabljeni odločiti, da želi nekaj spremeniti. Do te odločitve jih največkrat pripeljejo posledice spolne zlorabe, ki jih doživljajo na različnih področjih. V večini primerov nihče moči teh posledic ne pripisuje spolni zlorabi ampak sebi, kot tistemu, s katerim je nekaj zelo narobe, saj sicer ne bi doživljal toliko stisk in kriz. Slutijo pa, da tu in tam travmatičen dogodek prispeva del bolečin, vendar nikakor v takem obsegu in s tako intenzivnostjo, kot se kasneje v procesu terapije izkaže.

b.) KOGNITIVNO DOJEMANJE POSLEDIC SPOLNEGA NASILJA: ko terapevt začne na srečanjih vrednotiti te stiske in iskati izvor sedanjih doživljanj, ki se le prebujajo in prihajajo od drugod, začnejo zlorabljeni postopno kognitivno dojemati resnične posledice spolne zlorabe, samo dinamiko, povezave s preteklostjo in sedanjostjo … Čutenja, ki se jim ob tem pojavljajo, in telesne senzacije, ki jih doživljajo, najprej le racionalno povezujejo z zlorabo, in normalno je, da občasno celo podvomijo in zanikajo, da bi lahko zaradi te travme doživljali toliko groz.

c.) GLASNO PRIZNANJE, DA SE JE SPOLNO NASILJE ZARES ZGODILO: velik korak naprej je, ko lahko posameznik glasno pred vsemi v skupini (če je terapija skupinska) ali le pred terapevtom (če je terapija inividualna) po nekaj srečanjih izgovori in prizna, da je bil spolno zlorabljen. Običajno na začetku besedne zveze spolna zloraba sploh še ne zmorejo izreči, če pa že, jo izrečejo zelo disociirano – brez čutenj, bolj racionalno, kot da to ni nič tako zelo hudega.

Zaradi krutosti dogodka je povsem normalno, da se občasno pri posameznikih pojavljajo dvomi, da si morda vse samo domišljajo, da so kaj takega sanjali in da se jim to sploh ni res zgodilo. Vendar se v času terapije prej ali slej pokaže, da grozne sanje pustijo drugačen pečat in ne vplivajo tako močno in tako intenzivno na človekovo doživljanje, vsakdanje življenje. Prav tako se tudi izkaže, da so dvomi

le posledica bolečine in stisk ter da obstajajo tudi resnična dejstva, ki ovržejo dvom in potrdijo resničnost doživljanj. Terapevt mora pri tem verjeti, da se bo v času terapije izkazalo in da je možno prepoznati, začutiti, kdaj bi bil nek spomin implantiran ali izmišljen.

Č.) RAZKRITJE SPOLNEGA NASILJA TUDI DRUGIM: občutek olajšanja, da zlorabljeni ne nosi več sam svoje temne skrivnosti, nekatere vzpodbudi, da tvegajo povedati o spolni zlorabi tudi komu drugemu izven terapije – morda enemu od družinskih članov (izkušnje kažejo, da najlažje sestri ali bratu, šele potem mami, očetu redkokdo), ali nekomu od prijateljev. Pomembno pri tem je, da žrtev zlorabe čuti, da je ta oseba vredna zaupanja in da bo razumela zlorabo – da ne bo obsojala, minimalizirala dogodka, opravičevala storilca … Dogaja se tudi ravno obratno – da posamezniki, preden pridejo na terapijo, vsem brez razmejitev razlagajo, kaj se jim je zgodilo, saj na ta način regulirajo bolečino in sproščajo napetost. V času terapij pa kmalu vidijo, da so izbirali predvsem take osebe, ki njihovih klicev na pomoč niso slišali. Tako začnejo postopno postavljati meje in svoje zgodbe ne pripovedujejo več vsakemu, ki jih bolj ko ne iz vljudnosti vpraša, kako so, saj se s tem samo še bolj ranijo in razvrednotijo.

d.) OZAVEŠČANJE TRAVME SPOLNEGA NASILJA IN DVOJNO PREPOZNAVANJE – SPOMINJANJE: naravna pot procesa je, da začne v času terapij postopno prihajati na dan vedno več potlačenih čutenj, afektov, telesnih reakcij … Pri nekaterih se te vsebine prebujajo v odnosih z drugimi ljudmi, pri drugih prek sanj, branja neke knjige, lahko prek močnih telesnih odzivov, gledanja kakega filma, vedenj, ki si jih ne znajo razložiti … Vse te reakcije mora terapevt nenehno vrednotiti in skupaj z udeležencem terapije prek povezav iskati, kam to spada, kje se je že tako počutil, od kod to prihaja … Prve tedne ali celo mesece zlorabljeni te povezave razumejo le na kognitivni ravni – verjamejo, da je to res tako, lahko najdejo povezave, si to ponavljajo, vendar jim to še ne prinaša nekega olajšanja, saj se še niso zmogli tudi čustveno umiriti, telesno sprostiti. Pri tem je bistveno in nujno potrebno **vztrajanje, nenehno ponavljanje in utrjevanje teh povezav** dvojnega prepoznavanja tega, kar se je zgodilo v preteklosti in se sedaj samo prebuja, se ne dogaja več. Spolno zlorabljeni vsa čutenja in afekte, ki so povezani s travmo spolne zlorabe in se v neki sedanji situaciji samo prebujajo, doživljajo na način, kot da se travma dogaja **ravno sedaj** [17, 22]. Njihovo telo, čustva, mimika, ton glasu, vedenje, vse se odzove na podoben način, kot da jih prav sedaj nekdo spolno zlorablja. Ko se jim podobna občutja pojavljajo v okolici, pogosto povedo, da jih drugi označujejo kot »preobčutljive«, »zakomplicirane«, »neuravnovešene«, »čudne« … Prav ponavljanje in utrjevanje teh povezav pa postopno začne spreminjati tudi čustvene in telesne odzive. V pomoč pri tem je lahko tudi obrazec dvojnega prepoznavanja [2, 20], da se prek njega tudi izven terapevtskih srečanj utrjujejo novi vzorci. Na začetku pa udeleženci na terapijah le kognitivno dojemajo dejstva in se čustveno ter telesno stvari kaj bistveno še ne spreminjajo, kar pa je povsem razumljivo; če je namreč nekdo živel dvajset, trideset ali več let z vzorci in posledicami spolne zlorabe, se to čez noč nikakor ne more začeti spreminjati.

V tem kontekstu lahko rečemo, da v prvih tednih ali mesecih na terapijah čustva in afekte dojemajo kognitivno, kar pomeni, da že vedo, da niso bili oni krivi, vendar se krivde še ne morejo znebiti. Prav tako šele postopno lahko priznajo, da se jim je zgodila krivica, vendar jeze na začetku sploh še ne zmorejo začutiti, če pa že, jo usmerjajo na »napačne« osebe in izražajo na neprimerne načine. S strani terapevta korak za korakom v začetnih fazah dojemajo, da je v takem dejanju s strani storilca čutiti ogromno sramu, gnusa, prezira, vendar večinoma ostajajo še indiferentni, včasih prav depresivni in v svojem svetu – ne da bi telesno začutili omenjena čutenja. Običajno ta čutenja začnejo doživljati veliko prej v vsakdanjem življenju, ob osebah, ki niso bile originalen storilec, postopno celo ob terapevtu. V tem se jasno vidi transfer, prenos čutenj prek projekcijsko-introjekcijske identifikacije na neko drugo osebo, kar se za začetek izkaže kot bolj varno, manj tvegano, kot če bi to čutili ob storilcu. Na to kaže tudi najpogostejši odgovor na vprašanje, koliko so jezni in besni na tistega, ki jih je spolno zlorabljal. Ob tem večina samo skomigne z rameni, da jeze ne čutijo, čeprav bi se jim zdelo prav, da bi jo čutili – razumsko torej vedo, da bi »morali« čutiti jezo, gnus, bes, vendar čutenjsko tega še niso »sposobni« oziroma še ne smejo začutiti. Z drugimi besedami, ni še dovolj varno in je v terapiji potrebno prej odkriti še druge plasti, obrambne mehanizme, afektivne psihične konstrukte in simptome kompulzivne retravmatizacije, da lahko pridejo do samega jedra bolečine.

e.) DOŽIVLJANJE KRIZ: spomini in čutenja, ki se začnejo intenzivno prebujati, drugačno gledanje na okolico, ljudi okrog sebe, izvirno družino ter nove in nove teme, ki se še odpirajo, sprožajo obdobja kriz. Posamezniki bodo v takih trenutkih povedali, da se jim zdi, da je sedaj še slabše, da je bilo lažje živeti prej, ko niso ničesar vedeli, da sedaj nenehno razmišljajo samo o zlorabah, da nimajo nobene moči, da bi se ukvarjali s čim drugim, da se jim zdi, da bodo zapravili vse dobre priložnosti, ki se jim ponujajo, da niso sposobni videti drugega, razen svojih stisk in težav, da nimajo energije, da bi si sploh kaj lepega zase privoščili, da se jim zdi, da so nori, da bi se najraje zapili, za vedno zaspali, pustili službo, študij … V takih krizah klinične izkušnje kažejo, da najbolj pomaga, če imajo izven terapije ob sebi vsaj še eno osebo, na katero se lahko obrnejo, včasih tudi to, da si jasno strukturirajo dan, predvsem takrat, ko jim je najtežje in bi samo spali, ter da obvezno naredijo zase vsaj nekaj pozitivnega [23]. Prav tako nekaterim pomaga, če lahko ubesedijo svoje stiske in krize na list papirja in napišejo čisto vse, kar čutijo, brez »cenzure«, samo da dajo ven iz sebe. Taki zapisi se izkažejo za zelo pozitivne tudi kasneje, ko pride morda novo obdobje krize in ko z branjem podoživljajo, kaj vse so že prestali, preživeli, premagali ... Mnogi povedo, da so se najbolj umirili, ko so težka čutenja uspeli umestiti v čas in prostor, jih ovrednotiti kot sestavni del procesa, ki ne traja večno in vedno mine. Četudi se obdobja kriz običajno še ponovijo, nikoli niso več tako intenzivna oziroma so bolj obvladljiva in manj časa trajajo. Ko so posamezniki razmišljali o premaganem obdobju krize, so videli, da so prav zaradi takih »zdravilnih« kriz lahko kasneje šli hitreje in drugače naprej, kot bi šli sicer brez njih. Cilj pri razreševanju travme spolne zlorabe je iti *samo naprej,*

četudi včasih najprej korak nazaj, da lahko gre kasneje dva koraka naprej!

f.) POVEZOVANJE ELEMENTOV KOMPLEKSNEGA DOŽIVLJANJA – SIBAM: vsako kompleksno doživljanje, dogodek je iz različnih elementov. Levine [24] je za razreševanje travme razvil SIBAM model (S = sensations, I = image, B = behaviour, A = affect, M = meaning), ki temelji na predpostavki, da celoten spomin nekega dogodka vključuje priklic petih elementov: slik, senzacij, vedenja, afektov in pomena. Ko posameznik doživi travmatičen dogodek spolne zlorabe, ti elementi med seboj zaradi hudega stresa niso povezani, ampak so disociirani, zato tudi kasneje lahko doživlja prave napade panik in groz, pa niti ne ve, kaj jih je povzročilo. Slik in vedenja nima, pomena tudi ne, ima pa telesne senzacije in afekt. Pri izogibanju, zamrznitvi in drugih odzivih so prisotne druge kombinacije izmed teh petih elementov. Na terapijah poleg dvojnega prepoznavanja terapevt in udeleženec iščeta tudi povezave vseh petih elementov doživljanja in tako določene odseke spomina sestavljata v celoto. Najprej proces poteka na kognitivnem nivoju, potem pa tudi na čustvenem in telesnem. Potrebnih je veliko ponovitev in vztrajanja, da se začnejo pojavljati novi vzorci vedenja, večji nadzor nad prej na videz neobvladljivimi čutenji in navsezadnje tudi pozitivne spremembe.

g.) VREDNOTENJE SPREMEMB IN NAPREDKA: ob vseh teh doživljanjih, prebujanju bolečih spominov, vzponih in padcih pa je nujno potrebno veliko pozornosti posvetili tudi priznanjem in pohvalam osebi, ki je bila spolno zlorabljena in se trudi s pomočjo terapij v vsakdanjem življenju nekaj spremeniti. Še zlasti, ko pridejo obdobja kriz in takrat običajno človek ne vidi ničesar pozitivnega, je bistveno in pomembno, da terapevt zna podpreti ter ovrednotiti prehojeno pot, pohvaliti tisto, kar je posameznik že naredil, trud in vztrajanje, vložene napore in moči ter verjeti še naprej in še bolj, da bo šlo vse še na boljše in samo naprej. Na skupinskih terapijah se udeleženci podpirajo med seboj in na podlagi svojih izkušenj povedo, kaj jim najbolj pomaga, jih najbolj napolni z energijo, usmeri naprej, daje moč, da vztrajajo. Vsi pa povedo, da jim zelo veliko pomeni, da terapevt zaupa in verjame v njihov napredek, da pohvali in ovrednoti spremembe ter da lahko pokaže in pove, da je ponosen na njihovo »garanje« – prehojeno pot. Prav te spodbude so med drugim zelo ključne pri procesu okrevanja in razreševanja travmatične izkušnje spolne zlorabe.

5. ZAKLJUČEK

Klinične izkušnje terapevtov kažejo, da ne moremo nikoli reči, da pridemo kdaj na konec poti razreševanja travme spolnega nasilja – kajti, ko je dosežen en cilj, se pokažejo potrebe po novi rasti, po novem cilju, in proces postane sestavni del življenja. Kljub vsem posledicam, ki zaznamujejo posameznikovo življenje, lahko v terapevtskem odnosu posameznik tako napreduje, da ga travma nezavedno ne usmerja več, ko pa pridejo sprožilci, ki sprožijo določne organske senzacije, bo prav kognitivna raven tista, ki bo pomagala premostiti obdobja kriz in razmejiti, da se dogodek sedaj ne dogaja več, ampak se le prebujajo določeni segmenti. To kognitivno zavedanje

zlorabljenim ne pomaga le preživeti, ampak tudi živeti in zaživeti človeka vredno življenje.

6. REFERENCE

[1] Rozman, S. 2015. *Pogum, kako preboleti travmo.* Modrijan Založba.

[2] Repič Slavič, T. 2015. *Nemi kriki spolne zlorabe in novo upanje*. [2. izd.]. Celje: Celjska Mohorjeva družba: Društvo Mohorjeva družba. 2015.

[3] Repič Slavič, T. 2016. Sexual abuse in childhood and the unconscious repetition of pain in couple relationships. V: Józef Stala, ur., Jadranka Garmaz, ur. *Strengthening families*, 137-151. *The series Family studies,* 5. Kraków: The Pontifical University of John Paul II.

[4] Simonich, H. K., Wonderlich, S. A., Erickson, A. L., Myers, T. C., Hoesel, J., Wagner S. in Engel, K. 2015. A statewide trauma-focused cognitive behavioral therapy network: creating an integrated community response system. *Journal of Contemporary Psychotherapy* 45, št. 4:265-274.

[5] Vaillancourt-Morel, M. P., Godbout, N., Sabourin, S., Briere, J., Lussier, Y. in Runtz, M. 2016. Adult sexual outcomes of child sexual abuse vary according to relationship status. *Journal of Marital and Family Therapy* 42, št. 2:341-56.

[6] van der Kolk, B. A. 2002. Beyond the talking cure: Somatic experience and subcortical imprints in the treatment of trauma. V: Francine Shapiro, ur. *EMDR as an integrative psychotherapy approach: Experts of diverse orientations explore the paradigm prism,* 57-83. Washington: American Psychological Association.

[7] Yule, W. 2001. Posttraumatic stress disorder in the general population and in children. *Journal of Clinical Psychiatry.* 62. št. 17:23-8.

[8] Solomon, M. F. in Siegel, D. J. 2003. *Healing trauma: Attachment, mind, body, and brain.* New York: W. W. Norton & Company.

[9] Maniglio, R. 2011. The role of child sexual abuse in the etiology of substance-related disorders. *Journal of Addictive Diseases* 30, št. 3:216-28.

[10] Boonmann, C., Grisso, T., Guy, L. S., Colins, O. F., Mulder, E. A., Vahl, P., Jansen, L. M. C., Doreleijers, T. A. H. in Vermeiren, R. R. J. M. 2016. Childhood traumatic experiences and mental health problems in sexually offending and non-sexually offending juveniles. *Child Adolesc Psychiatry Ment Health,* št. 10:45.

[11] Cvetek, R. 2009. *Bolečina preteklosti: travma, medosebni odnosi, družina, terapija.* Celje: Celjska Mohorjeva družba: Društvo Mohorjeva družba.

[12] Gostečnik, C. 2008. *Relacijska paradigm in travma.* Ljubljana: Brat Frančišek in Frančiškanski družinski center.

[13] Henry, J., Sloane, M. in Black-Pond, C. 2007. Neurobiology and neurodevelopmental impact of childhood traumatic stress and prenatal alcohol exposure.

[14] Gostečnik, C. 2017. *Družinske terapije in klinična praksa.* Ljubljana: Brat Frančišek in Frančiškanski družinski inštitut.

[15] Knauer, S. 2002. *Recovering from sexual abuse, addictions, and compulsive behaviors: "Numb" survivors.* New York, NY: Haworth Press.

[16] Rothschild, B. 2000. *The body remembers: The psychophysiology of trauma and trauma treatment.* New York: W. W. Norton & Company.

[17] Gostečnik, C. 2017. *Relational family therapy: the systemic, interpersonal, and intrapsychic experience.* New York; London, Abingdon: Routledge.

[18] Dorais, M. 2002. *Don't tell: The sexual abuse of boys.* Quebec: McGill-Queen's University Press.

[19] Schore, A. N. 2003. Early relational trauma, disorganized attachment, and the development of a predisposition to violence. V: D. J. Siegel in M. F. Solomon, ur. *Healing trauma: Attachment, mind, body, and brain*, 107-67. New York: W. W. Norton & Company.

[20] Rothschild, B. 2006. *Help for the helper: The psychophysiology of compassion fatigue and vicarious trauma.* New York: W. W. Norton & Company.

[21] Repič Slavič, T. in Gostečnik, C. 2017. Relational family therapy as an aid toward resolving the trauma of sexual abuse in childhood in the process of separation in the couple relationship. *Journal of marital and family therapy 43*, št. 3:422-434.

[22] Scaer, R. 2005. *The trauma spectrum: Hidden wounds and human resiliency.* New York, NY: W. W. Norton & Company.

[23] Simonič, B. 2018. Nežnost in njen pomen v medosebnih odnosih in v pastorali. *Bogoslovni vestnik.78,* št. 1:209-218.

[24] Levine, P. A. 1997. *Waking the tiger: Healing trauma: The innate capacity to transform overwhelming experiences.* Barkeley, CA: North Atlantic Books.

Medgeneracijski prenos nasilja

Prof. dr. Christian Gostečnik
Teološka fakulteta, katedra za
Zakonsko in družinsko terapijo ter
psihologijo in sociologijo religije
Poljanska 4, Ljubljana

Frančiškanski družinski inštitut
Prešernov trg 4, Ljubljana

christian.gostecnik@teof.uni-lj.si

POVZETEK

Vsak emocionalni odziv na stresne situacije vedno vpliva na celotno psiho-organsko strukturo posameznika. Vsekakor pa le-to najbolj zaznamujejo nasilna, travmatična doživetja, ki se lahko zaradi svojih destruktivnih vplivov, prenašajo iz generacije v generacijo. Nasilna, travmatična doživetja s senzacijami in afekti se namreč usidrajo v posameznikovo psihosomatsko, psiho-organsko strukturo oziroma v njegov implicitni spomin in tam nezavedno bivajo ter lahko dramatično vplivajo na naslednje rodove. V tem prispevku bomo zato skušali prvenstveno pokazati, da težja ko je travma, nasilje, močneje posameznika sili k zunanjemu izrazu ali ekspresiji in to še prav posebej tedaj, ko gre za PTSD, katerih posledice pa lahko zato doživljajo tudi bodoče generacije.

Ključne besede

travmatična doživetja, implicitni spomin, prenos nasilja, PTSD, relacijska družinska terapija

1. UVOD

Mnogi v sodobnosti, kot npr. [1-5], v svojih raziskavah odkrivajo transgeneracijski prenos travme v naslednje generacije in lahko celo na otroke, ki se sploh še niso rodili, ko se je nasilje zgodilo [6]. Nekateri avtorji [7, 8] pri tem ocenjujejo, da npr. približno ena tretjina spolno zlorabljenih oseb naprej zlorablja, približno dve tretjini žrtev pa ne bo nikoli spolno zlorabljajo naprej, predvsem če bodo imele podporo in ljubezen v družini. Johnson [9] ugotavlja, da imajo tisti, ki so bili v preteklosti žrtve čustvene, fizične ali spolne zlorabe, 6-krat večjo možnost, da bodo tudi sami nadaljevali zlorabo, ki so jo izkusili na lastni koži. Druge raziskave Collin-Vezina [10] ugotavlja, da je bila polovica mater, katerih otroci so bili spolno zlorabljeni, tudi sama žrtev spolne zlorabe. In tudi če se dejanje nasilja ne prenaša naprej, to še ne pomeni, da bodo otroci staršev, ki so doživeli nasilje, varni pred osebami, ki izvajajo nasilje.

2. RELACIJSKA DRUŽINSKA TERAPIJA

Relacijska družinska teorija [11-14] govori o nepredelanih afektih, psiho-organskih vsebinah, ki so pri nasilju predvsem strah, prezir, gnus, sram in jeza, ki se globoko vtisnejo v psiho-organsko strukturo posameznika, le-te se prek mehanizma projekcijsko-introjekcijske identifikacije vertikalno prenesejo s travmiranega starša na otroka [2, 3, 15-17]. Četudi skuša ta starš otroka opozoriti na vse nevarnosti nasilja [18], sam pri sebi pa ne bo v stiku z nepredelanimi afekti ter se zato tudi ne bo znal zavarovati in postaviti razmejitev, bo veliko večja verjetnost, da otrok postane žrtev nasilja [11, 13, 14, 19]. Podobno pravi Millerjeva v knjigi Upor telesa [20], ko piše, da se zloraba v otroštvu razrešuje na dva načina: odrasli, ki je bil kot otrok spolno zlorabljen, svoja nepriznana čustva, psiho-organske vsebine prenese na svojega otroka oziroma na druge ljudi okoli sebe, lahko pa posledice plačuje telo zlorabljenega, in sicer s psihosomatiko ali kroničnimi boleznimi [21]. Tako sekundarna travmatizacija kot tudi transgeneracijski prenos pripomoreta k pogostemu razvoju nizke samopodobe otrok [4]. Prav to je raziskovalce še bolj motiviralo, da niso raziskovali le mehanizmov in učinkov prenosa travme, temveč tudi nezavedno in zavedno komunikacijo med starši in otroki. Ugotovili so, da se je komunikacija lahko razlikovala od skoraj popolne tišine do odprtega razpravljanja in pripovedovanja najbolj krutih zgodb in izkušenj. Otroci so v obeh primerih zatirali in potlačili svoja čustva in niso bili gotovi, kaj se je v resnici dogajalo z njihovimi starši. [2, 3, 15, 22, 23] Poleg so tudi ugotovili, da je prekomerna komunikacija in z njo povezano razlaganje vseh izkušenj pri otrocih povzročala manj depresije in anksioznosti, vendar več krivde. Krell [24] trdi, da je bilo za otroke še bolj strašljivo, če starši o svojih izkušnjah niso govorili, saj so si otroci v svojih glavah s pomočjo fantazije ustvarili določene slike, ki so bile zelo patogene. Tako so otroci nezavedno uprizarjali in ponovno doživljali usodo staršev, in če jim je ob tem primanjkovalo informacij in zgodb, je bilo to zanje lahko še hujše in intenzivnejše [5]. Vsi ti prenosi čutenj in afektov, psiho-organskih vsebin pa se odvijajo prek mehanizmov projekcijsko-introjekcijske identifikacije in kompulzivnega ponavljanja.

3. TRAVMA IN NJENA SIMPTOMATOLOGIJA

Tipični znaki travme se odražajo (2, 3, 5, 15-17, 23, 25, 26] predvsem hipervzburjenju v AŽS, le-ta pa je rezultat travmatičnega doživljanja. Tu govorimo predvsem o hitrejšem bitju srca, srčni palpitaciji, hladno-vročih potenjih, hitrejšem, a bolj plitvem dihanju, hiperbudnosti ali pazljivosti, mravljinčavosti, strahu, jezi, avtodestrukciji itd. V kroničnih stanjih pa lahko ti simptomi vodijo v motnje spanja, depresijo, izgubi apetita, anksioznosti, spolnim disfunkcijam, zamrznjenosti, disociaciji in velikokrat lahko ima ta posameznik težave s koncentracijo in posledično s spominom [2, 3, 15, 23, 27-29]. Tako te žrtve lahko ponovno doživljajo travmatični dogodek v senzornih oblikah ali »flashbackih« in se zato začno izogibati vsemu, kar bi jih lahko spominjalo na travmo, ali pa doživljajo kronično hipervzburjenje v avtonomnem živčnem sistemu [16, 17, 25, 30-33].

Z drugimi besedami, na ta način lahko proces disociacije odcepi tako narativne komponente doživetja kakor tudi sosledje dogajanja, vključno s fiziološkimi in psihološkimi skratka psiho-

organskimi reakcijami, kar pa pomeni, da ostaja implicitni oziroma organski spomin globoko vrisan v posameznikov organizem in le-ta se lahko prenese v naslednjo generacijo [4]. Govorimo torej o amneziji, ki je v različnih oblikah, poleg ostalih simptomov, lahko vsekakor najpogostejši pojav pri disociativni dinamiki, zato se to nezavedno, psiho-organsko prenaša na naslednje generacije. Posameznik lahko doživi anestezijo in ne čuti nobene bolečine, lahko vsa občutja in afekte odcepi, zopet drugi lahko izgubi zavest ali pa čuti, kakor da je izven telesa. Lahko pa vse to pozneje v najrazličnejših oblikah zaznavajo v naslednjih rodovih [5, 25, 34].

4. PONOVNI VZNIK TRAVME V NASLEDNJIH GENERACIJAH

Podoživljanje travmatičnih dogodkov pa vedno znova omogočajo »flashbacki«, in sicer lahko v celoti ali pa samo v določenih delih; le-ti se lahko pojavljajo v naslednjih generacijah in to brez očitnih travm, ki bi jih ti posamezniki v naslednjih generacijah doživeli [2, 3, 15, 23]. Ti fenomeni so lahko izredno moteči, saj se posameznik, ki je doživel travmatični dogodek, ob teh »flashbackih« počuti tako, kot da se celoten travmatični dogodek ponovno in v vsej intenziteti ponavlja, čeprav se je travma že zdavnaj končala. Še veliko bolj pa so vznemirjajoči za posameznika v naslednji generaciji, ki nima nobenega spomina, da bi se mu karkoli tega zgodilo, občuti pa simptome travme. Drugače povedano, posameznik, ki ni doživel travme, loči, ki ne nujno pozna PTS in PTSD-travmatičnih spominov, saj so le-ti odcepljeni že v samem originalu, se pa zato enostavno lahko pojavljajo v vsakem času in se v obliki »flashbackov« vedno lahko vrinejo v sedanjost [16, 25, 35]. Te »flashbacke« ponavadi spremlja tudi zelo močna anksioznost, ki je v originalu spremljala tudi travmatično doživetje posameznika. Še posebej močno moteči pa so ob tem še panični napadi, ki lahko žrtev tudi po dolgih letih, ko travme ni več, še vedno spremljajo in velikokrat čisto ohromijo tudi naslednjo generacijo.

Pri tem gre za poskus posameznika (njegovega organizma in tudi njegove psihe), da bi pobegnil, ko boj ali upor nista več mogoča [16, 25, 36]. Vsekakor pa gre lahko tu tudi za zamrznitev, ki je še tretja oblika odziva na travmo. Tudi v tem primeru gre za disociacijo, in sicer največkrat v najglobji obliki, saj se žrtev v teh primerih zelo malo ali pa sploh ne spominja, kaj se je v resnici zgodilo. V tem primeru lahko gre za zelo globoko stanje amnezije ali, kot rečemo, čistega odklopa, vendar pa v posamezniku kljub temu ostajajo elementi travme, in sicer v najrazličnejših afektih, kot so strah, sram, jeza, ki se vedno lahko prebudijo ob podobnih situacijah, ki npr. samo bledo spominjajo na originalno travmo. V posamezniku ostajajo t. i. žarišča, ki še vedno tlijo, in to »tleče stanje« žrtev travme včasih še veliko bolj vznemirja, še posebej tedaj, ko se pojavlja v obliki paničnih napadov ter izjemno močnih anksioznosti ob določenih spodbujevalcih ali sprožilcih, ki so spremljali originalno travmo [37, 38] in se lahko, kot rečeno, ponavlja v naslednjih generacijah.

Če se povrnemo k fenomenu disociacije, potem vidimo, da žrtve med travmatičnimi doživetji izkušajo najrazličnejša občutja: da so zapustile telo, da se je čas ustavil, da so zamrznile, omrtvele in ničesar več doživljale. Žrtev je lahko tudi zelo omejeno videla puško, mučilno orodje, samo silhueto rablja ali napadalca, luč ali avto, ki sta se silovito približevala, in se potem ničesar več ne spomni do prebujenja v bolnišnici, ko je bilo okrog nje polno ljudi v belih haljah itd. [2, 3, 15-17, 23, 25]. Ostaja pa organski, implicitni spomin, ki se lahko vedno znova prebudi v najrazličnejših oblikah vedenj, čutenj, fobij in paničnih napadov. Žrtev je ponavadi tudi po travmatičnem doživetju še vedno

disociirana in še dolgo potem, ko se je travma zgodila, še vedno čuti, da je »izven sebe«. To stanje pa se, kot rečeno, lahko nadaljuje še dolge mesece in leta, včasih pa se zgodi, da se šele čez leta prebudi, in sicer v obliki otopelosti, zamrznjenosti, anksioznosti, »flashbackov«, depersonalizacije, delne ali popolne amnezije, doživljanja, kot da je žrtev izven telesa, nezmožnosti kar koli čutiti, pa tudi v obliki nerazumljivih vedenjskih in emocionalnih reakcij, ki nimajo nobenega pravega razloga v konkretnih situacijah. Vse bolj pa raziskave ugotavljajo, da je disociacija tudi vedno zvesti spremljevalec PTS- in PTSD-simptomov in s tem v zvezi paničnih napadov ter napadov anksioznosti [2, 3, 15-17, 23, 25, 39], kar ponuja izjemno možnost, da se travme preteklosti, zaradi svoje intenzitete, ki formira PTSD simptomatologijo, ponovno prebudijo v naslednjih generacijah, in sicer v najrazličnejših oblikah in intenzitetah.

5. ZAKLJUČEK

Za sklep lahko rečemo, da vsako zavestno doživljanje posameznika sestavljeno iz več komponent. Celoten spomin nekega določenega travmatičnega dogodka zato vedno vključuje integrirani priklic vseh elementov posameznikovega doživljanja, ki ga sestavljajo: senzacije, slike, vedenje, afekti in pomen. Kadar situacija ni preveč stresna, vsekakor pa, kadar je prijetna, se ti elementi medsebojno povežejo oziroma ostajajo nedotaknjeni v spominu. Tako se npr. posameznik lahko v vsem spominja lepih in vzhičenih trenutkov svojega življenja, spominja pa se tudi čisto navadnih dogodkov, npr. izleta, kosila, obiska itd. Pri tem se posameznik lahko spomni vseh pomembnejših dejanj oziroma vedenj, prikliče si sliko, še pozna vse bistvene senzacije, znani so mu afekti, ki so jih ta doživetja ustvarila oziroma prebudila, ve pa tudi za pomen – bil je npr. sproščen dan, poln lepega, svežega, sklenili so nova znanstva.

Kadar pa je stres prevelik oziroma ko govorimo o nasilju in travmah, se ti elementi med sabo cepijo oziroma disociirajo. Seveda pa pri tem ni nujno, da vsaka, tudi zelo stresna situacija povzroči disociacijo. Vendar, vedno kadar je stres ali travma v svoji intenziteti premočna, se aktivira dissociativni sistem, ki zavaruje posameznika pred dekompenzacijo. Tedaj lahko govorimo o PTSD simptomatologiji, ki pa lahko vedno pomeni žarišče travmatičnih organskih spominov, ki se neverjetno silovitostjo prenašajo v naslednje generacije.

6. Reference

[1] Epstein, H. 1979. *Children of the Holocaust: Conversations with sons and daughters of survivors*. New York, NY: Penguin.

[2] Barocas, H. A. in Barocas, C. B. 1973. *Manifestations of Concentration Camp Effects on the Second Generation. American Journal of Psychiatry, 130(7), 820–821.*

[3] Lev–Wiesel, R. 2007. Intergenerational Transmission of Trauma across Three Generations: A Preliminary Study. Qualitative Social Work: *Research and Practice, 6(1)*, 75–94.

[4] Woollett, N. in Thomson, K. 2016. Understanding the intergenerational transmission of violence. *South African Medical Journal 106(11):*1068-1070.

[5] Dekel, R. in Goldblatt, H. 2008. Is There Intergenerational Transmission of Trauma? The Case of Combat Veterans' Children. *American Journal of Orthopsychiatry.78(3).* 281-289.

[6] Schore, A. N. 2016. *Affect regulation and the origin of the self: The neurobiology of emotional development*. New York, NY: Taylor & Frances/Routledge.

[3] Schore, A. N. 2012. *The science of the art of psychotherapy*. New York, NY: W. W. Norton & Company.

[4] Steele, K. Boon, S. in van der Hart, O. 2017. *Treating trauma-related dissociation: A practical, integrative approach*. New York, NY:: W. W. Norton & Company.

[5] Burton, M. S., Feeny, N. C., Connell, A. M. in Zoellner, L. A. 2018. Exploring evidence of a dissociative subtype in PTSD: Baseline symptom structure, etiology, and treatment efficacy for those who dissociate. *Journal of consulting and clinical psychology* 86, št. 5:439.

[6] Catherall, R. D. 2004. *Handbook of stress, trauma, and the family*. New York, NY: Brunner-Routledge.

[7] Kaufman, G. 1989. *The psychology of shame: Theory and treatment of shame-based syndromes*. New York, NY: Springer Publishing Co.

[8] Hanson, R. K. in Slater, S. 1988. Sexual victimization in the history of sexual abusers: A review. *Annals of Sex Research* 1, št. 4:485-499.

[9] Johnson, S. M. 2002. *Emotionally focused therapy with trauma survivors: Strenghtening attachment bonds*. New York, NY: Guilford Press.

[10] Collin-Vézina, D. in Cyr, M. 2003. La transmission de la violence sexuelle: description du phénomène et pistes de compréhension. *Child Abuse & Neglect* 27, št. 5:489-507.

[11] Gostečnik, C. 2017. *Relational family therapy: The systemic, interpersonal, and intrapsychic experience*. New York, NY: Routledge/Taylor & Francis Group.

[12] Gostečnik, C. 2011. *Inovativna relacijska družinska terapija: inovativni psiho-biološki model*. Ljubljana: Brat Frančišek in Frančiškanski družinski inštitut.

[13] Gostečnik, C. 2005. *Psihoanaliza in religiozno izkustvo*. Ljubljana: Brat Frančišek in Frančiškanski družinski inštitut.

[14] Gostečnik. 2004. *Relacijska družinska terapija*. Ljubljana: Brat Frančišek in Frančiškanski družinski inštitut.

[15] Schore. A. N. 2003. *Affect dysregulation and disorders of the self*. New York, NY: W. W. Norton & Company.

[16] Rothschild, R. 2017. *The body remembers volume 2: Revolutionizing trauma treatment*. New York, NY: W. W. Norton & Company.

[17] Rothschild, B. 2010. *8 keys to safe trauma recovery*. New York, NY: W. W. Norton & Company.

[18] Howell, E. F. 2002. "Good Girls," Sexy "Bad Girls," and Warriors: The Role of Trauma and Dissociation in the Creation and Reproduction of Gender. *Journal of Trauma & Dissociation* 3, št. 4:5-32.

[19] Gostečnik, C. 2008. *Relacijska paradigma in travma*. Ljubljana: Brat Frančišek in Frančiškanski družinski inštitut.

[20] Miller, A. 2005. *Upor telesa: telo terja resnico*. Ljubljana: Tangram.

[21] Ogden, P. in Fisher J. 2015 *Sensorimotor psychotherapy: Interventions for trauma and attachment*. New York, NY: W. W. Norton & Company.

[22] Okner, D. F. in Flaherty, J. F. 1979. Parental communication and psychological distress in children of Holocaust survivors: A comparison between the U.S. and Israel. *International Journal of Social Psychiatry* 35, št. 3:265–273.

[23] Schore, A. 2003. *Affect regulation and the repair of the self*. New York, NY: W. W. Norton & Company.

[24] Krell, R. 1979. Holocaust families: The survivors and their children, *Comprehensive Psychiatry* 20, št. 6:560-568.

[25] Rothschild, B. 2006. *Help for the helper: The psychophysiology of compassion fatigue and vicarious trauma*. New York, NY: W. W. Norton & Company.

[26] Rothschild, B. 2003. *The body remembers casebook: Unifying methods and models in the treatment of trauma and PTSD*. New York, NY: W. W. Norton & Company.

[27] Cohen, J. A., Mannarino, A. P. in Deblinger, E. 2006. *Treating trauma and traumatic grief in children and adolescents*. New York, NY: Guilford Press.

[28] Kitayama, N., Quinn, S. in Bremner, D. J. 2006. Smaller volume of anterior cingulate cortex in abuse-related posttraumatic stress disorder, *Journal of Affective Disorders* 90, št. 2-3:171–174.

[29] Haagen, J. F., van Rijn, A., Knipscheer, J. W., van der Aa, N. in Kleber, R. J. 2018. The dissociative post-traumatic stress disorder (PTSD) subtype: A treatment outcome cohort study in veterans with PTSD. *British Journal of Clinical Psychology* 57, št2:203-222.

[30] Siegel, D. J. 2007. *Mindsight: The new science of personal transformation*. New York, NY: Bantam Books. 2011.

[31] D. J. Siegel. *The mindful brain: Reflection and attunement in the cultivation of well-being*. New York, NY: W. W. Norton & Company.

[32] Siegel, D. J. 2017. *Mind: A Journey to the Heart of Being Human*. W.W. Norton.

[33] Siegel, D. J. 2015. *The developing mind: How relationships and the brain interact to shape who we are*. Guilford Publications.

[34] Austern, D. J. 2017. Beyond Binary Thinking: Providing Best Practice Treatment to Veterans with PTSD, *Pragmatic Case Studies in Psychotherapy* 13, št. 2:154-172.

[35] Hyland, P., Shevlin, M., Hansen, M., Vallières, F., Murphy, J. in Elklit, A. 2016. The temporal relations of PTSD symptoms among treatment-seeking victims of sexual assault: a longitudinal study. *Journal of loss and trauma* 21, št. 6:492-506.

[36] Åkerblom, S., Perrin, S., Fischer, M. R. in McCracken, L. M. 2017. The impact of PTSD on functioning in patients seeking treatment for chronic pain and validation of the posttraumatic diagnostic scale. *International journal of behavioral medicine* 24, št. 2:249-259.

[37] Shipman, K. L.,Schneider, R., Fitzgerald, M. M., Sims, C., Swisher, L. in Edwards, A. 2007. Maternal emotion socialization in maltreating and non-maltreating families: Implications for children's emotion regulation. *Social Development* 16, št. 2:268–285.

[38] Lokshina, Y. in Liberzon, I. 2017. Enhancing efficacy of PTSD treatment: role of circuits, genetics, and optimal timing. *Clinical Psychology: Science and Practice* 24, št. 3:298-301.

[39] Johnson, E. M. Possemato, K., Barrie, K. A., Webster, B., Wade, M. in Stecker, T. 2017. Veterans' Thoughts About PTSD Treatment and Subsequent Treatment Utilization. *International Journal of Cognitive Therapy* 10, št. 2:138-160.

The Other Side of Neurotechnologies

Kristian Elersič

ekristian238@gmail.com

ABSTRACT

As a response to the overly-optimistic view of technology, I will present some of the obstacles when using new neurotechnologies, specifically cognitive enhancers. Those obstacles come in a form of physical limitations, such as not being able to develop methods sophisticated enough that they would have no (or minimal) negative side-effects, and in a form of ethical limitations, since even if these new technologies were physically possible, the extent to which we should use these technologies is not fully explored. To fully consider such obstacles, I propose an outline of the ethical framework that should be used when thinking of applying new, future or even old cognitive technologies. Furthermore, I pose some questions for further analysis of the intertwining of techno-scientific progress with social dynamics.

Keywords

Neurotechnology, transhumanism, enhancement, cognition, ethics

1. INTRODUCTION

Neurotechnology, technology that has a fundamental impact on how we understand the brain, cognitive processes and consciousness, is rapidly developing. Some of the examples of neurotechnologies are: imaging technology such as magnetic resonance imaging (MRI), electroencephalogram (EEG) or positron emission tomography (PET), brain stimulations such as transcranial magnetic stimulations (TMS) or transcranial direct current stimulation (tDCS), implant technologies, pharmaceuticals, gene editing, cell therapy and brain-computer interfaces. In this article I will be focusing on technologies that aim to improve the cognitive processes - cognitive enhancers. There is not a specific technology that suits this purpose only. It is rather a combination of different technologies that are believed to alter cognitive processes: gene editing, pharmaceuticals and brain stimulations.

My whole stance is formed as a response to the overly-optimistic view of technology – which is that technology can save many of our problems, if not all of them. At no point in this paper will I go into detailed discussion of any topic, but will rather open a broad field of questions that are important to consider when analysing neurotechnology and human enhancement. In the first section of this paper, I will describe what I mean with the term "overly-optimistic view of technology". In the second section, I will describe shorty what are the cognitive enhancers. Then I will continue by presenting some of the physical limitations of the enhancers, namely the limitations for gene editing, pharmaceuticals and brain stimulations. Then I will set the ethical framework that will serve the purpose of considering or questioning any intervention. Finally, I will try to think of the possible connections of the techno-scientific progress with society.

2. OVERLY-OPTIMISTIC VIEW OF TECHNOLOGY AND ENHANCEMENT

As a student of science, I am mainly surrounded by scientific literature and scientists. What I have found is that there is a general lack of interest in interdisciplinary discussion of impacts that techno-scientific progress has on a society. It almost seems that there prevails an opinion of that technology and science cannot be misused – over-optimism or maybe even ignorance. To clarify further, this opinion comes from my student surroundings – University of Ljubljana, Biotechnical faculty. I haven't done any analysis of other universities. Of course, this interdisciplinary topic is not strictly a scientific one, but nonetheless it is important. Therefore, I think that discussions of this kind should be included in the scientific community, especially when it comes to biotechnology that directly addresses and modifies the human condition. On the other hand, beside lack of interest of my academic surroundings, there is also an emerging movement that is overly-optimistic about technology. It is called transhumanism. Max More, philosopher and futurist, defined it:

> " … [as] both a reason-based philosophy and a cultural movement that affirms the possibility and desirability of fundamentally improving the human condition by means of science and technology. Transhumanists seek the continuation and acceleration of the evolution of intelligent life beyond its currently human form and human limitations by means of science and technology, guided by life-promoting principles and values." [1]

A few of the common characteristics of the mentioned philosophy are: a) optimistic view of progress of science and technology, b) aiming at the advancement of humans - to alleviate our biological constrains so that we could live longer and healthier, c) using science and technology as the main means of advancement and d) claiming that enhancement will transform the whole meaning of being human. In other words, they state that technology will help us as a humanity to overcome our current problems, it will help us to be more ethical, healthy, well in both economic and ecologic sense. Transhumanists also believe that by doing so we will transform ourselves into some future form where we will no longer be able to state we are the same as we were before as species – we will become post-human in a transhumanist sense.

For this purpose I want to analyse and present some questions related to the human bio-enhancement, specifically cognitive enhancement. My aim is not to discredit latest scientific development. New knowledge and new technologies are appreciated – but with them comes a responsibility of how and when to use them. My aim is to stimulate an interdisciplinary discourse of the mentioned subject – human enhancement and neurotechnologies. In other words, I want that we take a closer look at the other, darker side of new neurotechnologies that aim at enhancing humans, because the bright side is known to well.

3. COGNITIVE ENHANCERS

Cognitive enhancers are technologies that enhance cognitive processes. Like doping in sports that enhances physical

performance, cognitive enhancers stimulate memory, perception, concentration, motivation, learning, problem solving and computation. [2 3] Most widely known are the so-called "smart drugs" – cognitive enhancers in the form of pharmaceuticals (substances). Those are already used for medicinal and enhancing purposes. [3 4] Besides that, scientists are developing methods for cognitive enhancement in the form of genetic engineering and brain stimulations. In a broader sense, even traditional cognitive training can be classified as one form of cognitive enhancers, but I will rather focus on technological enhancement.[1]

4. PHYSICAL LIMITATIONS

When describing neurotechnological interventions for enhancement purposes, we have mentioned three types of modifications: a) genetic engineering, b) drug intake and c) brain stimulation. In this section I will present some of the (possible) limitations of mentioned technologies.

4.1 Genetic Engineering

Genetic engineering started in 1973 and has very quickly spread onto many fields of industry and research. In our case, genetic engineering would be used for manipulation of certain cognitive trait like attention, IQ, memory and so on. For example, we could genetically modify embryo's DNA, or later in life insert some gene products to tissues or organs (RNA or protein insertions), or "knock out", that is, make inoperative, some genes (gene knockout with DNA cutting proteins), or we could select one out of many embryos with the most preferable genome (prenatal diagnostic or embryo selection). But when it comes to genetics, we often misinterpret that there is a direct link between a gene and a trait. We have to bear in mind that only a few traits are monogenetic, meaning that there is only one gene that influences a certain trait. In most cases, there are multiple genes that affect one trait. Besides that, in many cases one gene also affects many traits. [6] Due to this phenomenon (complex genetic traits) it is very hard to predict the total outcome of such genetic manipulation. This is nicely illustrated by the study when researchers improved mouse's memory and simultaneously increases its sensitivity to pain. [7 8 9]

4.2 Pharmaceuticals

Enhanced cognition can be also achieved with numerous substances, even with dietary supplements, such as caffeine and nicotine, but for our purpose I will describe pharmaceuticals (which are classified as technological interventions). One of such pharmaceuticals is methylphenidate, commercially known as Ritalin. Ritalin is nowadays widely used not just for treating the attention deficit hyperactivity disorder (ADHD), but is popular as a cognitive enhancer among students/academics with the purpose of achieving better study or academical results. [3 4] Other similar substances are: amphetamines, modafinil, atomoxetine, reboxetine, donepezil, galatamine, rivastigmine and memantine. When prescribing these drugs it is important to take into account obstacles that may come with the use of these substances. One such obstacle

is that substances often have side effects such as nausea, vomiting or cramps. The second obstacle is that brain responds to the chemicals in such a way that a too high or too low concentration of substance might not produce optimal brain functioning effects.[2] In some individuals one dose may cause a positive effect while in the others the effect would be negative. [3] The third obstacle is that a drug may positively affect one brain function while simultaneously negatively affect the other.[3] Last but not least, it is important to add that we lack long term studies that would tell us how safe is the use of these drugs – in other words, what are the long-term side effects of such enhancers. [3 10 11]

4.3. Brain Stimulation

Brain stimulation is a technology using electrical currents or magnetic field (which induces electric currents) to stimulate certain parts of brain tissue. Methods of brain simulation can be further divided into invasive and non-invasive ones. The invasive method requires application of currents directly to the brain's surface or brain's core. In order to achieve this, we have to be invasive - the skull has to be opened and sometimes even the brain must be cut. Using the non-invasive methods, the currents or magnetic fields are applied over the head's surface. This is called non-invasive brain stimulation (NBS), examples of which are transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tDCS). For our further discussion it is not of a great importance what kind of technique is used.

Opposition to the overly optimistic view towards this form of cognitive enhancement may come from understanding brain's functioning. Key concept in this opposition is net zero-sum framework proposed by Brem et al. (2014). [12] Authors state that this concept is grounded on the physical principle of conservation of energy in a closed system. Furthermore, they clarify that, in this notion, brain operates within the constraints of a finite amount of energy and processing power. Not meaning that all the brains and the same brain through the life span do have equal amount of energy and processing power, but that an x-amount of energy and processing power is distributed throughout the brain at any given time. When demands shift, so does the distribution of energy and processing power. Most importantly, total sum of brain's activity stays the same. In other words, brain operates in net zero-sum framework. If this prediction is correct, NBS would enhance some brain function, but at the same time degrade other (compensation/cost). Some evidence accounting for this hypothesis comes from the phenomena of paradoxical functional facilitation. [12 13] This phenomenon appears when damage to an intact area of the brain normalizes a previously reduced level of functioning (compensation for previous loss – re-establishing balance between functional brain networks) or when a patient with damage performs a certain task better than before or better than healthy control subjects (energy and processing power of damaged brain tissue was distributed to other brain parts that are now enhanced). Authors conclude with raising bioethical concerns: "Is it acceptable to

---

[1]A special kind of cognitive enhancers are moral enhancers that are more strictly concerned with improving our moral behaviour. [5]

[2] This applies for future drugs too. This molecular dynamics is called U-shaped relationship between substance and receptor. [3]

[3] Example: when rigastimin improves learning on a motor task and making associations between symbols and digits, but can at the same time impair verbal and visual episodic memory. [3 10]

improve certain brain functions at the cost of others and can we take the responsibility for its impact on the individual and on society?" (p.14-15) They also state that "current neuroenhancment studies emphasize positive outcome of specific functions and concentrate on individual improvements, while related topics such as risk and safety, as well as social and moral factors are neglected or restricted to specific inquiries." (p.15) [12]

5. ETHICAL FRAMEWORK AND INTERTWINING OF TECHNO-SCIENTIFIC PROGRESS WITH SOCIETY

In the previous section I have described some physical limitations that we may encounter when applying neuroenhancing technologies. Now I want to state that even if we do develop better technology, we still have to consider questions from a very broad spectrum of topics.

Firstly, I want to propose an ethical framework consisting of three points that we have to consider in a discussions about any human interventions such as bio-enhancing technologies. (1) First is about personal freedom. (2) Second is about the line between long-term sustainability and short-term solutions. And (3) third is about ecological integrity/sustainability. For example: 1) do enhancers that aim to limit our immoral behaviours, limit our personal freedom? Or, do cognitive enhancers prescribed to children for enhancement of a specific trait, limit their freedom to choose/develop other characteristic that they would love to have? Or 2) are costs and side-effects of bio-enhancing accounted for? Are those short term solutions sustainable or do they just prolong the problem or maybe even worsen the situation? Where is the boundary between enhancement and a disease – former without the need (but with the want), latter with the immediate need? And 3) is this technology ecologically sustainable? Does it produce a non-sustainable mentality – that which is not able to think in a non-selfish way or in an ecological way or in terms of long-term solutions? What will this technology do to overpopulation, species extinction, how many waste will it produce? Those are just a few questions stated to further clarify the mentioned ethical framework.

Secondly, I want to briefly mention how broad the field of investigation in this and similar topics should be. Nature of questioning and regulating techno-scientific progress, especially when we think of interfering with our nature (enhancement), exceeds fields of natural sciences. It includes (or should include) questions from philosophy of science (what is knowledge, what is science, its place and role in modern and future society), philosophy of progress (genealogy of progress, what is progress), intertwining of the techno-scientific progress and the economic system (how is modern free market system influencing techno-scientific progress, business of science), intertwining of the market and the values/ideological system (how does the market with advertising influence the values and how do these values afterwards influence the market) and the application of technology in political systems (possibility of new technologies for achieving political goals, possibility of manipulation and control). Furthermore, we have to ask ourselves why do we want enhancement: is it a by-product of economic system (competition – to be better and to get a better job), or is it a by-product of modern values (for example, value of ideal

life – long and healthy with no negativity), or a product of advertisement or political viewpoint. We ought to analyse all of the mentioned subfields before interfering with human nature with the use of enhancers to the extent that would limit our capacity undo the damage.

6. CONCLUSION

Development of new technologies opens up many questions – from physical limitations to ethical considerations. I have focused on neurotechnologies, specifically cognitive enhancers. My general claim is that there is an overly-optimistic view of technology that states that technology can solve many, if not all, of our problems. Hence, we can enhance human beings to be better. Doubt arises, not only because of physical limitations of modifying technologies, such as in genetic engineering, drug intake and brain stimulation, but even more so because of ethical issues. Human condition is a very complex one and it has been, even if this does not seem so, analysed too little for us to drastically modify our nature. In this manner, we need an interdisciplinary and public discourse for discussions about mentioned topics – intertwining of modern values, human condition, society and techno-scientific progress.

7. REFERENCES

[1] Max More: *True Transhumanism: A reply to Don Ihde*, Transhumanism and Its Critics (2010) p.138

[2] N.Bostrom and A. Sandberg, *Cognitive Enhancement - Methods, Ethics, Regulatory Challenges*, Sci Eng Ethics (2009) 15:311-341

[3] Husain and Metha (2011) *Cognitive enhancement by drugs in health and disease*. Trends in Cognitive Sciences, January 2011, vol. 15, No.1

[4] Greely, H et al. *Towards responsible use of cognitive-enhancing drugs by the healthy*. Nature, 456(7223), 702–705 (2008).

[5] Juslian Savulescu adn Igmar Persson. *Moral Enhancement, Freedom and the God Machine*. Monist. 2012 July; 95(3): 399-421

[6] Griffiths AJ, Wessler SR, Carroll SB, Doebley JF. *Introduction to genetic analysis (Eleventh ed.): The Inheritance of Complex Traits (p.715)*. New York, NY. (2015)

[7] Simmons, D. *Genetic inequality: Human genetic engineering. Nature Education* 1(1):173, 2008.

[8] Wei, F., et al. *Genetic enhancement of inflammatory pain by forebrain NR2B overexpression*. Nature Neuroscience 4, 164–169 (2001)

[9] Tang, Y. P., et al. *Genetic enhancement of learning and memory in mice*. Nature 401, 63–69 (1999)

[10] Wezenberg E. *Modulation of memory and visuospatial processes by biperiden and rivastigmine in elderly healthy subjects*. Psychopharmacology (Berl.) 2005;181:582–594.

[11] Chris Zarpentine: *'The Thorny and Ardeous Path of Moral Progress': Moral Psychology and Moral Enhancemen*. Neuroethics (2013) 6:141-153

[12] Brem et al. *Is neuroenhancement by noninvasive brain stimulation a net zero-sum proposition?* Neuroimage 2014 Jan 15; 85(0 3): 1058-1068

[13] Kapur N. *Paradoxical functional facilitation in brain-behaviour research*. A critical review. Brain. 1996; 119(Pt 5):1775–1790.

Braincrafting: why playing video games is good for you

Katarina Fatur
katarina.fatur@gmail.com

ABSTRACT

In the article I present cognitive, emotional and social benefits of playing video games, suggest simple guidelines for ethical evaluation of gaming, and offer an overview of findings which support the use of video games for cognitive enhancement and social collaboration.

Keywords

video games, cognitive enhancement, attention, complex problem solving, brain plasticity, learning, consequentialism

1. INTRODUCTION

In the past, research has been mostly focused on the negative effects of video games, but in the last decade, a growing number of studies has supported the view that playing video games may significantly enhance our cognitive abilities and agility. These contradicting facts fuel the discussion on whether playing video games is harmful or not. Here I focus on the benefits of gaming and offer an overview of findings which support using video games as tools for cognitive enhancement and social collaboration, while also outlining the ethical framework which is to be followed in order to avoid negative aspects of gaming.

2. PROPOSED ETHICAL FRAMEWORK TO EVALUATE VIDEO GAMES

Up until a few years ago, research has been mainly focused on the potential adverse effects of video games, such as increased aggression [1] [2], addiction to gaming [3] and increased obesity risk [4]. While I do not believe that gaming is inherently bad, I can understand the reservation and caution. For better or worse, playing video games significantly changes our brain [5], which carries serious implications. With our growing understanding of the brain comes greater responsibility about how to apply this knowledge in a benevolent manner. If we consider video games as a tool for self-improvement and cognitive enhancement, they should align with (neuro)ethical guidelines which govern any other means for cognitive therapy. Ethics is a set of principles which prescribe what is right or wrong in terms of rights, obligations, or benefits for individuals and society. The subfield of neuroethics, which is focused particularly on the issues in neuroscience, is less than 20 years old. The concept was first introduced in 2002, when William Safire defined it as "the examination of what is right and wrong, good and bad about the treatment of, perfection of, or unwelcome invasion of and worrisome manipulation of the human brain" [6]. Video games possibly constitute a case of 'worrisome manipulation' and open a window for 'unwelcome invasion' of the brain, but at the same time they can also act as one of the most pleasurable and safe brain enhancement tools. Do benefits warrant the dangers?

To do good, we first have to determine what is good. The ethical framework to help us evaluate the 'goodness' of video games in this article is an epicurean-like teleological/consequentialist theory of ethics with a bit of pragmatism, where we pursue good and pleasurable outcomes, where an action's value is based on the outcome of that action, and where the evaluation whether the outcome is good or bad might shift when new scientific discoveries are made. It appears that this ethical framework fits neuroscience research better than other normative ethical theories, simply because the subjective virtuous conduct of game-designers (virtue ethics) or a scientists' firm belief that undertaking experiments is the right thing to do (deontological ethics) is of little or no importance in comparison to unwanted, but nonetheless possibly detrimental effects a person might suffer as the consequence of her or his invaded and changed brain.

The last conclusion partly rests on the notion of 'fairness' – it is not fair that a player would be harmed for the sole purpose of advancing science, or because the game designer made a mistake despite the designer's moral character being impeccable. Fairness can be fostered by assuming equality of all involved parties and by clear and honest communication, where the gaming industry candidly informs the player about possible effects. Informing the player is also the prerequisite for the player to give their consent to undergo a benevolent cognitive manipulation when playing a game. At the same time, the player should provide feedback, so that the industry can align and improve their practice. This also involves the issue of privacy protection, especially in the case of collecting neurodata in neurogaming[1].

Simply put, for gaming to be ethical, it has to benefit the player and it has to fulfill three criteria: it should result in positive, useful outcome for the gamer and/or society, it should respect and protect gamer's privacy, and the gamer must participate knowingly and willingly. Since new technologies pose difficult ethical questions and could be easily used for ill intentions, it is important to show, why is it worth to undergo the gaming

---

[1] In neurogaming, the player's heart rate, brain activity, facial expressions, voice, skin conductance, eye movement, pupil dilation, and similar indicators are measured in order to use the input to dynamically adjust the gaming content and provide a completely personalized, immersive gaming experience, which takes into account the player's emotional and cognitive state [7]. In this way, the desired (or feared) cognitive manipulation is perfectly suited to the player. Some game engines allow the players to navigate the game merely by directing their gaze or blinking. Future development also includes different techniques which will simulate certain sensations, e.g. the player holding an object in their hand while there is no object there [7]. The ethical concerns are significant, but in this article I mainly focus on the benefits of gaming and mention neurogaming merely as an interesting new branch of the industry, and my purpose is to show, why it is worthwhile that we acknowledge and tackle these ethical issues to be able to benefit from our gaming experience as much as we can.

experience despite the risk, or even better, while avoiding the risk. In the following sections I offer an overview of beneficial uses of video games and discuss how they affect the brain.

3. BENEFITS OF VIDEO GAMES

Video games have become increasingly diverse and complex, (hyper)realistic, and socially engaging. It has been shown that playing certain types of video games may enhance perceptual, motor, social, emotional and cognitive learning [8]. While scientists still do not know exactly how different brain regions interact, or what is the role of each and every particular neuroanatomical element, some hypotheses have been made in studies researching the impact of gaming on brain plasticity. Below I present the supporting studies in more detail.

3.1 Learning

At a neuroethics conference in 2013 C. Shawn Greene explained that the adult brain does not want to learn anything new, because learning entails costly brain tissue restructuring [9]. So, how do games facilitate learning and foster neuroplasticity? An incentive for the brain to more readily learn something new is the possibility of a reward at the end [9]. A study by Shaowen Bao and his team has shown that when stimulation of the dopaminergic system occurs at the same time as a particular tone is played, over time a larger area of auditory cortex is allocated to process that particular frequency of sound [10]. The brain structurally changes itself in order to more accurately and quickly process the input that will reward it with dopamine release. Games foster learning because they activate the reward system.

Circumstances which bathe the brain in rewarding hormones are also excitement and exploratory behavior [11], which both abound in challenging, but ultimately safe virtual environment of your favorite game. Learning is further encouraged by the fact that you can safely fail – paired with novelty and flexibility of the play, safe environment promotes creativity [12]. In addition, the immediate feedback evaluating the player's game (leveling-up, points, new abilities and tools), present a perfect ground to acquire an 'incremental theory' of players' ability, where the players believe that they can change and actively improve their skills by investing more time and effort [8]. How is this useful in real-life scenarios? The learned skills can be transferred, as illustrated by the examples below.

3.2 Visual attention

Green explains that action games exhibit particular qualities, which sharpen a gamer's visual attention [9], such as: complex 3D environment, fast motion, transient visual stimuli, and heavy perceptual load with several events occurring simultaneously, where not all events are important. 15 years ago Green and Bavelier demonstrated that selective visual attention improves with playing action video games, with effects lasting for 6 months after the last training [13]. In more complex multiplayer first person shooter games, there is also a heavy cognitive load: the player has to observe certain strategy and keep track of other players' actions. While primary visual focus is centrally on the screen, peripheral vision is fully engaged as well, since important elements tend to pop up from the sides.

The causal relationship between playing action games and improved players' ability to coordinate visual input with their

motor control was shown in 2016 [14]. Players who played action games for 5 hours per week over 6 months were later re-tested for their driving abilities and they exhibited better lane-keeping and visuomotor-control skills. Cognitive gains acquired in gaming are thus transferable to real-world visual ability.

Another example of induced neural plasticity in the players' visual system is the case of adult amblyopia ('lazy eye') patients, whose fundamental visual functions improved from 16-54% after playing action video games [15]. It was a small scale pilot study, but researchers are optimistic about the possibility to apply the principles to treating other cortical dysfunctions as well. Additional benefits of playing action games in relation to vision are improved visual attentional skills in dyslexic children [16]. Without any direct reading training, the reading abilities of these children improved significantly after only 12 hours of playing action video games.

3.3 Multi-tasking

Each time we switch our attention from one task to another we pay a small cognitive cost and playing action games helps to reduce this cost [9]. In our everyday life we often have to multitask, and the cumulative time spent in task switching is relevant. The positive effects that action games training has on the brain is lasting and can still be observed months after the last gaming session. This effect is also applicable in serious, complex real-life events such as surgery. When novice surgeons played action games for prescribed period of time, the result was a significant improvement in their ability to perform surgery [17]. Even more surprisingly, surgical skills could be better predicted by their screen time than by the number of surgeries performed, or even the amount of time spent practicing the surgery.

3.4 Social and emotional support

Multiplayer online games which engage thousands of people from all over the globe are probably the most exciting, rewarding and productive development in the realm of video games. Immersive social contexts not only provide a playground to develop social skills, but also serve as a solid source of emotional support and resilience, which I present in more detail below.

Massive multiplayer online games (MMOGs) encourage the development of online communities, with associated attachments and social rituals [18]. These communities represent a suitable model for a variety of human societies [19], because they exhibit a strong sense of social connection, identification with other members of the group and a sense of distinguishing themselves from other groups, as well as a strong commitment to the group [18], [19]. On the continuum from *Gemeineschaft* (communal society) to *Gesellschaft* (associational society) [20], where the first originates from solidarity, social union and spontaneous expressions of emotions in personal relationships, and the second is established on rational self-interest with less emphasis on kinship and personal relations, MMOG communities show a greater degree of *Gemeineschaft* [18]. Such social environment promotes cohesion among the members and increases social proximity among them [18], which in turn creates a safe space for establishing personal relationships which offer similar support as real-life interaction [21]. Research on social support from MMOG relationships and associated levels of depression in players demonstrated that higher social

involvement in MMOG communities resulted in a greater level of perceived social support, which consequentially led to lower levels of depression [21].

On a different note, but still relevant, because studies have shown that emotional and physical pain share a lot of common properties [22], [23], playing video games might also have an analgesic effect. In patients undergoing treatment for severe burns, playing video games proved successful in their pain management [24]. When conducting a study in treatment of combat-related burns where pharmacological therapy was complemented with VR gaming, the perceived pain intensity score fell from 6.25 to 4.50 out of 10, unpleasant feelings related to pain dropped from 6.25 to 2.83 out of 10, and the time spent thinking about the pain dropped from 76% during no VR to 22% during VR-coupled treatment [25]. It is suggested that inputs from intense, fast-paced immersive virtual environment distract the brain from processing the pain by disturbing the descending pain-control system, which results in pain-attenuating effects [26].

3.5 Collaboration and problem solving

In addition, in the last decade a subset of these games has tried to tackle some of the many problems which trouble our society and attempted to gamify real life problems. One of the first games which was successful in harnessing the gamers' superior spatial and problem-solving skills was Foldit in 2008 [27]. The objective of the game was to predict the native structure of various proteins to help the scientists to better understand the molecules. Over 50,000 players collaboratively interacted with protein models using user-friendly manipulation tools, and at the end of 2010 their results matched or outperformed algorithmically computed solutions. Foldit players achieved 'an epic win' in 2011, when they successfully deciphered the structure of an enzyme which is critical for reproduction of the HIV/AIDS virus – a problem which had been previously unsolved for 15 years [28]. Following in Foldit's steps is Eterna, a similar online gaming tool, which asks players to solve puzzles in order to design molecular medicines in collaboration with Stanford's School of Medicine [29]. The mechanisms which are responsible for the success of gamification in solving difficult tasks are probably the same learning-promoting mechanisms described in 3.1.

4. CONCLUSION

I summarized important benefits of playing video games, which can be mainly characterized as improved motor, cognitive, social, emotional and problem solving skills. Is there something as a benevolent brain manipulation? I dare say there is, and as studies suggest, playing games moderately and in a responsible way can serve as a therapeutic and cognitive enhancement tool despite the dangers inherent to gaming technologies.

5. REFERENCES

[1] Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., Rothstein, H.R., Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in Eastern and Western countries: A meta-analytic review. *Psychological Bulletin*, *136*(2), 151–173. DOI= http://doi.org/10.1037/a0018251

[2] Gentile, D. A., Lynch, P. J., Linder, J. R., & Walsh, D. A. (2004). The effects of violent video game habits on adolescent hostility, aggressive behaviors, and school performance. *Journal of Adolescence*, *27*(1), 5–22. DOI= http://doi.org/https://doi.org/10.1016/j.adolescence.2003.10.002

[3] Stockdale, L., & Coyne, S. M. (2018). Video game addiction in emerging adulthood: Cross-sectional evidence of pathology in video game addicts as compared to matched healthy controls. *Journal of Affective Disorders*, *225*, 265–272. DOI= http://doi.org/10.1016/j.jad.2017.08.045

[4] Turel, O., Romashkin, A., & Morrison, K. M. (2016). Health Outcomes of Information System Use Lifestyles among Adolescents: Videogame Addiction, Sleep Curtailment and Cardio-Metabolic Deficiencies. *PLOS ONE*, *11*(5), e0154764. Retrieved from https://doi.org/10.1371/journal.pone.0154764

[5] Palaus, M., Marron, E. M., Viejo-Sobera, R., & Redolar-Ripoll, D. (2017). Neural Basis of Video Gaming: A Systematic Review. *Frontiers in Human Neuroscience*, 11, 248. DOI= http://doi.org/10.3389/fnhum.2017.00248

[6] Safire, W. (2002). Our New Promethean Gift. In S. J. Marcus (ed.), *Neuroethics: Mapping the Field*. San Francisco: Dana Press. Retrieved September 1, 2018, from http://dana.org/Cerebrum/2002/Neuroethics__Mapping_the_Field/.

[7] Frank, A. (2013). The Future Of Gaming - It May Be All In Your Head. *Forbes*. Retrieved September 7, 2018, from https://www.forbes.com/sites/singularity/2013/06/03/the-future-of-gaming-it-may-be-all-in-your-head/#7839776b7790

[8] Granic, I., Lobel, A., & Engels, R. C. M. E. (2014). The benefits of playing video games. *The American Psychologist*, *69*(1), 66–78. DOI= http://doi.org/10.1037/a0034857

[9] Green, C. S. (2013, December 16). Neurogaming: What's Neuroscience and Ethics Got To Do With It? - Exploring Ethics. [video file] Retrieved September 7, 2018, from https://www.youtube.com/watch?v=apFE6AR9214

[10] Bao, S., Chan, V. T., & Merzenich, M. M. (2001). Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature*, *412*(6842), 79–83. DOI= http://doi.org/10.1038/35083586

[11] DeYoung, C. (2013). The neuromodulator of exploration: A unifying theory of the role of dopamine in personality. *Frontiers in Human Neuroscience*, *7*, 762. DOI= http://doi.org/10.3389/fnhum.2013.00762

[12] Davies, D., Jindal-Snape, D., Collier, C., Digby, R., Hay, P., & Howe, A. (2013). Creative learning environments in education—A systematic literature review. *Thinking Skills and Creativity*, *8*, 80–91. DOI= http://doi.org/https://doi.org/10.1016/j.tsc.2012.07.004

[13] Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*(6939), 534–537. DOI= http://doi.org/10.1038/nature01647

[14] Li, L., Chen, R., & Chen, J. (2016). Playing Action Video Games Improves Visuomotor Control. *Psychological Science*, *27*(8), 1092–1108. DOI= http://doi.org/10.1177/0956797616650300

[15] Li, R. W., Ngo, C., Nguyen, J., & Levi, D. M. (2011).

Video-game play induces plasticity in the visual system of adults with amblyopia. *PLoS Biology*, *9*(8), e1001135. DOI= http://doi.org/10.1371/journal.pbio.1001135

[16] Franceschini, S., Gori, S., Ruffino, M., Viola, S., Molteni, M., & Facoetti, A. (2013). Action video games make dyslexic children read better. *Current Biology : CB*, *23*(6), 462–466. DOI= http://doi.org/10.1016/j.cub.2013.01.044

[17] Schlickum, M. K., Hedman, L., Enochsson, L., Kjellin, A., & Fellander-Tsai, L. (2009). Systematic video game training in surgical novices improves performance in virtual reality endoscopic surgical simulators: a prospective randomized study. *World Journal of Surgery*, *33*(11), 2360–2367. DOI= http://doi.org/10.1007/s00268-009-0151-y

[18] Simpson, J. M., Knottnerus, J. D., & Stern, M. J. (2018). Virtual Rituals: Community, Emotion, and Ritual in Massive Multiplayer Online Role-playing Games—A Quantitative Test and Extension of Structural Ritualization Theory. *Socius*, *4*, 2378023118779839. DOI= http://doi.org/10.1177/2378023118779839

[19] Szell, M., & Thurner, S. (2010). Measuring social dynamics in a massive multiplayer online game. *Social Networks*, *32*(4), 313–329. DOI= http://doi.org/https://doi.org/10.1016/j.socnet.2010.06.001

[20] Tönnies, F., & Hollis, M. (2001). *Community and Civil Society*. (J. Harris, ed.) (1st ed.). Cambridge: Cambridge University Press. DOI= http://doi.org/10.1017/CBO9780511816260.006

[21] Dupuis, E. C., & Ramsey, M. A. (2011). The relation of social support to depression in massively multiplayer online role-playing games. *Journal of Applied Social Psychology*, 41(10), 2479–2491. DOI=http://doi.org/10.1111/j.1559-1816.2011.00821.x

[22] Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science*, *302*(5643), 290–292. DOI= http://doi.org/10.1126/science.1089134

[23] Kross, E., Berman, M. G., Mischel, W., Smith, E. E., & Wager, T. D. (2011). Social rejection shares somatosensory representations with physical pain. *Proceedings of the National Academy of Sciences*, *108*(15), 6270–6275. DOI= http://doi.org/10.1073/pnas.1102693108

[24] Hoffman, H. G., Doctor, J. N., Patterson, D. R., Carrougher, G. J., & Furness, T. A. 3rd. (2000). Virtual reality as an adjunctive pain control during burn wound care in adolescent patients. *Pain*, *85*(1–2), 305–309.

[25] Maani, C. V, Hoffman, H. G., Morrow, M., Maiers, A., Gaylord, K., McGhee, L. L., & DeSocio, P. A. (2011). Virtual reality pain control during burn wound debridement of combat-related burn injuries using robot-like arm mounted VR goggles. *The Journal of Trauma*, *71*(1 Suppl), S125-30. http://doi.org/10.1097/TA.0b013e31822192e2

[26] Gold, J. I., Belmont, K. A., & Thomas, D. A. (2007). The Neurobiology of Virtual Reality Pain Attenuation. *CyberPsychology & Behavior*, *10*(4), 536–544. http://doi.org/10.1089/cpb.2007.9993

[27] Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popović, Z. and Foldit players (2010). Predicting protein structures with a multiplayer online game. *Nature*, *466*(7307), 756–760. DOI= http://doi.org/10.1038/nature09304

[28] Foldit. In *Wikipedia*. Retrieved September 7, 2018, from https://en.wikipedia.org/wiki/Foldit

[29] Eterna Game. [Web page] Retrieved September 9, 2018, from https://eternagame.org/

In search of the authentic self: explaining phenomenology of authenticity

Maša Urbančič
Independent researcher
Štefanova 13 (telo.si)
1000 Ljubljana
masa.urbancic@gmail.com

ABSTRACT

There are moments in most people's lives when they feel more or less themselves. This experience is usually caught in people sayings that they need to find themselves or just be who they 'really are'. The purpose of this paper is to provide an explanation for these feelings of authenticity. I first explore this concept and consequently introduce necessary conditions for the phenomenological experience of authenticity. This is followed by the examination of two problems that authenticity faces and two possible ways of explaining the experience of authenticity, which are discovery and creation. I then assess three discovery views and show why they are unsatisfactory in explaining feelings of authenticity. In response I then provide my own creation model of authenticity, which argues that feelings of authenticity are a result of a convergence between our current and ideal self. In this sense, authentic self is created. What we are is a matter of choice.

Keywords

Authentic self, true self, authenticity, phenomenology, ideal self

1. INTRODUCTION

The idea of an authentic or a true self[1] seems to have been present since ancient times, from the Greek "Know Thyself" to Shakespeare's "To thine own self be true" to modern advice to "just be yourself" [11]. One often hears people speaking of their need to find themselves or to discover who they "really are" (ibid). I take it that most people agree with Bialystok that "almost everyone has had an intuitive experience of authenticity that seems to reveal a glimmer of one's true identity" [1]. However, the problem arises how it is possible that a person sometimes feels more or less like herself, when she is never another person, but only herself. However, if that is the case, how can these feelings of authenticity be explained? Do these moments of experiencing authenticity reveal a deeper, truer identity of a person, an unchanging inner core that lies beneath, an authentic self? Or do they reveal something else?[2]

---

[1] For clarity purposes, only authentic self is used in this paper, not true or deep self.

[2] I do not differentiate between identity and self and I only take appropriation as a necessary condition for personal identity or self – that we see certain actions, beliefs, etc as ours. For a discussion on personal identity, self (and true self) from a phenomenological point of view, see Jacobs [5].

The aim here is to provide a model for specifically explaining the phenomenological (i.e. lived) experience of authenticity, which I will call the authentic-like experience. The main argument will be that feelings of authenticity are (usually positive) momentary experiences, which are a result of a convergence between the current self (beliefs about oneself) and the ideal self (what one wants to be, thinks she should be and what is important to her). On this account, the authentic self comes about as a consequence of creation and not as a consequence of discovery.

2. AUTHENTICITY AND ITS PROBLEMS

The purpose of this section are the following four points: 1) to introduce the concept of authenticity, 2) to examine necessary conditions for its phenomenology 3) to present two problems that authenticity faces and 4) to present two ways of thinking about the explanation of an authentic-like experience.

The concept of authenticity is defined in multiple ways (see [1], [3], [12]). Here, I take authenticity to denote "a convergence between what something is and what it claims it is" [1]. In other words, when a person is being authentic, her statements and thoughts, as well as her beliefs and actions are in accordance with one another.

For a person to undergo an authentic-like experience, I argue there are two necessary conditions that need to be met: 1) a certain degree of enjoyment or dislike of a certain trait or ability and 2) like Bialystok [1] argues, a convergence with *a set of criteria* where, when the convergence with those criteria takes place, the result will be an authentic-like experience. In other words for the second condition, if a person is sometimes authentic, there must be times when she is inauthentic [1] so we need to have a standard that defines when the person is being authentic. Whenever a person's current state converges with those criteria, an authentic-like experience will take place. I will call those criteria the authentic self* (and later on in my model, the ideal self), which specifically denotes *the thing* with which the person converges *with* in an authentic-like experience. It can be thought of as discovered or created. On the other hand, the authentic self denotes the *result* of a convergence between the person and the authentic self*, i.e. the authentic-like experience and what people refer to as their authentic self.

The convergence condition or the authentic self* presents a problem in the discussion of authenticity because it is unclear what the authentic self* represents, which will be the discussion for the rest of the paper. In other words, it is unclear what makes some parts count as the authentic self and how to separate it from the inauthentic parts (see [1] for this issue). Authenticity denotes the harmony between one's beliefs, thoughts and actions (see previous paragraph), but a person rarely identifies with all her

thoughts and beliefs and declares them as her authentic self (e.g., rarely do we encounter a person that would identify having a thought of strangling somebody who is annoying as her authentic self). However, as Bialystok [1] points out, in current understanding of authenticity the authority on the authentic self is found within ourselves, which is why it is to an extent a matter of self-interpretation. There are no objective criteria against which it could be determined which parts of the person are authentic and which not or any empirical facts that could confirm a person's authentic self [1]. It will therefore always come down to what the person declares as her authentic self. I do not hold that the matter of authentic self depends on self-interpretation is problematic as Bialystok [1] does, since the final decision with what we identify ultimately lies in our hands. However, the problem of separating authentic from inauthentic parts still persists.

There is a second further issue with authenticity, which Bialystok [1] labels as the hard case of authenticity. More specifically, it is impossible to determine when the person is being authentic where one's feelings, thoughts or beliefs change throughout life. To imagine such scenarios, Bialystok refers to religious conversions, career change, divorce, change in party affiliations or simply changes in personality [1]. It is not difficult to imagine such cases where one is different from how one used to be in the past. The dilemma however, is, when is the person then authentic? For instance, is a person authentic before or after a religious conversion? Intuitively one might say that one is authentic after the change, but most people claim they were also authentic before. If a person is authentic in both cases, then authenticity seems impossible [1]. But if a person is inauthentic in both cases, when is he then authentic at all if not in the beliefs that constitute a greater part of his identity (ibid)? I believe this argument mistakenly presupposes that a person's identity should be constant in order to constitute the authentic self. One can still act in accordance with what one feels in that moment, without assuming the authentic self* is constant. Nevertheless, any explanation of authenticity needs to account for such cases as well.

This brings us to two ways of thinking how the authentic self* is to be understood. These are discovery and creation, which I take from Waterman's [14] distinction between different ways of understanding identity formation. The metaphor of discovery denotes that something, which has already *existed before*, has now become known or understood [14]. The discovery entails the process of causing the unknown to become known (ibid). In this case, a person is discovering a prior authentic self* and every time she converges with this authentic self*, she undergoes an authentic-like experience. Creation, on the other hand, involves creating something that has never existed before. In this process, one starts off from a point of unlimited possibilities and among those makes a choice and brings together parts to create something of value [14]. On this view, the assumption would be there is no prior not-yet-discovered authentic self* and it is only a product of our choices [12]. In this way, the authentic-like experience can be either thought of as being a result of a convergence with a discovered prior authentic self* or as a result of a convergence with a created authentic self*. In this way, an authentic self* can therefore be something that has already existed before and is becoming known or understood in moments of experiencing authenticity (or most probably parts of it). Alternatively, the authentic self* has not existed before and is created from unlimited possibilities as a matter of choice (see e.g. [13]). Moments of experiencing authenticity are therefore moments of experiencing something that we have created. I will

argue that it is the latter case that takes place when authenticity is felt. In the next part, I will present different ways one can think about the authentic self* as being discovered and why these do not offer a satisfactory explanation for the phenomenal experience of authenticity.

3. AUTHENTIC SELF* AS DISCOVERY

The first question that needs to be asked is if the authentic self* (the criteria) is to be discovered, what exactly is it that is being discovered? I will consider three possible answers to this problem, which are 1) an unchanging ontological substance 2) a unique personality or 3) a daimon. These can be regarded as manifestation of these criteria (or the authentic self*).

The first possibility of what the authentic self* represents is connected to philosophical debates on the concept of self. Metzinger defines it as an enduring individual entity, an unchanging essence or an ontological substance that could exist independently in the world and would represent a basic constituent of reality [6] [7]. He goes on to deny the existence of such a self and argues there is currently no substantive empirical evidence or any theoretical argument that would necessitate the existence of a self [7]. Zahavi, on the other hand, rejects Metzinger's definition and characterizes the self rather as being constituted through the first-person givenness of different experiences [15]. In other words, there is no separate entity, but a self exists simply due to the fact that the world is being experienced from a first-person perspective. Olson [8] goes even so far as to state we should cease to speak of selves due to lack of a common definition and its causing dilemmas that would otherwise be avoidable.

Any of these definitions do not provide any insight into the authenticity of the self. If the authentic self* is understood in Zahavi's terms, such a definition presents a problem for authenticity, since the givenness of a first-person's perspective is always present, regardless whether a person feels authentic or inauthentic. Just having a first-person experience does not solve the issue of why a person sometimes feels more or less like herself. If the authentic self* is understood in the way that Metzinger defines it, the authentic self* is then understood as an unchanging ontologically independent core that is distinct from any characteristic, perception or trait that constitute the person. Furthermore, as Metzinger points out, there is currently no empirical evidence for this notion of self [7]. Additionally, there is also the issue that even if there seems to be the experience of an unchanging self from a phenomenological standpoint (i.e. constant unity of consciousness), we are nevertheless unable to directly access this inner core self if it is a separate and distinct entity, as Hume initially pointed out (see [4]). Therefore, if there is an entity as an authentic self* that is separate and *distinct* from all characteristics or perceptions that constitute the person it still does not help address the problem of authenticity. It is not connected to something that would be over and above the abilities and characteristics that constitute him and which provide him with an authentic-like experience. Hence, feelings of authenticity cannot be disclosing a convergence with an independent entity-like authentic self*, but must be presenting something else. For these reasons it seems viable to conclude that conceptualizing the authentic self* as a separate entity cannot satisfactorily explain the phenomenology of authenticity.

The second model avoids the Hume worry in that it conceptualizes the authentic self* as a unique personality, i.e. a set of characteristics and traits that are unique to every individual and not as something distinct. This model presents the modern

conception of authenticity that is especially prevalent in contemporary self-help literature (see [2]). In opposition to older conceptions of authenticity, where the task was to become what one can be, i.e. realizing her potentials and purpose, the contemporary ideal of authenticity requires that the person realizes and becomes what she *already is*, the set of characteristics and traits that are *already situated within* the person ([2, 13], my emphasis). The main idea of this understanding of authenticity is that everyone possesses a deep, authentic self* within – a so-called Real Me – that is distinct from everything else that is not really me. In other words, there is great emphasis on the difference between the inner and outer, where *the outer is false and the inner true* ([2], my emphasis). Authentic self* represents the collection of all feelings, needs, desires, capacities, dispositions and abilities that constitute a person's unique personality and our task is to get in touch with this personality (e.g. through introspection) and to express it (see [2] on this). Here therefore, whenever the person is acting in accordance with this unique personality, she is feeling authentic. The problems here are the assumptions our personality has been shaped at birth and there seems to be disregard for personal experiences on shaping a person's self because this understanding of the authentic self* assumes the authentic self* is within us and not influenced by the environment (the outer), neither it is changed throughout life. With the inner-outer distinction, there is also disregard for societal influences on our identity.

The third model envisions a daimon (= authentic self*) that represents the potentialities and talents within each person which are not yet realized but whose realization represents the greatest fulfilment in life (an ideal) [14]. In other words, the person aims to reach her authentic self* or her daimon by realizing her potentials, not by becoming a person she already is within. The difficulty with this model is still the persisting assumption that daimon is fixed since birth (if daimon is to be discovered with time and lived, it is difficult to know what is being discovered if it consistently changes). And as the hard case of authenticity highlights, nothing about a person's personality appears to be permanent enough that it would remain fixed upon discovery. The person does not stay the same after unearthing the daimon. It might be objected that the daimon is only a set of potentialities and abilities. It presents us with the limitations of our capabilities or reveals in what activities or tasks we find enjoyment in. It is not to be understood as a set of fixed properties. This argument holds some merit, but it must nevertheless be argued that we might become aware of our capabilities, interests and talents, but we are not equal to them. Feelings of authenticity can be out of alignment with one's abilities as in, for example, when a person is highly talented for running but does not enjoy it. Therefore, potentials and abilities are not a sign of an authentic* self.

4. AUTHENTIC SELF* AS CREATION: MY OWN MODEL

The purpose of this section is to present my own model for explaining feelings of authenticity, which views the authentic self* as created. The idea that the authentic self is created is mostly found in existentialist writings. Generally put, according to existentialism, there is no pre-existent essence or a "true self", but one's identity is created through personal choices [14]. Sartre states that "man first of all exists /…/ and defines himself afterwards" and that every person chooses herself [10] (see also Nietzsche and Foucault).

This model is ultimately compatible with the existentialist view of the authentic self, but whereas they focus more broadly on the

ontology and the nature of self, I focus specifically on the explanation of feelings of authenticity. I argue this phenomenology of authenticity is best explained by a model on which feelings of authenticity occur as a result of a convergence between the current and the ideal self. As argued above, there must be a convergence between a person and her authentic self* for authenticity to occur. From now on I will refer to the authentic self* as the ideal self in my model. According to this model, feelings of authenticity are (usually positive) momentary emotional experiences. Feelings of authenticity are usually connected to positive traits because the ideal self is generally viewed as intrinsically good (see [8]). However, there can also be instances where the ideal self is connected to negative traits (e.g. unsociable, not funny, etc).

Current self denotes all beliefs one holds about herself at the present time, for example being athletic, not funny enough, etc. The ideal self can be thought of as a collection of all desires of how one wants to be, norms of how one should be and values that constitute what is important in a person's life. It is important to note that what will constitute the ideal self will be highly influenced by what is desirable within the society as well. Both the current and ideal self-beliefs can be either conscious or non-conscious. When one considers one's ideal self, there are four possible scenarios that can take place. First is the scenario where a convergence between a positive current self belief (e.g. I am funny) and the ideal self (I want to be funny) occurs. In those cases, feelings of authenticity will be experienced. In the second scenario there is a lack of convergence between a person's positive belief about the current self and her ideal self. In this case, the person most likely experiences a questioning of her behaviour, or (if such situations happen often enough) an identity crisis if the belief is central to her identity. The third scenario is a case of negative current self belief (e.g. I am not smart) and a lack of convergence with the ideal self. In this case, a person believes she is lacking certain ability or trait (current self) that she wants to possess (ideal self) and situations confirm that it is the case she is lacking them (confirmation of existing current belief). In this scenario, there is no self-acceptance present of how she is at the present time (and oftentimes no self-awareness of what constitutes her ideal self). In such scenarios, possible phenomenology might be one of sadness or disappointment, but for the purposes of this paper it is important that no feelings of authenticity will be experienced. The fourth scenario presents the most complicated scenario where the negative current self belief initially does not converge with the ideal self but can gradually converge with the ideal self if a) negative current self beliefs change to positive or b) ideal self changes (to negative belief). The current negative beliefs can change into positive in two ways: 1) by changing behaviour or 2) by gradually discovering that the negative beliefs were inaccurate. For illustration of 1 and 2, a person named Peter believes he is unsociable (negative current belief) and wants to be sociable (ideal self). There is currently no convergence and no authentic-like experience is present. He can change his behaviour to being sociable, which will in turn change his negative current self belief of unsociable to sociable and a convergence with the ideal self will occur. In those moments, Peter will have an authentic-like experience. Alternatively, Peter is exposed to situations where he receives feedback that he is indeed sociable and like his ideal self, despite his thinking otherwise. In this case, Peter's belief of being unsociable will gradually transform into current self belief of being sociable and a convergence between his current and ideal self will take place. Again, authenticity will be present in such a scenario. In terms of changing the ideal self

(option b), Peter comes to accept he is unsociable and realizes that his enjoyment of his alone time is more important to him than being sociable (or some other reason for preferring not to be sociable). In this case, a shift within the ideal self takes place in that it now includes not being sociable. Therefore, if Peter is being reproached for not being sociable, he says "This is who I am" and experiences authenticity because being unsociable is now part of his ideal self.[3] If Peter might be disrespecting the latter desire, this could lead to scenario two.

This model avoids the problems faced by the previous presented models. Firstly, it avoids the problem of having to find an enduring ontologically distinct entity that is over and above the traits and abilities of a person, because the ideal self is constituted by them. Secondly, it avoids the problems of the ideal self being pre-determined and separate from the society. The ideal self is a collection of values, norms and desires that continually develops and changes throughout life. The ideal self is influenced and constituted in interaction with and by society and shaped by personal experiences. Furthermore, the presented model of the ideal self can help address the two problems with authenticity raised in section two. First, it can help explain the hard case of authenticity since I argue that the bar for authenticity is behaving in accordance with the ideal self, not constancy in one's identity. To the second problem, namely how to distinguish between authentic and inauthentic parts, my answer is that it is the convergence with the ideal self (plus enjoyment or dislike as discussed before) that makes the difference. In this way, it could be researched what people aspire to and move away from (i.e. what is part of the ideal self and what not) and how they feel about these characteristics (enjoyment or dislike) by which we could better understand how they perceive their authentic selves and if people act in accordance with the scenarios presented.

Finally, it might be objected that the phenomenology of becoming more and more self-aware with time pushes us more towards the discovery view. There is a sense of "peeling-off" layers and coming to know ourselves better during life. However, I would argue that we are not discovering an authentic self* that is waiting to be discovered in those moments. We are becoming more aware of our ideal self and the latter is becoming more specific and detailed with time. The one other concern is how does one distinguish what is part of the ideal self and what not? I would argue that what the ideal self contains is largely a matter of choice. There is a matter of making the ideal self conscious and more detailed. But once that becomes the case, a person can make a choice in what the ideal self will be. In that way, we decide what we want to be. In that sense this does not simply push the authentic self one step further with the ideal self (instead of an authentic self* we are discovering an ideal self) because the ideal self is not a constant entity, but a collection of our values, beliefs, desires and influences that is simply becoming more and more detailed with time. The more detailed it is, the greater choice we have in what we want to become, i.e. our authentic self.

5. CONCLUSION
The aim of this paper was to provide a preliminary model for explaining phenomenology of authenticity. On this account, authentic-like experiences are a result of a convergence between our current self and ideal self. The ideal self is a collection of our

beliefs, thoughts and desires at the present time. The task for further research would be to look deeper into the structure and creation of the ideal self. It would be further useful to empirically examine which situations elicit feelings of authenticity and to investigate what drives people's experiences of authenticity in those moments.

7. REFERENCES
[1] Bialystok, Lauren. Authenticity and the Limits of Philosophy. *Dialogue,* 53: 271-98. 2014.

[2] Guignon, Charles. *On Being Authentic*. USA, Routledge, 2004.

[3] Guignon, Charles. Authenticity. *Philosophy Compass,* 3 (2): 277-90, 2008.

[4] Hume, David. *A Treatise of Human Nature (2nd Edition)*, L.A. Selby-Bigge and B.H. Nidditch (eds). Oxford, Oxford University Press, 1739-1740/2014.

[5] Jacobs, Hanne. Towards a Phenomenological Account of Personal Identity. In *Philosophy, Phenomenology, Sciences: Essays in Commemoration of Edmund Husserl,* Phaenomenologica 200. Springer, New York, 2010.

[6] Metzinger, Thomas. *Being No One: The Self-Model Theory of Subjectivity*. USA, MIT Press, 2003.

[7] Metzinger, Thomas. The No-Self Alternative. In S.Gallagher, editor, *The Oxford Handbook of the Self*, pages 279-96. Oxford, Oxford University Press, 2011.

[8] Newman, E. George, Bloom Paul and Joshua Knobe. Value Judgments and the True Self. *Personality and Social Psychology Bulletin,* 40 (2): 203-16, 2014.

[9] Olson, T. Eric. There is no problem of the self. *Journal of Consciousness Studies,* 5: 645-57, 1998.

[10] Sartre, Jean-Paul. *Existentialism and humanism*, trans. Philip Mairet. UK, Methuen London, 1980.

[11] Schlegel, J. Rebecca, Hicks, A. Joshua, Arndt Jamie and Laura A. King. Thine Own Self: True Self-Concept Accessibility and Meaning in Life. *Journal of Personality and Social Psychology,* 96 (2): 473-90, 2009.

[12] Vannini, Phillip and Alexis Franzese. The Authenticity of Self: Conceptualization, Personal Experience, and Practice. *Sociology Compass,* 2 (5): 1621-37, 2008.

[13] Varga, Somogy. Existential choices: to what degree is who we are a matter of choice? *Continental Philosophy Review,* 44: 65-79, 2011.

[14] Waterman S., Allan. Identity Formation: Discovery or Creation? *The Journal of Early Adolescence,* 4 (4): 329-41, 1984.

[15] Zahavi, Dan. Unity of Consciousness and the Problem of Self. In S. Gallagher, editor, *The Oxford Handbook of the Self*, pages 316-35. Oxford: Oxford University Press, 2011.

---

[3] I would argue that if a characteristic is seen as fixed and cannot be changed, it is scenario three. If the said characteristic is more a choice, then it is scenario four.

Exploring Features of
Cognitive Science as Natural Epistemology

Tine Kolenik
Department of Intelligent Systems
Jožef Stefan Institute
Jamova cesta 39
1000 Ljubljana, Slovenia
tine.kolenik@ijs.si

ABSTRACT

Historically as well as conceptually, epistemology and cognitive science seem to be intimately connected. The overlap between them seems to be rather extensive, and this overlap can be dubbed as natural epistemology, which has been similarly proposed by a few other researchers. In the work, I define natural epistemology as the study of epistemological questions using scientific methods. Due to the particular nature of studying knowing, which is what cognitive science largely does, I identify other important features, hidden behind the definition. To explore such features, I apply Kuhn's 'essential tension' between convergent and divergent thinking to how progress in cognitive science had historically played out, especially regarding its paradigm shifts (cognitivism, connectionism, embodied/embedded cognition, enactivism). I propose that 'essential tension' that fuels progress in cognitive science manifests as a loop between knowing of cognitive science (epistemology of cognitive science) and knowing within cognitive science (epistemological insights in cognitive science). I describe the knowledge loop's importance and relate it to the history of cognitive science. In the end, I discuss future work, pointing to the problem of what is epistemological or not as well as the problem of potential convergence of the knowledge loop with the ideas of second-order cybernetics.

Keywords
cognitive science, epistemology, Kuhn, philosophy of cognitive science, philosophy of science

1. INTRODUCTION

Cognitive science (CS) has been by many characterized by its strong historical [1] and contemporary connection to epistemology [2], despite some opposition to the possibility of expanding philosophical ideas to scientific disciplines [3]. Kvanvig [4] indicates that the question of the relationship between mind and world (and more broadly of 'what it is to know') is the deepest question of epistemology, and some [2] have claimed that this same relationship is the foremost investigative matter of CS. Both CS and epistemology ask certain questions that seem to be the same, especially in terms of sources of knowledge of the external world, e.g., perception [5], where epistemology offers answers through philosophical analysis [6], while CS uses different scientific methods for its answers [7]. For example, both are interested in perception, being one of the sources of knowledge, and how it is used to construe the world. Berkeley [6] argues against material objects using philosophical analysis, while one of the constituent disciplines of CS, artificial intelligence, uses computer vision to examine different ways that the world is constructed. Namely, in computer vision, material objects are largely construed and

identified by continually updating the image and comparing it with stored images of that same object [7] – this happens especially when identifying the same object from different perspectives. It therefore seems that CS can offer some sort of answers to epistemological questions. I believe the insight into the deep intertwinings between CS and epistemology is important to be able to discern an overlapping research endeavor that studies epistemological questions with natural scientific methods. This research endeavor, which will be dubbed as natural epistemology, is strongly embodied in CS. By realizing the epistemological background of the latter, it is important to delve into its features as natural epistemology to understand its particular nature, organization and scientific workings, which have significant influences on how research and progress in natural epistemology is manifested.

2. NATURAL EPISTEMOLOGY

While gaining insight from studying epistemological questions with contemplative philosophical work is an extremely valuable endeavor, philosopher and logician Quine followed what Dennett [8, p. 134] colorfully described with the following words: "Just as you cannot do very much carpentry with your bare hands, there is not much thinking you can do with your bare brain." Quine [9] described a view wherein epistemology uses scientific methods and subsequently coined the phrase 'naturalized epistemology'. In parallel, similar views arose from cybernetics and second-order cybernetics circles as well as from the sociological domain. Bateson [10] used the terms 'empirical epistemology' and 'experimental epistemology' when describing an endeavor of researching perceivers' presuppositions built into the product or the final image of perceiving. Keeney [11] used the term 'natural epistemology' to research how living beings make distinctions in the world that construct it while at the same time knowing that these distinctions are wholly dependent on these same living beings. Luhmann [12] used 'natural epistemology' to characterize that knowledge depends on an observer's inner processes and not on the question of what is 'true' or 'false' as such. He notes that there is a certain circularity in this, observing self-reference in natural epistemology, which is hinted by Keeney as well – we supposedly know that the construction of the world is what living beings do, which we want to research, but at the same time we are aware that we are these same living beings who construct the world in the first place. I therefore propose that the base definition of natural epistemology is this: Natural epistemology is the study of epistemological questions with the use of natural scientific methods. However, due to the particular nature of the study, stemming from the fact that the construction of knowledge is studied by constructors of knowledge or that knowing is investigated by others' 'knowings', natural epistemology as manifested in CS possesses some unique

features. I try to ascertain some of them in this work. There are some similarities of the ascertained features to those of second-order cybernetics, which is "the study of the organization in autopoietic machines, which are able to build their own components, and possess cognizance which allows them to observe and control other machines" [13, p. 73]. This insight is self-referenced in natural epistemology and it thus determines the progress in CS as such. To identify the features and their influence in the progress of CS, I describe and apply Kuhn' concept of 'essential tension' between convergent and divergent thinking in science[1] [14] to CS.

3. KUHN'S ESSENTIAL TENSION

Kuhn [14] proposed that progress in science is typified by the interplay between convergent thinking and divergent thinking which emerges as 'essential tension'. According to Kuhn, convergent thinking in science is what scientists usually do and the way they normally operate, where "the research scientist is not an innovator but a solver of puzzles, and the puzzles upon which he concentrates are just those which he believes can be both stated and solved within the existing scientific tradition" [14, p. 234]. The usual research scientist, "the practitioner of a mature science, from the beginning of his doctoral research, continues to work in the regions for which the paradigms derived from his education and from the research of his contemporaries seem adequate" [14, p. 245] and tries to "elucidate topographical detail on a map whose main outlines are available in advance" [Ibid.]. Kuhn believes, partly opposed to usual views on the topic, that convergent thinking is extremely important as it ensures safe, steady, sure and stable production the likes of which cannot be achieved with divergent thinking as prevalent; it is also why, apart from being the norm, convergent thinking is so abound and common. It is exactly its exuberance that eventually leads to encountering "a problem in which the anticipated does not occur, a problem that goes wrong to ways suggestive of a fundamental weakness in the paradigm itself" [Ibid.]. Convergent thinking in established sciences is therefore a prelude to divergent thinking, which has to happen in order to find a solution to the described problem. However, scientists do not always go beyond the prelude as they do not identify the problem as paradigm-breaking. They believe that continuing research in the established paradigm will exhibit desired outcomes. Alas, according to Kuhn, the desired outcome cannot occur, as convergent thinking is "neither intended nor likely to produce fundamental discoveries or revolutionary changes in scientific theory" [14, p. 233]. The scientist therefore has to have "the ability to recognize trouble when confronted by it" [14, p. 235] and to recognize "that something has gone wrong with existing knowledge and beliefs" [Ibid.]. This modus operandi continues into divergent thinking, which means that the scientist must "rearrange the intellectual and manipulative equipment he has previously relied upon, discarding some elements of his prior belief" [14, p. 226] and "lack prejudice" [Ibid.], which is imposed by the established paradigm, "to a degree where [the scientist] can look at the most `self-evident'

---

[1] Kuhn described and published his work on 'essential tension' as a result of lecturing at a conference on scientific creativity, where the knowledge of creativity, which is gained by scientists who study creative talent, was applied to scientists themselves.

facts or concepts without necessarily accepting them" [Ibid.]. This is the last step before "one of those shifts in fundamental theory, in problem field, and in scientific standards […], referred as scientific revolutions" [14, p. 234].

I believe that understanding Kuhn's concept of 'essential tension' and its role in scientific progress is necessary to articulate how the self-reference of natural epistemology in CS plays out. This articulation will follow after a short examination of how progress worked historically in CS in relation to Kuhnian scientific revolutions or paradigm shifts.

4. SHORT EXAMINATION OF PROGRESS IN COGNITIVE SCIENCE

Froese [15] distinguishes (roughly) four paradigm shifts in CS so far, which can be seen in the figure below.
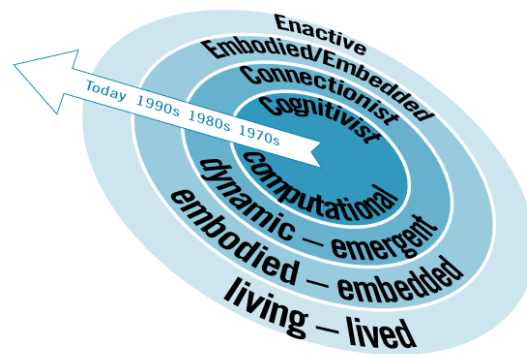

Figure 1. Froese's interpretation of paradigm shifts in CS [from 15, p. 76]. Circles do not represent the end of a paradigm, only emergence of a new one.

The four paradigms are cognitivism, connectionism, embodied/embedded cognition and enactivism. Following Kuhn's 'essential tension', convergent thinking was researchers' modus operandi within these paradigms, while divergent thinking ultimately spawned a paradigm shift. The short overview below is meant as a conceptual validation of Kuhn's ideas rather than a comprehensive and nuanced treatise of the history of CS. It is more focused on the role of artificial intelligence (AI) in it, but it has to be noted that AI was only one wave in the necessary tsunami that caused a paradigm shift. This means that ideas from many other consitutent fields of CS that directly contributed to paradigm shifts are not mentioned.

The insight that cognition is supposedly computation with arbitrary symbols started the cognitive revolution and the era of cognitivism in the 1950s. This idea on cognition has been manifested in most research since. A few decades long prosper, especially in AI, signalled to scientists that their idea about cognition was correct. One of the most publicly known examples of this prosper is the AI chess player, as it symbolized human higher-order cognitive abilities. The artificial chess player's skills grew exponentially, and in 1996, the reigning world champion of the time Kasparov was beaten by it. However, the idea of the top-down, centralized cognition, embodied in the AI of the time proved to have many problems. For example, large domains were a big issue [16], which could not be solved by convergent thinking. It was recognized "that something ha[d] gone wrong with existing knowledge and beliefs" [14, p. 235],

and in the 1980s, the idea on cognition shifted to it being parallel, distributed and bottom-up, which was largely synonymous with the method of artificial neural networks (ANN). Connectionism opened doors to new research in many phenomena where the cognitivist paradigm failed (e.g., natural language processing [16]). ANNs are still predominantly used today and exhibit numerous achievements in mimicking human abilities[2]. However, many scientists [17] felt that connectionism was missing a key point in what cognition supposedly is and how organisms perceive and act by dismissing the constitutive role of bodies in cognition and focusing solely on the brain. Again, convergent thinking shifted the view on cognition that started to incorporate organisms' bodies and how they constitute organisms' cognition. By being aware that cognition is embodied as well as embedded in the environment, research in robotics (as well as using robots to investigate cognition) started to flourish like never before, convergent thinking leading the way again.

The last paradigm shift to enactivism will not be covered here, as its manifestation seems to be a bit more unclear than the previous ones as well as it being harder to delineate from the embodied cognition paradigm [18]. The short overview was meant to be exemplary as to apply Kuhn's 'essential tension' to historical progress in cognitive science as natural epistemology.

In the next chapter, I will build on the manifestation of Kuhn's 'essential tension' in CS by noting the role the circularity of knowing (see chapter 2) plays and relating it to convergent and divergent thinking. In the chapter I will show how, when conceptualized in this way, CS's particular organization works.

5. THE KNOWLEDGE LOOP OF COGNITIVE SCIENCE AS NATURAL EPISTEMOLOGY

From the way the progress in CS was described in the previous chapter, I can outline an abstract step-by-step template:

1. Scientists research a particular cognitive phenomenon with their existing knowledge (Kuhn's 'prior beliefs' or presuppositions from an established paradigm) about knowing.

2. Scientists make great advances, but then encounter an insurmountable problem, which cannot be solved with convergent thinking (current epistemological presuppositions).

3. If the problem is identified as such, scientists have to work through a different set of epistemological presuppositions and new ideas about cognition, therefore exhibiting divergent thinking.

4. The problem is solved, which opens the door to numerous new research not envisioned as such before. A paradigm shift within which research is conducted occurs, changing how the discipline sees knowing and changing epistemological presuppositions of the scientists.

This outline forms a loop, as the last step inevitably leads into the first. The figure below summarizes this process:
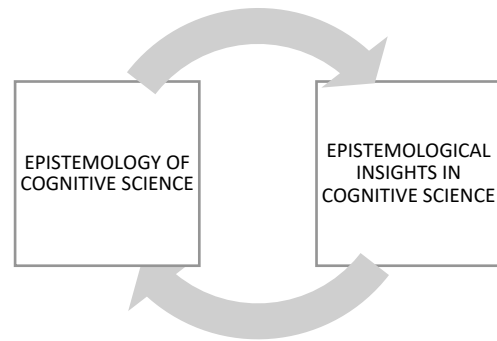

Figure 2. The loop between epistemological insights in cognitive science and epistemology of cognitive science.

Relating this loop to Kuhn's 'essential tension', convergent thinking represents epistemological insights in CS or research on knowing and cognition that occurs on daily basis. Divergent thinking represents epistemology of CS, as this is where a shift has to occur to overcome the problems in epistemological insights in CS. The interplay of knowings is apparent here: epistemology of CS may be dubbed as knowing of cognitive science, while epistemological insights in CS may be dubbed as knowing within cognitive science. One refers to the established paradigm and idea on cognition, from which the other stems. To overcome paradigm-breaking problems in knowing within cognitive science, scientists have to become aware of epistemological presuppositions they hold on knowing and shift them. Knowing within cognitive science can therefore only advance at crucial points when something in knowing of cognitive science happens, and vice versa. The knowledge loop emphasizes the importance researchers play in researching and defining cognition and knowing.

6. DISCUSSION AND CONCLUSION

In this work, I identify and specify certain features of cognitive science as natural epistemology. This is only the first step in a process to try to characterize how scientific studying of knowing manifests itself. There are a few issues that I will explore in future work. First, the issue of what is epistemological and what is not will have to be addressed. Are the crucial questions in cognitive neuroscience like "What is the complete connectome of the human brain?" [19, p. 173] epistemological questions or not[3]? Second, the knowledge loop and the analysis of the interplay between knowing within cognitive science and knowing of cognitive science will have to be taken further to incorporate stronger ideas of second-order cybernetics, as it seems that this is where the concept of Kuhn's 'essential tension' in cognitive science as natural epistemology is gravitating towards. Hopefully, the endeavor will not completely converge into ideas of second-order cybernetics and offer something new to philosophy of CS; if not, the connection between Kuhn's 'essential tension' and second-order cybernetics was somewhat established.

---

[2] And, interestingly, being used to model game players in more and more complex domains outside of strict-rules based ones as chess (e.g., OpenAI in Dota 2).

[3] Interestingly, some researchers [2] are sceptical of putting neuroscience under the interdisciplinary umbrella of CS for precisely this reason.

As CS is considered as a too loose of a research program by some [20], especially because it lacks its own methods, it seems important to characterize it by its own, unique properties. I believe that the knowledge loop between knowing within cognitive science and knowing of cognitive science may be one such property. What does seem certain is that epistemology works as a sort of glue between the constituents of CS, which gives validity to the latter's interdisciplinary nature.

7. REFERENCES

[1] Fabricius, W. V. 1983. Piaget's theory of knowledge: Its philosophical context. *Hum Dev*. 26, 6 (1983), 325–334.

[2] Ó Nualláin, S. 2002. *The Search For Mind: A new foundation for Cognitive Science*. Intellect Books, Bristol.

[3] Hacker, P. M. S. 2001. An orrery of intentionality. *Lang Commun*. 21, 2 (2001), 119–141.

[4] Kvanvig, J. 2003. *The Value of Knowledge and the Pursuit of Understanding (Cambridge Studies in Philosophy)*. Cambridge University Press, Cambridge.

[5] Alston, W. 1999. Perceptual Knowledge. In *The Blackwell Guide to Epistemology*. J. Greco, E. Sosa, Eds., Blackwell, Oxford, 223–242.

[6] Berkeley, G. 1710/1982. *A treatise concerning the principles of human knowledge*. Hackett. (Original work published in 1710).

[7] Marr, D. 1982. *Vision*. Freeman, San Francisco.

[8] Dennett, D. 1996. *Kinds of Minds*. Basic Books, New York.

[9] Quine, W. V. O. 1969. Epistemology naturalized. In *Ontological Relativity and Other Essays*. W. V. O. Quine, Ed., Columbia University Press, New York, 69–90.

[10] Bateson, G. 1979. *Mind and Nature*. A Necessary Unity. E. P. Dutton, New York.

[11] Keeney, B. P. 1983. *Aesthetics of Change*. Guilford Press, New York.

[12] Luhmann, N. 1996. *Social Systems* (J. Bednarz, Jr. & D. Baecker, Trans.). Stanford University Press, Stanford, CA.

[13] Mancilla, R. 2011. Introduction to Sociocybernetics (Part 1): Third Order Cybernetics and a Basic Framework for Society. *J Sociocybern*. 42, 9 (2011), 35–56.

[14] Kuhn, T. S. 1959. The essential tension. Tradition and innovation in scientific research. In *The third (1959) University of Utah research conference on the identification of scientific talent*. C. W. Taylor, Ed., University of Utah Press, Salt Lake City, 162–174.

[15] Froese, T. 2010. From Cybernetics to Second-Order Cybernetics: A Comparative Analysis of Their Central Ideas. *Construct Found*. 5, 2 (2010), 75–85.

[16] Rumelhart, D. E. and McClelland, J. 1986. On learning the past tense of English verbs. In *Parallel Distributed Processing, Volume 2*. D. E. Rumelhart, J. L. McClelland, Eds., MIT Press, Cambridge, Mass, 216–271.

[17] Varela, F. J., Thompson, E. and Rosch, E. 1991. *The embodied mind: cognitive science and human experience*. MIT Press, Cambridge, MA.

[18] Ward, D., Silverman, D. and Villalobos, M. 2017. Introduction: The varieties of enactivism. *Topoi*. 36, 3 (2017), 365–375.

[19] Adolphs, R. 2015. The unsolved problems of neuroscience. *Trends Cogn Sci*. 19, 4 (2015), 173–175.

[20] Coulter, J. 1982. Theoretical problems of cognitive science. *Inquiry*. 25, 1 (1982), 3–26.

Učinki telesne poze na prag termične bolečine za vroče

Novaković Dušanka
MEi: CogSci, Univerza v Ljubljani
1000 Ljubljana
nd333@protonmail.com

Puh Urška
Zdravstvena fakulteta
Univerza v Ljubljani
Zdravstvena pot 5, 1000 Ljubljana
urska.puh@zf.uni-lj.si

Vidmar Gaj
Univerzitetni rehabilitacijski inštitut
Republike Slovenije – Soča
Linhartova 51, 1000 Ljubljana
IBMI, Medicinska fakulteta
Univerza v Ljubljani
Vrazov trg 2, 1000 Ljubljana
FAMNIT, Univerza na Primorskem
Glagoljaška 8, 6000 Koper
gaj.vidmar@ir-rs.si

Moharić Metka
Univerzitetni rehabilitacijski inštitut
Republike Slovenije – Soča
Linhartova 51, 1000 Ljubljana
metka.moharic@ir-rs.si

Bresjanac Mara
Inštitut za patološko fiziologijo
Medicinska fakulteta
Univerza v Ljubljani
Vrazov trg 2, 1000 Ljubljana
maja.bresjanac@mf.uni-lj.si

POVZETEK

Obvladovanje bolečine je pomembno področje vsakodnevnega življenja, pa tudi klinične in raziskovalne prakse. Ker so novi načini obvladovanja bolečine vedno dobrodošli in je povezava med telesno pozo ter bolečinskim pragom slabo raziskana, smo se namenili preveriti, ali zavzetje naključno dodeljene dominantne ali submisivne poze lahko povzroči akuten dvig ali znižanje praga termične bolečine. Uporabili smo najvišje in najnižje ocenjeno pozo moči iz predhodno izvedene spletne raziskave ter ponovili in nadgradili eksperiment Bohnsove in Wiltermutha [1]. Dobljeni rezultati so pokazali, da je dominantna poza statistično značilno vplivala na dvig praga bolečine za vroče, kar je skladno z izsledki izvirne raziskave [1].

Ključne besede

Neverbalna komunikacija, poza visoke moči (dominantna poza), poza nizke moči (submisivna poza), termični bolečinski prag.

1. UVOD

Moč in nadvlado izražamo poleg besedne govorice tudi skozi nebesedno govorico telesa [2], ki za sporočanje poleg intonacije in barve glasu uporablja telesno držo, geste in obrazno mimiko. Posameznik kaže lastno moč skozi nadrejeno, ekspanzivno držo in asertivne gibe ter podrejenost skozi zaprto, kolabirano držo in skromne gibe.

Nekatere raziskave [1 - 3] nakazujejo, da odnos med nebesedno govorico in počutjem lahko poteka tudi v obratni smeri: namerno zavzeta dominantna poza spodbudno vpliva na dojemanje lastne moči, namerno zavzeta submisivna poza pa ima nanj negativen učinek. Izraz poza moči uporabljamo, ko govorimo o telesnih pozah, ki posnemajo dominantno ali submisivno držo telesa. Glede na njeno sporočilnost in občutenje moči osebe, ki jo vzdržuje, jih imenujemo tudi poze visoke moči (dominantne poze) ali poze nizke moči (submisivne poze), glede na držo telesa ter položaj rok in nog pa tudi ekspanzivne (odprte) in kolabirane (zaprte) poze. Pri dominantnih pozah so udi običajno stran od telesa, hrbtenica je zravnana in brada nekoliko privzdignjena. Pri submisivnih pozah so udi običajno tesno ob trupu ali prekrižani, brada je nemalokrat nekoliko spuščena in drža je sključena [1, 2]. Čeprav številne raziskave iz zadnjih treh desetletij navajajo učinke telesnih poz na različne parametre [4 - 6], vedno več novejših raziskav [3, 7, 8] oporeka izsledkom predhodnih.

Zaradi kontroverznih izsledkov, vedno dobrodošlih novih načinov obvladovanja bolečine in neraziskanega področja morebitnega vpliva telesnih poz na bolečinski prag smo se namenili ponoviti in nadgraditi predhodni eksperiment. V navedeni raziskavi [1] so udeležencem pred in po vzdrževanju ene od poz izmerili bolečinski prag za ishemično bolečino, ki je bila izzvana z manšeto za merjenje arterijskega tlaka. Rezultati so pokazali, da je dominantna poza zvišala bolečinski prag, pri submisivni pozi pa statistično značilnega učinka ni bilo.

Naša raziskava je potekala tako, da smo najprej izbrali najbolj dominantno in najbolj submisivno pozo, nato pa izbrani pozi uporabili pri preverjanju povezave med naključno dodeljeno pozo in spremembo v pragu termične bolečine. Predhodno raziskavo smo nadgradili z večjim vzorcem, zanesljivejšo metodo merjenja ter z zagotovitvijo nepristranskosti in ponovljivosti navodil za zavzemanje poze.

2. METODOLOGIJA

2.1 Udeleženci

V raziskavi je sodelovalo 130 udeležencev, 56 moških in 74 žensk, starih od 19 do 62 let. En udeleženec je bil naknadno izločen zaradi neupoštevanja navodil.

2.2 Pripomočki in merski instrumentarij

Za merjenje praga termične bolečine za vroče smo uporabili napravo Medoc Pathway Pain & Sensory Evaluation System (PSES, model ATS). S pomočjo spletnega vmesnika je udeleženec odgovoril na nekaj vprašanj, si kasneje ogledal naključno dodeljeno fotografijo poze, sledil navodilom za njeno zavzemanje ter jo ob koncu ocenil glede na zaznani občutek moči oziroma nemoči.

2.3 Postopek

Udeleženec je ob prihodu podpisal soglasje o sodelovanju in prek spletnega vmesnika odgovoril na nekaj demografskih vprašanj. Za zagotovitev anonimnosti podatkov mu je bila avtomatsko dodeljena štirimestna koda, pod katero so bili obravnavani vsi

nadaljnji podatki. Sledile so testne meritve s termodo na koži palčne kepe nedominantne roke. Po testni meritvi je sledila prva eksperimentalna meritev na enakem mestu dominantne roke.

Meritve bolečinskega praga za vroče so bile izvedene trikrat, pri čemer je bila izhodiščna temperatura termode vedno 32 °C. Po meritvah je bil udeleženec povabljen pred namizni računalnik. Naročeno mu je bilo, naj po odhodu eksperimentatorja iz prostora na spletnem vmesniku nadaljuje na naslednjo stran, kjer mu bodo podana nadaljnja pisna navodila, fotografija naključno dodeljene poze ter posnetek glasovno predvajanih navodil za zavzetje poze. Pozo je udeleženec izvêdel in v njej vztrajal 3-krat po 20 sekund, pri čemer je glasovno navodilo označilo iztek 20 sekund, začetek in konec premora, ponovno zavzemanje ter konec preizkusa.

Po vrnitvi v prostor je eksperimentator povabil udeleženca k drugi eksperimentalni seriji meritev na predelu palčne kepe dominantne roke, ki so potekale enako kot prej. Vrednosti eksperimentalnih meritev so se shranjevale preko spletnega vmesnika ter bile kasneje povprečene za vsak sklop treh vrednosti.

Po drugi meritvi je udeleženec ocenil, s kolikšnim občutkom moči oziroma nemoči ga je na lestvici od 1 do 11 navdala zavzemana poza.

2.4 Analiza podatkov
Učinek poze smo analizirali z univariatno primerjavo razlik v srednji vrednosti med skupinama. Za analizo podatkov smo uporabili programje IBM SPSS Statistics 23 za okolje Windows.

3. REZULTATI
Sprememba v pragu bolečine za vroče se je med skupinama z dominantno in submisivno pozo pred in po vzdrževanju poz statistično značilno razlikovala (t(127) = -3,11, p = 0,002). Pri udeležencih z dodeljeno dominantno pozo je bila povprečna vrednost temperature praga bolečine za vroče na začetku 44,79 °C (*SD* 3,55 °C) in po vzdrževanju poze 45,97 °C (*SD* 3,51 °C). Pri udeležencih z dodeljeno submisivno pozo je bila povprečna vrednost temperature praga bolečine za vroče pred vzdrževanjem poze 46,40 °C (*SD* 3,14 °C) in po vzdrževanju 46,34 °C (*SD* 3,17 °C).

Udeleženci, ki so vzdrževali submisivno pozo, so na lestvici od 1 do 11 v povprečju poročali o manjšem občutku moči (*M* = 3,8, *SD* = 1,4) v primerjavi z udeleženci, ki so vzdrževali dominantno pozo (*M* = 8,7, *SD* = 1,5; t(127) = -19,06, p < 0,001). Rezultati so primerljivi z ocenami udeležencev v spletni raziskavi za izbrani dve pozi.

4. RAZPRAVA
Namen naše raziskave je bil ugotoviti, ali imata izbrani pozi učinek na dvig oziroma spust praga termične bolečine za vroče. Rezultati so pokazali, da je dominantna poza vplivala na dvig praga za bolečino in da submisivna poza nanj ni imela učinka. Tovrstni izsledki so skladni s predhodno raziskavo, v kateri so prag bolečine merili z izzivanjem ishemične bolečine [1]. Avtorja slednje sta domnevala, da je višji bolečinski prag dominantne poze povezan z višjo aktivacijo mišic med njenim vzdrževanjem ter da to lahko vpliva na fiziološke procese v telesu, ki so povezani z zaznavo bolečinskega praga. V naši raziskavi mišične aktivnosti nismo neposredno merili, zato o tovrstnem učinku ne moremo argumentirano razpravljati. Lahko pa izpostavimo, da je bila izbrana dominantna poza iz naše raziskave v primerjavi s

submisivno razmeroma enostavna za zavzetje, bila je bolj stabilna in je verjetno zahtevala manj vzdrževane mišične aktivnosti. Zato dvomimo, da bi odsotnost vpliva submisivne poze in prisotnost vpliva dominantne poze na prag termične bolečine za vroče lahko pojasnili z večjo aktivnostjo mišic pri vzdrževanju dominantne poze.

Nekateri raziskovalci [2] so poročali, da je vzdrževanje dominantne poze privedlo do zvišanja testosterona v slini in vzdrževanje submisivne poze do njegovega znižanja. Drugi so poročali o višjem bolečinskem pragu po zvišani ravni testosterona pri živalskem modelu [9]. Raziskovalci študije o učinkih testosterona na moške, ki so bili razdeljeni v skupine glede na kronično izpostavljenost eksogenemu ali endogenemu testosteronu, pa so poročali o zmanjšani občutljivosti na taktilne dražljaje pri skupini moških z najvišjo ravnjo testosterona [10]. Iz tovrstnih raziskav bi lahko sklepali o povezavi med vzdrževanjem poz, zvišanjem ravni testosterona in zaznavanjem bolečine, vendar novejše raziskave takšnim povezavam oporekajo [3, 7]. V naši raziskavi se namenoma nismo odločili za meritve koncentracij hormonov. Z nepristransko metodo smo želeli preveriti domnevo o učinku poze na bolečinski prag in se pri tem namenoma izognili dodatnim virom variabilnosti. Poleg tega so bili v raziskavo vključeni udeleženci obeh spolov in med spoloma razlik v spremembi praga nismo zaznali. Čeprav bi nevroendokrini dejavniki lahko prispevali k vplivu, bi za vpogled v njihovo vpletenost morali izvesti novo, kompleksnejšo raziskavo, kjer bi na ustrezen način vrednotili tudi morebitno spreminjanje nevroendokrinih parametrov.

6. VIRI
[1] Bohns, V. K. in Wiltermuth, S. S. 2012. It hurts when I do this (or you do that): Posture and pain tolerance. *Journal of Experimental Social Psychology*. 48,1, 341–345. DOI= https://doi.org/10.1016/j.jesp.2011.05.022.

[2] Carney, D. R., Cuddy, A. J. C. in Yap, A. J. 2010. Power Posing: Brief Nonverbal Displays Affect Neuroendocrine Levels and Risk Tolerance. *Psychological Science*. 21,10, 1363–1368. DOI= https://doi.org/10.1177/0956797610383437.

[3] Ranehill, E., Dreber, A., Johannesson, M., Leiberg, S., Sul, S. in Weber, R. A. 2015. Assessing the Robustness of Power Posing. *Psychological Science*. 26,5, 653–656. DOI= https://doi.org/10.1177/0956797614553946.

[4] Carney, D. R., Cuddy, A. J. C. in Yap, A. J. 2015. Review and Summary of Research on the Embodied Effects of Expansive (vs. Contractive) Nonverbal Displays. *Psychological Science*, 26, 5, 657–663. DOI= https://doi.org/10.1177/0956797614566855.

[5] Nair, S., Sagar, M., Sollers, J., Consedine, N. in Broadbent, E. 2015. Do slumped and upright postures affect stress responses? A randomized trial. *Health Psychology*, 34, 6, 632–641. DOI= https://doi.org/10.1037/hea0000146.

[6] Peña, J. in Chen, M. 2017. Playing with power: Power poses affect enjoyment, presence, controller responsiveness, and arousal when playing natural motion-controlled video games.

Computers in Human Behavior. 71, 428–435. DOI= https://doi.org/10.1016/j.chb.2017.02.019.

[7] Smith, K. M. in Apicella, C. L. 2017. Winners, losers, and posers: The effect of power poses on testosterone and risk-taking following competition. *Hormones and Behavior*. 92, 172–181. DOI= https://doi.org/10.1016/j.yhbeh.2016.11.003.

[8] Garrison, K. E., Tang, D. in Schmeichel, B. J. 2016. Embodying Power. *Social Psychological and Personality Science*, 7, 7, 623–630. DOI= https://doi.org/10.1177/1948550616652209.

[9] Rao, S. S. in Saifi, A. Q. 1981. Effect of testosterone on threshold of pain. *Indian Journal of Physiology and Pharmacology*, 25(4), 387–388.

[10] Burris, A. S., Gracely, R. H., Carter, C. S., Sherins, R. J. in Davidson, J. M. 1991. Testosterone therapy is associated with reduced tactile sensitivity in human males. *Hormones and Behaviour*, 25, 2, 195–205.

The effect of distractors in lateralized change detection working memory task

Anka Slana Ozimič

Mind and Brain Lab, Department of Psychology, Faculty of Arts, University of Ljubljana
Aškerčeva 2
1000 Ljubljana
anka.slana@ff.uni-lj.si

ABSTRACT

The capacity of working memory to maintain visual information is highly limited and varies significantly across individuals. An important research effort is to understand the mechanisms of its limitation, one being the efficient selection of the relevant items from the immediate external environment to encode and maintain in working memory stores, while preventing the irrelevant items to occupy it's capacity. Recently, a series of EEG studies using lateralized change detection task, in which the participants only have to maintain items presented on one visual hemifield, while irrelevant items are also presented on the opposite hemifield, identified a neurophysiological correlate of storage capacity in the form of contralateral delay activity (CDA) wave. Moreover, studies revealed that low-capacity participants maintain irrelevant items along the target items, when both are presented in the same visual hemifield, indicating a reduced ability to filter irrelevant stimuli from visual working memory. These studies, however, do not consider the possibility that participants might also maintain the irrelevant items presented to the opposite visual hemifield. To address this concern, we designed an experiment in which we directly manipulated the presence of distractors in the irrelevant visual hemifield to estimate and control for their effect. Twenty-eight participants took part in a visual working memory experiment in which they were asked to maintain orientation of items presented to the left or right visual hemifield, while the distracting items were either present or absent in the opposite visual hemifield. The results revealed significantly lower estimates of the capacity in the presence vs. absence of distractors, suggesting that participants were not able to ignore the distracting items presented to the opposite visual hemifield, challenging the validity of the estimates of visual working memory capacity in CDA and other studies employing lateralized change detection task.

Keywords

Change detection task, capacity, contralateral delay activity, distractors, visual working memory.

1. INTRODUCTION

In everyday life, access to immediate external environment obtained through different sensory systems is crucial not only to respond to the sources of threat, being one of the basic functions of sensory systems, but also to guide complex mental tasks needed to carry out goal directed behavior. For example, when driving a car, one must have access to the visual environment at all times, in order to follow the driving rules and avoid potential obstacles, such as pedestrians and other vehicles in the traffic. Frequently, however, sensory input from the relevant visual information is either interrupted by short events, such as eye-blinks or saccades, or redirected to irrelevant visual stimuli, such as a phone screen when receiving a call. In such cases, a temporary memory buffer known as visual working memory (VWM) [1], which allows us to actively maintain and integrate relevant visual information in the absence of the external environment, is needed to carry out the tasks at hand.

An important effort of VWM research is to understand the reasons for the highly limited capacity to maintain visual representations, which spans from 3 to 5 meaningful items in young adults, varies significantly across individuals and declines in healthy aging [2]. Recently, a series of EEG studies [5, 3] focused on the neural substrates of VWM processes and identified a neurophysiological index of storage capacity in the form of the slow negative ERP difference wave, known as contralateral delay activity (CDA). Studies have shown that CDA is sensitive to the number of objects maintained in VWM [3]—its amplitude increases as the number of objects maintained in VWM increases, but reaches an asymptote at around 3-4 items (Figure 1F), depending on each individual's memory capacity. Interestingly, the extent of the increase in amplitude in CDA when working memory load is increased from two to four items was found to strongly correlate with individual's VWM capacity and is therefore considered a neurophysiological index of VWM capacity [5]. Moreover, studies have shown [3] that low-capacity participants find it hard to ignore irrelevant distractors when presented either concurrently with or successively to the target items [6]. The authors suggested that irrelevant items burden the limited VWM capacity, leading to lower working memory performance.
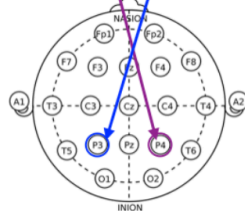
A typical paradigm used to study CDA is lateralized change detection task [5] (Figure 1A), in which participants are presented with a number of items on both sides of the screen, but only have to focus on the side that was previously indicated with a cue (usually an arrow pointing either left or right), encode and maintain the relevant item information (e.g. colors or orientations), while keeping their eye-gaze focused at the fixation point in the middle of the screen. Such tasks evoke an increase in electrical activity above the parietal-occipital cortex of the hemisphere contralateral to the visual hemifield in which the target items were presented (Figure 1C). CDA is computed as the difference in EEG activity between contralateral and ipsilateral hemispheres (Figure 1E). In this way, any nonspecific, task-general bilateral activity such as perceptual responses (Figure 1D), is removed and the remaining signal should be solely related to working memory processes [3].

While this logic seems well reasoned for eliminating task-irrelevant activity, we have identified a potential issue in the experimental design that could lead to false conclusions. When testing VWM capacity, researchers usually don't consider the possibility that participants might be either distracted by or also encode and maintain the items presented to the irrelevant hemifield. Encoding and maintenance of irrelevant distractors was demonstrated in previous studies [6, 3], however, in these studies the irrelevant distractors were present in the same visual hemifield interspersed between relevant targets. In this study we considered the possibility that the distractors significantly affect VWM performance even when presented to the irrelevant visual
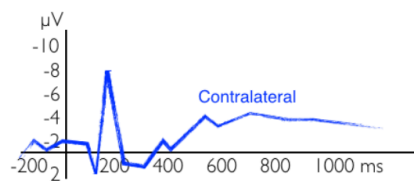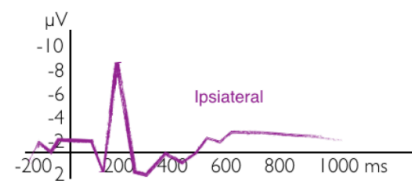
Figure 1: Contralateral delay activity as induced by a change detection task. A: Lateralized change detection task. B: EEG recording of the signal obtained whil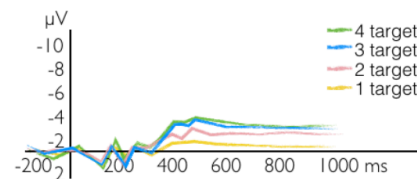e a participant is performing the task. C: EEG response to the target items in the contralateral hemisphere. D: EEG response in the ipsilateral hemisphere. E: Contralateral delay activity (CDA), computed as the difference in EEG signal between contralateral and ipsilateral hemisphere. F: CDA increases as a function of memory load and reaches an asymptote around participant's working memory capacity. Adopted from [3, 5].
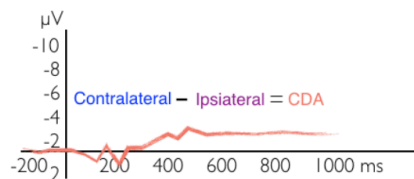
hemifield. In this case, the participants might try to encode items from both hemifields, spreading the VWM resources across both relevant and irrelevant items and reducing the ability to maintain the items from the relevant visual hemifield. If that is the case, in such studies the results would underestimate the VWM capacity for the relevant hemifield.

To address this concern, we designed an experiment in which we directly manipulated the presence of distractors in the irrelevant visual hemifield, which enabled us to estimate and control for their effect.

2. METHOD

2.1 Participants

Twenty-eight students (22 females) aged between 19 and 25 (M = 20.5, SD = 1.6) signed an informed consent to participate in an 1.5-hour experimental session. Five participants were excluded from the analysis due to inadequate performance in the VWM task (accuracy was lower than chance in one of the task conditions).

2.2 Task and procedure

Participants completed multiple trials of the VWM task (Figure 2). On each trial they were shown a brief array of black rectangles of different orientations (*targets*) presented either on one or both sides of the screen and were asked to remember the items presented in either left or right hemifield only, as indicated with an arrow (*initial cue*). Following a brief delay a *second cue*, which matched the initial cue and reminded the participants which objects they will need to base their response on, was presented. After another delay, the *probe* items were shown on both sides of the screen and the participants had to indicate by a button-press, whether there was a change in the orientation of any of the rectangles on the relevant, previously indicated side of the screen, ignoring a possible change on the other side of the screen.

Two main factors were manipulated: *side*, which corresponded to the side of the screen from which the participants had to remember orientation of the presented items (*left, right*) and *distractors,* corresponding to whether the items were presented to both or the relevant visual hemifield only (*distractor* and *no-distractor condition*). Overall we tested 4 experimental conditions: 4 target items presented to the left (L4N) or right (R4N) visual hemifield with no distractors on the opposite
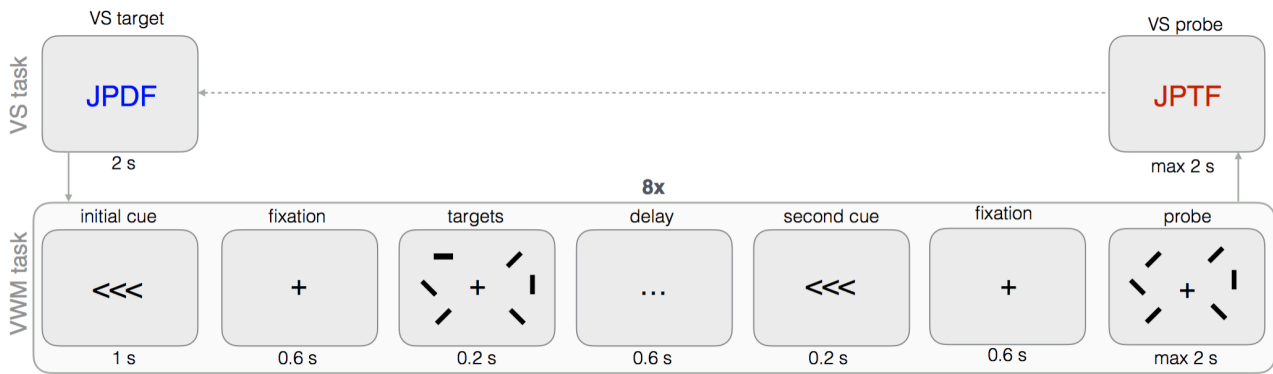
Figure 2: Task progression and trial structure. See main text for full description.

hemifield; and 4 target items presented to the left (L4D) or right (R4D) visual hemifield, with distractors present.

In order to hinder verbal recoding of the visually presented stimuli, in addition to the VWM task, participants also performed a concurrent verbal suppression (VS) task, which was embedded between the trials of the main VWM task (Figure 2). Before the start of the VWM task, a suppression stimulus (VS target) consisting of blue-colored letters was presented for 2 s. Participants were instructed to maintain the letter sequence while performing the VWM task. After every eight VWM trials a VS probe was presented in red and the participants had to indicate by pressing the appropriate key, whether the probe was the same or different from the target.

Data were collected in a sound-isolated room, allowing participants to attend to the task without any interruption. To control for eye-movements, we used an EyeLink 1000 system.

2.3 Data analysis

Statistical analyses are based on estimates of working memory capacity (K). The capacity estimates were computed with Pashler's formula [4]:

$$K = N \frac{h - f}{1 - f},$$

where h and f are the observed *hit* and *false alarm* rates and N is the number of to-be-remembered items.

3. RESULTS

The goal of this study was to assess the effect of distractors in lateralized change detection task on the VWM capacity. A repeated measures ANOVA with within-subject factors *side* (left vs. right) and *distractor* (distractors vs. no-distractors) revealed a main effect of distractor, $F(1, 22) = 19.7$, $p < .001$, $\eta^2 = .06$, reflecting lower capacities in the presence of distractors (see Figure 3). Neither the effect of side, $F(1, 22) = 0.754$, $p = .394$, nor its interaction with distractor, $F(1, 22) = 1.36$, $p = .254$, were significant.

4. DISCUSSION

The aim of this experiment was to identify a potential shortcoming of lateralized change detection tasks used in the CDA studies in assessing VWM capacity, by manipulating the presence of distractors in the irrelevant visual hemifield. Specifically, our concern was, that the presence of distractors could result in reduced estimates of VWM capacity, when defined as the number of remembered target items.
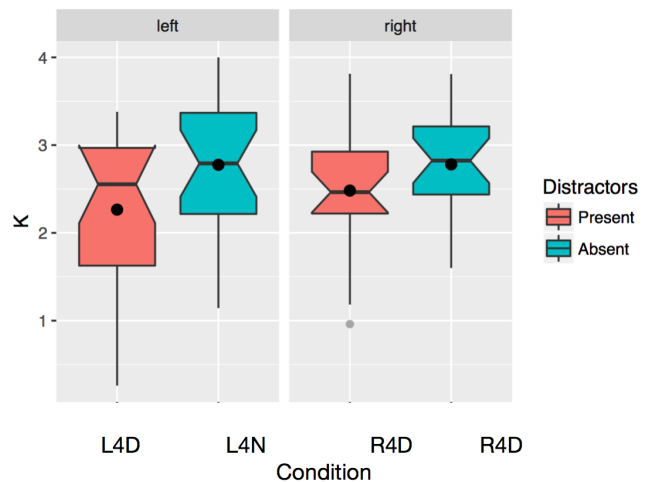

Figure 3: Estimated number of successfully maintained items (K) in the four experimental conditions. L4D: four targets presented to left visual hemifield with distractors present in the right hemifield, R4D: four targets presented to the right visual hemifield with distractors present in the left visual hemifield, L4N and R4N, four targets presented to the left and right visual hemifield, respectively, with no distractors shown in the contralateral visual hemifield. Box plot: The bottom and top of the box plot are Q1 and Q3, the line between them is the median. The lower whisker is Q1 +1.5*IQR. The upper whisker: Q3 + 1.5*IQR. The black circle is the mean.

Whereas the estimated VWM capacity did not differ between left and right hemispheres, the results indeed revealed significantly lower estimates of VWM capacity in the presence vs. absence of distractors. This suggests that participants were not able to ignore the distracting items in the irrelevant contralateral visual hemifield, which resulted in lower estimate of successfully remembered items (K) in the distractor present conditions. These results are in line with prior studies using CDA [3], which have shown that participants find it hard to ignore irrelevant distractors when presented along the target items [6]. These studies, however, assessed the effect of distractors presented in the same visual hemifield, while not taking into consideration the possibility that the distractors can significantly affect working memory performance even when presented to the irrelevant visual hemifield.

The effect of the distractors could be explained either by a reduced ability to encode the target items, or—as suggested by previous research [3]—by leading the participants to encode and maintain also irrelevant distractors, thereby reducing the effective capacity for encoding of the relevant items.

Though confirming the negative effect of the distractors, the results do not enable unequivocal explanation of the mechanism underlying their effect. One possibility is that, similar to the observations of the CDA studies [3], participants fail to ignore the irrelevant stimuli and maintain them along the relevant items occupying limited working memory resources and leaving less of them available to successfully encode and maintain target items. The second possibility is that distractors do not occupy the limited capacity working memory store, but rather disrupt the initial encoding of the relevant items, so that the active maintenance system fails to successfully engage and then sustain the activity of the representations in the first place.

Whereas our experiment does not provide sufficient information to distinguish between these two possibilities, future studies tracking the magnitude of the CDA in the presence and absence of distractors in the contralateral visual hemifield, could help resolve the dilemma.

6. REFERENCES

[1] Baddeley, A., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), Recent Advances in Learning and Motivation (Vol. 8, pp. 47-90). New York: Academic Press.

[2] Cowan, N. (2010). The Magical Mystery Four: How is Working Memory Capacity Limited, and Why? *Current Directions in Psychological Science*, *19*(1), 51–57.

[3] Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. *Neuroscience and Biobehavioral Reviews*, *62*, 100–108.

[4] Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review*, *18*(2), 324–330.

[5] Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428(6984), 748–751.

[6] Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, 438(7067), 500–503.

Zmota luči v hladilniku in značilnosti opazovanja fenomenov z obrobja zavesti

Urban Kordeš
Univerza v Ljubljani
Pedagoška fakulteta
Kadeljeva pl. 16, Ljubljana, Slovenija
+38615892200
urban.kordes@pef.uni-lj.si

Viktorija Lipič
Univerza v Ljubljani
Pedagoška fakulteta
Center za kognitivno znanost
+38631460174
viktorija.lipic@gmail.com

POVZETEK

V prispevku predstavljamo delne rezultate študije, v kateri smo poskušali odgovoriti na vprašanje, kaj se zgodi s fenomeni, ki se nahajajo na obrobju zavedanja, ko vanje usmerimo luč pozornosti. Osredotočamo se na fenomen srži, ki je del širšega doživljanja udejanjanja znanja. V uvodnem poglavju predstavimo tako imenovano zmoto luči v hladilniku – vprašanje, ali so doživljajska stanja, ki se jih zavedamo, ko o doživljanju reflektiramo, prisotna tudi takrat, ko o njih ne reflektiramo. Na podlagi empiričnih podatkov o doživljanju fenomena srži in transformacijah tega doživljanja pod lučjo refleksivnega preiskovanja, poskušamo sklepati o možnostih reflektiranja fenomenov na obrobju in veljavnosti empiričnih podatkov o fenomenih iz predrefleksivne dimenzije zavesti.

Ključne besede

Doživljanje, empirična fenomenologija, refleksija

1. UVOD

Raziskovanje doživljanja, torej subjektivne (oziroma fenomenološke) dimenzije zavesti, predstavlja za uveljavljene paradigme empiričnega raziskovanja kopico problemov. Med njimi je verjetno največji pa tudi epistemološko najbolj zanimiv, problem, ki ga Dehaene in sodelavci [1] imenujejo *zmota luči v hladilniku* (angl. refrigerator light fallacy)[1].

Po Dehaeneju, vsakič, ko se vprašamo o doživljanju, metaforično odpremo hladilnik in pogledamo ali je luč še vedno prižgana. Takoj, ko na določen del doživljajskega polja usmerimo pozornost, lahko opazimo bogato doživljajsko pokrajino, naseljeno z čustvenimi, miselnimi, telesnimi in drugimi fenomeni. O opaženem doživljanju lahko razmišljamo in poročamo. Vemo, da hladilnikova luč ugasne, ko vrata zapremo. Kako je s tem pri doživljanju? So doživljajska stanja, ki se jih zavedamo, ko o njih reflektiramo, prisotna tudi takrat, ko o njih ne reflektiramo? Je akt raziskovanja (refleksivni akt) samo odprtina, skozi katero opazujemo doživljanje takšno *kot je*, ali refleksija nekaj »naredi« doživljanju – prispeva k konstrukciji bogate fenomenološke pokrajine?

Vprašanje, ki si ga zastavljamo, ni novo. Navajajo ga že avtorji kot so Comte [2] in James [3], ki tok zavesti deli na na osrednjo, jasno vsebino in območje, ki to glavno vsebino spremlja – t. i. obrobje. Poskus opazovanja obrobja James [3, str. 244] primerja s snežinko

ujeto v tople roke: »snežinka namreč ob stiku z rokami spremeni svojo kvaliteto – ni več kristal, temveč le še kapljica«.

V nadaljevanju vprašanje naslovimo iz vidika treh raziskovalnih področij: kognitivne znanosti, fenomenološke filozofije in empiričnega raziskovanja doživljanja.

Kognitivna znanstvenica Blackmore, podobno kot Dehaene, uporablja metaforo luči v hladilniku [4]. Blackmoreova meni, da je naš občutek tekočega, neprekinjenega toka zavestnega doživljanja »velika iluzija« [4, str. 26] - vsakič ko se vprašamo »sem zdaj zavesten?« se seveda ponudi pozitiven odgovor. Napaka, ki jo delamo je, da iz tega sklepamo, da smo zavestni tudi, ko tega ne preverjamo. Refleksivni akt oz. *preizkušanje* (angl. probing) je vedno na voljo, zaradi tega ni nikoli lukenj v našem (iluzornem) občutku neprekinjenega toka zavesti.

Po drugi strani, Rosenthal [5] ne vidi nobene iluzije. Njegovo stališče je, da zavestno izkustvo sovpada z zmožnostjo poročanja o njem. Ob tem pa poudarja, da je pogosto prisotna tudi vsebina, ki ni zavestna in o kateri udeleženci ne morejo poročati. Rosenthal zagovarja t.i. plitko konceptualizacijo zavesti [6], katere stališče je, da se je možno zavedati le doživljanja, ki je prej že bilo predmet refleksije. To kar je 'zavestno' je vselej neposredno doživeto med procesom refleksije.

Sodobni fenomenologi, na primer Zahavi [7], so po večini mnenja, da je zavestno izkustvo takoj na voljo, ko se po njem vprašamo, ravno zaradi predrefleksivnega samo-zavedanja. Pojasnjuje, da refleksivno samo-zavedanje temelji na predrefleksivnem samo-zavedanju in da ravno zaradi tega lahko neposredno poročamo o svojem izkustvu. Za ponazoritev predlaga naslednji primer:

»Če sem vključen v določeno zavestno dejavnost, kot je na primer branje zgodbe, moja pozornost ni usmerjena name ali na aktivnost branja, temveč je na zgodbi. Če bi me nekdo v tem trenutku povprašal o mojem početju, bi takoj lahko odgovoril, da sem (že kar nekaj časa) bral. Samo-zavedanje, na podlagi katerega lahko odgovorim na to vprašanje, ni nekaj, kar je bilo pridobljeno v tistem trenutku, temveč je bilo prisotno v sami zavesti že ves čas« [7, str. 21]. Zahavi trdi, da se vselej predrefleksivno zavedamo lastnega doživljanja in se lahko ponavadi takoj odzovemo »brez sklepanja ali opazovanja, če nas kdo vpraša, kaj počnemo, razmišljamo ali vidimo, ali kaj smo občutili neposredno pred vprašanjem« [prav tam].

---

[1] Zmotna domneva, da je luč v hladilniku prižgana tudi, ko so vrata hladilnika zaprta – saj je vedno prižgana, ko odpremo vrata in pogledamo.

Za Zahavija je ključno sklicevanje na trenutek refleksivnega samo-zavedanja, ki razkriva obstoj (vselej prisotnega) predrefleksivnega samo-zavedanja. Jasno zanika možnost, da akt refleksije spremeni doživljanje. Zanj akt refleksije le razpira to, kar je vselej že bilo implicitno (predrefleksivno) zaznano.

1.1. Vloga akta opazovanja v raziskovanju doživljanja

Kot vidimo, filozofska razprava o vplivu akta refleksije razkriva problematiko luči v hladilniku, in z njo povezane različne epistemološke možnosti pogleda na problematiko. Problem vpliva akta opazovanja na opazovano pa ni le teoretski – razumevanje vloge akta refleksije pri opazovanju doživljanja je ključno v empiričnem fenomenološkem raziskovanju, saj sta z njim bistveno povezani vprašanji vloge raziskovalnega aparata pri raziskovanju doživljanja in narave s takšnim raziskovanje dobljenih (fenomenoloških) podatkov.

Vodilni predstavniki prvoosebnega raziskovanja doživljanja priznavajo, da združevanje akta opazovanja z vsebinsko komponento opazovanja predstavlja pomemben izziv za empirično raziskovanje doživljanja [8]. Podobno kot vodilni predstavniki sodobne fenomenološke filozofije, večina empiričnih raziskovalcev doživljanja verjame, da je mogoče pridobiti doživljajske podatke o pristnem, z refleksijo neomadeževanem doživljanju.

Sodobno empirično prvoosebno raziskovanje vključuje celo vrsto metodoloških pristopov opazovanja doživljanja, ki poskušajo ujeti neokrnjeno, pristno (angl. pristine) doživljanje (npr. [9]). Raziskovalci, povezani s tehniko elicitacijskega in mikro-fenomenološkega raziskovanja [10, 11] verjamejo, da njihova metoda omogoča razširitev pozornosti na predrefleksivno dimenzijo doživljanja. Z drugimi besedami, verjamejo, da lahko vedno artikulirajo vsebine, ki se jih udeleženec v trenutku, ko so nastale ni zavedal (oziroma se jih je zavedal samo v predrefleksivnem »obrobju«). Dotični raziskovalni pristop Froese, Gould in Seth [6] imenujejo globoka konceptualizacija zavesti, kar pomeni, da sta raziskovalno dostopni tako predrefleksivna kot refleksivna dimenzija zavesti.

Mikro-fenomenološki intervju poskuša kultivirati posameznikovo refleksijo in retrospekcijsko poročanje o doživljajskih podatkih. Predrefleksivno dimenzijo opisujejo kot skrito, obrobno, skoraj že nezavedno raven zavesti, refleksivno pa kot zaznano vsebino, na katero je usmerjena pozornost [10]. Skozi proces izvedbe intervjuja raziskovalec vodi udeleženca do ponovnega podoživetja izkustva in pozornost udeleženca poskuša preusmeriti z eksplicitnega, refleksivnega dela, na implicitni, predrefleksivni, »obrobni« del doživljanja [12].

Avtorji torej svojo metodo gradijo na predpostavki, da je možno reflektirati doživljanje s predrefleksivne dimenzije. Njihovo razumevanje razdelitve zavesti in vloge akta opazovanja je bistveno za razumevanje podatkov, ki jih producira njihova metoda. Je vera avtorjev v sam obstoj predrefleksivne dimenzije zavesti in v možnost razširitve pozornosti na to dimenzijo upravičena?

2. ŠTUDIJA: FENOMEN SRŽI

Vprašanje veljavnosti in pomena prvoosebnih podatkov seveda presega okvire tega prispevka. Upamo pa, da bo opis rezultatov empirične študije lahko pomenil droben korak na poti k razumevanju fenomenalne zavesti. V tem razdelku predstavljamo izsek iz raziskave, usmerjen na vprašanje, kaj akt opazovanja (refleksije) »naredi« fenomenom, ki se navadno izmikajo fokusu pozornosti.

Predstavljen izsek je del širše študije doživljanja udejanjanja znanja, v kateri smo z uporabo tehnike drugoosebnega globinskega fenomenološkega raziskovanja [13] zasledili fenomen *srži*. Gre za posebno vrsto doživljanja, ki ga Kordeš in Demšar [14] opišeta kot občutek, da imamo na voljo védenje o določeni stvari oziroma občutenje bistva vsebine. Odgovor še ni jasno opredeljen v zavesti – imamo pa občutek, da je na voljo in da vemo kam je treba usmeriti pozornost, da se vsebina razkrije. Bistvena značilnost srži je, da se nahaja na obrobju zavedanja – kot taka je idealni poskusni zajček za raziskovanje vpliva akta refleksije. Kaj se zgodi, če na ta obrobni fenomen posvetimo z lučjo pozornosti? Odgovor na to vprašanje lahko pomaga pri oceni veljavnosti predpostavk raziskovalnega pristopa globoke konceptualizacije zavesti (ki trdi, da je možno raziskovati tudi fenomene, ki se po svoji naravi nahajajo na predrefleksivni dimenziji zavesti).

2.1. Opis metode

Osnovo drugoosebnega globinskega fenomenološkega raziskovanja predstavljata mikro-fenomenološki intervju, opisan v Petitmengin [10] in metodološki obrat, pri čemer udeleženec postane raziskovalec svojega doživljanja. S tem se njegova vloga spremeni iz udeleženca v soraziskovalca. Bistvena značilnost raziskovanja te vrste je, da udeleženec ni več samo vir informacij, ampak postane dejaven sodelavec v študiji. Metodološki obrat udeleženca v soraziskovalca pa je možen samo, če udeleženca zanima raziskovanje in opazovanje lastnega doživljanja. Globinsko raziskovanje postane zares »globinsko« šele, ko udeleženci vzamejo raziskovalno vprašanje za svoje in se raziskovani tematiki posvetijo. Takrat postanejo soraziskovalci [15].

Raziskava je razdeljena v dve fazi. Prva faza vključuje začetne intervjuje in je primarno namenjena urjenju udeležencev v samoopazovanju lastnega doživljanja. Udeleženci prve faze, ki skozi izvedbo začetnih intervjujev ugotovijo, da jih raziskovalna tema in opazovanje doživljanja zanimata, postanejo aktivni soraziskovalci in preidejo v drugo fazo raziskovanja. Na tej točki so soraziskovalci že dovolj izurjeni v samoopazovanju in lahko začnejo samostojno raziskovati doživljanje fenomena v vsakodnevnem življenju. V drugi fazi se tako začenja razvijati dialog v smislu skupnega participatornega sodelovanja med raziskovalcem in soraziskovalcem [13].

2.2. Poskusna teorija

Osredotočamo se na dobljene doživljajske opise, ki se nanašajo na primere, kjer je soraziskovalec poskušal pri sebi opaziti in opredeliti srž v intervjujih.

2.2.1. Neoprijemljivost srži

Soraziskovalci so poročali o velikih težavah pri poskusih reflektiranja na srž. Pogosto so se pojavljali opisi, da se srž izmika pozornosti, da jo poskus opazovanja ne ujame, ali da izginja s poskusom opazovanja. Tipična poročila soraziskovalcev so bila v obliki občutkov, ki jih niso znali konkretizirati. Soraziskovalec

K01[2] je občutek neoprijemljivosti srži opisal kot: »bolj kot ga skušam opisat, bolj ga ne morem«, »bolj ko ga poskušam najt, bolj mi izginja«. V istem intervjuju je v nadaljevanju poročal, da se doživljanje srži začenja pojavljati, ko srži aktivno ne poskuša iskati in opazovati: »ko pa neham tko na en način to iskat, ko pa doživljam, pa se pojavi«.

2.2.1.1. Opisi srži med udejanjanjem znanja
Nekateri soraziskovalci so poročali tudi o srži znotraj udejanjanja znanja – kako se je občutek, da imajo védenje na voljo, kazal med procesom udejanjanja (pojavljanja znanja v zavesti).

Bistven skupni imenovalec soraziskovalcev, ki so poročali o srži med udejanjanjem znanja, je zaznavanje srži na obrobju, ko je pozornost usmerjena na pojavljanje vsebin v zavesti. Srž se kaže kot občutek še neoprijemljivih možnosti oziroma nejasnih vsebin v ozadju. Pozornost je že na določeni vsebini, obenem pa so soraziskovalci zmožni zaznati srž kot oblak potencialnih novih vsebin, v katere se lahko misli nadaljujejo: »Torej poleg tega, da je bil zvok grmenja, je bil hkrati kokr v ozadju blazno ene vsebine, take neoprijemljive. [...] Kot neke ne zares predstave, ampak možnosti. Gist [srž] možnih predstav, kaj bi ta zvok lahko bil. [...] Takrat je bilo samo kot neko ne zares oprijemljivo ozadje« (F01).

Soraziskovalec F01 je poročal o udejanjeni vsebini, ki je bila v središču zavedanja in občutku prisotnosti srži, ki jo je zaznaval na robu »kot neko ne zares oprijemljivo ozadje«, občutek, da je »v backgroundu [ozadju] že nekaj«. Podobno so drugi soraziskovalci govorili o občutku, »kot da se nek proces dogaja zadaj«.

2.2.1.2. Razpiranje srži v vsebino
Naslednji primer soraziskovalca F01 ilustrira, da z usmeritvijo pozornosti na občutek védenja vsebina začne postajati vedno bolj dostopna. Soraziskovalec obenem nima občutka, da sam aktivno povzroča to razpiranje vsebine, ampak da vsebina sama nastaja, vse, kar je potrebno, je ohranjanje pozornosti:

»Da sam usmerim fokus nanj in potem se mi itak razvije samo od sebe. [...] To razpiranje srži. Ko iz srži nastaja neka vsebina. Kjer v bistvu nimam občutka, kokr da jaz to nekaj razpiram, da jaz povzročam to. Jaz samo držim pozornost na tem in pol se zadeva razvije in pol, ker imam pozornost na tem, vem, kaj to je.«

Prepletanje srži in pojavljanja vsebin (udejanjanja znanja) opiše soraziskovalec F01 v naslednjem primeru:

»Kokr, da začne se samo z eno vsebino, enim vedenjem. In pol se to začne razvijat v neko artikulirano misel. Ampak hkrati, ko se to počas artikulira, ko se iz tega razvijajo neke besede, ki jih nek moj notranji glas izgovarja, je hkrati skoz še ta, ta srž tega, kar prihaja pol.«

Opis sledi poteku udejanjanja, ki se začenja s sržjo in nadaljuje v udejanjanje znanja in njegovo artikulacijo. Srž se v podanem primeru opazi kot občutek možnosti vsebine, ki je lahko zanj na voljo, kot prisotnost nečesa, kar še prihaja. Vsebina, ki je na voljo in se lahko pojavi v zavesti, je zaznana na obrobju kot »srž tega, kar prihaja pol« in lahko preide v središče pozornosti, kjer bo ilustrirana in se bo lahko razvila »v neko artikulirano misel«.

3. ZNAČILNOSTI OPAZOVANJA OBROBNIH FENOMENOV
Ugotovitve o srži predstavljajo težavo predpostavki, na kateri temeljijo svoja spoznanja raziskovalni pristopi, ki trdijo, da sta raziskovalno dostopni tako predrefleksivna kot refleksivna dimenzija zavesti (tj. globoka konceptualizacija zavesti [6]).

Naši podatki kažejo, da je refleksivna dimenzija zlahka dostopna poročanju, na predrefleksivno pa ima raziskovanje transformirajoč učinek. Srž so soraziskovalci namreč uspeli zaznati, ampak se je razprla v vsebino ali izginila takoj, ko je prešla v fokus pozornosti. Ko so se soraziskovalci pričeli osredotočati na fenomen srži, ga niso bili več zmožni opisati, saj ga je pozornost izničila ali spremenila. Pokazalo se je, da reflektiranje vodi v spreminjanje fenomena – namesto jasnejšega zaznavanja fenomena srži se zgodi razpiranje vsebin v obliki slik, besed itd.

Reflektiranje izbranega izkustva ne omogoča jasne zaznave ali razširitve pozornosti na predrefleksivno dimenzijo, temveč ustvari novo vrsto doživljanja. To vzbuja dvom v konceptualizacijo zavesti, ki jo zagovarja Petitmengin [10, 12], da je možno pozornost preusmeriti na predrefleksivno doživljanje in le-to opazovati nedotaknjeno s strani refleksije.

Najbližje poskusu opisa predrefleksivnega doživljanja smo prišli z opisi »neoprijemljivosti srži«, kamor spadajo poročila soraziskovalcev, ki so srž zaznali, a se je ta izmikala opazovanju. Srž, zaznano v predrefleksivni dimenziji, so soraziskovalci opisali, kot da izginja, ko poskušajo nanjo usmeriti pozornost. Doživljanje srži pa se spet pojavi, ko soraziskovalec preneha vanjo posegati z aktom refleksije.

Opisi srži med udejanjanjem znanja in razpiranjem srži v vsebino kažejo, da ni jasnega prehoda med fenomenom na obrobju in reflektiranjem o tem fenomenu. Nekateri soraziskovalci so uspeli med reflektiranjem in poročanjem o vsebini na obrobju zaznati fenomen srži, a le, ko je bila pozornost osredotočena na drugih vsebinah v zavesti.

4.VIRI

[1] Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J. in Sergent, C. 2006. Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in cognitive sciences, 10*(5), 204-211.

[2] Overgaard, M. 2006. Introspection in science. *Consciousness and Cognition, 15*, 629-633.

[3] James, W. 1890. *The Principles of Psychology*. Baltimore: Holt.

[4] Blackmore, S. 2002. There Is No Stream of Consciousness. What is all this? What is all this stuff around me; this stream of experiences that I seem to be having all the time?. *Journal of Consciousness Studies, 9*(5-6), 17-28.

[5] Rosenthal, D. 2005. *Consciousness and Mind*. Oxford University Press. New York.

[6] Froese, T., Gould, C. in Seth, A. K. 2011. Validating and calibrating first-and second-person methods in the science of consciousness. *Journal of Consciousness Studies, 18*(2), 38.

---

[2] Šifriranje soraziskovalcev jemljemo iz širše študije, kjer je vsak kodiran z določeno črko in zaporedno številko po principu X01.

[7] Zahavi, D. 2006. *Subjectivity and selfhood: Investigating the first-person perspective*. Cambridge, Mass: MIT Press.

[8] Depraz, N., Varela, F. J. in Vermersch, P. 2003. *On becoming aware*. John Benjamins.

[9] Hurlburt, R. T. 1992. *Sampling Normal and Schizophrenic Inner Experience.* Plenum Press, New York.

[10] Petitmengin, C. 2006. Describing one's subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*, *5*(3-4), 229-269.

[11] Vermersch, P. 1994. *L'entretien d'explicitation*. Paris: ESF.

[12] Petitmengin, C. 2007. Towards the source of thoughts: The gestural and transmodal dimension of lived experience. *Journal of consciousness Studies, 14*(3), 54-82.

[13] Kordeš, U. in Klauser, F. 2016. Second-person in-depth phenomenological inquiry as an approach for studying enaction of beliefs. *Interdisciplinary Description of Complex Systems, 14*(4), 369–377.

[14] Kordeš, U. in Demšar, E. 2018. Excavating belief about past experience: experiential dynamics of the reflective act. *Constructivist Foundations 13*(2): 219–229.

[15] Kordeš, U. in Smrdu, M. 2015. *Osnove kvalitativnega raziskovanja*. Založba Univerze na Primorskem.

Zemljevid prepričanj –
od propozicije do doživljanja

Florian Klauser
Univerza v Ljubljani
Pedagoška fakulteta
Kardeljeva pl. 16, Ljubljana, Slovenija
florian.klauser@pef.uni-lj.si

Urban Kordeš
Univerza v Ljubljani
Pedagoška fakulteta
Kardeljeva pl. 16, Ljubljana, Slovenija
urban.kordes@pef.uni-lj.si

POVZETEK

V prispevku predstavimo nekaj izsledkov iz še odvijajoče se raziskave doživljanja udejanjanja prepričanj, natančneje iz ustvarjanja t.i. zemljevida prepričanj, ki naj bi služil razjasnitvi pojma prepričanje in s prepričanji povezanih ekoloških situacij, za namene empirične raziskave. Na splošno opredelimo pogled analitične filozofije in ga primerjamo z opisi vsakodnevnih situacij, kot jih soraziskovalci beležijo v dnevnik. Na koncu predstavimo, katera vprašanja o naravi prepričanj in možnosti njihovega empiričnega raziskovanja je zemljevid prepričanj razčistil, in katera ostajajo še odprta.

Ključne besede

prepričanje, doživljanje, propozicija, fenomenologija

1. UVOD

V prispevku predstavimo nekaj izsledkov iz še odvijajoče se raziskave doživljanja udejanjanja prepričanj. Celotna raziskava temelji na predpostavki, da se med pozivom po nekem prepričanju in njegovo manifestacijo odvija vrsta vsebinsko bogatih doživljajskih procesov. Naš raziskovalni cilj je te doživljajske procese raziskati. Na primer, če nas nekdo vpraša (pozove po prepričanju), kaj je voda, se odgovor ne pojavi takoj, temveč se morda najprej izzove spomin na učbenik kemije iz osnovne šole, v katerem piše »voda je tekočina brez vonja in okusa«; ali pa morda predstava kozarca vode in pričakovanje, da ob njenem zaužitju ne bo ne vonja ne okusa. Drug primer: v raziskavi je soraziskovalec[1] poročal o tem, da je poslušal predavanje v angleščini kjer je bil omenjen organ »adrenal gland« (slov. nadledvična žleza). Občutil je radovednost po tem, kje se ta organ v telesu nahaja in hkrati tudi zavedanje, da on to vé, četudi se mu 'odgovor' v tem trenutku ne pojavi v zavesti[2]. Z vztrajnim upiranjem pozornosti v ta občutek se mu je le pojavila predstava trupa in občutek določene lokacije v njem.

Osnovna metodologija, ki jo uporabljamo za raziskovanje je intervju, kjer poskušamo pridobiti iz konkretnih primerov čim bolj natančne doživljajske podatke o procesu manifestacije prepričanj (več o metodologiji v Kordeš in Klauser [1]). Pri načrtovanju raziskave smo se tako soočili z vprašanjem, kakšni

---

[1] Izraz soraziskovalec nadomesti sicer uveljavljen izraz udeleženec. Izbiro pojasnjujeta Kordeš in Klauser [1].

[2] Občutek, ki je v tovrstnih doživljajskih procesih tako pogost, da si je zaslužil svoje ime – srž. Več o srži v nekaterih drugih izsledkih raziskave v Kordeš in Demšar [2], glej tudi Kordeš in Klauser [1] ter Kordeš in Lipič [3] v tej publikaciji.

morajo biti ti konkretni primeri, oziroma katere vrste prepričanj je za nas najbolj smiselno raziskovati. Hitro je postalo jasno, da vprašanje izbora vrste prepričanj ni le tehničen problem, ampak da kaže na širšo (in morda usodno) nejasnost: kaže nam, da je – pred empiričnim preverjanjem – dobro razčistiti, kaj sploh mislimo s pojmom *prepričanja*. Prvi poskusi določitve nabora ustreznih vrst prepričanj so namreč pokazali na ogromno širino situacij, ki jih lahko povežemo s prisotnostjo prepričanj. Nabor možnih primerov je tako raznolik, da utemeljuje dvom v smiselnost obravnavanja pojma prepričanje kot jasno opredeljenega (kljub vsakodnevnim intuicijam, skozi katere ga navadno vidimo kot neproblematičnega ali celo očitnega).

Pričujoč prispevek predstavlja del našega poskusa odgovoriti na to vprašanje. Poudarjamo, da iščemo odgovor za potrebe fenomenološke raziskave (torej: kaj so prepričanja z vidika empirične fenomenologije in katera od teh so najustreznejša za raziskovanje?) in ne splošnega odgovora, ki bi predstavljal dokončno zamejitev pojma.

2. SPLOŠNA PREDSTAVITEV POGLEDA ANALITIČNE FILOZOFIJE

V zadnjem stoletju je bilo brez dvoma največ povedanega o prepričanjih v sklopu analitične filozofije, zaradi tega smo predvidevali, da bi to področje moralo služiti kot dobra izhodiščna točka za naše iskanje (za raziskovalno delo) uporabne razmejitve pojma.

Schwitzgebel [4] predstavi pregled splošno sprejetih pogledov in glavnih tokov razmišljanja analitične filozofije o prepričanjih, ki jih opredeli kot *propozicionalno naravnanost* – kot to, ko neko propozicijo vzamemo za resnično. Te propozicionalne naravnanosti pa služijo kot vzrok za vedênje. Schwitzgebel opozori, da sam obstoj propozicionalnih naravnanosti kot vzrokov za delovanje ni splošno sprejeto gledišče. Izpostavi, da je polje razmišljanja o prepričanjih zelo široko in se ukvarja z različnimi povezanimi problematikami [4].

Ena od teh je, na primer, vprašanje *reprezentacij*: ali so prepričanja propozicije, ki so na kakšen način reprezentirane v umu; ali bi bilo bolj primerno misliti prepričanja kot samo stanje, da je dana propozicija reprezentirana v umu. Ne glede na to, ali gre za reprezentacije prepričanj, ali prepričanja kot reprezentacije sveta, je vprašanje, v kakšni obliki se te reprezentacije nahajajo v umu. Ali se pojavljajo v obliki notranjega jezika misli, ali gre morda za bolj kompleksne sisteme zemljevidov. Med tem ko se z reprezentacijami in njihovo strukturo ukvarja struja reprezentacionalistov, drugo strujo, dispozicionalisti (in interpretacionalisti), bolj zanimajo vzorci dejanskega in potencialnega vedênja. Ti t.i. dispozicionalisti (in

interpretacionalisti) analizirajo, kakšno vedênje mora oseba izkazovati, da zanjo lahko rečemo, da je prepričana. Nekateri predstavniki – podstruja liberalnih dispozicionalistov – dopušča tudi zasebne mentalne epizode (doživljanje?) kot relevantne za določanje prepričanja [4].

Pojavi pa se tudi vprašanje o tem, ali prepričanja res obstajajo (kot to zanikajo eliminitavisti), in ali so, tudi če ne obstajajo, prepričanja še vedno smiseln (ali celo nepogrešljiv) koncept za razlaganje vedênja [4].

Schwitzgebel [4] predstavi tudi pogosto sprejeto tipifikacijo prepričanj v *pojavne* (angl. occurent) in *dispozicijske* (angl. dispositional). Če se prepričanje le nahaja v spominu ali t.i. »škatlici prepričanj« je to prepričanje dispozicijsko (i.e. subjekt ima dispozicijo ali nagnjenost k temu, da izrazi to prepričanje). Ko pa subjekt prepričanje prikliče iz spomina za uporabo pri načrtovanju, sklepanju ali aktivnemu izjavljanju, pa je to prepričanje pojavno. Schwitzgebel [4] navede primer subjekta Harryja, ki meni, da so kariraste kravate grde. O tem le redko razmišlja, tako da je to prepričanje večino časa dispozicijsko. Ko pa le privre na površino njegovega uma, bi ga označili kot pojavno.

Schwitzgebel [4] izpostavi tudi podobno delitev na *eksplicitna* in *implicitna* prepričanja. Če je dani subjekt eksplicitno prepričan, da P, se v njegovi »škatlici prepričanj« nahaja reprezentacija z vsebino P. Če pa je subjekt prepričan implicitno, tovrstne raprezentacije tam ni, a bi to prepričanje lahko hitro izpeljal iz že obstoječih. Schwitzgebel [4] to ilustrira s primerom števila planetov. Subjekt S sicer ni eksplicitno prepričan, da je planetov manj kot 9, in manj kot 10, in manj kot 11; a ker ima eksplicitno prepričanje, da je planetov 8, lahko hitro izpelje vsa omenjena prepričanja o tem, da je število planetov manjše od katerega koli števila, ki je večje od 8. Ko za dano prepričanje to stori, prav to prepričanje sicer postane eksplicitno, a pred tem ostaja implicitno.

Iz analitične filozofije pridobljeno znanje nam že pomaga zožiti nabor prepričanj, ki bi bila primerna za empirično raziskovanje – dispozicijskih ali implicitnih prepričanj, na primer, ni mogoče direktno preiskati, saj jih naša na ozaveščanju in eksplikaciji doživljanja temelječa metoda takoj spremeni v pojavna ali eksplicitna.

Pojem (propozicionalna) *naravnanost* vsekakor namiguje o doživljajskem fenomenu (naravnanosti do določenega intencionalnega sistema; [5]). Ni pa jasno kakšne vrste doživljanje naj bi takšna naravnanost bila, še manj pa – če je možno v doživljajskem polju opaziti intencionalni sistem ali vsebino propozicije do katerega/katere smo naravnani. Po drugi strani mnogi analitični avtorji podajajo vedênjske primere za svoje razprave o prepričanjih – nekdo vzame s seboj dežnik, ker je prepričan, da bo deževalo. Predhodni izsledki raziskav [1][2] kažejo na to, da eno in isto vedênje lahko spremlja mnogo različnih doživljanj in obratno – kateri vidik je torej relevanten za raziskovanje? Če pri blagajni sežem v žep po denarnici, katero prepričanje se tu manifestira: prepričanje, da imam v žepu denarnico; prepričanje, da je v denarnici denar; prepričanje, da žep in/ali denarnica obstajata; prepričanje, da svet obstaja; prepričanje, da jaz obstajam; prepričanje, da blagajničar od mene pričakuje plačilo;…? Po večini pregledanih razprav sklepamo, da je odgovor: prepričanje, ki ga v opisanem vedênju želi videti avtor opisa.

Z vidika analitične filozofije je očitno pomembna predvsem vsebina propozicije – tisti element torej, ki ga lahko preiskujemo z

orodji logike. Zgoraj omenjen govor o t.i. implicitnih prepričanjih, na primer, da vedeti, da ima vsak posameznik nešteto prepričanj (da planetov ni 10, da planetov ni 11 itd.). To pomeni, da so (vsaj implicitna prepričanja) konstrukt logike in ne morejo imeti (ne fizične, ne mentalne) reprezentacije – z vidika empirije jih torej ni. V analitičnih razpravah (e.g. [4], [5]) nismo našli jasnejše nastavke za možno empirično preizpraševanje predstavljenih argumentov in ugotovitev.

Ni tudi povsem jasno, kaj predstavlja enoto prepričanja. Analitična filozofija se najpogosteje sklicuje na eno prepričanje kot povezano z eno propozicijo. A, kot že Schwitzgebel [4] sam izpostavi, se lahko dve propozicionalno izraženo identični prepričanji med seboj razlikujeta v kontekstu spremljajočih prepričanj. Tako Ani kot Sanjay sta, na primer, prepričana, da so lososi ribe. A Ani je poleg tega prepričana, da so tudi kiti ribe, med tem ko je Sanjay prepričan, da kiti niso ribe. Sta prepričanji od Ani in od Sanjaya o tem, da so lososi ribe, potem identični? Mnenja se delijo: zagovorniki holizma bi to zanikali, zagovorniki atomizma pa pritrdili [4].

3. ZEMLJEVID PREPRIČANJ

Ker pregled razprav analitične filozofije ni zadovoljil naše potrebe po razmejitvi prepričanj za raziskovalne namene, smo se odločili sami zgraditi zemljevid situacij, povezanih s prepričanji. Namen je bil preiskati čim več čim bolj ekološko veljavnih tipov primerov situacij, ki jih soraziskovalci povezujejo z obstojem prepričanj. Dobljenemu naboru, ki smo ga kasneje poskušali na različne načine razmejiti, pravimo *zemljevid prepričanj*.

Primere za zemljevid smo začeli zbirati s pomočjo skupine soraziskovalcev – kolegov, ki jim je tema znana in so bili pripravljeni opazovati ter beležiti dogodke iz svojega življenja, za katere sodijo, da govorijo o prisotnosti prepričanja. V zemljevid smo dodali tudi nekatere primere iz tekstov analitične filozofije (predvsem iz Schwitzgebel[4]), za katere sklepamo, da so predstavniki posameznega tipa situacij, povezanih s prepričanji.

Vsak primer smo razdelili v tri komponente: situacijsko, ki opisuje širši kontekst dogajanja in/ali vedênje osebe; fenomenalno, ki opisuje doživljanje osebe; in propozicionalno, ki (skladno s tradicijo analitične filozofije) navede vsebino prepričanja v obliki povedi 'S je prepričan, da P'. Če iz primera katera od komponent ni bila razvidna, je njen razdelek ostal prazen. Vsakemu primeru smo tudi dodelili ime, za lažje kasnejše nanašanje.

Nad zbranimi primeri smo izvedli več kategorizacij in tako izdelali zemljevid z grupiranji opisov situacij, ki kažejo na razne vrste prepričanj. Zemljevid smo razdelili na več podzemljevidov, ki predstavljajo različne kategorizacije primerov, kot na primer na implicitna in eksplicitna prepričanja, na primere z opisano fenomenološko komponento in primere brez. Edina razdelitev, ki skoraj v celoti pokrije ves prostor zbranih primerov je delitev na »občutek, artikulacija, vedênje«, ki jo predstavljamo v nadaljevanju.

Delitev je osnovana na treh deskriptorjih. Vsak deskriptor predstavlja eno možno raziskovalno perspektivo na prepričanja: fenomenalno, verbalno in vedênjsko. Fenomenalni deskriptor *občutek* označuje prisotnost »občutka, da tako je.« Verbalni deskriptor *artikulacija* označuje prisotnost artikulacije prepričanja (v propozicionalni obliki) – »izrekanja, da…«. Vedênjski deskriptor *vedênje* pa označuje prisotnost »vedênja, kot da…« – torej obnašanja, ki namiguje na vpletenost določenega

prepričanja[3]. Vsak primer lahko vsebuje od enega do vseh treh deskriptorjev.

Tabela 1. Primer za samo *občutek*:

PRIMER »Bambus«

Vedenjski in kontekstualni opis:	Hodim po cesti in za sekundo pogledam desno v neka drevesa in nato spet drugam.
Fenomenološki opis:	Vidim dve tanjši debli, desno od njiju pa še dve zeleni podolgovati navpični stvari – ena ima vodoravno rumenkasto črto. Ob trenutku kot preblisk védenje, da je to tak debel in visok bambus. Občutek triumfa.
Propozicionalni opis:	

Tabela 2. Primer za samo *artikulacija*:

PRIMER »Robot«

Vedenjski in kontekstualni opis:	Človeku podoben robot ima v sebi shranjeno, v strojnem jeziku, stavek katerega prevod je »kemijska formula za vodo je H2O«. Če ga vprašamo, »iz katerih elementov je sestavljena voda?«, Robot dostopa do omenjenega stavka in ga manipulira tako, da poda odgovor, kot bi ga podal človek.
Fenomenološki opis:	
Propozicionalni opis:	*Robot je prepričan, »da je kemijska formula za vodo H2O.«*

Tabela 3. Primer za samo *vedênje*:

PRIMER »Bakterija«

Vedenjski in kontekstualni opis:	V vodi živeča bakterija ima v svojem telesu magnet, ki se odziva na Zemljino magnetno polje. Na severni polobli magnetni sever kaže navzdol, tako se bakterija, pod vodstvom magneta, giblje navzdol proti globlji vodi in sedimentu, stran od kisika, ki je zanjo škodljiv.
Fenomenološki opis:	
Propozicionalni opis:	*Bakterija je prepričana, »da je v smeri, v katero jo vodi magnet, manj kisika.«*

Tabela 4. Primer za *občutek* in *artikulacija*:

PRIMER »Voda«

Vedenjski in kontekstualni opis:	Janeza nekdo vpraša, kaj je voda. Janez odgovori s »tekočina brez vonja in okusa.«
Fenomenološki opis:	Janezu se pojavi vizualna predstava telesa vode, na katerem so valovi – ki ga pa še nikoli ni videl. Skupaj s to predstavo se pojavi splošen občutek o lastnostih vode, od teh se najbolj izpostavijo, nebesedno, »tekoče«, »brez vonja«, »brez okusa«.
Propozicionalni opis:	*Janez je prepričan, »da je voda tekočina brez vonja in okusa.«*

Tabela 5. Primer za *občutek* in *vedênje*:

PRIMER »Deževalo bo«

Vedenjski in kontekstualni opis:	Janez se odpravlja od doma. Na nebu vidi temne oblake. S seboj vzame dežnik.
Fenomenološki opis:	Ko že stopi ven opazi temne oblake in čuti védenje, da bo deževalo. Nameni se vzeti dežnik.
Propozicionalni opis:	*Janez je prepričan, »da bo deževalo.«*

Tabela 6. Primer za *artikulacija* in *vedênje*:

PRIMER »Filozofski zombi«

Vedenjski in kontekstualni opis:	Filozofski zombi je po vedenju identičen navadni osebi, le da ničesar ne doživlja. Tako kot Janez v prejšnjem primeru se zombi od doma odpravlja z dežnikom. Ko ga vprašamo, zakaj s seboj nosi dežnik, odvrne: »ker bo deževalo.«
Fenomenološki opis:	
Propozicionalni opis:	*Filozofski zombi je prepričan, »da bo deževalo.«*

Tabela 7. Primer za vse tri:

PRIMER »Deževalo bo 2«

Vedenjski in kontekstualni opis:	Janez se od doma odpravlja z dežnikom. Ko ga vprašamo, zakaj ga nosi s seboj, odvrne: »ker bo deževalo.«
Fenomenološki opis:	Janez se ob vprašanju spomni na občutek, ki ga je čutil ob pogledu na temne oblake – na védenje, da bo deževalo.
Propozicionalni opis:	*Janez je prepričan, »da bo deževalo.«*

4. ZAKLJUČEK: V KOLIKŠNI MERI PROPOZICIJE RES ODRAŽAJO PREPRIČANJA

V zaključku navajamo nekaj spoznanj, ki nam jih je prinesel razmislek o možnostih empiričnega raziskovanja prepričanj. Lista je razdrobljena in ni popolna – gre bolj za nastavke oziroma smeri za nadaljnjo razpravo.

Prvo presenečenje, ki smo ga opazili po razmejitvi prostora situacij, povezanih s prepričanji je, da so primeri, ki vključujejo propozicionalno komponento skoraj izključno tisti, ki prihajajo iz teoretskih logov analitične filozofije. Izjemno malo ekološko veljavnih primerov (torej tistih, nabranih iz dnevniških zapiskov soraziskovalcev) ima jasno definirano propozicionalno vsebino. Fenomenalna komponenta tovrstnih primerov največkrat ne vključuje artikuliranih propozicij (e.g. PRIMER »Bambus«). Propozicionalni opis prepričanja bi lahko bil venomer podan šele post hoc in arbitrarno – v skladu s pričakovanji in razumevanjem interpreta.

Zanimivo vprašanje povezano z ugotovitvijo, da ne iz vedênja, ne iz doživljanja, ne moremo enoznačno sklepati na propozicionalno vrednost prepričanja je: kakšno vrednost ima razprava o propozicijah za empirična razmišljanja o duševnosti? Sestavljanje zemljevida prepričanj je pokazalo na izziv fenomenološkega aspekta tovrstnih fenomenov (in morda s tem pojasnilo izogibanje tej temi v analitični filozofiji) – fenomenologija drže (angl. *attitude*), še bolj pa fenomenologija tistega, na kar se drža nanaša (torej propozicije), odpira zelo zapletena vprašanja. Kako se lotiti raziskovanja doživljanja fenomena, ki ga bistveno (izključno?) določa vsebina (oz. propozicionalna vrednost) ni jasno. Kljub temu ocenjujemo, da se odgovor skriva v natančni empirični preiskavi.

Kako nam lahko pomaga dobljeni zemljevid pri nadaljnjem raziskovanju doživljajskih procesov, povezanih z manifestacijo prepričanj? Vsekakor smo jasno razmejili področja, ki niso dosegljiva našim raziskavam (npr. celotno področje z vedênjem določenih situacij). Ostajata še področji artikuliranih prepričanj in tistih situacij, ki jih opredeljuje doživljajski občutek, »da tako je« (oz. – z besedami analitične filozofije – propozicionalna naravnanost).

Pomembno je opaziti, da med tema področjema ni popolnega prekrivanja. Ne le, da mnogokrat propozicionalna naravnanost ni artikulirana (ampak je prisoten le občutek takšnosti); še bolj zanimivo je neujemanje v nasprotno smer: velikokrat je enaka artikulacija pospremljena z različnimi doživljajskimi vsebinami. Vrnimo na primer odgovarjanja na vprašanje »kaj je voda?«, kjer smo predstavili dva možna doživljajska procesa, ki lahko spremljata manifestacijo odgovora »tekočina brez vonja in okusa.« Recimo, da Janez doživlja predstavo kozarca vode in pričakovanje, da ob njenem zaužitju ne bo ne vonja ne okusa; in Jože doživlja spomin na učbenik kemije iz osnovne šole, v katerem piše »voda je tekočina brez vonja in okusa.«

Pripisovanje enake propozicionalne vrednosti obema primeroma (»voda je tekočina brez vonja in okusa«) bi bilo najverjetneje napačno. Če bi že morali pripisati propozicionalno vrednost (zavedajmo se, kako arbitrarno je to dejanje), bi bili dejanskemu stanju še najbližje naslednji propoziciji:

1) Janez je prepričan, da je voda tekočina brez vonja in okusa.

2) Jože je prepričan, da je ustrezen odgovor na vprašanje »kaj je voda?« zapisan v učbeniku kemije iz osnovne šole in se glasi »voda je tekočina brez vonja in okusa.«

V obeh primerih je možno govoriti o propozicionalni naravnanosti. Razlika je, da je prva naravnana na stališče do narave elementa sveta. Druga pa na stališče o izjavi avtoritete (v takšni ali drugačni obliki).

Takšna analiza odpre nova vprašanja, nekatera zelo praktična in povezana z vsakodnevnim življenjem. Na primer vprašanje o naravi in možnostih preverjanja znanja (imajo res vsi, ki opravijo test z enakim rezultatom, enako znanje – oziroma bolje: enako vrsto znanja?).

5. VIRI

[1] Kordeš, U., in Klauser, F. 2016. Second-person in-depth phenomenological inquiry as an approach for studying enaction of beliefs. *Interdisciplinary Description of Complex Systems: INDECS*, 14(4), 369-377.

[2] Kordeš, U., in Demšar, E. 2018. Excavating belief about past experience: experiential dynamics of the reflective act. *Constructivist Foundations*, 13(2): 219-229.

[3] Kordeš, U., in Lipič, V. 2018. Zmota luči v hladilniku in značilnosti opazovanja fenomenov z obrobja zavesti. *Zbornik 21. mednarodne multikonference Informacijska družba – IS 2018.*

[4] Schwitzgebel, E. 2015. Belief. *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition). Edward N. Zalta (ur.), URL=https://plato.stanford.edu/archives/sum2015/entries/belief/.

[5] Fodor, J. A. 1990. *A theory of content and other essays*. The MIT press.

Indeks avtorjev / Author index

Kognitivna znanost / Cognitive Science

Toma Strle, Olga Markič