

Rocznik Kognitywistyczny tom 7/2014

ZESPÓŁ REDAKCYJNY

Aleksander Smywiński-Pohl, Marta Siedlecka, Tomasz Konik, Adrianna Smurzyńska

RECENZENCI

dr Krzysztof Chodasewicz, dr Łucja Domańska, dr Marta Falkowska, dr Mateusz Gola, dr Grzegorz Grzegorzczak, dr Waław Janikowski, dr Tomasz Komendziński, dr Izabela Krejtz, dr Piotr Przybysz, dr Adriana Schetz, dr Piotr Styrkowiec, dr Remigiusz Szczepanowski, dr hab. Michał Wierchoń, dr Dorota Żelechowska

REDAKTOR JĘZYKOWY

Mirosław Ruszkiewicz

KOREKTOR

Agnieszka Toczko-Rak

PROJEKT OKŁADKI

Agnieszka Winciorek

Na okładce wykorzystano grafikę Joanny Bernat

© Copyright by Aleksander Smywiński-Pohl, Marta Siedlecka, Adrianna Smurzyńska

& Wydawnictwo Uniwersytetu Jagiellońskiego

Wydanie I, Kraków 2014

All rights reserved

Niniejszy utwór ani żaden jego fragment nie może być reprodukowany, przetwarzany i rozpowszechniany w jakikolwiek sposób za pomocą urządzeń elektronicznych, mechanicznych, kopiujących, nagrywających i innych oraz nie może być przechowywany w żadnym systemie informatycznym bez uprzedniej pisemnej zgody Wydawcy.

ISSN 1689-927X

e-ISSN 2084-3895

Publikacja dofinansowana przez Uniwersytet Jagielloński ze środków Wydziału Zarządzania i Komunikacji Społecznej

Pierwotną wersją czasopisma jest wersja online publikowana w sposób ciągły w internecie na stronie www.ejournals.eu



www.wuj.pl

Wydawnictwo Uniwersytetu Jagiellońskiego

Redakcja: ul. Michałowskiego 9/2, 31-126 Kraków

tel. 12-631-18-81, 12-631-18-82, tel./fax 12-631-18-83

Dystrybucja: tel. 12-631-01-97, tel./fax 12-631-01-98

tel. kom. 506-006-674, e-mail: sprzedaz@wuj.pl

Konto: PEKAO SA, nr 80 1240 4722 1111 0000 4856 3325

Spis treści

KAROLINA FINC <i>Pułapki i wyzwania w badaniach nad skutecznością komputerowych treningów poznawczych</i>	1
MARCIN KOCULAK, WERONIKA KAŁWAK <i>Opisać świadomość na skali od zera do jeden. Perturbacyjny Indeks Złożoności jako naukowa próba pomiaru świadomości na poziomie indywidualnym</i>	9
TOMASZ KORBAK <i>Dlaczego reprezentacje nie trzymają się modeli dynamicznych?</i>	21
MARTA ŁUKOWSKA <i>Poza ciało: różnice indywidualne w podatności na eksterioryzację</i>	31
MAREK MUSZYŃSKI, JOANNA DURLIK, MARCIN BUKOWSKI <i>„Sunny side is up”, but is the cloudy side down? Znaczenie wymiaru wertykalnego w konceptualizacji emocji</i>	43
MARCIN URBANIAK <i>Rozpoznawanie stanów emocjonalnych jako biologiczne źródło zmysłu moralności</i>	59
KINGA WOŁOSZYN, DARIUSZ ASANOWICZ <i>Niezależność czy współpraca? Zagadnienie interakcji wzbudzeniowego i wykonawczego systemu uwagi</i>	77

Od Redakcji

Oddajemy w Państwa ręce kolejny numer „Rocznika Kognitywistycznego”. Publikujemy w nim artykuły młodych naukowców, którzy opisują badania własne lub prezentują przegląd najnowszych trendów badawczych w nauce o poznaniu. Poruszane w tym numerze zagadnienia dotyczą styku neuronauki, psychologii poznawczej i klinicznej, a także filozofii i językoznawstwa. Niektórzy Autorzy wychodzą poza klasycznie rozumianą kognitywistykę, odwołując się na przykład do koncepcji ucieleśnionego poznania lub dynamicznych modeli umysłu.

Od tego numeru „Rocznik Kognitywistyczny” będzie wydawany w trybie otwartego dostępu, a teksty można będzie ściągnąć ze strony Wydawnictwa Uniwersytetu Jagiellońskiego. Mamy nadzieję, że w ten sposób publikacje będą miały szansę dotrzeć do szerszego grona odbiorców. Chcielibyśmy również zachęcić Państwa do wzięcia udziału w dyskusji z Autorami tekstów i przesyłania komentarzy na temat opublikowanych artykułów – będziemy je publikować na stronie „Rocznika Kognitywistycznego”.

Redakcja „Rocznika Kognitywistycznego”

Pułapki i wyzwania w badaniach nad skutecznością komputerowych treningów poznawczych

Pitfalls and challenges in studies on the effectiveness of computerized cognitive training

Abstract: In recent years, in both science and business, we observe an increased interest in using of various types of software and computer games to improve human cognitive abilities. Paradoxically, together with an increasing number of studies proving the effectiveness of this type of computerized cognitive training, is also growing number of reports about the lack of its beneficial effects. In this situation many doubts are raised by methodology used in training studies, especially poorly matched control groups, and inadequate monitoring of the placebo effect. In this article I will present the most important issues and challenges facing today's research on computerized cognitive training.

Key words: cognitive training, intervention studies, video games, placebo effect

W ostatnich latach zarówno w nauce, jak i w biznesie obserwujemy wzmożone zainteresowanie wykorzystaniem różnego rodzaju programów i gier komputerowych do poprawiania zdolności poznawczych człowieka. Paradoksalnie, wraz ze zwiększającą się liczbą badań potwierdzających skuteczność tego typu komputerowych treningów poznawczych, rośnie również ilość doniesień o braku ich korzystnego wpływu. W tej sytuacji wiele wątpliwości budzi przede wszystkim metodologia stosowana w badaniach treningowych, a zwłaszcza źle dobrane grupy kontrolne oraz niewystarczająca kontrola efektu placebo. W tym artykule przedstawię najważniejsze problemy oraz wyzwania stojące przed współczesnymi badaniami nad komputerowym treningiem poznawczym.

Wprowadzenie

Komputerowe treningi poznawcze stają się obecnie biznesem przynoszącym wielomilionowe dochody. Łatwość dostępu i możliwość wykonywania treningów w domu powodują, że coraz więcej osób ma szansę na polepszenie swoich zdolności umy-

słowych. Do programów potencjalnie poprawiających funkcje poznawcze można zaliczyć zarówno proste zadania przypominające gry przeglądarkowe, jak i bardziej złożony trening z użyciem gier wideo. Większość programów treningowych przyciąga zainteresowanych możliwością usprawnienia pamięci, uwagi, szybkości przetwarzania informacji oraz rozwiązywania problemów [Hardy i Scanlon 2009]. Przykładowo jedna z najbardziej popularnych platform tego typu, amerykańska Lumosity, oferująca „wyzwanie dla mózgu z wykorzystaniem naukowo zaprojektowanego treningu”, przekroczyła obecnie 50 milionów użytkowników, pochodzących ze 182 państw z całego świata [Lumosity 2014]. Strony o podobnym charakterze, obiecujące podobne rezultaty, możemy również znaleźć w Europie. Największą z nich jest obecnie Cogmed [2014], specjalizujący się w treningu pamięci operacyjnej. Jego skuteczność została udowodniona przede wszystkim w grupie dzieci ze zdiagnozowanym ADHD. Skomputeryzowany trening proponowany przez Cogmed, poprzez zwiększenie pojemności pamięci operacyjnej, ma poprawiać różne aspekty uwagi, kontrolę poznawczą, zdolności społeczne, a także umiejętność rozumowania [Cogmed 2014].

Obecnie uwaga znacznej części naukowców skupia się na badaniach gier wideo. Badania eksperymentalne wskazują, że mogą one mieć korzystny wpływ na wiele aspektów naszego funkcjonowania poznawczego [np. Green i Bavelier 2003; Feng, Spence i Pratt 2007; Anguera, Boccanfuso, Rintoul, Al-Hashimi, Faraji, Janowich, Kong, Larraburo, Rolle, Johnston i Gazzaley 2013; Kühn, Gleich, Lorenz, Lindenberger i Gallinat 2014].

Zachwył i optymizm w stosunku do wizji ulepszenia naszych funkcji poznawczych różnymi treningami komputerowymi nie trwały jednak długo. Coraz częściej w badaniach uwidacznia się tendencja do zachowania pewnego sceptycyzmu, w ramach którego wypunktowuje się rażące błędy w metodologii badań potwierdzających skuteczność tego typu ćwiczeń [Boot, Blakely i Simons 2010; Shipstead, Hicks i Engle 2012; Boot, Simons, Stothart i Stutts 2013]. Moja praca ma na celu przedstawienie najważniejszych zarzutów wobec aktualnych badań dotyczących treningu poznawczego. Wraz z krytyką proponuję kierunek rozwoju badań nad komputerowymi treningami poznawczymi.

Specyficzność uczenia się i transfer na inne umiejętności

Największym wyzwaniem dla badań nad treningiem poznawczym jest tworzenie zadań zapewniających transfer trenowanych umiejętności na inne zadania, niezależnie od kontekstu [Schmidt i Bjork 1992]. Chodzi więc o to, aby trening nie przynosił rezultatów izolowanych do jednego zadania i by jego skutki były widoczne na gruncie praktycznym – w codziennym życiu.

Wiele ostatnich badań sugeruje, że proste treningi poznawcze oferowane na stronach takich jak Lumosity i Cogmed poprawiają wyniki tylko i wyłącznie w zadaniach podobnych do tych stosowanych w treningu [Owen, Hampshire, Grahn, Stenton, Dajani, Burns, Howard i Ballard 2010; Shipstead, Hicks i Engle 2012; Hulme i Melby-Lervåg 2012]. Ma to związek z tym, że uczenie się jest wysoce specyficzne w stosun-

ku do obecnych w treningu bodźców, kontekstu i zadania [Bavelier, Green, Pouget i Schrater 2012].

Coraz częściej w badaniach obserwuje się, że znacznie większy transfer na inne umiejętności mogą zapewnić bardziej złożone treningi, przykładowo te z wykorzystaniem gier wideo. Najnowsze badania pokazują, że popularne gry rozrywkowe, jak na przykład gry akcji z gatunku *first person shooter* (powszechnie określane jako „strzelanki”), mogą usprawniać szeroką gamę funkcji poznawczych, nie ograniczając się wyłącznie do poprawy trenowanej umiejętności [np. Green i Bavelier 2003; Feng, Spence i Pratt 2007; Green, Sugarman, Medford, Klobusicky i Bavelier 2012]. Inne badania wskazują na korzyści z treningu z wykorzystaniem gier wideo dla poprawy funkcji poznawczych osób starszych [np. Anguera, Boccanfuso, Rintoul, Al-Hashimi, Faraji, Janowich, Kong, Larraburo, Rolle, Johnston i Gazzaley 2013], a także potwierdzają, że trening z wykorzystaniem gier wideo może prowadzić do strukturalnych zmian w mózgu [np. Kühn, Gleich, Lorenz, Lindenberger i Gallinat 2014].

W celu sprawdzenia generalizacji umiejętności nabytych w trakcie treningów, badacze powinni weryfikować wyniki w testach angażujących inne funkcje niż sam trening, a także monitorować usprawnienie funkcjonowania poznawczego w codziennych sytuacjach.

Co więcej, wykorzystanie technik neuroobrazowania, takich jak fMRI czy EEG, do badania zmian strukturalnych i funkcjonalnych po treningu poznawczym może mieć duży wkład w wyjaśnienie podstawowych mechanizmów obserwowanej poprawy i wnioskowanie o jej ewentualnym transferze na inne funkcje. Na przykład obserwując efektywniejsze działanie sieci neuronalnych w trakcie wykonywania zadania jednego typu, możemy rejestrować, czy tego samego typu usprawnienie zachodzi w trakcie wykonywania innego rodzaju zadania [Bavelier, Achtman, Mani i Föcker 2012]. Obserwowane zmiany strukturalne, takie jak na przykład zwiększenie ilości istoty szarej w różnych obszarach mózgu, może także świadczyć o bardziej fundamentalnych oraz długotrwałych rezultatach treningu [Kühn i in. 2014].

Ustalanie związku przyczynowo-skutkowego

Kolejnym wyzwaniem, jakie stoi przed badaczami zajmującymi się treningiem poznawczym, jest zaprojektowanie eksperymentu w taki sposób, aby z jak największą pewnością można było wnioskować o związku przyczynowo-skutkowym pomiędzy treningiem a obserwowaną poprawą. Coraz częściej badacze jednak zwracają uwagę, że w dużej części eksperymentów wnioski przyczynowo-skutkowe mogą być wyciągane przedwcześnie [Boot, Blakely i Simons 2010; Bisoglio, Michaels, Mervis i Ashinoff 2014].

Wprowadzenie nowego leku na rynek wymaga długotrwałego i ściśle kontrolowanego testowania jego skuteczności. Za „złoty standard” w badaniach klinicznych uznaje się eksperymenty z wykorzystaniem podwójnie ślepej próby. W najprostszym schemacie eksperymentalnym osoby badane są przydzielane losowo do dwóch grup: eksperymentalnej – przyjmującej lek, i kontrolnej – przyjmującej tabletkę składającą się zwykle z samego cukru (placebo). W tego typu eksperymentach ani osoby badane, ani badacze podający lek oraz wykonujący badania przed i po terapii nie zdają sobie

sprawy, jakiego rodzaju interwencja została zastosowana. Pozwala to całkowicie wykluczyć możliwość, że obserwowany rezultat terapii mógłby wynikać z oczekiwań badaczy lub osób badanych [Sainani i Popat 2011; Boot, Simons, Stothart i Stutts 2013].

W badaniach psychologicznych – przykładowo w omawianych w tej pracy badaniach nad skutecznością komputerowych treningów poznawczych – podwójnie ślepa próba jest znacznie trudniejsza do utrzymania. Osoba badana świadomie uczestniczy w określonym treningu poznawczym, więc jego skutki mogą w dużej mierze być uzależnione od jej oczekiwań.

W celu ustalenia przyczynowości wyniki treningu poznawczego także są porównywane z wynikami grupy kontrolnej. Ze względu na trudność w zagwarantowaniu podwójnie ślepej próby to, jakie grupy kontrolne zostaną wybrane do eksperymentu, ma ogromne znaczenie.

Pasywne i aktywne grupy kontrolne

W eksperymentach badających skuteczność komputerowych treningów poznawczych spotykamy się z dwoma rodzajami grup kontrolnych: pasywnymi i aktywnymi. W pasywnej grupie kontrolnej osoba badana nie uczestniczy w treningu ani nie jest poddawana innym oddziaływaniom. Bierze za to udział, podobnie jak osoby z grupy eksperymentalnej, w badaniach funkcji poznawczych na początku badania (pre-test) oraz po czasie, który w grupie eksperymentalnej był przeznaczony na trening (post-test) [Boot, Simons, Stothart i Stutts 2013].

Przykładowo w badaniach Jaeggi i współpracowników [2008], aby odpowiedzieć na pytanie, czy trening pamięci operacyjnej może wpływać korzystnie na inteligencję, losowo przydzielono osoby badane do dwóch grup: (1) eksperymentalnej, w której badani przez określony czas byli poddawani intensywnemu treningowi z wykorzystaniem zadania n-back o zwiększającym się poziomie trudności, oraz (2) kontrolnej, w której badani nie byli poddawani żadnej interwencji, uczestnicząc tylko w pre- i post-tescie. Pasywne grupy kontrolne pojawiły się także w najnowszych badaniach sprawdzających wpływ treningu z wykorzystaniem gier wideo na funkcje poznawcze [Anguera i in. 2013; Kühn i in. 2014]. W badaniu Kühn i współpracowników [2014] u osób po treningu z wykorzystaniem gry *Super Mario 64* stwierdzono – w porównaniu z osobami z pasywnej grupy kontrolnej – zwiększenie ilości istoty szarej w hipokampie, prawej korze przedczołowej oraz mózdzku. Z kolei Anguera i współpracownicy [2013] wykazali rezultaty treningu z wykorzystaniem specjalnie zaprojektowanej gry *NeuroRacer*, z elementem wielozadaniowości, na zmniejszenie kosztu podzielności uwagi, oraz zmiany w aktywności mózgu u osób starszych. Badaniem objęto dwie grupy kontrolne: pasywną (bez kontaktu z grą) i aktywną (z wykorzystaniem gry *NeuroRacer* bez elementu wielozadaniowości) [Anguera i in. 2013].

Według Boota i współpracowników [2013] porównywanie wyników uzyskanych w grupie eksperymentalnej z wynikami pasywnych grup kontrolnych nie wnosi wiele do wyjaśnienia relacji przyczynowo-skutkowej pomiędzy stosowanym treningiem a obserwowaną poprawą. W takim schemacie badawczym mamy do czynienia z wieloma czynnikami dodatkowymi, które mogą odpowiadać za poprawę wyników.

Kontakt społeczny, czas spędzany przed komputerem, motywacja, a także różne oczekiwania osób należących do różnych grup – to tylko niektóre z tych czynników [Boot i in. 2013].

Pasywne grupy kontrolne mogą pomóc w wyjaśnieniu, jaki procent poprawy w testach poznawczych wynika z powtórnego pomiaru z wykorzystaniem tych samych zadań. Kolejnym argumentem za tworzeniem tego typu grup mogą być względy pragmatyczne – wychodzenie z domu, aktywne spędzanie czasu, uczestniczenie w treningu mogą być lepsze dla naszego systemu poznawczego od całkowitej stagnacji i braku inicjatywy. Nie odpowiadamy jednak na wcześniej postawione pytanie: czy określony trening poznawczy naprawdę działa? Czy jest on lepszy niż jakakolwiek inna aktywność umysłowa?

Aby zniwelować problem zbyt dużej liczby zmiennych dodatkowych mogących odpowiadać za poprawę funkcji poznawczych, coraz częściej prowadzi się badania z wykorzystaniem aktywnych grup kontrolnych. Osoby badane w tych grupach także są poddawane oddziaływaniu, jednak znacznie bardziej ograniczonemu, niezwiązanemu ze stawianym pytaniem badawczym. Aktywne grupy kontrolne możemy często znaleźć w badaniach nad wpływem gier akcji na funkcje poznawcze [np. Green i Bavelier 2006; Green i in. 2012]. Rezultat treningu z wykorzystaniem gry akcji (np. *Unreal Tournament*, *Medal of Honor*) porównuje się zazwyczaj z rezultatem treningu z użyciem gier, które elementu akcji nie mają (np. *Tetris*, *The Sims*).

Różnice w oczekiwaniach grupy kontrolnej i eksperymentalnej

Wątpliwości wśród badaczy budzi brak kontroli różnic w oczekiwaniach poprawy funkcji poznawczych u osób należących do tych dwóch grup, np. grających w grę o zwiększonym tempie akcji oraz w grę *The Sims* czy *Tetris*. Nie pozwala to wykluczyć możliwości, że zaobserwowana poprawa może być wynikiem efektu placebo [Boot i in. 2013]. Aby to sprawdzić, Boot i współpracownicy [2013] przeprowadzili prosty eksperyment, w którym podzielono osoby badane na trzy grupy i każdej z nich prezentowano krótkie filmy przedstawiające gry najczęściej stosowane w tego badaniach z wykorzystaniem treningu z grami akcji: *Unreal Tournament* (gra akcji), *Tetris* i *The Sims* (gry kontrolne). Następnie wszystkie badane osoby oglądały filmy pokazujące, jak wykonywane są różne testy poznawcze, a potem musiały ocenić: wykonanie których z nich może ulec usprawnieniu w wyniku grania we wcześniej pokazane gry wideo. Okazało się, że mimo iż respondenci nie uczestniczyli ani w treningu, ani w badaniach, ich oczekiwania wynikające wyłącznie z obserwacji odpowiadały w dużym stopniu dotychczasowym wynikom badań nad treningiem z użyciem gier akcji. Bardzo ważnym wnioskiem, jaki możemy wysnuć z tego badania, jest to, że stosowanie aktywnej grupy kontrolnej wciąż nie pozwala nam wykluczyć efektu placebo. Nie pozwala tym samym na wyciąganie wniosków przyczynowo-skutkowych, jeśli chodzi o rodzaj treningu i obserwowaną poprawę funkcji poznawczych.

Problem ten można znacznie zminimalizować poprzez staranny dobór gry dla aktywnej grupy kontrolnej oraz poprzedzenie właściwych badań nad skutecznością treningu badaniami różnic w oczekiwaniach każdej z grup. Coraz częściej stosowane

są także aktywne grupy kontrolne, w których wykonywane zadanie jest tego samego rodzaju, co te zaproponowane grupie eksperymentalnej – z jedną różnicą: otóż w grupie kontrolnej poziom trudności zadania nie adaptuje się wraz z postęпами osób badanych [np. Klingberg i współpracownicy 2005].

Minimalizacja różnic w oczekiwaniach stanowi duże wyzwanie dla badań nad treningiem poznawczym oraz może przyczynić się do dostarczania znacznie rzetelniejszych wyników w tej dziedzinie [Boot i in. 2013].

Podsumowanie

Badania nad komputerowym treningiem poznawczym niosą duże nadzieje na opracowanie innowacyjnych metod zarówno do celów neurorehabilitacyjnych, jak i usprawniających funkcje poznawcze zdrowego człowieka. Niemniej potrzeba jeszcze większej liczby projektów na dużą skalę, aby móc wyciągać wnioski przyczynowo-skutkowe z prowadzonych badań. Przede wszystkim naukowcy powinni dążyć do projektowania treningów poznawczych, których rezultaty nie ograniczają się wyłącznie do ekranu monitora, lecz ulegają generalizacji na nietrenowane funkcje oraz poprawiają nasze funkcjonowanie w codziennych sytuacjach. Prócz tego, aby wyciągać pełnoprawne wnioski, badacze muszą znacznie bardziej kontrolować stosowaną metodologię, powoływać aktywne grupy kontrolne oraz monitorować oczekiwania osób badanych w zależności od dobieranych treningów poznawczych. Oprócz tego badacze powinni starać się dookreślić, czy komputerowy trening umysłowy może być dla nas bardziej korzystny niż trening tradycyjny, fizyczny lub codzienna aktywność umysłowa. Sprostanie tym wyzwaniom oraz minimalizacja obecnych problemów metodologicznych mogą stanowić kamień milowy w badaniach nad treningami poznawczymi.

BIBLIOGRAFIA

- Anguera J.A., Boccanfuso J., Rintoul J.L., Al-Hashimi O., Faraji F., Janowich J., Kong E., Larraburo Y., Rolle C., Johnston E., Gazzaley A. (2013). *Video game training enhances cognitive control in older adults*. „Nature” 501 (7465), s. 97–101.
- Bavelier D., Achtman R.L., Mani M., Föcker J. (2012). *Neural bases of selective attention in action video game players*. „Vision Research” 61, s. 132–143.
- Bavelier D., Green C.S., Pouget A., Schrater P. (2012). *Brain plasticity through the life span: Learning to learn and action video games*. „Annual Review of Neuroscience”, s. 391–416.
- Bisoglio J., Michaels T.I., Mervis J.E., Ashinoff B.K. (2014). *Cognitive enhancement through action video game training: Great expectations require greater evidence*. „Frontiers in Psychology” 5 (2), s. 136.
- Boot W.R., Blakely D.P., Simons D.J. (2011). *Do action video games improve perception and cognition?* „Frontiers in Psychology” 2, s. 226.
- Boot W.R., Simons D.J., Stothart C., Stutts C. (2013). *The pervasive problem with placebos in psychology why active control groups are not sufficient to rule out placebo effects*. „Perspectives on Psychological Science” 8 (4), s. 445–454.

- Cogmed (2014). Program. URL: <http://www.cogmed.com/program> (data pobrania: 31.03.2014).
- Feng J., Spence I., Pratt J. (2007). *Playing an action video game reduces gender differences in spatial cognition*. „Psychological Science” 18 (10), s. 850–855.
- Green C.S., Bavelier D. (2003). *Action video game modifies visual selective attention*. „Nature” 423, s. 534–537.
- Green C.S., Sugarman M., Medford K., Klobusicky E., Bavelier D. (2012). *The effect of action video game experience on task-switching*. „Computers in Human Behavior” 28 (3), s. 984–994.
- Green C.S., Bavelier D. (2006). *Effect of action video games on the spatial distribution of visuospatial attention*. „Journal of Experimental Psychology: Human Perception and Performance”, s. 1465–1468.
- Hardy J., Scanlon M. (2009). *The Science behind Lumosity*. San Francisco, CA: Lumos Labs.
- Hulme C., Melby-Lervåg M. (2012). *Current evidence does not support the claims made for CogMed working memory training*. „Journal of Applied Research in Memory and Cognition” 1 (3), s. 197–200.
- Jaeggi S.M., Buschkuhl M., Jonides J., Perrig W.J. (2008). *Improving fluid intelligence with training on working memory*. „Proceedings of the National Academy of Sciences of the United States of America” 105, s. 6829–6833.
- Klingberg T., Forssberg H., Westerberg H. (2002). *Training of working memory in children with ADHD*. „Journal of Clinical and Experimental Neuropsychology” 24, s. 781–791.
- Klingberg T. et al. (2005). *Computerized training of working memory in children with ADHD—a randomized, controlled trial*. „Journal of the American Academy of Child & Adolescent Psychiatry” 44 (2), s. 177–186.
- Kühn S., Gleich T., Lorenz R.C., Lindenberger U., Gallinat J. (2014). *Playing Super Mario induces structural brain plasticity: Gray matter changes resulting from training with a commercial video game*. „Molecular Psychiatry” 19, s. 265–271.
- Lumosity (2014). URL: <http://www.lumosity.com/> (data pobrania: 31.03.2014).
- Owen A.M., Hampshire A., Grahn J.A., Stenton R., Dajani S., Burns A.S., Howard J.R., Ballard C.G. (2010). *Putting brain training to the test*. „Nature” 465 (7299), s. 775–778.
- Sainani K.L., Popat R.A. (2011). *Understanding study design*. „American Academy of Physical Medicine and Rehabilitation” 3, s. 573–577.
- Schmidt R.A., Bjork R.A. (1992). *New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training*. „Psychological Science” 3 (4), s. 207–217.
- Shipstead Z., Hicks K.L., Engle R.W. (2012). *Cogmed working memory training: Does the evidence support the claims?* „Journal of Applied Research in Memory and Cognition” 1 (3), s. 185–193.

Opisać świadomość na skali od zera do jeden. Perturbacyjny Indeks Złożoności jako naukowa próba pomiaru świadomości na poziomie indywidualnym

Describing consciousness on scale from zero to one. Perturbational Complexity Index as a scientific attempt on measuring consciousness on an individual level

Abstract: Scientific study of consciousness entered a new era with research focused on finding its neural correlates. The article depicts reasoning underlying this paradigm and controversies it arises using a novel tool designed to measure one's level of consciousness by means of analyzing EEG signal. Perturbational Complexity Index, created by team led by Giulio Tononi, is used to present theoretical debate concerning the possibility of creating an objective consciousness measure. Stemming from reflections on mechanisms of loss of consciousness, authors of the index infer the informational nature of the phenomena, which is said to be demonstrated by experiments with participants under anesthesia, in deep sleep or suffering from disorders of consciousness. The extraordinary capability of the index to measure the level of consciousness on an individual level and its practical applications are discussed in context of possibility of experimental access to subjective experience.

Key words: consciousness, disorders of consciousness, EEG, diagnosis

Zagadnienie świadomości, uznawane za jedno z najtrudniejszych problemów współczesnej nauki, weszło w nową erę dzięki paradygmatowi poszukiwania jej neuronalnych korelatów. Niniejszy artykuł przedstawia tę tematykę na przykładzie najnowszej propozycji zmierzającej do eksperymentalnego pomiaru, a następnie oszacowania poziomu świadomości przy użyciu metod analizujących dane EEG. Bazując na refleksji teoretycznej dotyczącej możliwości stworzenia narzędzia mierzącego świadomość, autorzy prezentują koncepcję Perturbacyjnego Indeksu Złożoności, stworzonego na podstawie teorii Giulio Tononiego. Twórcy wspomnianego Indeksu, wychodząc od badań nad mechanizmami utraty świadomości podczas snu, przekonują o informacyjnej naturze tego fenomenu. Swoje tezy podpierają wynikami badań z udziałem osób w głębokim śnie, poddanych działaniu środków znieczulających czy pacjentów

cierpiących na zaburzenia świadomości. Przekonują przy tym, że pierwszy raz możliwe jest diagnozowanie poziomu świadomości nie tylko w wyniku analizy danych grupowych, ale również na poziomie indywidualnym, czyniąc z indeksu narzędzie do użytku klinicznego i rehabilitacyjnego. Nowatorstwo zaproponowanego narzędzia jest dyskutowane w kontekście eksperymentalnego dostępu do subiektywnego doświadczenia, które według wielu badaczy jest nierozzerwalnie związane z fenomenem świadomości.

Trudny problem ze świadomością

Prestizowy magazyn „Science”, obchodząc w 2005 roku swoje 125-lecie, ogłosił listę najważniejszych pytań naukowych, które zdaniem redakcji powinny zajmować badaczy przez kolejne lata. Wśród dwudziestu pięciu zagadnień drugie miejsce, po pytaniu o naturę wszechświata, zajęła kwestia poszukiwań biologicznych podstaw świadomości [Miller 2005]. To podkreślenie wagi problemu przez redakcję „Science” nie tylko wydaje się wyrazem uznania dla złożoności tego zagadnienia oraz potencjalnych korzyści postępu w pracy nad nim – pokazuje również niezwykłą trudność towarzyszącą wpisaniu badań nad świadomością w krąg dociekań ściśle naukowych (w tym wypadku z zakresu biologii).

Trudności te wynikają między innymi z tego, że przez wieki badania nad naturą świadomości były prowadzone głównie przez filozofów posługujących się odmiennymi pojęciami i sposobami wyjaśniania od tych stosowanych obecnie w naukach empirycznych. Z tego powodu współczesne naukowe próby zrozumienia tego fenomenu muszą sobie radzić z konceptualizacjami problemu świadomości, których nie da się rozstrzygnąć za pomocą metod stosowanych we współczesnej nauce. Kolejną przeszkodą jest obecna w obiegu naukowym wielość teorii świadomości, uniemożliwiająca stworzenie spójnego poglądu na temat jej istoty i funkcji, wspólnego dla wszystkich badaczy. W rezultacie nie ma możliwości jednoznacznego wskazania jej obserwowalnych przejawów, na podstawie których można byłoby budować jej naukową – empiryczną – koncepcję. Jednym ze słynniejszych sformułowań tych wątpliwości było wystąpienie podczas konferencji w Tucson w 1996 roku filozofa Davida Chalmersa [1998]. Przyniósł on wówczas z sobą na mównicę suszarkę do włosów, którą zaprezentował jako „świadomościomierz”. Chalmers zastosował ten żartobliwy wybieg, argumentując, że świadomość ze swej natury leży poza możliwością bezpośredniej obserwacji. Dlatego suszarka do włosów nie gorzej służy do badania świadomości niż na przykład metody neuroobrazowania.

Chalmers w swoim wystąpieniu skrytykował przede wszystkim bezrefleksyjne podejście części badaczy z dziedziny neuronauki do tzw. neuronalnych korelatów świadomości (*neural correlates of consciousness* – NCC). Wskazał dwa zasadnicze problemy w badaniu świadomości za pomocą poszukiwania jej korelatów. Po pierwsze, jest ona badana poprzez pomiar aktywności mózgu – tę z kolei uznaje się za korelat świadomości wyłącznie na mocy arbitralnie przyjętego konstruktu teoretycznego (np. utożsamienie jej z obecnością oscylacji neuronalnych w korze mózgowej w paśmie 40Hz zaproponowane przez Cricka i Kocha, 1990). Dlatego pomiar

świadomości jest w rzeczywistości jedynie pomiarem jej z góry założonego korelatu, a wyciągane z tego pomiaru wnioski w sposób pewny odnoszą się jedynie do samego korelatu, a nie do świadomości. Chalmers uważa to za niedostatecznie satysfakcjonujące, skoro obiektem zainteresowania jest świadomość. Problem ten nie może jego zdaniem zostać przewyższony, ponieważ wynika z ontologicznej autonomiczności świadomości, czyli jej istotnej, jakościowej odmienności od każdego istniejącego lub możliwego fizykalnego zjawiska.

Po drugie, poznawcza niedostępność, czyli brak możliwości jakiegось rodzaju bezpośredniego poznania samej świadomości – jej obserwacji lub intelektualnego ujęcia i opisanie – uniemożliwia skonstruowanie jakiegokolwiek dodatkowego narzędzia, które mogłoby być stosowane jako miara kontrolna metod neuroobrazowania. Taka kontrolna miara miałaby pozwolić na obserwację zależności między aktywnością mózgową (czyli neurokorelatem) a samą świadomością. Chalmers uważa, że NCC mogą być przydatne w poszukiwaniach natury świadomego doświadczenia, wątpi jednak, by można było zrealizować ten cel jedynie przy ich użyciu.

Wskazywane przez filozofa trudności są szczególnie widoczne w badaniach naukowych prowadzonych w nurcie pozytywistycznym, w których nie można korzystać z niektórych narzędzi, np. raczej nie opiera się rozważań i wniosków na introspekcji. Poznanie z perspektywy pierwszoosobowej, uznawane czasem za jedyne udzielające bezpośredniego dostępu do świadomości, nie jest uważane za odpowiednio rygorystyczne postępowanie eksperymentalne. By badać świadomość, konieczne zatem jest ustalenie obserwowalnych przejawów bycia świadomym, na podstawie analizy których dałoby się skonstruować narzędzia do pomiaru świadomości. Dlatego właśnie współczesna nauka o świadomości jest oparta właściwie w całości na badaniu jej korelatów: reakcji motorycznych, zachowań językowych, a także aktywności mózgu. Wydaje się więc, że to, co Chalmers zarzuca nauce o świadomości, jest w nią wpisane na trwałe i tak długo będzie uważane za problem, jak długo subiektywny aspekt świadomości będzie uznawany za istotny.

Zmusza to badaczy do zejścia ze stanowiska skrajnie fizykalistycznego redukcjonizmu właściwego naukom ścisłym na rzecz poszukiwania rozwiązań uwzględniających subiektywne doświadczenie osób badanych. Przekłada się to na nowe rozwiązania metodologiczne, wprowadzające np. elementy introspekcji do procedur eksperymentalnych. Najprostszym sposobem jest oczywiście pomiar tego, czy dana treść została uświadomiona, czy nie – bazuje on na dychotomicznym rozróżnianiu bodźców uświadomianych i nieuświadomianych [np. Holender 1986]. Tak konstruowane obiektywne miary odwołują się do zdolności osób badanych do rozpoznania, co było treścią ich świadomego doświadczenia [Wierzchoń 2013]. Wyniki uzyskuje się za pomocą statystycznej analizy poprawności (wykorzystując głównie teorię detekcji sygnału) lub czasów reakcji, interpretując zmiany w zakresie tych parametrów jako pośrednie miary uświadomienia.

Takie miary całkowicie ignorują subiektywne doświadczenie osób badanych, redukując rozumienie świadomości do warunku koniecznego, by człowiek mógł funkcjonować w środowisku. Współczesne badania pokazują jednak, że możliwe są poprawne zachowania oparte na oddziaływaniach bodźców nieuświadomionych [Dehaene 2008]. Sama reakcja może natomiast zależeć od wielu czynników niezwiązanych bez-

pośrednio z poziomem świadomości, np. uwagi, typu bodźca czy zależności czasowej pomiędzy bodźcem a reakcją [Vermeiren i Cleeremans 2012].

W ciągu ostatnich kilkunastu lat rozwijano narzędzia łączące obiektywny pomiar poprawności wykonania zadania z próbą ilościowego opisu subiektywnego doświadczenia badanych osób. Dokonuje się tego przy użyciu samoopisowych skal, z pomocą których badany ma dostarczyć wyrażonego w sposób ilościowy opisu swojego stanu wewnętrznego. Do wcześniejszych procedur eksperymentalnych dodano element samoopisowy – po każdej zgodnej z instrukcją reakcji dyskryminującej osoba badana jest proszona o ocenę swojego doświadczenia (np. jak dobrze widziała prezentowany bodziec). Tak powstały miary wykorzystujące teorię detekcji sygnału, np. d' drugiego typu, gdzie za postrzegane świadomie uznaje się tylko reakcje poprawne, które dodatkowo zostały wysoko ocenione na skali widzialności [Maniscalco i Lau 2012]. Istnieje także kilka skal subiektywnego doświadczenia odwołujących się do różnych aspektów świadomego przeżycia [ich przegląd: Wierzchoń, Asanowicz, Paulewicz i Cleeremans 2012]. Motywem ich tworzenia była chęć ustalenia, kiedy i czego w rzeczywistości jest świadoma osoba badana, oraz stworzenia miar bardziej czułych na zmiany w świadomym doświadczeniu pojedynczych osób.

Badacze korzystający z tych narzędzi, zarówno obiektywnych, jak i subiektywnych, badają świadomość, którą rozumieją jako subiektywne doświadczenie pewnej określonej stymulacji zewnętrznej. Taka konceptualizacja, konieczna z perspektywy eksperymentalnej, ma pewne wady. Przede wszystkim osoba badana biorąca udział w eksperymencie jest cały czas świadoma. Można przypuszczać, że podczas prób zakwalifikowanych później jako nieświadome badani zdawali sobie sprawę z czegoś (po prostu było to coś innego niż przygotowany przez naukowców bodziec). Szczególnie niekorzystnie mogłoby to wpłynąć na badania, w których wykorzystuje się neuroobrazowanie, ponieważ korelat świadomości byłby obecny we wszystkich warunkach, uniemożliwiając jego wyłowienie metodą kontrastu. Potrzeba innego sposobu myślenia o korelatach świadomości i innych metod ich poszukiwania.

Sen – królewska droga do świadomości

Literatura przedmiotu pełna jest różnych definicji świadomości nawiązujących do wiedzy potocznej, hipotetycznych funkcji, korzeni filogenetycznych czy funkcjonowania mózgu. Pośród tego teoretycznego wielogłosu trudno wyłowić konceptualizacje, które byłyby w stanie objąć całość doświadczeń uznawanych za świadome. Udaje się to w definicji Searle'a [1999], który rozumie świadomość jako „ten stan odczuwania i przytomności, który zaczyna się zwykle, gdy budzimy się rano ze snu [pozbawionego marzeń sennych – przyp. aut.], i trwa tak długo, aż znów zaśniemy wieczorem”. Trudno ukryć, że taka lakoniczność każe podać w wątpliwość przydatność definicji do celów eksperymentalnych. Bardzo dobrze jednak oddaje sposób myślenia o świadomości, odmienny od opisywanego wcześniej i powiązanego z uświadomieniem konkretnych treści.

Definicja Searle'a nie mówi w zasadzie nic o tym, czym świadomość sama w sobie jest. Wskazuje jednak precyzyjnie, kiedy najpewniej jesteśmy świadomości pozbawie-

ni – podczas głębokiego snu bez marzeń sennych. W tej definicji świadomość jest traktowana jako właściwość organizmu, którą w pewnych warunkach się posiada, a w innych traci, a nie jako stopień uświadomienia konkretnych treści. Nie jest już istotna treść doświadczenia, lecz raczej taka organizacja systemu przetwarzającego informacje, która umożliwi pojawienie się subiektywnych doznań. W badaniu tak właśnie rozumianej świadomości nie ma miejsca na porównywanie różnych rodzajów treści doświadczenia, ale raczej są z sobą zestawiane różne momenty, w których zgodnie z naszymi założeniami badany system (w tym wypadku mózg) jest lub nie jest zdolny takie doświadczenia wytworzyć. Wprawdzie tego typu podejście znacznie ogranicza możliwości manipulacji eksperymentalnej (przy braku świadomości prezentowanie uczestnikom bodźców i oczekiwanie odpowiedzi jest daremnym przedsięwzięciem), jednak pozwala ominąć problem precyzyjnego określania odpowiedniości treści subiektywnych stanów wewnętrznych człowieka.

Sen, będąc naturalnym stanem pozbawienia świadomości (w przeciwieństwie do wielu patologicznych przypadków, jak śpiączka, odurzenie substancjami chemicznymi czy zasłabnięcie), wydaje się też najlepszym kandydatem do szukania mózgowych korelatów świadomości. Obiecującą próbą zrealizowania badań opartych na tym pomysle okazała się praca zespołu badawczego skupionego wokół Giulio Tononiego oraz Marcello Massiminiego, o której można przeczytać w ich szeroko cytowanej publikacji [Massimini i in. 2005]. Badacze postawili sobie za cel rozpoznanie mechanizmów odpowiedzialnych za zapadanie człowieka w sen. Mózgi osób przytomnych oraz osób we śnie wystawiono na działanie przezczaszkowej stymulacji magnetycznej (ang. *transcranial magnetic stimulation* – TMS). Procedura ta polega na wielokrotnym wzbudzaniu przepływu prądu przez neurony kory mózgowej za pomocą elektromagnesu umieszczonego przy czaszce osoby badanej. Jednocześnie rejestruje się u takich osób sygnał EEG w celu oceny wpływu stymulacji na funkcjonowanie mózgu.

Badacze poszukiwali różnic w sposobie przesyłania impulsów elektrycznych pomiędzy neuronami pod wpływem zewnętrznie wygenerowanego pobudzenia. Zauważyli, że mózgi osób w głębokim śnie NREM odpowiadały aktywnością o mniejszej sile oraz węższym zasięgu niż u osób przytomnych. Zespół Tononiego interpretował otrzymane wyniki w taki sposób, że istnieje zasadnicza różnica w sposobie przetwarzania informacji pomiędzy systemem świadomym a takim, który świadomości jest pozbawiony. Taki pogląd ma swoje korzenie w ogólnej teorii świadomości zaproponowanej na początku XXI wieku przez Edelmiana i Tononiego [2000]. W teorii tej kładzie się największy nacisk na aspekt informacyjny świadomego doświadczenia. Każde przeżycie, którego doświadczamy, jest według badaczy wysoce informatywnym odróżnieniem danego stanu od innych stanów. Wysoka informatywność oznacza, że świadomość pozwala nam odróżnić dany moment od każdego innego, nawet jeżeli są one do siebie niezwykle podobne. Ta właściwość nie tylko miałaby znaleźć swoje odbicie w pracy mózgu (właściwie nieskończonej liczbie kombinacji aktywacji neuronów), ale przede wszystkim byłaby rezultatem specyficznego sposobu przetwarzania informacji przez sieci neuronalne mózgu, dzięki któremu świadomość mogłaby zaistnieć.

Teoria zintegrowanej informacji

Edelman i Tononi [2000] wskazują na dwie główne cechy, które odróżniają świadome procesy od wszystkich innych zachodzących w ludzkim umyśle. Po pierwsze, każde takie doświadczenie jest zintegrowane – nie występuje jako prosta suma dopływających do nas zewnętrznych i wewnętrznych danych percepcyjnych, ale jest w pewnym sensie emergentne, to znaczy stanowi trwałe połączenie tych danych, jakościowo różne od pojedynczych składników. Łączenie z sobą tak wielu informacji w jedno skutkuje drugą własnością – zróżnicowaniem – która przekłada się na wyjątkowość każdego świadomego doświadczenia i możliwość odróżnienia go od innych, wcześniejszych lub późniejszych. Tononi i jego zespół nie ograniczają się jednak wyłącznie do takiego fenomenologicznego opisu, ale postulują, jakie te właściwości mają mieć odpowiedniki na poziomie neuronalnym i jak można je scharakteryzować za pomocą matematycznych równań przepływu informacji.

Teoria zintegrowanej informacji (ang. *Information Integration Theory*) postuluje, że każdy system, który potrafi zintegrować ogrom informacji sensorycznych w spójne i interpretowalne doświadczenie, jest zdolny do wytworzenia subiektywnego wrażenia [Tononi 2008]. Neurony, zlokalizowane w sieciach obejmujących korę oraz wzgórze, organizują się tymczasowo, tworząc tzw. dynamiczny rdzeń (ang. *dynamic core*) – komórki nerwowe powiązane funkcjonalnie w celu integracji stymulacji [Edelman 2003]. Spójność, czyli poziom wewnętrznego połączenia informacji, można zmierzyć – badacze operacjonalizują ją jako ilość informacji oddziałującej przyczynowo na aktywację pomiędzy dwoma najbardziej oddalonymi od siebie punktami w obrębie wspomnianego funkcjonalnego rdzenia. Odpowiednikiem na poziomie neuronalnym byłaby zdolność połączeń w układzie wzgórzowo-korowym do efektywnego wywołania określonej aktywności innych grup neuronów w sieci. Im bardziej aktywność mózgową byłaby wewnętrznie powiązana (a mniej przypadkowa i lokalna), tym większy byłby w danym systemie poziom świadomości [Tononi 2004, 2010].

Rozumienie świadomości jako specyficznego trybu aktywności połączeń o dalekim zasięgu w układzie wzgórzowo-korowym czyni z niej własność względnie niezależną od innych cech systemu przetwarzającego informacje (np. metabolizmu czy pojawiania się konkretnych oscylacji). Obecność funkcjonujących w ten sposób sieci neuronalnych jest zdaniem Tononiego warunkiem koniecznym do pojawienia się świadomości, choć może nie być wystarczającym. Kluczowe jest tu założenie, że brak oznak działania wspomnianej sieci wzgórzowo-korowej powinien charakteryzować wszystkie przypadki, w których człowiek jest pozbawiony świadomości.

Teoria Tononiego wykracza poza kontekst danego eksperymentu, określoną modalność zmysłową czy treściowy charakter dostarczanej stymulacji. Aspiruje do określenia, jak – idealnie – powinien funkcjonować system, aby mogła zaistnieć świadomość, oraz proponuje narzędzia, które powinny pozwolić ocenić podobieństwo interesującego nas systemu (np. mózgu) do takiego idealnego modelu. Dzięki temu badacze nie muszą przejmować się doborem bodźców, tym, czy są zauważalne, chęcią współpracy osób badanych czy zmianami poziomu uwagi. Mózg powinien zachować swój specyficzny dla świadomego systemu sposób organizacji i przetwarzania informacji niezależnie od tego, czego w danym momencie doświadcza osoba badana.

Wykorzystując stymulację TMS oraz narzędzia matematycznej analizy sygnału EEG, powinno dać się zaobserwować różnicę w sposobie przekazywania sobie stymulacji neuronów osób mających w danym momencie świadome doznania i tych, którzy są ich pozbawieni. Hipoteza ta została przetestowana w serii kolejnych eksperymentów dotyczących zarówno naturalnych, jak i indukowanych stanów braku świadomości. Powtórzono pierwotne badanie na śpiących osobach i uzyskano wyniki potwierdzające wcześniejsze obserwacje dotyczące stanu głębokiego snu [Massimini i in. 2007] oraz odkryto podobne zależności u osób poddanych ogólnemu znieczuleniu [Alkire, Hudetz i Tononi 2008; Boly i in. 2012; Ferrarelli i in. 2010], u pacjentów znajdujących się w stanie wegetatywnym lub w stanie minimalnej świadomości [Rosanova i in. 2012], a także w fizjologicznej bądź farmakologicznej śpiączce [Noirhomme i in. 2010]. Jednocześnie podobne badania podczas snu REM (snu, w czasie którego występują marzenia sennie), kiedy aktywność kory jest wzmożona, nie wykazały żadnych różnic pod względem funkcjonalnych połączeń w porównaniu z osobami przytomnymi [Massimini i in. 2010].

Tononi wraz z zespołem przekonują, że takie wyniki wspierają hipotezę o istnieniu jednego mechanizmu umożliwiającego pojawienie się świadomości, opartego na połączeniach o dalekim zasięgu pomiędzy wzgórzem i różnymi obszarami kory mózgowej. Załamanie się efektywności tych połączeń miałyby zawsze skutkować utratą świadomości, zarówno podczas naturalnych procesów, jak i głębokiego snu, oraz w przypadku wydarzeń nienaturalnych, takich jak śpiączka, stan wegetatywny czy znieczulenie ogólne. Ostatnie działania badaczy zmierzają ponadto do opracowania nowej miary świadomości, która wykorzystywałaby wymuszoną aktywność mózgu (np. dzięki stymulacji TMS) do szacowania wskaźnika, który miałby charakteryzować poziom świadomości u pojedynczego człowieka w momencie badania.

Psychofizjologiczne miary świadomości

Idea miary wykorzystującej bioelektryczną aktywność ludzkiego mózgu nie jest jednak nowa. Literatura przedmiotu obfituje w propozycje przeróżnych wskaźników, które miałyby służyć do wnioskowania na temat procesów psychicznych, w tym świadomości. Najbardziej znanym pomysłem tego typu jest klasyczne wyróżnienie stadiów snu (obecnie trzy fazy NREM i faza REM) i czuwania na podstawie procentowego składu poszczególnych typów fal mózgowych w zapisie EEG [Kales i Rechtschaffen 1968]. Choć kryteria dotyczą statystycznego rozkładu częstotliwości pracy właściwie całego mózgu, posiadają one korelaty behawioralne powiązane z konkretnymi funkcjami (np. jedynie podczas fazy NREM3 osoba śpiąca nie reaguje na własne imię).

Podobne zabiegi analizujące różne aspekty aktywności elektrycznej mózgu stosowano już w przeszłości do uchwycenia zmian zachodzących w świadomości. Casali i inni [2013] dzielą ten zbiór istniejących w literaturze markerów neurofizjologicznych na dwie zasadnicze grupy. Pierwszy typ stanowią metody próbujące szacować „bogactwo” rejestrowanego sygnału lub ilość niesionej w nim informacji. Oceniają one, czy rejestrowana aktywność kory (zarówno spontaniczna, jak i wywołana stymulacją) jest wystarczająco różnicująca (np. czy sygnał EEG ulega zmianie w zależ-

ności od typu bodźca, czy przeciwnie – kora reaguje stereotypowo na szerokie spektrum stymulacji). Mogą też oceniać złożoność sygnału, np. jak wiele częstotliwości fal można w nim wyodrębnić. Na takiej analizie spektrum bazuje popularny indeks bispektralny [Rosow i Manberg 2001], służący anestezjologom do pomiaru poziomu sedacji pacjenta podczas zabiegów medycznych. Innymi wskaźnikami tego typu może być np. szacowanie entropii [Pincus, Gladstone i Ehrenkranz 1991] lub entropii widma sygnału [Johnson 1987], których zadaniem jest przybliżenie, na ile rejestrowana aktywność zdradza faktyczne przetwarzanie informacji, a nie jedynie redundantne przekazywanie tej samej treści.

Druga grupa metod wykorzystuje przestrzenną analizę rozprzestrzeniania się sygnału lub synchronizację półkul mózgowych. Należy do niej m.in. badanie późnych potencjałów wywołanych [Plourde i Picton 1991] czy propagacji stymulacji elektrycznej wywołanej za pomocą TMS [Ragazzoni i in. 2013]. Cytowani badacze śledzili, jak aktywność wywołana stymulacją przemieszcza się po mózgu, wiążąc większą sprawność kory (oraz idący za tym wyższy poziom świadomości) ze zdolnością do propagacji sygnału do dalszych obszarów. Do tej grupy zaliczają się również próby szacujące poziom synchronizacji różnych obszarów mózgu podczas analizy różnej stymulacji. Dobrym przykładem są tu miary łączliwości korowej [Boly i in. 2011], szukające związków przyczynowych pomiędzy aktywacją w różnych rejonach kory. Niestety w próbach tych, nawet jeśli można było rozróżnić poziomy świadomości, było to możliwe jedynie na poziomie analiz danych grupowych – nie dało się poprawnie ocenić stanu pojedynczego człowieka (co jest konieczne w przypadku diagnozowania osób z zaburzeniami świadomości).

Perturbacyjny Indeks Złożoności

Zaproponowany przez badaczy skupionych wokół Tononiego Perturbacyjny Indeks Złożoności (ang. *Perturbational Complexity Index* – PCI) ma w założeniu zawierać najlepsze cechy dostępnych do tej pory metod, a być pozbawionym ich trudności i niedostatków. Autorzy PCI wskazują na zasadnicze różnice pomiędzy ich indeksem a innymi miarami i wcześniejszymi próbami podejmowanymi przez ich własny zespół [Sarasso i in. 2014].

Przede wszystkim PCI jest ugruntowany w teorii zintegrowanej informacji – wykorzystuje pomiar kluczowych charakterystyk systemu zdolnego do podtrzymania świadomości. Istotne, zdaniem Tononiego, fenomenologiczne cechy świadomości, czyli zróżnicowanie i zintegrowanie, mają być odzwierciedlane w charakterystyce sygnału EEG u osoby, której mózg jest zewnątrznie stymulowany za pomocą TMS. Wspomniane wcześniej dwie grupy metod psychofizjologicznych mierzą, według Tononiego, dwie podstawowe własności świadomego systemu – wysokie zróżnicowanie oraz integrację przetwarzanych informacji.

Zróżnicowanie odpowiadałoby „bogactwu” informacyjnemu zawartemu w rejestrowanym sygnale. Można nim zatem mierzyć redundantność aktywności kory (im byłaby mniejsza, tym zawartość informacyjna byłaby większa). Integracja przejawiałaby się natomiast w zdolności mózgu do przekazywania aktywacji pomiędzy odległy-

mi rejonami kory oraz do synchronizacji pracy w celu szybszego i skuteczniejszego przetworzenia otrzymanych informacji.

Procedura eksperymentalna polegała na stymulacji kory osoby badanej za pomocą TMS wraz z jednoczesną rejestracją aktywności mózgu poprzez EEG. W celu zwiększenia skuteczności procedur statystycznych zarówno stymulacja TMS, jak i pomiar odpowiedzi mózgowej wykonywane były wielokrotnie, aby następnie uśrednić wyniki otrzymane z każdej próby [Casali i in. 2013]. Analizie poddawano pierwsze 300ms surowego zapisu EEG po każdym uderzeniu impulsu z TMS, algorytmicznie przetwarzając uzyskany sygnał w binarną matrycę (gdzie na osi x zostały umieszczone punkty czasowe pomiaru, natomiast na osi y – wszystkie elektrody rejestrujące sygnał). Następnie używając algorytmu Lempela-Ziva, za pomocą którego szacuje się złożoność sygnału, dokonywano kompresji otrzymanej matrycy. Użycie algorytmu kompresji informacji powodowało, że im większe fragmenty wspomnianej matrycy były do siebie podobne, tym niższy współczynnik złożoności był jej przypisywany. Ostatni krok stanowiła normalizacja uzyskanego wyniku, której rezultatem była pojedyncza wartość liczbowa pomiędzy zero a jeden. Oznacza to tyle, że bardziej skomplikowany wzór aktywności (wskazujący, jak mówi teoria Tononiego, na wyższy poziom świadomości) otrzymywał wynik bliższy jedynce.

Walidację metody postanowiono przeprowadzić na grupach ochotników oraz różnych grupach pacjentów, co do których można było mieć najwyższą możliwą pewność, że ich poziom świadomości odbiega od zdrowych przytomnych ludzi [Casali i in. 2013]. Okazało się, że na podstawie pomiaru PCI można odróżnić grupy osób przytomnych (indeks pomiędzy 0,44 a 0,67) od tych, które pogrążone były w głębokim śnie (indeks wynosił od 0,18 do 0,28). Autorzy uważają, że różnice pomiędzy grupami są na tyle znaczące, że można mieć nadzieję, iż PCI będzie użyteczny w diagnozowaniu poziomu świadomości u pojedynczych osób.

Wniosek ten jest wzmacniany przez kolejne wyniki – tym razem uzyskane dla osób zdrowych będących pod wpływem różnych anestetyków: midazolamu (PCI od 0,18 do 0,28), propofolu (PCI 0,13–0,30) czy xenonu (PCI 0,12–0,31). Ostatnią przebadaną grupą były osoby z diagnozą zaburzeń świadomości o różnym stopniu nasilenia. Tutaj również indeks PCI potwierdził przewidywania co do swojej czułości, gdy wynik dla osób w stanie wegetatywnym (PCI 0,12–0,31) był podobny do wyniku dla osób w głębokim śnie i w znieczuleniu ogólnym. Osoby badane będące w stanie minimalnej świadomości uzyskiwały wyniki wyższe (PCI 0,32–0,49), jednak ciągle istotnie poniżej wyników zdrowych i przytomnych ochotników.

Jak słusznie zauważają w swoim komentarzu Sitt, King, Naccache i Dehaene [2013], opracowana przez Casaliego i współpracowników metoda wydaje się charakteryzować nieosiągniętą nigdy wcześniej czułością na zmiany w świadomości, co więcej – na poziomie analizy danych indywidualnych, a nie tylko grupowych. Nowe narzędzie może przynieść szczególne korzyści dla diagnozowania pacjentów z zaburzeniami świadomości, poprzez precyzyjny pomiar zdolności kory mózgowej do utrzymywania osoby w stanie świadomości. Należy jednak mieć na względzie, że badanie zostało przeprowadzone na stosunkowo niewielkiej próbie osób i konieczne byłyby kliniczne testy efektywności tej metody. Nie jest ona również pozbawiona potencjalnych technicznych trudności – żeby przeprowadzić badanie, trzeba wygenerować setki impul-

sów TMS w celu otrzymania stabilnego wskazania PCI. Tak długa procedura eksperymentalna może się okazać męcząca i nieprzyjemna dla osób, które są jej poddawane. Ponadto jej wyniki mogą być zakłócone przez dynamicznie zmieniający się poziom świadomości, np. u pacjentów w stanie minimalnej świadomości.

Czy PCI to już świadomościomierz?

Indeks PCI wydaje się narzędziem najbardziej z dotychczasowych zbliżonym do ideału uniwersalnego świadomościomierza, z którego stroił sobie żarty Chalmers w słynnym wystąpieniu. I choć nie wystarczy tego narzędzia po prostu wycelować, jak suszarkę Chalmersa, w stronę człowieka, by od razu odczytać poziom jego świadomości, to jednak możliwość przełożenia złożonej aktywności neuronalnej na jeden, zaskakująco czuły, parametr wydaje się pociągająca. Skuteczność indeksu PCI w określaniu poziomu świadomości byłaby również najmocniejszym dowodem poprawności teorii świadomości rozwijanej przez Tononiego i jego zespół. Istotna jest jednak również jego krytyka. Sami jego autorzy wskazują na zintegrowane i zróżnicowane przetwarzanie informacji jako warunek konieczny, lecz nie zawsze wystarczający do zaistnienia świadomości.

Oczywiście stosuje się do niego podstawowy zarzut filozofa – nie mierzy bezpośrednio świadomego doświadczenia osoby, lecz jedynie konkretną psychofizjologiczną aktywność mózgu, która na podstawie przyjętej z góry teorii uznawana jest za korelat świadomości. Zarzut ten można pociągnąć dalej: wydaje się, że indeks PCI jako narzędzie oraz podbudowująca go teoria, żeby przyczynić się do badań nad świadomością – których podstawowym celem jest wykazanie, czy świadomość istnieje, a jeśli tak, to czym jest – muszą istnienie świadomości założyć na wstępie, by takie badanie z ich pomocą w ogóle było możliwe. Co więcej, zakładając ją, w zasadzie jej nie definiuje, tylko wskazuje: teraz świadomość jest, a teraz nie. Taka ostensywna definicja świadomości nie wykracza poza to, co już wiadomo i jako punkt wyjścia w badaniu, którego w zasadzie nie da się przekroczyć, znacznie utrudnia zdobycie nowej i wartościowej wiedzy.

Wydaje się, że korzystając z tego narzędzia, nie potrafimy odpowiedzieć na teoretyczne i filozoficzne pytania dotyczące istnienia i natury świadomości: ani czym jest, ani czy w ogóle istnieje. Pytania, na które możemy odpowiedzieć, są pytaniami raczej praktycznymi i technicznymi. Jednak potencjalne korzyści teoretyczne i badawcze płynące z posiadania takiego narzędzia, a także z jego krytyki, schodzą na dalszy plan, gdy pomyśli się o jego potencjalnych walorach praktycznych. Gdyby indeks PCI dowiódł swojej przydatności, zastąpiłby prawdopodobnie dotychczasowe metody diagnozowania pacjentów z zaburzeniami świadomości. Dałoby to większą szansę na prawidłowe rozpoznanie stanu takiej osoby, przez to na odpowiednią rehabilitację, a w niektórych przypadkach – szczęśliwe odzyskanie pełni świadomości.

BIBLIOGRAFIA

- Alkire M.T., Hudetz A.G., Tononi G. (2008). *Consciousness and anesthesia*. „Science” 322 (5903), s. 876–880.
- Boly M., Garrido M.I., Gosseries O., Bruno M.-A., Boveroux P., Schnakers C., Friston K. (2011). *Preserved feedforward but impaired top-down processes in the vegetative state*. „Science” 332 (6031), s. 858–862.
- Boly M., Moran R., Murphy M., Boveroux P., Bruno M.-A., Noirhomme Q., Friston K. (2012). *Connectivity changes underlying spectral EEG changes during propofol-induced loss of consciousness*. „The Journal of Neuroscience: the official journal of the Society for Neuroscience” 32 (20), s. 7082–7090.
- Casali A.G., Gosseries O., Rosanova M., Boly M., Sarasso S., Casali K.R., Massimini M. (2013). *A theoretically based index of consciousness independent of sensory processing and behavior*. „Science Translational Medicine” 5 (198).
- Chalmers D.J. (1998). *On the search for the neural correlate of consciousness* [w:] S.R. Hameroff, A.V. Kaszniak, A. Scott (red.), *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates* (s. 219–229). Cambridge: MIT Press.
- Crick F., Koch C. (1990). *Towards a neurobiological theory of consciousness*. „Seminars in the Neurosciences” 2, s. 263–275.
- Dehaene S. (2008). *Distinct forms of evidence accumulation?* [w:] Ch. Engel, W. Singer (red.), *Better Than Conscious?: Decision Making, the Human Mind, and Implications for Institutions*. Cambridge: MIT Press.
- Edelman G.M. (2003). *Naturalizing consciousness: a theoretical framework*. „Proceedings of the National Academy of Sciences of the United States of America” 100 (9), s. 5520–5524. doi:10.1073/pnas.0931349100.
- Edelman G.M., Tononi G. (2000). *A Universe of Consciousness: How Matter Becomes Imagination*. New York: Basic Books.
- Ferrarelli F., Massimini M., Sarasso S., Casali A., Riedner B.A., Angelini G., Pearce R.A. (2010). *Breakdown in cortical effective connectivity during midazolam-induced loss of consciousness*. „Proceedings of the National Academy of Sciences of the United States of America” 107 (6), s. 2681–2686.
- Holender D. (1986). *Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal*. „Behavioral and Brain Sciences” 9 (01), s. 1–23.
- Johnson R.W. (1987). *Relative-entropy minimization with uncertain constraints: Theory and application to spectrum analysis* [w:] C.R. Smith, G.J. Erickson (red.), *Maximum-Entropy and Bayesian Spectral Analysis and Estimation Problems* (s. 57–73). Springer Netherlands.
- Kales A., Rechtschaffen A. (1968). *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. U.S. National Institute of Neurological Diseases and Blindness, Neurological Information Network.
- Mascalco B., Lau H. (2012). *A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings*. „Consciousness and Cognition” 21 (1), s. 422–430.
- Massimini M., Ferrarelli F., Esser S.K., Riedner B.A., Huber R., Murphy M., Tononi G. (2007). *Triggering sleep slow waves by transcranial magnetic stimulation*. „Proceedings of the National Academy of Sciences of the United States of America” 104 (20), s. 8496–8501.
- Massimini M., Ferrarelli F., Huber R., Esser S.K., Singh H., Tononi G. (2005). *Breakdown of cortical effective connectivity during sleep*. „Science” 309 (5744), s. 2228–2232.
- Massimini M., Ferrarelli F., Murphy M., Huber R., Riedner B.A., Casarotto S., Tononi G. (2010). *Cortical reactivity and effective connectivity during REM sleep in humans*. „Cognitive Neuroscience” 1 (3), s. 176–183.

- Miller G. (2005). *What is the biological basis of consciousness?* „Science” 309 (5731), s. 79.
- Noirhomme Q., Soddu A., Lehembre R., Vanhauzenhuysse A., Boveroux P., Boly M., Laureys S. (2010). *Brain connectivity in pathological and pharmacological coma.* „Frontiers in Systems Neuroscience” 4. doi: 10.3389/fnsys.2010.00160.
- Pincus D.S.M., Gladstone I.M., Ehrenkrantz R.A. (1991). *A regularity statistic for medical data analysis.* „Journal of Clinical Monitoring” 7 (4), s. 335–345.
- Plourde G., Picton T.W. (1991). *Long-latency auditory evoked potentials during general anesthesia: N1 and P3 components.* „Anesthesia and Analgesia” 72 (3), s. 342–350.
- Ragazzoni A., Pirulli C., Veniero D., Feurra M., Cincotta M., Giovannelli F., Miniussi C. (2013). *Vegetative versus minimally conscious states: a study using TMS-EEG, sensory and event-related potentials.* „PloS one” 8 (2). doi: 10.1371/journal.pone.0057069.
- Rosanova M., Gosseries O., Casarotto S., Boly M., Casali A.G., Bruno M.-A., Massimini M. (2012). *Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients.* „Brain: A Journal of Neurology” 135 (Pt 4), s. 1308–1320.
- Rosow C., Manberg P.J. (2001). *Bispectral index monitoring.* „Anesthesiology Clinics of North America” 19 (4), s. 947–966.
- Sarasso S., Rosanova M., Casali A.G., Casarotto S., Fecchio M., Boly M., Massimini M. (2014). *Quantifying cortical EEG responses to TMS in (un)consciousness.* „Clinical EEG and Neuroscience” 45 (1), s. 40–49.
- Searle J.R. (1999). *Umysł, język, społeczeństwo: filozofia i rzeczywistość.* Warszawa: Wydawnictwo CiS.
- Sitt J.D., King J.-R., Naccache L., Dehaene S. (2013). *Ripples of consciousness.* „Trends in Cognitive Sciences” 17 (11), s. 552–554.
- Tononi G. (2004). *An information integration theory of consciousness.* „BMC Neuroscience” 5 (1), s. 42.
- Tononi G. (2008). *Consciousness as integrated information: A provisional manifesto.* „The Biological Bulletin” (December), s. 216–242.
- Tononi G. (2010). *Information integration: its relevance to brain function and consciousness.* „Archives Italiennes de Biologie” 148 (3), s. 299–322.
- Vermeiren A., Cleeremans A. (2012). *The validity of d' measures.* „PloS one” 7 (2), e31595. doi: 10.1371/journal.pone.0031595.
- Wierchoń M. (2013). *Granice świadomości.* Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- Wierchoń M., Asanowicz D., Paulewicz B., Cleeremans A. (2012). *Subjective measures of consciousness in artificial grammar learning task.* „Consciousness and Cognition” 21 (3), s. 1141–1153.

Dlaczego reprezentacje nie trzymają się modeli dynamicznych?

Why representations don't stick with dynamic models?

Abstract: In this paper I investigate thesis, embraced by proponents of dynamicism in cognitive science, that mind is not representational and explanation of cognition can go without representations. This claim has received serious criticism from cognitive scientists and philosophers of mind, who accuse dynamical explanation of being satisfying only for a narrow class of simple cognitive phenomena. Thus, genuine, representation-free explanation of cognition will always be incomplete. I espouse another strategy and present two arguments saying that the language of pure dynamical systems theory is not rich enough to define any nontrivial notion of representation. If I am right, then at least these phenomena dynamical explanation deals well with are not representational and representation talk can in no way help us understand them.

Key words: philosophy of cognitive science, dynamical explanation, representation, cognitive models, Clark, van Gelder

1. Na czym polega wyjaśnianie dynamiczne?

Idea wyjaśnienia dynamicznego jest urzekająco prosta: to pewien formalizm (choćby tak prosty, jak równanie algebraiczne z dwiema niewiadomymi), który opisuje, jak manipulacja wartością jednej zmiennej (może być nią np. czas) modyfikuje wartości pozostałych. Choć historia wyjaśnień dynamicznych jest równie długa jak historia nowożytnej nauki i szybko doczekała się spektakularnych sukcesów, takich jak mechanika newtonowska, to jej zastosowanie do badań umysłu jest stosunkowo nowym trendem. Pierwsze próby dynamicznego myślenia o procesach psychicznych podejmowali krótko po drugiej wojnie światowej cybernetycy w Stanach Zjednoczonych i Związku Radzieckim [Wiener 1948], a następnie – już po narodzinach kognitywistyki, jako przedstawiciele jej drugiej fali – koneksjoniści, proponując modele poznawcze o rozproszonej architekturze, takie jak sztuczne sieci neuronowe [Smolensky 1988]. Modele koneksjonistyczne oferowały aparat formalny do badania pewnych na ogół ignorowanych wcześniej aspektów procesów poznawczych (takich jak ich wymiar czasowy czy nieliniowe zależności), zachowując przy tym tradycyjny obraz poznania jako procesu obliczeniowego. Dopiero w latach 90. niektórzy filozofowie

i kognitywiści poszli o krok dalej, zaprzęgając wyjaśnienia dynamiczne do ustanowienia nowego paradygmatu, rozwijanego w opozycji do głównego nurtu kognitywistyki.

Klasycznym przykładem takiego podejścia są badania nad rozwojem poznawczym i motorycznym niemowląt Lindy Smith i Esther Thelen [2003]. W ich ujęciu rozwój polega na wyłanianiu się ze złożonych, nieliniowych współoddziaływań nowych wzorców (*patterns*) zachowania. Przykładowo pojawienie się około 12. miesiąca życia umiejętności postrzegania stałości przedmiotów nie da się wytłumaczyć, postulując pojawiające się *ex machina*, zgodnie z genetycznym programem, nowe kompetencje mentalne; ta umiejętność to wzorzec zachowania wyłaniający się wraz z rosnącą sprawnością motoryczną (chwywanie przedmiotów) i odpowiednią stymulacją środowiskową. Badaczki proponują psychologicznie wiarygodny i matematycznie elegancki model rozwoju umiejętności percypowania stałości przedmiotów jako pokonywania przez system poznawczy pewnej trajektorii w przestrzeni stanów, która to trajektoria podlega nieustannym modyfikacjom ze względu na lokalne wzorce zachowania (które z kolei współokreśla sama trajektoria).

2. Hipoteza dynamicystyczna

Stosowanie wyjaśnień dynamicznych w praktyce badawczej psychologów, robotyków, lingwistów i neuronaukowców to fakt historyczny. Bardziej interesująca z punktu widzenia naszej pracy jest filozoficzna interpretacja nadawana tej praktyce badawczej. Dynamicyzm urósł bowiem w kognitywistyce, jak wspomnieliśmy, do rangi autonomicznego paradygmatu, który postuluje szereg silnych tez empirycznych i teoretycznych (metodologicznych, ontologicznych) i staje w szranki z dominującym paradygmatem obliczeniowym. Moim punktem wyjścia będą manifesty jednego z pierwszych filozoficznych zwolenników dynamicyzmu, Tima van Geldera [1995, 1998; zob. też van Gelder i Port 1995]. Swoje stanowisko nazywa on „hipotezą dynamicystyczną” i formułuje następująco:

(HD) Systemy poznawcze (a) są systemami dynamicznymi i (b) najlepiej opisuje je aparat teoretyczny teorii układów dynamicznych [van Gelder 1998].

Przez system rozumie tu zbiór wzajemnie zależnych od siebie (*coupled*) zmiennych, a przez system *dynamiczny* – system zmieniający swój stan zgodnie z dobrze określoną regułą przejścia, którego przestrzeń stanów jest przestrzenią metryczną (tzn. dla każdej pary stanów jest określona ich odległość od siebie). Bardziej zgodne z duchem dynamicyzmu byłoby może dodanie zastrzeżenia, że chodzi o regułę przejścia opisaną funkcją ciągłą i wielokrotnie różniczkowalną, co presuponuje między innymi, że system może przybierać nieprzeliczalnie nieskończenie wiele stanów, ale van Gelder nie uznaje ciągłości przestrzeni stanów za cechę definicyjną systemu dynamicznego.

HD składa się z (a) komponentu ontologicznego (umysł *jest* systemem dynamicznym) i (b) metodologicznego (umysł *należy opisywać* jako system dynamiczny). Zdaniem van Geldera [1995] HD implikuje co najmniej trzy tezy szczegółowe:

- (1) Poznanie (a) nie jest procesem obliczeniowym i (b) nie należy go opisywać na sposób obliczeniowy.
- (2) Umysł (a) nie ma charakteru reprezentacyjnego i (b) poznanie można rozumieć bez postulowania istnienia wewnętrznych reprezentacji otoczenia¹.
- (3) Poznanie (a) odbywa się pod presją czasu i (b) w modelach procesów poznawczych należy uwzględnić czas rzeczywisty (ciągły) jako istotną eksplanacyjnie zmienną.

Warto uzupełnić tę listę o jeszcze dwie ważne tezy, nieformułowane *explicitie* przez van Geldera: tezę o aktywnym charakterze i ucieleśnieniu poznania. Postulują je takie nurty kognitywistyki, jak robotyka behawioralna [Brooks 1991], podejście ucieleśnione [Varela, Thompson i Rosch 1991] i enaktywizm [Thompson 2010], które są naturalnymi kontynuacjami gelderowskiego dynamicyzmu.

- (4) System poznawczy (a) nie jest mechanizmem odpowiadającym na egzogenne stymulacje według zaprogramowanego genetycznie algorytmu, ale samorganizującą się i autonomicznie eksplorującą otoczenie całością. (b) Nie należy go rozumieć jako statycznej struktury, ale jako proces związany obustronnymi i obukierunkowymi relacjami przyczynowymi ze sobą i z otoczeniem.
- (5) System poznawczy (a) jest systemem fizycznym i jego materialna konstytucja zarówno ogranicza jego możliwości poznawcze, jak i podsuwa mu gotowe rozwiązania pewnych problemów. (b) Należy uwzględniać niskopoziomowe cechy anatomii lub architektury systemu poznawczego w wyjaśnianiu wysoko-poziomowych procesów poznawczych².

¹ Stosunek van Geldera do tezy (2) (antyreprezentacjonizmu) jest ambiwalentny. Jako jej asercję można interpretować użycie przez niego metafory maszyny Watta jako urządzenia zdolnego do precyzyjnego kontrolowania swojego stanu, a przy tym niewątpliwie niereprezentacyjnego. Chwilę później czytamy jednak: „there is nothing preventing dynamical systems from incorporating some form of representation; indeed, an exciting feature of the dynamical approach is that it offers opportunities for dramatically reconceiving the nature of representation in cognitive systems, even within a broadly noncomputational framework” [van Gelder 1995, s. 376]. W późniejszym artykule [1998, s. 622] van Gelder wyjaśnia, że choć systemy poznawcze nie są ze swej istoty (*inherently*) reprezentacyjne, to jednak postulowanie reprezentacji jest legalnym i czasem pożądanym posunięciem. W szczególności dynamicysta dysponuje szerszą paletą rodzajów reprezentacji, może np. postulować reprezentacje *dynamiczne* na wyższych poziomach organizacji, przykładowo w przestrzeni stanów (por. krytykę tego pomysłu w sekcji 4.1). Mimo ambiwalencji van Geldera co do antyreprezentacjonizmu, na potrzeby tego tekstu uznajemy tezę (2) za istotny składnik HD, ponieważ po pierwsze, bez uznania tej tezy ciężko byłoby dynamicyzmowi określić swoją odrębność od głównego nurtu kognitywistyki oraz, po drugie, akceptuje ją większość innych dynamicystów, takich jak Anthony Chemero, Alva Noë i Evan Thompson (wśród filozofów) czy Linda Smith, Esther Thelen, Kevin O’Regan i Rodney Brooks (wśród badaczy praktyków).

² Teza (5) nie jest oryginalna dla dynamicyzmu i wydaje się, że można uczynić jej zadość przy zaprzeczeniu tez (1)–(4), np. w ramach wyjaśnienia obliczeniowego uwzględniającego mechanizm implementacji obliczeń. Nawet jeśli jej oponenci należą już do rzadkości, (5) pozostaje istotną, empirycznie testowalną konsekwencją HD.

Traktując HD jako falsyfikowalną hipotezę, van Gelder kładzie nacisk raczej na jej artykulację niż poparcie argumentami, a te ostatnie opiera bardziej na praktyce badawczej niż analizach pojęciowych. I tak koronnym argumentem za prawdziwością HD są sukcesy eksplanacyjne badań, które przyjęły HD jako hipotezę roboczą (np. badania Smith i Thelen; dla przykładów z innych obszarów kognitywistyki zob. antologię: Port i van Gelder 1995). Argument ten wzmacnia fakt, że wyjaśnienie dynamiczne jest kanonicznym modelem wyjaśnienia naukowego i odnosi sukcesy także (lub przede wszystkim) poza kognitywistyką, np. w naukach fizycznych: HD oferowałyby więc atrakcyjną filozoficznie unifikację kognitywistyki z innymi naukami przyrodniczymi w ramach wspólnego schematu eksplanacyjnego.

Przykładowa linia rozumowania, wychodząca od sukcesów badań Smith i Thelen [2003], może przebiegać następująco. Zakładając, że (1b) kompetencji poznawczych nie powinno się traktować jako zaprogramowanych genetycznie oraz że należy uwzględnić ich (3b) czasowy wymiar i (4b) tendencje do samoorganizacji, badaczki stworzyły model, który pozwala skutecznie przewidywać, jak zachowania eksploracyjne podejmowane przez dziecko wpływają na dynamikę jego dalszego rozwoju. Podobnie Thelen i Smith rozwijają klasyczną teorię rozwoju poznawczego Jeana Piageta, (2b) negując jednocześnie jego tezę, że nabywanie umiejętności percypowania stałości przedmiotu można wyjaśnić przez nabywanie umiejętności operowania na reprezentacjach mentalnych. Okazuje się, że zbudowanie niereprezentacyjnego i spełniającego wcześniejsze warunki modelu dynamicznego pozwala lepiej wytłumaczyć różnice indywidualne rozwoju badanych i gwałtowne zmiany wzorców zachowania (tzw. bifurkacje) przy drobnej rearanżacji sytuacji eksperymentalnej. Wreszcie zakładając, że (5b) rozwój poznawczy jest nierozzerwalnie związany z rozwojem motorycznym, Thelen i Smith tłumaczą, jak rozwój umiejętności chwytania wpływa na rozwój percepcji stałości przedmiotów. Ponieważ przyjęcie postulatów metodologicznych (1b–5b) prowadzi do sukcesów eksplanacyjnych, rozsądne jest wytłumaczenie tych sukcesów przyjęciem tez ontologicznych (1a)–(5a), które są konsekwencjami HD.

Argumentacja sukcesu eksplanacyjnego nie jest, z czego van Gelder zdaje sobie sprawę, rozstrzygająca. Nie przesądza bowiem, że nie istnieje inny zestaw postulatów metodologicznych, który prowadzi do jeszcze lepszych przewidywań, albo że nie ma alternatywnej ontologicznej podbudowy tych postulatów. Rzecz jasna istnieją też niezależne racje empiryczne za tezami (1)–(5) (van Gelder powołuje się np. na istotową czasowość poznania), ale ich omówienie przekracza ramy tego tekstu. Jak zresztą stwierdzono, van Gelder argumentuje nie tyle *explicite* za HD, ile raczej za tym, że nie ma powodów, by *a priori* wykluczyć HD. W szczególności nie jest naszym celem przeprowadzenie systematycznej obrony HD, lecz wykazanie pewnych jej konsekwencji, które skutecznie blokują pewien rodzaj krytyki.

3. Pięć argumentów przeciw hipotezie dynamicystycznej

Zacniemy od przytoczenia pięciu argumentów wysuwanych przeciw HD przez bardziej zachowawczo nastawionych badaczy i filozofów kognitywistyki.

(A1) Van Gelder traktuje systemy koneksjonistyczne jako podklasę systemów dynamicznych. Ale systemy koneksjonistyczne mają charakter obliczeniowy i reprezentacyjny, więc teza o dynamicznym charakterze poznania nie pociąga za sobą tezy o jego nieprezentacyjnym i nieobliczeniowym charakterze [Eliasmith 1997].

(A2) Wyjaśnienie dynamiczne nie jest pełnym wyjaśnieniem, bo ignorując strukturę wewnętrzną systemu poznawczego, nie potrafi zdać sprawy z tego, dlaczego niektóre systemy wykonują te same zadania wydajniej niż inne. Formułująca ten argument Melanie Mitchell [1998] powołuje się na swoje badanie nad binarnymi automatami komórkowymi. Ewoluuując za pomocą algorytmu genetycznego różne strategie rozwiązania pewnego zadania (wprowadzenia wszystkich komórek automatu w stan 0 lub 1, który dominował w warunkach początkowych), Mitchell zauważyła, że klasy abstrakcji różnych rozwiązań ze względu na złożoność obliczeniową nie pokrywają się z klasami abstrakcji ze względu na dynamikę. Wyjaśnienia obliczeniowe i dynamiczne są zatem wzajemnie nieredukowalne i komplementarne, więc wyjaśnienie dynamiczne jest niepełne.

(A3) Krytyka reprezentacjonizmu na gruncie zarówno dynamicyzmu [Smith i Thelen 2003], jak i robotyki behawioralnej [Brooks 1991] opiera się na ekwiwokacji pojęcia reprezentacji: ma przesłanki co najwyżej przeciw globalnie dostępnym i manipulowalnym reprezentacjom symbolicznym (*à la* Newell i Simon 1976), a w konkluzji odwołuje się do wszelkich typów reprezentacji. Przykładowo konstrukt taki jak pętla emulacyjna Grusha [2004] stanowi reprezentację w całkiem silnym sensie, a można ją postulować w pełnej zgodzie z tezą o ucieleśnieniu i usytuowaniu poznania, więc jest odporna na argument z biologicznej nieadekwatności reprezentacji i falsyfikuje HD [Clark i Toribio 1994; Clark 1997, 2001].

(A4) Przypadki, na które powołują się dynamicyści i robotycy behawioralni, nie są akurat problemami, które domagają się postulowania reprezentacji, co nie znaczy, że wyjaśnienie wszystkich aspektów poznania może się bez nich obejść. Wrażliwość systemu poznawczego na aktualnie niedostępne percepcyjnie przedmioty lub abstrakcyjne cechy przedmiotów jest na przykład fenomenem, dla którego nie podano dotąd przekonującego wyjaśnienia dynamicznego. HD jest zatem co najmniej przedwczesna [Clark i Toribio 1994; Clark 2001].

(A5) Opis aktywności systemu poznawczego w kategoriach normatywnych zakłada istnienie w nim komponentów o własnościach semantycznych, których przykładem są właśnie reprezentacje. Opis dynamiczny niepostulujący reprezentacji nie potrafi więc odróżnić systemu dysfunkcyjnego od działającego poprawnie, a zatem jest niepełny [Dennett 1998].

Przed ustosunkowaniem się do tej krytyki spróbujemy wyklarować pozycję teoretyczną, z której będziemy się przed nią bronili. Dennett [1998] słusznie zauważa, że zwa wewnątrznie niespójnych manifestów dynamicyzmu wyłania się właściwie dwóch van Gelderów. Pierwszy z nich, van Gelder bezkompromisowy, jest konsekwentnym

antyreprezentacjonistą, drugi, van Gelder ugodowy, argumentuje tylko na rzecz zastąpienia reprezentacji symbolicznych w stylu GOFAI czymś bardziej biologicznie realistycznym. Bezkompromisowy van Gelder spotyka się z przytoczoną surową krytyką, przed którą nie umie obronić się inaczej, niż wcielając się w swoje ugodowe *alter ego* – które równie dobrze mogłoby uosabiać panujący w kognitywistyce konsens. Wydaje się, że rozterkę van Geldera można rozwiązać, podważając przesłankę, na której się opiera: żeby wyjaśnienie dynamiczne mogło być wartościowe, musi być w stanie wyjaśnić wszystkie aspekty poznania. Nie będziemy tu przedstawiali systematycznej krytyki tak sformułowanego absolutyzmu eksplanacyjnego³, pokażemy po prostu, że zakładając roboczo, iż jest fałszywy, można uratować dynamizm bezkompromisowego van Geldera, nie osłabiając go ani trochę.

Spróbujemy więc zinterpretować HD na tyle radykalnie, by była w stanie obronić się przed zarzutami lub z podniesioną głową odpowiedzieć *touché*. To ostatnie wydaje się jedyną rozsądną reakcją na zarzuty (A1), (A2) i (A5) o tyle, o ile van Gelder jak jego bliźniak nie chce rozcieńczać dynamizmu w hybrydowy program badawczy. Zarówno wysokopoziomowe opisy dynamiki sieci koneksjonistycznych, jak i niskopoziomowe opisy ich architektury (postulujące reprezentacje rozproszone) mogą być wartościowe w wyjaśnianiu pewnych aspektów zachowania sieci, a bezradne wobec innych. To, że opis dynamiczny nie może wyjaśnić czegoś, co wyjaśnia opis obliczeniowy, nie jest więc argumentem przeciw dynamizmowi, lecz co najwyżej przeciw absolutyzmowi eksplanacyjnemu. Zabezpieczywszy się tak przed zarzutami typu „bez reprezentacji nie ma wyjaśnienia”, w następnej części tekstu pokażę, że w ramach wyjaśnienia dynamicznego w istocie nie da się mówić o reprezentacjach w żadnym nietrywialnym sensie.

4. Reprezentacje w schemacie wyjaśniania dynamicznego

Zarzuty Clarka (A3 i A4) są w istocie głosami w długim sporze o istnienie reprezentacji. Ponieważ Clark (słusznie) sprzeciwia się konceptualizacji tego sporu jako dychotomii, przytoczymy rozważane przez niego [1997] cztery sensy pojęcia reprezentacji i sformułujemy dwa argumenty, jakie bezkompromisowy van Gelder mógłby przeciwko nim wytoczyć.

(R1) Reprezentacja to stan wewnętrzny systemu poznawczego. (System może zmieniać swój stan pod wpływem przeszłych stymulacji z otoczenia i jego stan może wpływać na reakcje na przyszłe stymulacje)⁴.

(R2) Reprezentacja to stan wewnętrzny systemu poznawczego skorelowany z jego otoczeniem (lub jego wycinkiem).

³ Celną krytykę absolutyzmu eksplanacyjnego w kognitywistyce (w kontekście wyjaśniania obliczeniowego) przedstawia np. Marcin Miłkowski [2013].

⁴ Jak zauważa Clark, reprezentacja w sensie R1 nie ma żadnej wartości eksplanacyjnej. Dalej, mówiąc o nietrywialnych reprezentacjach, będę miał na myśli reprezentacje w sensach R2–R4.

(R3) Reprezentacja to stan wewnętrzny systemu poznawczego, którego funkcją właściwą jest bycie skorelowanym z otoczeniem. (Funkcją właściwą rozumiemy jako określoną przez historię filogenetycznej ewolucji systemu poznawczego i stanowiącą o jego dostosowaniu).

(R4) Reprezentacja to stan wewnętrzny systemu poznawczego skorelowany z otoczeniem, który umożliwia systemowi przewidywanie i analizowanie otoczenia podczas jego sensorycznej niedostępności.

4.1. Argument z braku struktury

Nietrywialne pojęcia reprezentacji zakładają istnienie co najmniej jednego inwariantu: struktury, która jest nośnikiem swoich stanów wewnętrznych i może stanowić statyczny układ odniesienia do szukania korelacji pomiędzy nimi a otoczeniem. Struktura w teorii układów dynamicznych jest własnością emergentną całego systemu (funkcją wielu zmiennych, w tym czasu), a nie czymś gotowym. Żeby mówić o strukturze (w momencie t_1), najpierw musimy mieć gotowy model dynamiczny, czyli właściwie gotowe wyjaśnienie – a skoro takie mamy, postulowanie reprezentacji to tylko przyklejanie etykietek, które nie doda do niego niczego, czego już nie wiemy. W szczególności budowanie redukcyjnej teorii reprezentacji na modelu dynamicznym będzie poznawczo jałowe, bo nie jest spełniona przesłanka pragmatyczna postulowania obiektu redukowalnego: że nie mamy efektywnej procedury przewidywania stanów systemu na najniższym poziomie opisu. W modelu dynamicznym taką efektywną procedurą jest manipulacja równaniem stanu układu, w ramach której z definicji można przewidzieć wszystko, co dane jest do przewidzenia (bez potrzeby wchodzenia na wyższy, reprezentacyjny poziom opisu).

Rozważmy strategię ugodowego van Geldera, polegającą na postulowaniu reprezentacji *dynamicznych* [van Gelder 1995]. Taką reprezentacją może być np. wzorzec trajektorii w przestrzeni fazowej czy region w przestrzeni fazowej (np. basen atrakcji) lub abstrahując jeszcze bardziej, region na diagramie bifurkacji. Tego typu reprezentacje są jednak zwykłymi nietrywialnymi reprezentacjami zlokalizowanymi w *stacyjnej* strukturze modelu dynamicznego: strukturą jest tutaj obiekt matematyczny – przestrzeń stanów danego układu dynamicznego. Samo pojęcie reprezentacji dynamicznej zasadza się więc na fundamentalnej ekwiwokacji modelu i modelowanego zjawiska: postulujemy byt drugiego rzędu na naszym modelu, a nie w modelowanym zjawisku. Przestrzeń fazowa nie jest niczym realnie istniejącym w świecie fizycznym – to geometryczna interpretacja pewnej struktury algebraicznej, która stanowi abstrakt; topologiczne własności przestrzeni fazowej nie muszą się więc przekładać na topologiczne własności modelowanego zjawiska. Równie dobrze moglibyśmy modelować system poznawczy za pomocą innego formalizmu, np. narracji w języku naturalnym, i argumentować, że nasz system poznawczy ma kompetencję językową, bo nasz model ma własności gramatyczne (jest ciągiem zdań złożonych języka polskiego). Reprezentacje drugiego rzędu, zlokalizowane w przestrzeni fazowej, o ile nie umiemy ich zlokalizować w fizycznej instancji modelu, są tylko artefaktem tego modelu.

4.2. Argument ze złożoności oddziaływań

Nietrywialne pojęcia reprezentacji zakładają skorelowanie stanów systemu poznawczego z otoczeniem. Załóżmy, że mamy układ dynamiczny złożony z agenta i jego środowiska, a w nim dwie skorelowane zmienne: x i y . Jeśli zmienne są skorelowane liniowo, to x jest funkcją liniową od y ($y = ax + b$, gdzie a i b to stałe). Wtedy jedna ze zmiennych jest redundantna: nie osiągamy korzyści eksplanacyjnej, modelując osobno x i osobno jego izomorficzną reprezentację y ⁵. Jeśli są skorelowane nieliniowo, to a i b w poprzednim równaniu też są funkcjami innych zmiennych, c i d , dla których przeprowadzamy analogiczne rozumowanie. Wnioskujemy, że jeśli co najmniej dwie zmienne są z sobą skorelowane nieliniowo, to wszystkie zmienne w modelu są z sobą parami skorelowane nieliniowo, więc relacja korelacji (zarówno liniowej, jak i nieliniowej) jest trywialna (wszystko jest ze sobą skorelowane) i nie pozwala zdefiniować nietrywialnej relacji reprezentacji.

Dodatkowe warunki nakładane przez (R3) i (R4) w niczym nie pomagają. Pierwszy (dostosowanie) dlatego, że zakłada adaptacjonizm, który nie cieszy się dobrą sławą wśród dynamicystów. Wolą oni myśleć o ewolucji jako o ko-ewolucji systemu poznawczego i środowiska, w którym dobór naturalny nie jest jedynym czynnikiem [Thompson 2010]. W takim ujęciu złożoność przyczynowa historii filogenetycznej danego systemu poznawczego nie pozwala efektywnie wyróżniać cech adaptacyjnych. Warunki nakładane przez (R4) nie zawężają w żaden sposób denotacji pojęcia reprezentacji, bo (z wymienionych już powodów) nie odwołują się wprost do anatomii bądź architektury systemu poznawczego. Nie mogą więc zapobiec temu, że reprezentacja „rozleje” się na cały system poznawczy.

Okazuje się więc, że fundamentalne dla pojęcia reprezentacji pojęcia struktury i korelacji okazują się nieadekwatne wtedy, kiedy mamy do czynienia z wyjaśnieniem dynamicznym. Argument z braku struktury opiera się na tym, że nie został spełniony warunek pozytywny sensowności definicji struktury (nie istnieje żadna struktura), a argument ze złożoności oddziaływań opiera się na tym, że nie został spełniony negatywny warunek sensowności definicji korelacji (nie istnieje taka para zmiennych, która nie jest skorelowana). Dochodzimy więc do konkluzji, że język teorii układów dynamicznych jest zbyt ubogi, by zdefiniować w nim jakiegokolwiek nietrywialne pojęcie reprezentacji.

Oczywiście do wyjaśnienia dynamicznego oprócz modelu matematycznego potrzebujemy też jego implementacji, o której można myśleć jako o funkcji ze zmiennej w fizyczną własność pewnej instancji modelu. Większość kognitywistów tutaj widziałaby zapewne odpowiedni moment na zapostulowane reprezentacji, ale nie byłoby to posunięcie *fair play*. Przede wszystkim dlatego sam model matematyczny nie określa warunków, kiedy jakiś element jego fizycznej instancji jest reprezentacją, a kiedy nie. Tak jak relacje wyższego rzędu między zmiennymi (domniemane reprezentacje dynamiczne) nie przechodzą na fizyczne instancje modelu, tak struktura fizycznej instancji nie przechodzi na zmienne (o ile nie modelujemy w skali molekularnej).

⁵ W pewnym sensie ten krok mojego argumentu to algebraiczne sformułowanie tezy Brooka, że świat jest swoim najlepszym modelem. Jest to sformułowanie życzliwe, zakładające, że bycie czymś różnym od swojego desygnatu nie stanowi cechy definicyjnej modelu.

larnej). Warunki istnienia reprezentacji musiałaby więc zostać określone przez jakąś heterogeniczną teorię – ale wtedy mamy do czynienia z wyjaśnieniem hybrydowym w stylu van Geldera kompromisowego: takie wyjaśnienie postuluje reprezentacje tylko o tyle, o ile uzupełniamy je wyjaśnieniem innego typu (np. koneksjonistycznym czy obliczeniowym). Sam układ dynamiczny jako model matematyczny zobowiązuje nas do uznania reprezentacji co najwyżej w sensie (R1), niezlokalizowanych nigdzie konkretnie w strukturze zjawiska i eksplanacyjnie jałowych.

5. Uwagi końcowe

Można zarzucać naszym argumentom, że tak naprawdę nie są argumentami za antyreprezentacjonizmem w sporze teoretycznym przeciw reprezentacjonizmowi, ale argumentami za pluralizmem w sporze *metateoretycznym* z absolutyzmem eksplanacyjnym. Jesteśmy skłonni uznać taką interpretację. Naszym głównym celem było obalenie tezy, że w *czystym* modelu dynamicznym jest miejsce na reprezentacje. Jeśli nie uda się odeprzeć zarzutów (A1), (A2) i (A5), to wynika stąd, że nie jest też prawdą, że model dynamiczny może wyjaśnić wszystkie aspekty poznania, co skłania do uznania pluralizmu eksplanacyjnego albo odrzucenia wyjaśnień dynamicznych. Nie widzimy ani jednego dobrego powodu, który mógłby przemawiać za tym ostatnim krokiem.

BIBLIOGRAFIA

- Brooks R. (1991). *Intelligence without representation*. „Artificial Intelligence” 47, s. 139–159.
- Clark A., Toribio J. (1994). *Doing without representing?* „Synthese” 101, s. 401–431.
- Clark A. (1997). *The dynamical challenge*. „Cognitive Science” 21, s. 461–481.
- Clark A. (2001). *Mindware: An Introduction to the Philosophy of Cognitive Science*. New York–Oxford: Oxford University Press.
- Dennett D. (1998). *Revolution, no! Reform, si!* „Brain and Behavioral Sciences” 21, s. 636–637.
- Eliasmith C. (1997). *Computation and dynamical models of mind*. „Minds and Machines” 7, s. 531–541.
- Gelder T. van (1995). *What might cognition be, if not computation?* „The Journal of Philosophy” 26, s. 345–381.
- Gelder T. van, Port R. (1995). *It's about time: an overview of the dynamical approach to cognition* [w:] R. Port, T. van Gelder (red.), *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge–London: MIT Press.
- Gelder T. van (1998). *The dynamical hypothesis in cognitive science*. „Behavioral and Brain Sciences” 21, s. 615–665.
- Grush R. (2004). *The emulation theory of representation: motor control, imagery, and perception*. „Behavioral and Brain Sciences” 27, s. 377–396.
- Miłkowski M. (2013). *Explaining the Computational Mind*. Cambridge–London: MIT Press.
- Mitchell M. (1998). *A complex-systems perspective on the “computation vs. dynamics” debate in cognitive science* [w:] M.A. Gernsbacher, S.J. Derry (red.), *Proceedings of the 20th Annual Conference of the Cognitive Science Society – CogSci 1998*, s. 710–715.

- Newell A., Simon H. (1976). *Computer science as empirical inquiry: symbols and search*. „*Communications of the ACM*” 19, s. 113–126.
- Port R., Gelder T. van (red.) (1995). *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge–London: MIT Press.
- Smith L., Thelen E. (2003). *Development as a dynamic system*. „*Trends in Cognitive Sciences*” 7, s. 343–348.
- Smolensky P. (1988). *On the proper treatment of connectionism*. „*Behavioral and Brain Sciences*” 11, s. 1–74.
- Thompson E. (2010). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Harvard: Harvard University Press.
- Varela F., Thompson E., Rosch E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge–London: MIT Press.
- Wiener N. (1948). *Cybernetics; or, Control and Communication in the Animal and the Machine*. Paris: Technology Press.



Poza ciało: różnice indywidualne w podatności na eksterioryzację

Individual differences in susceptibility to out-of-body experience

Abstract: During out-of-body experience (OBE) visual perspective and bodily awareness are altered. The phenomenon has been investigated in studies with brain-damaged patients, people exposed to gravitational overloads, as well as victims of traumatic events. Recently, virtual reality technology enabled evoking transient OBE in healthy people. Studies using full-body illusion (FBI) protocol point to incorrect visuo-vestibular integration as a potential mechanism evoking OBE. This conclusion is strengthened by neuroimaging and lesion studies suggesting important role of temporo-parietal junction (TPJ; multisensory associative cortical area damaged in patients suffering from OBE) for proper visuo-vestibular integration. Finally, potential mechanisms of individual differences in susceptibility to OBE ranging from simple perceptual to higher cognitive processes are discussed.

Key words: out-of-body experience, individual differences, bodily awareness, virtual reality, visuo-vestibular integration

Doświadczenie eksterioryzacji objawia się zniekształceniem perspektywy wzrokowej oraz poczuciem zlokalizowania poza własnym ciałem. Badania tego zjawiska prowadzono z udziałem pacjentów po uszkodzeniach mózgu, osób narażonych na przeciążenia grawitacyjne oraz ofiar traumatycznych wydarzeń. Ostatnio wykorzystanie technologii wirtualnej rzeczywistości umożliwiło wywołanie stanu eksterioryzacji u osób zdrowych. Badania w paradygmacie iluzji całego ciała sugerują, że OBE może wynikać z zaburzeń integracji informacji wzrokowych z przedsionkowymi. Dane wzmacniają wnioski z badań na pacjentach, które wskazują na rolę styku skroniowo-ciemieniowego dla prawidłowej perspektywy wzrokowej i postrzegania zlokalizowania siebie w przestrzeni. Ponadto zaobserwowano zróżnicowanie indywidualne w podatności na OBE, które może wynikać z różnic w przebiegu prostych procesów percepcyjnych oraz złożonych procesów poznawczych, w tym uwagi i motywacji.

Demitologizacja zjawiska eksterioryzacji

Eksterioryzacja (*out-of-body experience*) przez lata była analizowana jako zjawisko paranormalne, często wymykające się opisowi w języku nauki [Blackmore 1978]. Rozważano jej podłoże, odnosząc się do narracji ezoterycznej i religijnej, często przedstawiając OBE jako argument na rzecz istnienia pozacielesnego [Ustinova 2011]. W wielu tradycjach religijnych doświadczenie to stanowiło ważny element rytuałów, stąd rozwinęto szereg technik osiągania stanu postrzegania rzeczywistości z perspektywy pozacielesnej. Również współcześnie w kręgach poszukiwaczy zmienionych stanów świadomości oraz nietypowych doświadczeń percepcyjnych da się znaleźć zestaw metod umożliwiających osiągnięcie stanu eksterioryzacji. Wśród kluczowych można wymienić: zażywanie odpowiednich środków psychoaktywnych [Wilkins i in. 2011; Wilde 2011], techniki relaksacji połączone z technikami wizualizacji [Blackmore 1987; Braithwaite i Dent 2011], deprivację sensoryczną [Zubek 1969] czy wreszcie techniki osiągania stanów świadomego śnienia (*lucid dreaming*) [Stumbrys i in. 2012], które umożliwiają wywołanie paraliżu przysennego (*sleep paralysis*) [Cheyne i Girard 2009], często prowadzącego do OBE.

Ponadto zanotowano wystąpienie doświadczenia eksterioryzacji i innych zaburzeń samoświadomości cielesnej (*bodily self-consciousness*) [Blanke 2012] w sytuacjach bardzo dużego stresu, np. u ofiar tragicznych wypadków oraz wojen [Cardena i in. 1997]. Inną grupą często doświadczającą podobnych stanów są piloci, którzy podczas silnych przesileń grawitacyjnych mogą doznawać zmienionych stanów świadomości, w tym eksterioryzacji [SKYbrary 2013]. Kolejnym przykładem są osoby przebywające w stanie mikrogravitacji, które mogą błędnie lokalizować siebie w przestrzeni w wyniku deprivacji informacji docierających do organów ucha wewnętrznego [Kornilova 1997; Nagy i in. 2000] – otolitów (*otoliths*) i kanałów półkolistych (*semicircular canals*) – i dalej do podkorowych oraz korowych struktur układu przedsionkowego (*vestibular system*) [Lopez i Blanke 2011], odpowiedzialnego za percepcję zlokalizowania głowy w stosunku do wektora grawitacji, a tym samym zlokalizowania ciała w przestrzeni oraz utrzymywanie równowagi.

Często występowanie OBE jest związane z niepoprawnym funkcjonowaniem określonych obszarów mózgu, jak w przypadku pacjentów z lezjami w obszarze styku skroniowo-ciemiennego¹ (*temporoparietal junction*, TPJ) [Blanke i Arzy 2005]. Innym przykładem są pacjenci dotknięci padaczką skroniową (*temporal epilepsy*) [Heydrich i in. 2011], którzy podczas ataków bądź przed ich wystąpieniem mogą postrzegać rzeczywistość z lokalizacji poza ciałem. Bardziej powszechnym schorzeniem, które w swojej ostrej formie również może wywołać eksterioryzację, jest migrena [Podoll i Robinson 1999].

¹ Termin ten jest używany na określenie obszaru na styku tylnej części bruzdy skroniowej górnej (*superior temporal sulcus*, STS), płacika ciemiennego dolnego (*inferior parietal lobule*, IPL) oraz bocznej kory potylicznej (*lateral occipital cortex*) [Mars i in. 2012]. Obszar ten uznawany jest także za neuronalny korelat wielu funkcji poznawczych, w tym uwagi orientacyjnej (*orienting attention*) [Corbetta, Patel i Shulman 2008], przełączania między zadaniami (*task switching*) [Jimura i in. 2014], rozróżniania siebie od innych (*self-other distinction*) [Uddin i in. 2006], przyjmowania perspektywy oraz empatii (*perspective taking*) [Decety i Sommerville 2003] czy wreszcie teorii umysłu (*theory of mind*, ToM) [Decety i Lamm 2007].

Ponadto ostatnie badania sugerują, że także osobom niedotkniętym żadnym z powyższych schorzeń, nieposzukującym aktywnie zmienionych stanów świadomości, nieznajdującym się w żadnym z wyżej wymienionych kontekstów zdarza się sporadycznie doświadczać poczucia eksterioryzacji [Braithwaite i in. 2010]. Co więcej – częstotliwość doświadczenia tych stanów oraz podatność na nie są silnie zróżnicowane indywidualnie, tj. można wyróżnić pewien zestaw cech szczególnie predysponujących do tego typu doznań [Braithwaite i in. 2010; Ionta i in. 2011]. Jak dotychczas badania OBE rzadko były prowadzone w ujęciu różnic indywidualnych, badacze skupiali się raczej na zidentyfikowaniu ogólnego mechanizmu odpowiedzialnego za powstawanie tego stanu świadomości [Aspell i in. 2011; Blanke i in. 2005; Blanke i Dieguez 2009; Blanke i Mohr 2005; Blanke i Thut 2007; Brugger i in. 1997; Cheyne i Girard 2009; Easton i in. 2009; Ehrsson 2007; Ionta i in. 2011; Lenggenhager i in. 2007]. W tym celu rozwinęli szereg metod eksperymentalnych umożliwiających wywoływanie OBE w sposób kontrolowany i odwracalny u osób niedotkniętych żadnymi schorzeniami. W dalszej części niniejszego artykułu pokrótce zostaną przedstawione owe metody, jak również przytoczone najważniejsze wyniki badań poszukujących neuronalnych i poznawczych korelatów OBE. Na zakończenie zostanie poruszona kwestia potencjalnych źródeł różnej podatności na doświadczenie eksterioryzacji. Zanim jednak powyższe aspekty zostaną opisane, istotne jest bliższe przyjrzenie się fenomenologii doświadczenia eksterioryzacji.

Gdzie jest ciało poza ciałem?

Fenomenologia poczucia eksterioryzacji była przedmiotem licznych dociekań filozofów oraz badaczy empirycznych [m.in. Blackmore 1982; Blanke i Dieguez 2007; Brugger i in. 1997; Cheyne i Girard 2009; Monroe 1971; Wilde 2011]. Aktualne stanowisko zakłada, że osoba doświadczająca OBE postrzega rzeczywistość z lokalizacji poza własnym ciałem zazwyczaj z perspektywy uniesionej (*elevated perspective*) w stosunku do rzeczywistej orientacji w przestrzeni. Kolejne cechy doświadczenia OBE, które były dyskutowane w ostatnich latach, to zrotowanie perspektywy oglądu rzeczywistości, tj. postrzeganie z pozycji zrotowanej o 180 stopni w stosunku do rzeczywistej. Najintensywniej dyskutowana jest natomiast sugestia niektórych badaczy, że o OBE można mówić wyłącznie wtedy, kiedy osoba doświadczająca zmienionego stanu świadomości widzi swoje ciało z perspektywy pozacieleśnej. Zważywszy na różnorodność doświadczeń określanych w literaturze mianem OBE, niektórzy badawcze podają w wątpliwość, czy nazywanie ich wszystkich tym samym terminem jest uzasadnione, i postulują opracowanie ścisłej definicji tego zjawiska oraz sporządzenie szczegółowej taksonomii doświadczeń pokrewnych [Braithwaite i Dent 2011].

Ustalenie jednoznacznej definicji OBE jest o tyle istotne, że stanowi warunek zaprojektowania metod umożliwiających eksperymentalne wywoływanie eksterioryzacji u osób zdrowych. Wyniki badań z udziałem osób zdrowych mogą być porównywane z wynikami badań na pacjentach z uszkodzeniami mózgu, epileptykach, pilotach etc. pod warunkiem, że doświadczenie wywołane eksperymentalnie będzie zbliżone do doświadczenia OBE występującego w wyżej wymienionych specyficznych

kontekstach. Prowadzenie kontrolowanych badań eksperymentalnych dotyczących eksterioryzacji utrudnia intymny charakter tego zjawiska – fakt doświadczania iluzji dostępny jest wyłącznie osobie, która ją przeżywa. W tym wypadku miary obiektywne, takie jak lokalizacja lezji wywołujących OBE czy analiza aktywacji mózgu u osób raportujących eksterioryzację, nie pozwalają skutecznie przewidywać występowania tego stanu u innych osób. Do pomyslenia jest sytuacja, w której takie samo uszkodzenie mózgu czy taki sam stan aktywacji kluczowych dla wywołania OBE struktur mózgowych u jednej osoby zaowocuje doświadczeniem eksterioryzacji, podczas gdy u innej osoby nie.

Pomimo licznych problemów definicyjnych, w ostatnich latach zaproponowano szereg koncepcji demitologizujących zjawisko OBE i proponujących metody empirycznego badania mechanizmów leżących u podłoża tego fenomenu.

Eksperymentalne metody wywołania OBE

Jedną z pierwszych prób eksperymentalnego manipulowania poczuciem zlokalizowania siebie w przestrzeni podjął pod koniec XIX wieku George M. Stratton [1989, za: Blanke 2012]. Skonstruował on urządzenie składające się z zestawu lusterek, umożliwiające doświadczenie zmienionego postrzegania zlokalizowania swojego ciała w przestrzeni. Pionierskie badania Strattona stały się inspiracją dla wielu późniejszych rozwiązań, w tym eksperymentów dotyczących adaptacji pryzmatycznej (*prism adaptation*) [Redding i in. 2005] oraz okularów odwracających pole widzenia, zarówno w płaszczyźnie wertykalnej, jak i horyzontalnej [Linden i in. 1999]. Wpływ tego typu manipulacji ograniczał się jednak do zaburzania perspektywy wzrokowej, stąd iluzje w ten sposób wywołane można by nazywać iluzjami wzrokowymi. Zmianie ulegała wyłącznie perspektywa oglądu rzeczywistości, dlatego rzadko wpływało to na percepcję własnego ciała.

Krok dalej poszli Botvinick i Cohen, którzy w swoich pionierskich badaniach dotyczących iluzji gumowej ręki (*rubber hand illusion*, RHI) [Botvinick i Cohen 1998] wykorzystali tendencję układu nerwowego do scalania zdarzeń współwystępujących w czasie w jedno spójne doświadczenie percepcyjne. Podstawową manipulacją była synchroniczna stymulacja dotykowa gumowej ręki, widocznej dla osób badanych, oraz ręki rzeczywistej – ukrytej przed ich wzrokiem. W rezultacie wprowadzenia takiego konfliktu sensorycznego badani czuli dotyk na swojej fizycznej kończynie, lecz widząc go w tym samym czasie na ręce sztucznej, doświadczali iluzji, jakby dotykana gumowa ręka była częścią ich ciała. Eksperyment ten udało się wielokrotnie zreplikować, uzyskując wyniki potwierdzające, że synchroniczna (w porównaniu z asynchroniczną) stymulacja dotykowa może wywołać przejściowe i odwracalne zmiany w postrzeganiu własnego ciała. Ponadto udało się potwierdzić, że taka manipulacja wpływa na postrzegane zlokalizowanie własnego ciała w przestrzeni. W przypadku iluzji gumowej ręki efekt ten nazwano dryfem proprioceptywnym (*proprioceptive drift*) [Botvinick i Cohen 1998]. Zjawisko to polega na tym, że po synchronicznej stymulacji ręki gumowej i fizycznej badani systematycznie mylą się podczas oceny położenia swojej rzeczywistej kończyny w przestrzeni – lokalizują ją tak, jakby była

przesunięta w kierunku ręki gumowej. Co istotne, subiektywne miary poczucia iluzji nie zawsze korelują z miarami bardziej obiektywnymi [Rohde i in. 2011], co często używane jest jako argument na rzecz co najmniej dwóch rodzajów reprezentacji cielesnej – mniej świadomego, wykorzystywanego do planowania działania schematu ciała (*body schema*) oraz świadomego obrazu ciała (*body image*). Debata na temat różnych typów reprezentacji cielesnych nie jest jednak przedmiotem niniejszego artykułu, stąd osoby zainteresowane tym problemem odsyłam do licznych artykułów poświęconych tej kwestii [np. Holmes i Spence 2006; de Vignemont 2010].

W kolejnym kroku badacze połączyli idee zawarte w dwóch typach opisanych powyżej eksperymentów – manipulację perspektywą wzrokową zaproponowaną przez Strattona oraz manipulację wzrokowo-dotykową wykorzystaną w iluzji gumowej ręki. W rezultacie rozwinęli metody umożliwiające wywoływanie iluzji dotyczących całego ciała (*full-body illusion*, FBI) [Lenggenhager i in. 2007; Ehrsson 2007]. Wykorzystując hełmy do projekcji wirtualnej rzeczywistości (*head-mounted display*, HMD) oraz proste systemy robotyczne, udało się uzyskać między innymi efekt identyfikacji z wirtualnym ciałem oraz zaburzyć poprawne poczucie zlokalizowanie całego ciała w przestrzeni. Podczas klasycznego eksperymentu osoba badana widzi poprzez HMD wirtualne ciało bądź zdelokalizowane w stosunku do rzeczywistego położenia ciało własne. Następnie dotykana jest równocześnie w wybranym punkcie na ciele i widzi zsynchronizowany w czasie i przestrzeni dotyk na ciele iluzorycznym. Na skutek tego doświadczą przejściowych zmienionych stanów samoświadomości cielesnej [przełład: Blanke 2012] – między innymi poczucia oglądu rzeczywistości z perspektywy poza własnym ciałem.

Wirtualnie poza ciałem

Opisany powyżej paradygmat postanowiono wykorzystać do wywołania poczucia eksterioryzacji. Jak dotychczas zastosowano w tym celu dwa nieco odmienne podejścia, z których jedno rozwinęli szwajcarscy badacze z zespołu Olafa Blankego, a drugie szwedzcy badacze pod kierunkiem Hendrika Ehrssona. W pierwszym przypadku [Lenggenhager i in. 2007] osoba badana za pośrednictwem HMD oraz systemu kamer umieszczonego około dwóch metrów za jej plecami widzi swoje ciało, tak jakby znajdowało się ono przed nią (tj. widzi swoje plecy około dwóch metrów przed sobą). Następnie w celu wywołania iluzji znajdowania się poza własnym ciałem osoba badana dotykana jest przez eksperymentatora po plecach. W rezultacie widzi ona przed sobą własne ciało dotykane przez eksperymentatora i jednocześnie doświadczają zsynchronizowanego czasowo i przestrzennie dotyku na własnych plecach. Analogicznie do iluzji gumowej ręki manipulacja ta wywołuje iluzję jakoby widziane za pośrednictwem hełmu ciało było ich własnym, tj. wprowadza dysocjację pomiędzy faktycznym a odczuwanym zlokalizowaniem ciała. Innymi słowy, wywołano wrażenie postrzegania własnego ciała z perspektywy pozacielesnej, co zgodne jest z kluczowym kryterium fenomenologii OBE².

² „feeling of spatial separation of the observing self from the body” [Brugger 2002].

Alternatywny sposób eksperymentalnego wywołania poczucia eksterioryzacji zaproponował Ehrsson [2007]. Wprawdzie wykorzystał on także system kamer oraz hełm do projekcji wirtualnej rzeczywistości, niemniej w tym wypadku stymulacja dotykowa była dostarczana do klatki piersiowej osoby badanej. Również w tej wersji osoba badana za pośrednictwem HMD widziała przed sobą obraz swoich pleców, rejestrowany przez system kamer umieszczony za nią. Tym razem jednak badacz delikatnie uderzał w klatkę piersiową osoby badanej, jednocześnie wykonując zsynchronizowany w czasie ruch imitujący dotykanie wirtualnej klatki piersiowej poniżej pola widzenia kamery, z której obraz był prezentowany badanym. Skutkiem tego osoba badana miała wrażenie, że znajduje się poza własnym ciałem – w miejscu, w którym była zlokalizowana kamera – i z tej perspektywy ogląda swoje ciało. Ponownie zatem doświadczenie to zdaje się nosić znamiona kluczowego aspektu doświadczenia eksterioryzacji, jakim jest przestrzenna dysocjacja poczucia zlokalizowania ciała i perspektywy wzrokowej.

Podsumowując, w przypadku obydwu podejść rezultatem było doświadczenie postrzegania rzeczywistości spoza własnego ciała. Powstaje pytanie, czy rzeczywiście eksperymenty tego typu można uznać za odpowiednik doświadczenia eksterioryzacji, które dotyka między innymi pacjentów po urazach mózgu? Odpowiedź na to pytanie wymaga zdefiniowania zjawiska OBE, jak również wyczerpującego opisu fenomenologii eksterioryzacji. Niemniej zanim badacze ustalą wspólne stanowisko w tej kwestii, warto skupić się na dokładniejszej analizie mechanizmów, które odpowiadają za pojawienie się iluzji wywołanych eksperymentalnie, i porównać je z dotychczasową wiedzą na temat mechanizmów powstawania OBE w warunkach „naturalnych”.

Out-of-brain experience?

Jednym ze wspomnianych już powodów występowania OBE jest uszkodzenie obszarów kory zaangażowanych w scalanie informacji docierających z różnych zmysłów. Analiza lokalizacji lezji u pacjentów doświadczających eksterioryzacji wykazała, że obszarem szczególnie istotnym, którego dysfunkcja skutkuje występowaniem OBE, jest styk skroniowo-ciemieniowy (TPJ) [Blanke i in. 2004]. Dalszego potwierdzenia znaczenia tego obszaru dostarczyły badania, w których elektryczna stymulacja TPJ wywołała doświadczenie eksterioryzacji [Blanke i Thut 2007]. Także mniej inwazyjne badania z wykorzystaniem przeczaskowej stymulacji magnetycznej (*transcranial magnetic stimulation*, TMS) [Blanke i in. 2005] wykazały, że przejściowe zaburzenie funkcjonowania obszaru styku może skutkować problemami z poprawnym postrzeganiem zlokalizowania swojego ciała w przestrzeni.

Niedawno z wykorzystaniem metody zaproponowanej przez laboratorium Blankego powiodło się eksperymentalne wywołanie OBE w skanerze rezonansu magnetycznego, co umożliwiło zlokalizowanie obszarów kory specyficznie aktywowanych podczas doświadczania eksterioryzacji [Ionta i in. 2011]. W eksperymencie tym osoby badane widziały przed sobą plecy wirtualnego ciała, podczas gdy specjalnie skonstruowana platforma robotyczna [Duenas i in. 2011] dostarczała zsynchronizowanej czasowo i przestrzennie (z obserwowaną na wirtualnym ciele) stymulacji doty-

kowej. Manipulacja wywoływała doświadczenie identyfikacji z ciałem wirtualnym oraz poczucie oglądu rzeczywistości z perspektywy pozacielesnej. Wyniki ponownie potwierdziły istotne znaczenie TPJ, którego silniejszą aktywację zarejestrowano podczas doświadczenia eksterioryzacji w porównaniu z warunkiem kontrolnym.

Powtarzając się w badaniach z wykorzystaniem różnych metod i ujęć badawczych wyniki sugerujące, że zaburzenie funkcjonowania styku skroniowo-ciemieniowego może skutkować doświadczeniem eksterioryzacji, doprowadziły badaczy do wniosku, że obszar ten jest szczególnie istotny dla wytworzenia się poprawnego poczucia zlokalizowania ciała w przestrzeni. Zestawiając wyniki badań nad OBE z dotychczasową wiedzą na temat funkcjonalnego znaczenia TPJ, zasugerowano, że eksterioryzacja występująca po uszkodzeniu styku może być wynikiem zaburzenia funkcjonowania mechanizmu scalającego informacje docierające z kluczowych dla postrzegania ciała w przestrzeni zmysłów, tj. wzroku, dotyku, propriocepcji oraz układu przedsionkowego [Blanke i in. 2005; Cheyne i Girard 2009; Ionta i in. 2011]. W szczególności zwraca się uwagę na występujące po uszkodzeniu TPJ problemy z integracją danych wzrokowych z informacjami dopływającymi z organów ucha wewnętrznego – kanałów półkolistych i otolitów [Lopez i in. 2008]. Skutkiem nieprawidłowego scalania tych danych może być błędne poczucie zlokalizowania głowy w stosunku do wektora grawitacji, a w rezultacie zniekształcona perspektywa wzrokowa. Sugestie badaczy dotyczące mechanizmów odpowiedzialnych za wytworzenie doświadczenia eksterioryzacji znajdują odzwierciedlenie w opisaną wcześniej fenomenologii OBE. Osoby w tym stanie postrzegają z zaburzonej w stosunku do rzeczywistej perspektywy oraz często raportują doświadczenia unoszenia się, rotowania etc., które mogą być wynikiem błędów w przetwarzaniu informacji z układu przedsionkowego oraz ich scalaniu z danymi wzrokowymi [Cheyne i Girard 2009].

Czy każdy może uwolnić się od swojego ciała?

Dotychczasowe badania dotyczące mechanizmów powstawania OBE prowadzone były głównie w ujęciu psychologii ogólnej, tj. poszukiwano jednego mechanizmu wyjaśniającego proces powstawania tego stanu zmienionej świadomości. Niemniej badania wykazały, że poszczególne osoby w sposób istotny różnią się podatnością na eksperymentalnie wywołane stany eksterioryzacji [Braithwaite i in. 2013; Ionta i in. 2011]. Stąd w ostatnich latach badacze zaczęli uwzględniać indywidualne zróżnicowanie w tym zakresie, a nawet poszukiwać jego podstaw poznawczych, rozwojowych oraz neuronalnych.

Powstaje pytanie: jakie mogą być podstawowe źródła zróżnicowania tej podatności? Jak dotychczas kwestii tej poświęcono niewiele uwagi, a pierwsze próby odpowiedzi na to pytanie stanowią przedmiot badań kilku zespołów na całym świecie. W niniejszym artykule zaprezentowane zostaną trzy hipotetyczne źródła różnic indywidualnych w podatności na eksperymentalnie wywołane doświadczenie eksterioryzacji.

Pierwszym z nich może być wskazana przez Iontę i współpracowników [2011] preferencja do polegania w procesie lokalizowania siebie w przestrzeni na danych pochodzących z układu wzrokowego, ponad danymi docierającymi z układu przed-

sionkowego. Zważywszy na sposób, w jaki wywołuje się metodami eksperymentalnymi stany typu OBE, tj. poprzez wykorzystanie technologii VR, oddziałujących głównie na zmysł wzroku, można przypuszczać, że efekt, który próbuje się uzyskać, jest próbą wprowadzenia w błąd organów układu przedsionkowego odnośnie do realnego zlokalizowania ciała w stosunku do wektora grawitacji. Stąd osoby, które na co dzień bardziej polegają na danych wzrokowych, powinny ulegać silniejszym iluzjom wywołanym wskazanymi metodami eksperymentalnymi.

Kolejnym hipotetycznym powodem może być odmienna skuteczność mechanizmu integrującego dane pochodzące z układu przedsionkowego i wzrokowego. Innymi słowy, osoby mogą dysponować mechanizmem scalania o zróżnicowanej odporności na konflikt między danymi na temat zlokalizowania w przestrzeni docierającymi z tych dwóch zmysłów. Jednym z parametrów, którego silne zróżnicowanie międzypersoniczne zaobserwowano w badaniach dotyczących integracji polisensorycznej, jest szerokość okna czasowego (*temporal binding window*, TBW) [Stevenson i in. 2012], w ramach którego możliwe jest scalenie informacji docierających z różnych zmysłów w jeden percept. Innymi słowy – maksymalny czas pomiędzy prezentacją dwóch bodźców w dwóch modalnościach, przy którym obydwie bodźce wciąż są postrzegane jako współwystępujące w czasie i przestrzeni. Poparciem dla przypuszczenia, że zróżnicowana szerokość tego okna może mieć znaczenie dla podatności na eksperymentalnie wywołane zmienione stany świadomości cielesnej, są wyniki badań dotyczące osób z zaburzeniami ze spektrum autystycznego (*autism spectrum disorder*, ASD). Z jednej strony badania potwierdzają, że szerokość okna scalania czasowego jest u tych osób wydłużona w stosunku do populacji zdrowej [Foss-Feig i in. 2010]. Z drugiej zaś udało się potwierdzić, że są oni bardziej podatni na wspomnianą wcześniej iluzję gumowej ręki [Cascio i in. 2012].

Oprócz dwóch wymienionych mechanizmów, które można zaliczyć do bardziej podstawowych mechanizmów percepcyjnych, potencjalnym źródłem zróżnicowania może być także odmienna motywacja do osiągnięcia zmienionych stanów świadomości [Braithwaite i in. 2010]. Innymi słowy, niektóre osoby mogą aktywnie dążyć do osiągnięcia takich stanów, stąd być bardziej wyczulone na nieścisłości w przetwarzaniu danych sensorycznych i błędy integracji polisensorycznej. W rezultacie osoby takie mogą doświadczać OBE częściej i łatwiej.

Podsumowując, rozwój najnowszych metod badawczych umożliwił naukowe i systematyczne badania nad OBE z udziałem osób zdrowych, niepozostających pod wpływem środków psychoaktywnych. Obserwowane zróżnicowanie międzyludzkie w podatności na eksperymentalnie wywołane OBE prowokuje do pytań o podłoże tych odmienności. W niniejszym artykule zaproponowano trzy hipotetyczne źródła tegoż zróżnicowania, z których dwa dotyczą prostych procesów percepcyjnych, a trzecie – wyższych procesów poznawczych, w tym uwagi i motywacji. Na zakończenie należy jednak podkreślić, że powyższe propozycje dotyczące podstaw różnic indywidualnych w podatności na eksperymentalnie wywołane OBE mają charakter wysoce spekulacyjny i wyciągnięcie bardziej uzasadnionych wniosków wymaga przeprowadzenia cyklu badań weryfikujących te założenia.

BIBLIOGRAFIA

- Aspell J., Blanke O. (2009). *Understanding the out-of-body experience from a neuroscientific perspective* [w:] D. Murray (red.), *Psychological Scientific Perspectives on Out of Body and Near Death Experiences* (s. 73–88). New York: Nova Science Publishers Inc.
- Aspell J., Lenggenhager B., Blanke O. (2011). *Multisensory perception and bodily self-consciousness: from out-of-body to inside-body experience* [w:] M. Murray, M. Wallace, *The Neural Bases of Multisensory Processes, Frontiers in Neuroscience*, rozdz. 24 [online]. URL: <http://www.ncbi.nlm.nih.gov/books/NBK92870/> (data pobrania: 15.09.2013).
- Blackmore S. (1982). *Beyond the Body: An Investigation of Out-of-the-Body Experiences*. London: Heinemann (wyd. I).
- Blackmore S. (1978). *Parapsychology and Out-of-the-Body Experiences*. Hove: Transpersonal Books.
- Blackmore S. (1987). *Where am I? Perspectives in imagery and the out-of-body experience*. „Journal of Mental Imagery” 11 (2), s. 53–66.
- Blanke O. (2012). *Multisensory brain mechanisms of bodily self-consciousness*. „Nature Reviews Neuroscience” (July), s. 556–571.
- Blanke O., Arzy S. (2005). *The out-of-body experience: disturbed self-processing at the temporoparietal junction*. „The Neuroscientist” 11 (1), s. 16–24.
- Blanke O., Dieguez S. (2009). *Leaving body and life behind. Out-of-body and near-death experiences* [w:] S. Laureys, G. Tononi (red.), *The Neurology of Consciousness* (s. 303–325). London: Elsevier Ltd.
- Blanke O., Mohr C. (2005). *Out-of-body experience, heautoscopy, and autoscopic hallucination of neurological origin. Implications for neurocognitive mechanisms of corporeal awareness and self-consciousness*. „Brain Research. Brain Research Reviews” 50 (1), s. 184–199.
- Blanke O., Mohr C., Michel C.M., Pascual-Leone A., Brugger P. (2005). *Linking out-of-body experience and self processing to mental own-body imagery at the temporoparietal junction*. „The Journal of Neuroscience” 25 (3), s. 550–557.
- Blanke O., Thut G. (2007). *Inducing out-of-body experiences* [w:] S.D. Salla (red.), *Tall Tales about the Mind and Brain: Separating Fact from Fiction* (s. 425–439). Oxford: Oxford University Press.
- Botvinick M., Cohen J. (1998). *Rubber hands “feel” touch that eyes see*. „Nature” 391 (6669), s. 756.
- Braithwaite J.J., Brogna E., Bagshaw A.P., Wilkins A.J. (2013). *Evidence for elevated cortical hyperexcitability and its association with out-of-body experiences in the non-clinical population: New findings from a pattern-glare task*. „Cortex” 49 (3), s. 793–805.
- Braithwaite J.J., Dent K. (2011). *New perspectives on perspective-taking mechanisms and the out-of-body experience*. „Cortex” 47 (5), s. 628–632.
- Braithwaite J.J., Samson D., Apperly I., Brogna E., Hulleman J. (2010). *Cognitive correlates of the spontaneous out-of-body experience (OBE) in the psychologically normal population: evidence for an increased role of temporal-lobe instability, body-distortion processing, and impairments in own-body transformations*. „Cortex” 47 (7), s. 839–853.
- Brugger P. (2002). *Reflective mirrors: perspective-taking in autoscopic phenomena*. „Cognitive Neuropsychiatry” 7, s. 179–194.
- Brugger P., Regard M., Landis T. (1997). *Illusory reduplication of one's own body: Phenomenology and classification of autoscopic phenomena*. „Cognitive Neuropsychiatry” 2 (1), s. 19–38.
- Cardeña E., Krippner S., Powers S.M. (1997). *The Etiologies of Dissociation. Broken Images, Broken Selves*. New York, NY: Brunner/Mazel.
- Cascio C., Foss-Feig J., Burnette C., Heacock J., Cosby A. (2012). *The rubber hand illusion in children with autism spectrum disorders: delayed influence of combined tactile and visual input on proprioception*. „Autism” 16 (4), s. 406–419.

- Cheyne J.A., Girard T. (2009). *The body unbound: vestibular-motor hallucinations and out-of-body experiences*. „Cortex” 45 (2), s. 201–215.
- Corbetta M., Patel G., Shulman G.L. (2008). *The reorienting system of the human brain: from environment to theory of mind*. „Neuron” 58 (3), s. 306–324.
- Decety J., Lamm C. (2007). *The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition*. „The Neuroscientist” 13 (6), s. 580–593.
- Decety J., Sommerville J. (2003). *Shared representations between self and other: a social cognitive neuroscience view*. „Trends in Cognitive Sciences” 7 (12), s. 527–533.
- Duenas J., Chapuis D., Pfeiffer C., Martuzzi R., Ionta S., Blanke O., Gassert R. (2011). *Neuroscience robotics to investigate multisensory integration and bodily awareness*. Conference Proceeding IEEE Engineering in Medicine Biology Society, s. 8348–8352.
- Easton S., Blanke O., Mohr C. (2009). *A putative implication for fronto-parietal connectivity in out-of-body experiences*. „Cortex” 45 (2), s. 216–227.
- Ehrsson H.H. (2007). *The experimental induction of out-of-body experiences*. „Science” 317 (5841), s. 1048.
- Foss-Feig J.H., Kwakye L.D., Cascio C.J., Burnette C.P., Kadivar H., Stone W.L., Wallace M.T. (2010). *An extended multisensory temporal binding window in autism spectrum disorders*. „Experimental Brain Research” 203 (2), s. 381–389.
- Heydrich L., Lopez C., Seeck M., Blanke O. (2011). *Partial and full own-body illusions of epileptic origin in a child with right temporoparietal epilepsy*. „Epilepsy & Behavior” 20 (3), s. 583–586.
- Holmes N.P., Spence C. (2006). *Beyond the body schema. Visual, prosthetic and technological contributions to bodily perception and awareness*, [w:] G. Knoblich, I.M. Thornton, M. Grosjean, M. Shiffrar, *Human Body Perception from the Inside Out*. New York: Oxford University Press.
- Ionta S., Heydrich L., Lenggenhager B., Mouthon M., Fornari E., Chapuis D., Gassert R. i in. (2011). *Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective*. „Neuron” 70 (2), s. 363–374.
- Jimura K., Cazalis F., Stover E.R.S., Poldrack R. (2014). *The neural basis of task switching changes with skill acquisition*. „Frontiers in Human Neuroscience” 8 (May), s. 339.
- Kornilova L.N. (1997). *Orientation illusions in spaceflight*. „Journal of Vestibular Research” 7, s. 429–439.
- Lenggenhager B., Tadi T., Metzinger T., Blanke O. (2007). *Video ergo sum: manipulating bodily self-consciousness*. „Science” 317 (5841), s. 1096–1099.
- Linden D.E.J., Kallenbach U., Heinecke A., Singer W., Goebel R. (1999). *The myth of upright vision. A psychophysical and functional imaging study of adaptation to inverting spectacles*. „Perception” 28 (4), s. 469–481.
- Lopez C., Blanke O. (2011). *The thalamocortical vestibular system in animals and humans*. „Brain Research Reviews” 67 (1–2), s. 1–28.
- Mars R.B., Sallet J., Schüffelgen U., Jbabdi S., Toni I., Rushworth M.F.S. (2012). *Connectivity-based subdivisions of the human right “temporoparietal junction area”: Evidence for different areas participating in different cortical networks*. „Cerebral Cortex” 22 (8), s. 1894–1903.
- Monroe R. (1971). *Journeys Out of the Body*. New York: Doubleday (wyd. I).
- Nagy E., Bognár L., Csengery A., Almási A., Bencze G. (2000). *Effect of microgravitation on the human equilibrium*. „The International Tinnitus Journal” 6 (2), s. 120–123.
- Podoll K., Robinson D. (1999). *Out-of-body experiences and related phenomena in migraine art*. „Cephalalgia” 19, s. 886–896.
- Redding G.M., Rossetti Y., Wallace B. (2005). *Applications of prism adaptation: A tutorial in theory and method*. „Neuroscience and Biobehavioral Reviews” 29 (3), s. 431–444.
- Rohde M., di Luca M., Ernst M.O. (2011). *The rubber hand illusion: Feeling of ownership and proprioceptive drift do not go hand in hand*. „PLoS ONE” 6 (6). doi: 10.1371/journal.pone.0021659.

- SKYbrary (2013). *Vestibular System and Illusions. Operator's Guide to Human Factors in Aviation: Human Performance and Limitation* [online]. URL: [http://www.skybrary.aero/index.php/Vestibular_System_and_Illusions_\(OGHFA_BN\)](http://www.skybrary.aero/index.php/Vestibular_System_and_Illusions_(OGHFA_BN)) (data pobrania: 15.09.2013).
- Stevenson R.A., Zemtsov R.K., Wallace M.T. (2012). *Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions*. „Journal of Experimental Psychology: Human Perception and Performance” 38 (6), s. 1517–1529.
- Stratton G.M. (1899). *The spatial harmony of touch and sight*. „Mind” 8, s. 492–505.
- Stumbrys T., Erlacher D., Schädlich M., Schredl M. (2012). *Induction of lucid dreams: a systematic review of evidence*. „Consciousness and Cognition” 21 (3), s. 1456–1475.
- Uddin L.Q., Molnar-Szakacs I., Zaidel E., Iacoboni M. (2006). *rTMS to the right inferior parietal lobule disrupts self-other discrimination*. „Social Affective & Cognitive Neuroscience” 1, s. 65–71.
- Ustinova Y. (2011). *Consciousness alteration practices in the West from prehistory to late antiquity*, [w:] E. Cardena, M. Winkelman (red.), *Altering Consciousness. Multidisciplinary Perspectives* (s. 1–45). Westport: Praeger.
- Vignemont F. de (2010). *Body schema and body image – pros and cons*. „Neuropsychologia” 48 (3), s. 669–680.
- Wilde D.J. (2011). *Finding meaning in out-of-body experiences: an interpretative phenomenological analysis*. PhD thesis, the University of Manchester, Faculty of Medical and Human Sciences.
- Wilkins L.K., Girard T., Cheyne J.A. (2011). *Ketamine as a primary predictor of out-of-body experiences associated with multiple substance use*. „Consciousness and Cognition” 20 (3), s. 943–950.
- Zubek J.P. (red.) (1969). *Sensory Deprivation: Fifteen Years of Research*. New York: Appleton-Century-Crofts.

„Sunny side is up”, but is the cloudy side down? Znaczenie wymiaru wertykalnego w konceptualizacji emocji

„Sunny side is up”, but is the cloudy side down? Significance of vertical dimension in conceptualisation of emotions

Abstract: One of the key research areas in the embodied cognition field is role of metaphors in creating abstract notions. One of such metaphors is vertical dimension (UP-DOWN) used e.g. for conceptualising positive and/or negative emotions. Its importance has been confirmed by many empirical findings, but some of them arouse methodological concerns regarding the stimuli selection and the level to which observed patterns are universal. The main goal of the present study was to replicate findings of one of the most influential experiments in the field. Its results pointed that positive stimuli are processed faster when presented on the top of the screen while negative ones are processed faster when presented on the bottom. The results of our study yielded a slightly different pattern: positive stimuli were indeed processed faster on top of the screen, but we did not replicate faster processing of negative stimuli on the bottom of the screen. Possible explanations of such a pattern of results were discussed, along with ideas for follow up studies.

Key words: embodied cognition, metaphors, vertical dimension, conceptualisation of emotions

Wprowadzenie

Zagadnienie genezy i charakteru reprezentacji umysłowych znajduje się w centrum zainteresowania szeroko pojętych nauk kognitywnych. Przedstawiciele nurtu poznania ucieleśnionego (*embodied cognition*) uważają, że nawet reprezentacje pojęć abstrakcyjnych, rozumianych jako niemające fizycznych desygnatów [patrz: Barsalou 1999; Pecher, Boot i Van Dantzig 2011], powstają na bazie doświadczeń sensomotorycznych. Przejście od konkretnego doświadczenia do umysłowej reprezentacji abstrakcyjnego pojęcia wyjaśniane jest w tym ujęciu za pomocą teoretycznego konstruktów *schematu wyobraźniowego*. *Schemat wyobraźniowy* definiowany jest jako umysłowa reprezentacja powstała na bazie wielokrotnie przeżywanego zmysłowego doświadczenia, stanowi punkt odniesienia i nadaje strukturę pojęciu abstrakcyjnemu [Gibbs 2008; Johnson 1997; Talmy 1988]. Umysłowe reprezentacje, będące podstawą

procesów poznawczych, są więc na stałe powiązane ze swoją percepcyjną podstawą i zakorzenione w działaniu [Lakoff i Johnson 1999], co znajduje potwierdzenie w badaniach laboratoryjnych [np. Glenberg i Kaschak 2002; Pecher, Zeelenberg i Barsalou 2003; Pecher i in. 2011].

Większość badań prowadzonych w nurcie ucieleśnionego poznania skupiała się dotąd na reprezentacjach pojęć konkretnych, coraz częściej pojawiają się jednak prace badające związek między fizycznym doświadczeniem a reprezentacjami pojęć abstrakcyjnych. Podczas gdy związek pojęć konkretnych z ucieleśnionym doświadczeniem wydaje się intuicyjnie zrozumiały (poruszając się w przestrzeni, wchodzimy w interakcje z istniejącymi materialnie obiektami), to związek pojęć abstrakcyjnych, nieposiadających dostępnych zmysłom desygnatów, z doświadczeniem sensomotorycznym wydaje się zaskakujący [o roli kontekstu w powstawaniu pojęć: Yeh i Barsalou 2006]. Przedstawiciele nurtu twierdzą jednak, że ludzki umysł nie jest zdolny do poznawania w sposób całkowicie niezależny od świata materialnego.

Teoria metafory pojęciowej [Johnson 1987; Lakoff i Johnson 1999] doskonale wpisuje się w nurt ucieleśnionego poznania, bardzo mocno akcentując fakt, że reprezentacje wszelkich pojęć uzależnione są od budowy i funkcjonowania ludzkiego ciała oraz jego interakcji z otoczeniem. W myśl tej teorii abstrakcyjne pojęcia mają być pojmovalne przede wszystkim za pośrednictwem metafor, odnoszących i porównujących je do doświadczenia fizycznego [Załaźńska 2001]. Powstawanie połączeń pomiędzy konkretnymi i abstrakcyjnymi pojęciami dokonuje się od najwcześniejszego dzieciństwa, kiedy doświadcza się współwystępowania pewnych zjawisk [Johnson 1989; Boot i Pecher 2010]. Na przykład obserwacja, że podobne do siebie obiekty są często zgromadzone obok siebie – podobne klocki w pudełku, podobne kwiaty na klombie, podobne łyżki w szufladzie – tworzy podstawę konceptualizacji PODOBIENSTWO TO BLISKOŚĆ [Lakoff i Johnson 1999], na poparcie której Boot i Pecher [2010] zgromadziły dane eksperymentalne. Podobnie postrzeganie rodzica jako osoby jednocześnie wysokiej i posiadającej duży wpływ na dziecko pozwala na ukształtowanie się metafory WYSOKI/W GÓRĘ TO SILNY, POSIADAJĄCY WŁADZĘ [Lakoff i Johnson 1999]. Cykl eksperymentów przeprowadzonych przez Schuberta [2005] pokazał, że ludzie istotnie używają schematu wertykalnego, myśląc o relacjach społecznych, oraz że leży on u podstaw takich abstrakcyjnych pojęć, jak status społeczny, siła, władza. Co więcej, okazało się, że manipulowanie nierelevantną dla wykonania zadania zmienną (wskazówką wizualną, sposobem udzielania odpowiedzi) znacząco wpływało na szybkość i poprawność oceny bodźca: osoby badane szybciej i poprawniej reagowały na etykiety silnych grup społecznych prezentowane na górze ekranu komputera lub gdy prośzone były o udzielnie odpowiedzi za pomocą skierowanej w górę strzałki, a na etykiety grup słabych, kiedy te pokazywano na dole ekranu lub gdy reakcją miało być naciśnięcie strzałki skierowanej w dół.

Inspiracją do badań Schuberta były wyniki uzyskane przez Meiera i Robinsona [2004], dotyczące konceptualizacji GÓRA/W GÓRĘ TO DOBRO, DÓŁ/W DÓŁ TO ZŁO. W ich eksperymencie zadaniem badanych była ocena walencji słów (pozytywne *versus* negatywne), prezentowanych u góry bądź u dołu ekranu komputera. Zaobserwowano, że bodźce pozytywne pokazane u góry ekranu oceniano szybciej, niż gdy były one wyświetlane u dołu, natomiast reakcje na bodźce negatywne okazały

się szybsze w przypadku ich prezentacji u dołu ekranu w porównaniu z prezentacją u góry. Ponadto bodźce pozytywne były przetwarzane szybciej niż negatywne, niezależnie od lokalizacji (góra lub dół).

Jakkolwiek teoretyczne rozważania przedstawicielei nurtu poznania ucieleśnionego znajdują poparcie w danych eksperymentalnych, w ostatnim czasie rośnie liczba tekstów krytykujących owo podejście, proponujących alternatywne wyjaśnienia uzyskanych wyników i wskazujących na niedostatki niektórych eksperymentów [Adams 2010; Mahon i Caramazza 2008; Wilson 2002]. Jednym z najlepszych sposobów weryfikacji tez głoszonych przez badaczy reprezentujących nurt poznania ucieleśnionego wydają się więc prowadzone z dużą dbałością o poprawność metodologiczną próby replikacji przeprowadzanych przez nich eksperymentów [Wicherts i Bakker 2012; Wicherts, Borsboom, Kats i Molenaar 2006].

W niniejszym artykule prezentowane są rezultaty rozszerzonej replikacji eksperymentu Meiera i Robinsona [2004]. Ich praca traktuje o jednej z fundamentalnych metafor pojęciowych: GÓRA/W GÓRĘ TO DOBRO, DÓŁ/W DÓŁ TO ZŁO. Tekst przyciągnął uwagę środowiska (298 cytowań wg Google Scholar) i spotkał się z przychylnym odbiorem, można w nim jednak odnaleźć pewne niedostatki metodologiczne. Po pierwsze, nie kontrolowano długości i frekwencji bodźców użytych w eksperymencie. Z tym zaniedbaniem wiąże się ryzyko, że przynajmniej część zaobserwowanych efektów może wynikać z szybszego przetwarzania bodźców o wysokiej frekwencji [Brybaert 1996]. Po drugie, Meier i Robinson [2004] użyli przemieszanych z sobą bodźców konkretnych i abstrakcyjnych, przez co trudniej zweryfikować faktyczny związek pojęć wyłącznie abstrakcyjnych z fizycznym wymiarem wertykalnym. Po trzecie, choć Meier i Robinson podnoszą uniwersalność swoich wyników dla wszystkich kultur i języków [2004; 2005], nie wskazują badań, w których uniwersalność ta zostałaby potwierdzona. Po czwarte, w niektórych prezentowanych w swoim tekście eksperymentach Meier i Robinson [2004] uzyskują szybsze przetwarzanie bodźców pozytywnych w porównaniu z negatywnymi, a w innych wzór wyników jest odwrotny, lecz rozbieżność ta nie jest w żaden sposób skomentowana. Po piąte, walencja bodźców emocjonalnych jest tylko jedną z wielu charakterystyk mających wpływ na ich przetwarzanie [np. Smith i Ellsworth 1985, wyróżniają sześć ortogonalnych cech bodźców emocjonalnych, które mają wpływ na ich przetwarzanie]. W naszym eksperymencie uwzględniliśmy więc wymiar AKTYWNOŚĆ–PASYWNOŚĆ, by sprawdzić, czy powiązanie bodźców emocjonalnych z wymiarem GÓRA–DÓŁ zachodzi wyłącznie ze względu na ich walencję, czy też inne cechy tych bodźców również wiążą się z wymiarem wertykalnym. Wymiar AKTYWNOŚĆ–PASYWNOŚĆ można łączyć bądź z intensywnością motywacji, które wzbudzają dane emocje [Gable i Harmon-Jones 2010], bądź z pobudzeniem, które powodują [Vogt, De Houwer, Koster, Van Damme i Crombez 2008].

Badanie Meiera i Robinsona [2004] poddano replikacji w dwóch eksperymentach. W pierwszym użyto bodźców identycznych z wersją oryginalną, ale przetłumaczonych na język polski, będących mieszkanką wyrazów konkretnych i abstrakcyjnych. W drugim natomiast użyto tylko wyrazów abstrakcyjnych oznaczających stany emocjonalne, dodatkowo wprowadzając podział na podstawie wymiaru AKTYWNOŚĆ–PASYWNOŚĆ [Anooshian i Hertel 1994]. Wymiar ten jest wyraźnie obecny w języku naturalnym, ale nie był dotychczas przedmiotem wielu badań eks-

perymentalnych. Pomimo to wymiar jest obecny w kluczowych koncepcjach emocji, akcentując rolę wysiłku lub zaangażowania motywacji do zbliżania bądź oddalania się od źródła emocji [Gable i Harmon-Jones 2010; Smith i Ellsworth 1985]. Powiązanie wymiaru AKTYWNOŚĆ–PASYWNOŚĆ z wymiarem GÓRA–DÓŁ nie jest jednak oczywiste. Na przykład *złość* stanowi emocję negatywną, powinna więc być kojarzona „z dołem”, jednakże najczęściej skłania ona też do działania, co może wiązać się ze zbliżaniem i „górami”. Planując nasz eksperyment, założyliśmy, że wymiar AKTYWNOŚĆ–PASYWNOŚĆ może podlegać podobnej asymetrii przetwarzania jak wymiar DOBRY–ZŁY, w zależności od usytuowania bodźców w przestrzeni werbalnej (u podstaw tej asymetrii leżałaby wówczas metafora RUCH/AKTYWNOŚĆ TO DOBRO, BEZRUCH/PASYWNOŚĆ TO ZŁO). Procedura badawcza obu eksperymentów była identyczna z oryginalną.

Eksperyment 1

Eksperyment miał na celu zreplicowanie na bodźcach polskojęzycznych wyników uzyskanych przez Meiera i Robinsona [2004]. Przewidywano, że czasy reakcji dla bodźców pozytywnych będą średnio krótsze niż dla negatywnych oraz że zostanie zaobserwowana kluczowa dla eksperymentu interakcja walencji i lokalizacji. Założono, że reakcje na bodźce pozytywne będą szybsze wtedy, kiedy bodźce te zostaną zaprezentowane na górze, a nie na dole. Odwrotny wzorec wyników przewidywano dla bodźców negatywnych, które będą przetwarzane szybciej, jeśli zostaną zaprezentowane na dole ekranu, w porównaniu z warunkami, w których będą prezentowane na górze.

Metoda

Osoby badane

Zbadano 26 osób (w tym 16 kobiet, średni wiek: 22,4) będących studentami jednej z krakowskich uczelni i rodzimymi użytkownikami języka polskiego. Wszyscy badani byli ochotnikami, za udział w badaniu uczestnicy otrzymywali punkty potrzebne do zaliczenia nieobowiązkowego przedmiotu „Udział w badaniach naukowych w psychologii”.

Procedura

Każda próba eksperymentalna (*trial*) zaczynała się od prezentacji trzech krzyżyków fiksacyjnych. Każdy z nich był obecny na ekranie przez 300 ms: pierwszy wyświetlano na środku ekranu, kolejny pojawiał się poniżej lub powyżej poprzedniego, a ostatni tuż pod górną lub tuż nad dolną krawędzią ekranu. W ten sposób krzyżyki naprowadzały wzrok badanych na miejsce pojawienia się bodźca, zapobiegając losowym ruchom gałek ocznych, które mogłyby dodać dodatkową wariację błędów do czasów

reakcji [Meier i Robinson 2004; Tzelgov, Henik i Leiser 1990]. Po trzecim krzyżyku fiksacyjnym pojawiał się bodziec, który był obecny na ekranie przez 3000 ms lub do momentu udzielenia przez badanego odpowiedzi. Po bodźcu pojawiał się na 500 ms pusty ekran (*blank screen*), by przygotować badanych na kolejny bodziec. Każdy ze 100 bodźców pojawiał się dwukrotnie w dwóch lokalizacjach, cała procedura składała się więc z 400 prób (100 bodźców x 2 lokalizacje x 2 powtórzenia). Słowa były napisane czarną czcionką i prezentowane na białym tle. Zadaniem osób badanych było zdecydować, czy pokazane słowo jest pozytywne, czy negatywne. Badani udzielali odpowiedzi, naciskając odpowiedni przycisk na klawiaturze. Instrukcja kładła nacisk zarówno na szybkość, jak i poprawność odpowiedzi. Procedura została napisana i była odtwarzana w programie DMDX [Forster i Forster 2003].

Materiały

Jako bodźców do zadania użyto przetłumaczonych na język polski słów zaczerpniętych z badania Meiera i Robinsona [2004]. W oryginalnym badaniu prezentowano słowa w języku angielskim. Analizie zostały poddane długość i frekwencja przetłumaczonych słów, obliczona przy użyciu bazy Subtlex-PL [Mandera, Wodniecka i Keuleers 2012]. Słowa pozytywne miały w języku polskim wyższą frekwencję niż słowa negatywne ($t(95) = 3,01, p < 0,01$), były jednak porównywalnej długości, co słowa negatywne ($t < 1$; zob. Tabela 1). Lista przykładowych bodźców znajduje się w Aneksie 1 na końcu tekstu.

Tabela 1. Średnie dla poszczególnych grup bodźców. W nawiasie odchylenie standardowe

	Pozytywne	Negatywne
Frekwencja	0,96 (0,71)	0,51 (0,75)
Długość	7,66 (2,63)	7,72 (2,53)

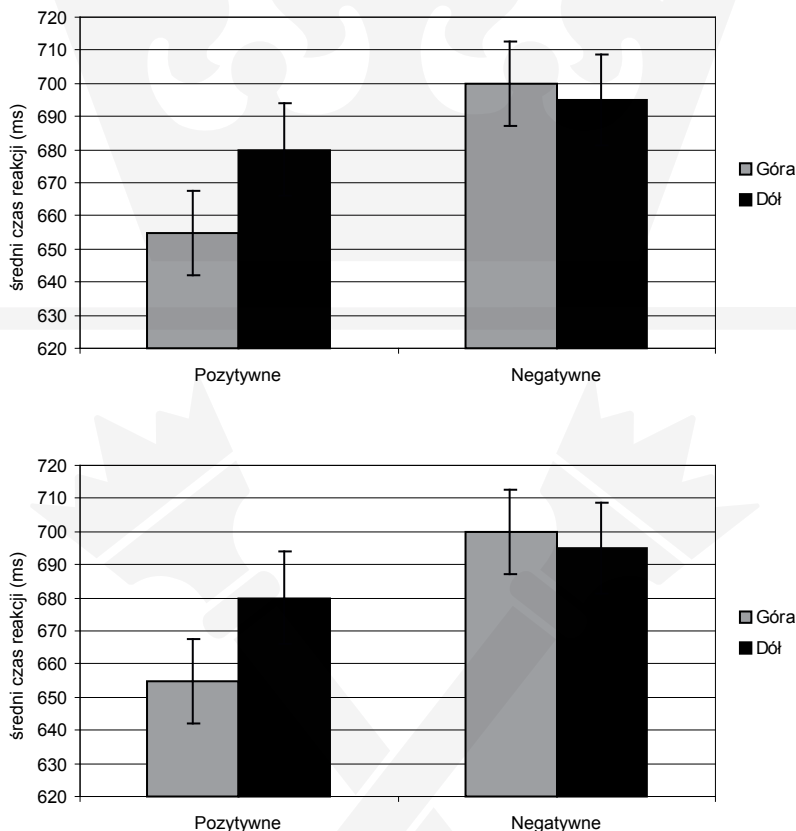
Frekwencja podana w log10 wystąpień na milion wyrazów w korpusie. Długość podana w liczbie liter.

Wyniki

Celem normalizacji rozkładu średnie czasy reakcji zlogarytmizowano, jednak dla ułatwienia interpretacji dane podane są w milisekundach. Czasy reakcji osób odstających od średniej dla wszystkich obserwacji i warunków (*grand mean*) o ponad 2,5 odchylenia standardowego nie zostały poddane analizie, co wyeliminowało pięć z 26 badanych osób [Meier i Robinson 2004; Ratcliff 1993]. Dalsze analizy zostały przeprowadzone na podstawie danych pochodzących od 21 osób.

Efekt główny walencji okazał się istotny ($F(1,20) = 40,69, p < 0,001, \eta^2_p = 0,67$), badani reagowali szybciej na bodźce pozytywne (662 ms) niż negatywne (704 ms). Efekt interakcji walencji i lokalizacji również był istotny ($F(1,20) = 11,62, p < 0,01, \eta^2_p = 0,37$). Analiza efektów prostych wykazała, że bodźce pozytywne prezentowane

na górze ekranu były przetwarzane szybciej niż te pokazywane na dole ($F(1,20) = 8,23$, $p < 0,01$, $\eta_p^2 = 0,29$), a bodźce negatywne były przetwarzane tak samo szybko, niezależnie od lokalizacji ($F < 1$). Efekt główny lokalizacji był nieistotny ($F < 1$). Badani popełnili znikomą liczbę błędów ($< 1\%$), nie analizowano więc poprawności reakcji.



Wykres 1. Średnie czasy reakcji jako funkcja walencji i lokalizacji w Eksperymentie 1. Słupki błędów wskazują na błąd standardowy średniej w każdym warunku

Dyskusja

Porównanie uzyskanych danych z replikowanym eksperymentem [Meier i Robinson 2004] pokazuje, że w nowym eksperymencie uzyskano właściwie bliźniacze wyniki poza jednym istotnym aspektem: w naszym eksperymencie bodźce negatywne nie były istotnie szybciej przetwarzane na dole niż na górze ekranu. Mała wielkość tego efektu ($\eta_p^2 = 0,04$) sugeruje, że brak istotnej różnicy między warunkami nie był spowodowany zbyt małą próbą osób badanych. Przy użyciu programu PS [Dupont

i Plummer 1998] obliczono, że uzyskana różnica między warunkami (7,1 ms) byłaby istotna, jeśli w eksperymencie wzięłyby udział 244 osoby (przy założeniu mocy testu $\pi = 0,80$ i oczekiwanego poziomu istotności różnicy $\alpha = 0,05$). Wynik ten pokazuje, że to raczej efekt jest bardzo nikły, a nie że zbadano zbyt mało osób, by go wykryć. Składnia to do zadania pytania, czy asymetria GÓRA–DÓŁ tak samo wpływa na bodźce o różnej walencji. Szybsze reakcje na wyrazy pozytywne niż negatywne są prawdopodobnie w pewnej mierze efektem ich wyższej frekwencji [np. Brysbaert 1996].

Eksperyment 2

Celem badania była replikacja rezultatów uzyskanych w Eksperymencie 1 na innym materiale leksykalnym oraz rzucenie nowego światła na interakcję walencji i lokalizacji, kluczową dla teorii znaczenia wymiaru wertykalnego. Powtórzenie eksperymentu z nowymi bodźcami miało na celu rozstrzygnięcie, czy wyrazy negatywne podlegają takiej samej asymetrii poznawczej przy użyciu konceptualizacji wymiaru GÓRA–DÓŁ, jak wyrazy pozytywne. Ponadto Eksperyment 2 miał odpowiedzieć na pytanie, czy wymiar AKTYWNOŚĆ–PASYWNOŚĆ podlega podobnej asymetrii przetwarzania na wymiarze GÓRA–DÓŁ jak wymiar DOBRO–ZŁO. Rozstrzygnięcie tej hipotezy badawczej jest istotne dla współczesnych teorii emocji, gdyż odnosi się do postulowanej wielowymiarowości bodźców emocjonalnych [Smith i Ellsworth 1985]. Przewidywano, że wyrazy odnoszące się do aktywnych stanów emocjonalnych („aktywne”) będą przetwarzane szybciej niż wyrazy odnoszące się do stanów pasywnych („pasywne”) oraz że wyrazy aktywne będą przetwarzane szybciej wtedy, kiedy zostaną zaprezentowane na górze ekranu, a wyrazy pasywne – gdy będą pokazywane na dole, co wyrazi się w interakcji aktywności i lokalizacji. Spodziewano się również interakcji aktywności z walencją, czyli szybszych reakcji na bodźce aktywne niż na pasywne.

Metoda

Osoby badane

Zbadano 42 osoby (w tym 26 kobiet, średni wiek: 23,2) będące studentami jednej z krakowskich uczelni i rodzimymi użytkownikami języka polskiego. Żadna z osób nie brała wcześniej udziału w Eksperymencie 1. Wszyscy badani byli ochotnikami, za udział w badaniu otrzymywali punkty potrzebne do zaliczenia nieobowiązkowego przedmiotu „Udział w badaniach naukowych w psychologii”.

Procedura

Procedura była analogiczna do tej użytej w Eksperymencie 1, wyjąwszy liczbę prób eksperymentalnych. W Eksperymencie 2 było ich 160, każdy z 40 bodźców prezentowano dwa razy w dwóch lokalizacjach (40 bodźców \times 2 lokalizacje \times 2 powtórzenia).

Materiały

Jako bodźców w zadaniu użyto 40 wyrazów wybranych z listy 80 polskich słów opisujących stany emocjonalne. Bodźce zostały wybrane na podstawie pilotażu: ankiety, w której wzięło udział 27 osób (w tym 15 kobiet, średnia wieku: 25,11) – studentów polskich uczelni, rodowitych użytkowników języka polskiego. Ankietowani oceniali emocje jako „pozytywne” lub „negatywne” oraz jako „wyzwalające chęć działania” lub „niewyzwalające chęci działania” [porównaj: Anooshian i Hertel 1994; Sutton, Altarriba, Gianico i Basnight-Brown 2007]. Następnie bodźce zostały przyporządkowane do czterech grup: emocji pozytywnych aktywnych (w skrócie POZA, np. euforia), pozytywnych pasywnych (POZP, np. duma), negatywnych aktywnych (NEGA, np. furia) i negatywnych pasywnych (NEGP, np. apatia), tak by stworzyć możliwie najbardziej jednorodną grupę. Średnie ocen „pozytywności” i „wyzwalania chęci działania” dla poszczególnych grup znajdują się w Tabeli 2. Oceny poszczególnych grup bodźców na obu wymiarach różnią się od siebie, jedynie emocje pozytywne aktywne i pasywne są takie same pod względem oceny pozytywności.

Tabela 2. Średnie z pilotażu dla poszczególnych grup bodźców. W nawiasie odchylenie standardowe. Obie zmienne podane w ocenach z kwestionariusza, skala 0–10

	POZA	POZP	NEGA	NEGP
Pozytywność	8,80 (0,49)	8,41 (0,68)	2,51 (1,32)	1,25 (0,92)
Aktywność	8,26 (0,58)	5,73 (1,06)	7,05 (0,60)	1,79 (0,73)

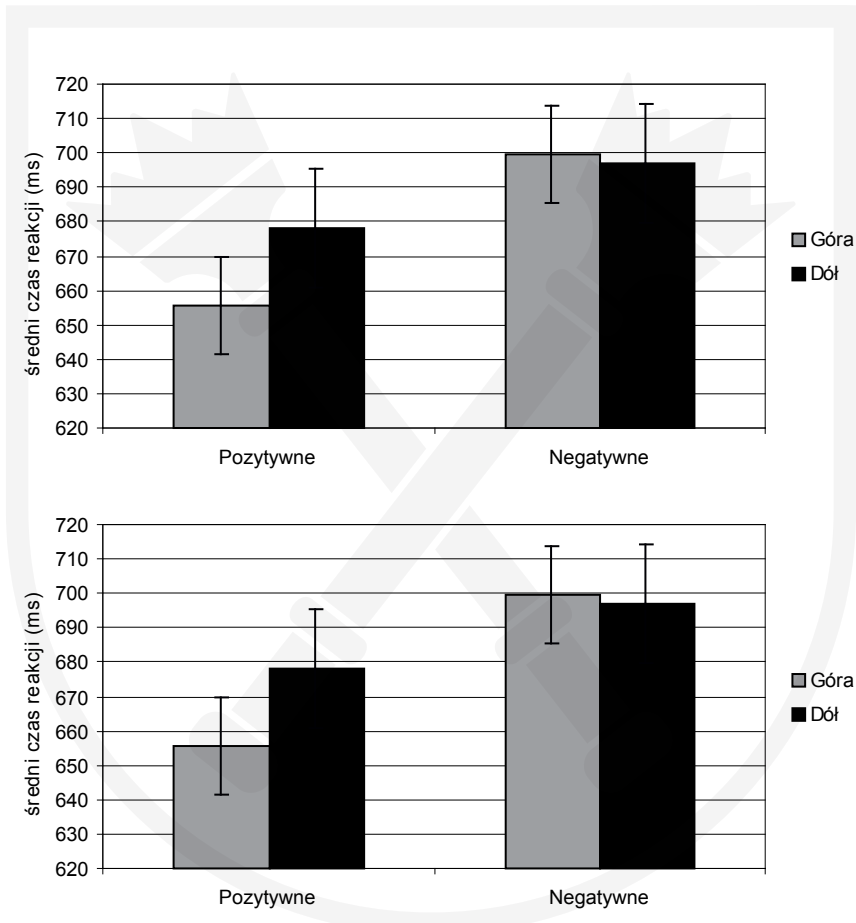
Zmierzono również średnią długość i frekwencję słów (użyto bazy Subtlex-PL; Mander i in. 2012). Jedynie średnia frekwencja dla bodźców NEGP różniła się istotnie od frekwencji dla bodźców NEGA. Pozostałe różnice pomiędzy grupami były nieistotne. Średnia frekwencja dla bodźców pozytywnych, bez względu na aktywność (POZ), nie różniła się istotnie od średniej frekwencji bodźców negatywnych (NEG) ($t < 1$). Podobnie frekwencja bodźców aktywnych (A), bez względu na ich walencję, nie była istotnie różna od frekwencji bodźców pasywnych (P) ($p > 0,1$), natomiast bodźce aktywne były krótsze od wyrazów pasywnych, choć różnica nie osiągnęła statystycznej istotności ($p = 0,07$). Pełne zestawienie frekwencji dla poszczególnych grup można uzyskać u pierwszego autora. Lista bodźców znajduje się w Aneksie 2 na końcu artykułu.

Wyniki

Celem normalizacji rozkładu średnie czasy reakcji zlogarytmizowano, jednak dla ułatwienia interpretacji dane podane są w milisekundach [Meier i Robinson 2004; Ratcliff 1993].

Efekt główny walencji okazał się istotny ($F(1,41) = 40,05, p < 0,001, \eta_p^2 = 0,49$), bodźce pozytywne (668 ms) były przetwarzane szybciej niż negatywne (699 ms). Podobnie istotny był efekt główny aktywności ($F(1,41) = 38,48, p < 0,001, \eta_p^2 = 0,48$), słowa

oznaczające emocje aktywne przetwarzane były szybciej (670 ms) niż słowa oznaczające emocje pasywne (698 ms). Również czas przetwarzania bodźców prezentowanych na górze (678 ms) istotnie różnił się od czasu przetwarzania bodźców pokazywanych na dole ekranu (688 ms), $F(1,41) = 7,83$, $p < 0,01$, $\eta_p^2 = 0,16$). Interakcja walencji i aktywności okazała się istotna ($F(1,41) = 24,47$, $p < 0,001$, $\eta_p^2 = 0,37$), jak wykazała analiza efektów prostych, wyrazy kojarzone z aktywnością były przetwarzane szybciej niż te kojarzone z pasywnością, dotyczyło to jednak w o wiele większym stopniu wyrazów pozytywnych ($F(1,41) = 66,02$, $p < 0,001$, $\eta_p^2 = 0,62$) niż negatywnych ($F(1,41) = 4,68$, $p < 0,05$, $\eta_p^2 = 0,10$). Kolejna interakcja: walencji i lokalizacji również okazała się istotna ($F(1,41) = 7,14$, $p < 0,05$, $\eta_p^2 = 0,15$). Analiza efektów prostych wykazała, że lokalizacja wpływała jedynie na czas reagowania na bodźce pozytywne, przyspieszając reakcje na bodźce prezentowane u góry ekranu ($F(1,41) = 15,33$, $p < 0,001$, $\eta_p^2 = 0,27$). Wpływ lokalizacji na czas reakcji na bodźce negatywne był nieistotny, podobnie jak pozostałe efekty w modelu (wszystkie $F < 1$). Badani popełnili znikomą liczbę błędów (< 1%), więc żadne analizy poprawności reakcji nie były przeprowadzane.



Wykres 2. Średnie czasy reakcji jako funkcja walencji i lokalizacji w Eksperymentie 2. Słupki błędu wskazują na błąd standardowy średniej w każdym warunku

Dyskusja

Uzyskane wyniki ponownie pokazały, że bodźce pozytywne przetwarzane są szybciej od negatywnych. Jak zaznaczono w opisie metody, słowa pozytywne i negatywne miały w naszym badaniu równą frekwencję i długość, można zatem sądzić, że rola frekwencji w szybkości oceny walencji wyrazów jest mniejsza niż rola znaku owej walencji.

Wyrazy aktywne były oceniane szybciej niż wyrazy pasywne, co wskazuje, że wymiar AKTYWNOŚĆ–PASYWNOŚĆ również ma wpływ na czas reakcji na bodźce. Nieistotna interakcja aktywności i lokalizacji wskazuje jednak, że wymiar ten nie podlega poznawczej asymetrii GÓRA–DÓŁ. Interakcja aktywności i walencji okazała się istotna: zgodnie z przewidywaniami akcelerujący wpływ wymiaru aktywności dotyczył zarówno wyrazów pozytywnych, jak i negatywnych, choć tych pierwszych w większym stopniu.

W Eksperymentcie 2 ponownie wykazano istotną interakcję walencji i lokalizacji, jednak znów jedynie bodźce pozytywne były przetwarzane szybciej na górze niż na dole. Reakcje na bodźce negatywne kolejny raz nie były istotnie różne w obu lokalizacjach. Minimalna wielkość tego efektu ($\eta_p^2 = 0,001$) sugeruje, że nie mamy tutaj do czynienia np. ze zbyt małą próbą, by otrzymać istotną różnicę. Wykazanie jej istotności wymagałoby przebadania tysięcy osób przy założeniu mocy testu $\pi = 0,80$ i oczekiwanego poziomu istotności różnicy $\alpha = 0,05$ [program PS; Dupont i Plummer 1998].

Wynik ten ponownie pokazuje, że brak istotnej różnicy związany jest z nikłą wielkością efektu, a nie ze zbyt małą liczbą przebadanych osób (małą mocą testu).

Dyskusja ogólna

W przedstawionych dwóch eksperymentach próbowano dokonać replikacji wyniku zaprezentowanego przez Meiera i Robinsona [2004], którzy dowiedli, że bodźce pozytywne są przetwarzane szybciej niż negatywne oraz że czasy reakcji na bodźce pozytywne są krótsze, jeśli bodźce te są prezentowane na górze ekranu, natomiast reakcje na bodźce negatywne przeciwnie – są szybsze, gdy bodźce negatywne są pokazywane na dole. Badacze tłumaczyli ten efekt zakorzenieniem reprezentacji abstrakcyjnych pojęć (np. emocji) w doświadczeniach sensomotorycznych, zwłaszcza związanych z przestrzenią, oraz używaniem powszechnie występujących w języku naturalnym metafor opartych na reprezentacjach świata fizycznego do komunikowania pojęć abstrakcyjnych [Lakoff i Johnson 1999; Meier i Robinson 2004].

W Eksperymentcie 1 dokładnie powtórzono procedurę użytą w badaniach Meiera i Robinsona [2004] i częściowo zreplikowano oryginalne wyniki. Wyjątkiem był brak pełnej asymetrii GÓRA–DÓŁ: reakcje na bodźce negatywne nie były szybsze podczas prezentacji na dole ekranu w porównaniu z prezentacją u góry. Ten sam efekt powtórzył się w Eksperymentcie 2, gdzie ponownie bodźce negatywne były nieistotnie szybciej przetwarzane na dole niż na górze. Podobny wzorzec wyników zaobserwował w swoich badaniach Schubert [2005] – efekt rozpoznawania słów związanych z wyso-

kim statusem społecznym na górze był silniejszy niż efekt rozpoznawania słów odnoszących się do niskiego statusu na dole. Takie wyniki, wskazujące na pewną asymetrię w przetwarzaniu słów prezentowanych na górze lub na dole, skłaniają ku zadaniu pytania, czy przyjęte przez Meiera i Robinsona wyjaśnienie o istnieniu zakorzenionej w doświadczeniu fizycznym poznawczej asymetrii GÓRA–DÓŁ, która kojarzy bodźce pozytywne z przestrzennym usytuowaniem w górze, a negatywne – z położeniem na dole, jest w pełni trafne. Jeśli zaakceptujemy taką hipotezę, należałoby jeszcze znaleźć wyjaśnienie znacznie mniejszego wpływu lokalizacji w przestrzeni dla przetwarzania bodźców negatywnych niż pozytywnych. Meier i Robinson [2004] nie podają w swoim artykule wielkości poszczególnych efektów ani nawet dokładnych średnich i odchyłek standardowych dla poszczególnych warunków eksperymentalnych, tak więc precyzyjne określenie różnic między bodźcami pozytywnymi i negatywnymi w ich badaniu nie jest możliwe. Jednakże analiza umieszczonych w artykule wykresów wskazuje na znacznie większą różnicę w średnim czasie reakcji dla bodźców pozytywnych (około 35 ms krótszy dla bodźców prezentowanych na górze) niż dla bodźców negatywnych (około 15 ms krótszy dla bodźców prezentowanych na dole). By spróbować wyjaśnić większe znaczenie wymiaru wertykalnego dla przetwarzania bodźców pozytywnych, należy odrębnie przeanalizować wpływ walencji i wertykalności na uzyskiwane wyniki.

Istnieje spora grupa wyników świadczących, że bodźce (słowa lub obrazki) pozytywne są przetwarzane szybciej niż bodźce negatywne [m.in. Carretié i in. 2008; Hinojosa, Méndez-Bértolo i Pozo 2010; Kanske i Kotz 2007]. Badania z wykorzystaniem EEG i neuroobrazowania pokazują, że bodźce negatywne przyciągają uwagę wcześniej niż bodźce pozytywne i neutralne oraz że „poświęcanie” im uwagi kończy się wcześniej [Carretié, Hinojosa, Martín-Loeches, Mercado i Tapia 2008]. Ponadto efekt ten utrzymuje się zarówno dla bodźców wysoko-, jak i niskofrekwencyjnych [Méndez-Bértolo, Pozo i Hinojosa 2011]. Oznacza to, że bodźce negatywne szybko ściągają uwagę (*attentional capture*), ale nie są głęboko przetwarzane, jeśli nie zostaną uznane za zagrażające, co ma swoje ewolucyjne uzasadnienie [Ohman, Flykt i Esteves 2001]. Potwierdzają to kolejne badania prowadzone z wykorzystaniem EEG, które pokazują, że semantyczne przetworzenie bodźca (ok. 500 ms po jego pojawieniu się), konieczne do określenia jego emocjonalnego znaczenia, może być opóźnione w przypadku bodźców negatywnych. Dzieje się tak z powodu braku facylitacji przetwarzania [Hinojosa i in. 2010] albo trudności z ponownym związaniem uwagi z bodźcem negatywnym, który ściąga uwagę w początkowych fazach przetwarzania (ok. 100 ms po pojawieniu się). Po upływie tego czasu jego przetwarzanie jest stłumione, co może dawać w rezultacie zjawisko analogiczne do hamowania powrotu (*inhibition of return*) [Posner i Cohen 1984; West i Holcomb 2002].

Warto również zastanowić się nad silnym interindywidualnym zróżnicowaniem reakcji na bodźce negatywne. Jak pokazują badania Robinsona i współpracowników [Robinson, Wilkowski, Meier, Moeller i Fetterman 2012], ludzie różnią się pod kątem przetwarzania bodźców negatywnych ze względu na poziom agresji, mierzonej kwestionariuszem Bussa i Perry’ego [1992]. U badanych z niskim poziomem agresji następowało wyhamowanie reakcji w zetknięciu z bodźcem negatywnym, natomiast u tych z wysokim jej poziomem takiego efektu nie zaobserwowano [Robinson i in. 2012]. Wynik ten jest łączony ze zwiększoną regulacją reakcji w odpowiedzi na bodźce nega-

tywne, czego potwierdzenie stanowi rosnąca liczba badań z użyciem fMRI [np. Kerns, Cohen, MacDonald, Cho, Stenger i Carter 2004]. Wolniejsze reagowanie na bodźce negatywne w porównaniu z pozytywnymi wydaje się związane ze specyfiką uwagowego przetwarzania bodźców emocjonalnych i zależy od wielu czynników, zarówno psychologicznych (jak np. natężenie pewnych cech osobowości), jak i związanych ze specyfiką procedur badawczych (np. występowanie bodźców pozytywnych i negatywnych w osobnych blokach lub przemieszanych w jednym, kolor czcionki etc.) [przeгляд: Sutton 2010]. Warto również wspomnieć, że osoby cierpiące na depresję zwykle szybciej wiążą uwagę z bodźcami negatywnymi i szybciej je identyfikują [przeгляд: Gotlib i Joormann 2010], co wskazuje, że zależność omawianych efektów od stanu emocjonalnego badanych jest znaczna i powinna być uwzględniana w badaniach.

Meier i Robinson uważają, że podobnym różnicom interindywidualnym podlega również przetwarzanie wymiaru GÓRA–DÓŁ, czyli wpływ lokalizacji bodźców na ich przetwarzanie. Osoby wykazujące się obniżonym nastrojem lub wysokim poziomem neurotyzmu mają tendencję do przetwarzania bodźców neutralnych (np. liter) szybciej, gdy te są prezentowane na dole ekranu, niż gdy pokazuje się je w jego górnej części [Meier i Robinson 2006].

Wyniki te sugerują, że większa liczba czynników ma wpływ na przetwarzanie bodźców negatywnych oraz wymiaru DÓŁ niż na reagowanie na bodźce pozytywne i na wymiar GÓRA. Klucz do rozwiązania zagadki braku wpływu wymiaru lokalizacji na przetwarzanie bodźców negatywnych wydaje się tkwić właśnie w większej liczbie procesów zaangażowanych w ich poznawczą „obróbkę” w porównaniu z bodźcami pozytywnymi. Ponadto na reagowanie na bodźce negatywne wpływa wiele cech osobowościowych, które nie mają normalnego rozkładu w populacji, brak jest natomiast tak silnych związków osobowości z przetwarzaniem bodźców pozytywnych [Gerevich, Bácskai i Czobor 2007].

Osobnym wytłumaczeniem słabszego związku bodźców negatywnych z wymiarem GÓRA–DÓŁ w przedstawionych badaniach byłyby różnice kulturowe i/lub językowe między językiem polskim a amerykańskim i angielskim. Różnice te wydają się szczególnie istotne w przypadku słów nazywających stany emocjonalne, bo repertuary takich określeń znacząco różnią się między językami, a zakresy znaczeń słów uważanych za ekwiwalenty tłumaczeniowe często nie pokrywają się w pełni [Matsumoto i Juang 2007; Wierzbicka 2007]. Takie rozstrzygnięcie wymagałoby jednak osobnych, szeroko zakrojonych badań, a dotychczasowa literatura przedmiotu sugeruje raczej uniwersalność wymiaru wertykalnego GÓRA–DÓŁ [Johnson 1987; przeгляд: Meier i Robinson 2005].

Ostateczne rozstrzygnięcie, czy bodźce negatywne są słabiej związane z wymiarem wertykalnym, czy nie, będzie możliwe tylko wówczas, gdy przeprowadzi się kolejne eksperymenty, tym razem kontrolując zmienne mogące wpływać na reakcje na bodźce negatywne, jak np. poziom agresji [Robinson i in. 2012], czy na lokalizację bodźców, jak np. nastroj i neurotyczność [Meier i Robinson 2006].

Wymiar AKTYWNOŚĆ–PASYWNOŚĆ wymaga dalszych badań, jednakże brak jego związku z wymiarem wertykalnym oraz przyspieszanie czasów reakcji bodźców afektywnych są zastanawiające i warte dalszego zgłębiania. Być może wymiar ten jest związany z przestrzenią w inny sposób, np. z wymiarem horyzontalnym: BLISKO–DALEKO.

Prezentowane eksperymenty zreplikowały wyniki uzyskane przez Meiera i Robinsona [2004] na bodźcach innych niż pozytywnie i negatywnie nacechowane słowa w języku angielskim, czyli zarówno na polskich tłumaczeniach słów użytych w oryginalnym badaniu (Eksperyment 1), jak i na nowych bodźcach: polskich słowach opisujących stany emocjonalne (Eksperyment 2). Uzyskane rezultaty potwierdziły postulatory o istnieniu asymetrii GÓRA–DÓŁ w przetwarzaniu wyrazów emocjonalnych, jednocześnie wskazując na znacznie słabszy związek bodźców negatywnych z wymiarem wertykalnym.

BIBLIOGRAFIA

- Adams F. (2010). *Embodied cognition*. „Phenomenology and the Cognitive Sciences” 9, s. 619–628.
- Anooshian J.L., Hertel P.T. (1994). *Emotionality in free recall: Language specificity in bilingual memory*. „Cognition and Emotion” 8, s. 503–514.
- Barsalou L.W. (1999). *Perceptual symbol systems*. „Behavioral and Brain Sciences” 22, s. 577–660.
- Boot I., Pecher D. (2010). *Similarity is closeness: Metaphorical mapping in a perceptual task*. „Quarterly Journal of Experimental Psychology” 63, s. 942–954.
- Brysbart M. (1996). *Word frequency affects naming latency in Dutch when age of acquisition is controlled*. „European Journal of Cognitive Psychology” 8, s. 185–193.
- Buss A.H., Perry M. (1992). *The aggression questionnaire*. „Journal of Personality and Social Psychology” 63, s. 452–459.
- Carretié L., Hinojosa J.A., Albert J., López-Martín S., de la Gándara B.S., Igoa J.M., Sotillo M. (2008). *Modulation of ongoing cognitive processes by emotionally intense words*. „Psychophysiology” 45, s. 188–196.
- Dupont W.D., Plummer W.D. (1998). *Power and sample size calculations for studies involving linear regression*. „Controlled Clinical Trials” 19, s. 589–601.
- Forster K.I., Forster J.C. (2003). *DMDX: A Windows display program with millisecond accuracy*. „Behavior Research Methods, Instruments & Computers” 35 (1), s. 116–124.
- Gable, P., Harmon-Jones E. (2010). *The motivational dimensional model of affect: Implications for breadth of attention, memory, and cognitive categorisation*. „Cognition and Emotion” 24 (2), s. 322–337.
- Gerevich J., Bácskai E., Czobor P. (2007). *The generalizability of the Buss-Perry aggression questionnaire*. „International Journal of Methods in Psychiatric Research” 16 (3), s. 124–136.
- Gibbs R.W. (2008). *Images schemas in conceptual development: What happened to the body?* „Philosophical Psychology” 21, s. 231–239.
- Glenberg A.M., Kaschak M.P. (2002). *Grounding language in action*. „Psychonomic Bulletin & Review” 9, s. 558–565.
- Gotlib I.H., Joormann J. (2010). *Cognition and depression: Current status and future directions*. „Annual Review of Clinical Psychology” 6, s. 285–312.
- Hinojosa J.A., Carretié L., Valcárcel M.A., Méndez-Bértolo C., Pozo M.A. (2009). *Electrophysiological differences in the processing of affective information in words and pictures*. „Cognitive, Affective and Behavioral Neuroscience” 9, s. 173–189.
- Hinojosa J.A., Méndez-Bértolo C., Pozo M.A. (2010). *Looking at emotional words is not the same as reading emotional words: Behavioral and neural correlates*. „Psychophysiology” 47, s. 748–757.
- Johnson M. (1987). *The Body in the Mind. The Bodily Basis of Meaning, Imagination and Reason*. Chicago: The University of Chicago Press.
- Kanske P., Kotz S.A. (2007). *Concreteness in emotional words: ERP evidence from a hemifield study*. „Brain Research” 1148, s. 138–148.

- Kerns J.G., Cohen J.D., MacDonald A.W., Cho R.Y., Stenger V.A., Carter C.S. (2004). *Anterior cingulate conflict monitoring and adjustments in control*. „Science” 303, s. 1023–1026.
- Lakoff G., Johnson M. (1999). *Philosophy in the flesh: The embodied mind and its challenges to western thought*. New York: Basic Books.
- Mahon B.Z., Caramazza A. (2008). *A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content*. „Journal of Physiology-Paris” 102, s. 59–70.
- Mandera P., Keuleers E., Wodniecka Z., Brysbaert M. (2014). *Subtlex-pl: subtitle-based word frequency estimates for Polish*. „Behavior Research Methods”. doi: 10.3758/s13428-014-0489-4 (online publication).
- Matsumoto D., Juang L. (2007). *Psychologia międzykulturowa*. Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- Meier B.P., Robinson M.D. (2006). *Does “feeling down” mean seeing down?: Depressive symptoms and vertical selective attention*. „Journal of Research in Personality” 40, s. 451–461.
- Meier B.P., Robinson M.D. (2005). *The Metaphorical representation of affect*. „Metaphor and symbol” 20 (4), s. 239–257.
- Meier B.P., Robinson M.D. (2004). *Why the sunny side is up. Associations between affect and vertical position*. „Psychological Science” 15 (4), s. 243–247.
- Méndez-Bértolo C., Pozo M.A., Hinojosa J.A. (2011). *Word frequency modulates the processing of emotional words: Convergent behavioral and electrophysiological data*. „Neuroscience Letters” 494, s. 250–254.
- Ohman A., Flykt A., Esteves F. (2001). *Emotion drives attention: detecting the snake in the grass*. „Journal of Experimental Psychology: General” 130, s. 466–478.
- Pecher D., Boot I., Van Dantzig S. (2011). *Abstract concepts: Sensory-motor grounding, metaphors, and beyond* [w:] B. Ross (red.), *The Psychology of Learning and Motivation*, vol. 54 (s. 217–248). Burlington: Academic Press.
- Pecher D., Zeelenberg R., Barsalou L.W. (2003). *Verifying different-modality properties for concepts produces switching costs*. „Psychological Science” 14, s. 119–124.
- Posner M.I., Cohen Y. (1984). *Components of visual orienting* [w:] H. Bouma, D. Bouwhuis (red.), *Attention and Performance X: Control of Language Processes* (s. 531–56). Hillsdale, NJ: Erlbaum.
- Ratcliff R. (1993). *Methods for dealing with reaction time outliers*. „Psychological Bulletin” 114 (3), s. 510–532.
- Robinson M.D., Wilkowski B.M., Meier B.P., Moeller S.K., Fetterman A.K. (2012). *Counting to ten milliseconds: Low-anger, but not highanger, individuals pause following negative evaluations*. „Cognition and Emotion” 26 (2), s. 261–281.
- Schubert T.W. (2005). *Your Highness: Vertical Positions as Perceptual Symbols of Power*. „Journal of Personality and Social Psychology” 89, s. 1–21.
- Smith C.A., Ellsworth P.C. (1985). *Patterns of cognitive appraisal in emotion*. „Journal of Personality and Social Psychology” 48 (4), s. 813–838.
- Sutton T.M. (2010). *The influence of emotion on attention: Examining the processing of negative and positive emotion words in the dot probe task*. Niepublikowana praca doktorska. Albany: State University of New York at Albany.
- Sutton T.M., Altarriba J., Gianico J.L., Basnight-Brown D.M. (2007). *The automatic access of emotion: Emotional Stroop effects in Spanish-English bilingual speakers*. „Cognition and Emotion” 21 (5), s. 1077–1090.
- Talmy L. (1988). *Force dynamics in language and cognition*. „Cognitive Science” 12, s. 49–100.
- Tzelgov J., Henik A., Leiser D. (1990). *Controlling the Stroop interference: Evidence from a bilingual task*. „Journal of Experimental Psychology: Learning, Memory and Cognition” 16 (5), s. 760–771.
- Vogt J., De Houwer J., Koster E.H.W., Van Damme S., Crombez G. (2010). *Allocation of spatial attention to emotional stimuli depends upon arousal and not valence*. „Emotion” 8 (6), s. 880–885.

- West W.C., Holcomb P.J. (2002). *Event-related potentials during discourse-level semantic integration of complex pictures*. „Brain Research. Cognitive Brain Research” 13 (3), s. 363–375.
- Wicherts J.M., Bakker M. (2012). *Publish (your data) or (let the data) perish! Why not publish your data too?* „Intelligence” 40 (2), s. 73–76.
- Wicherts J.M., Borsboom D., Kats J., Molenaar D. (2006). *The poor availability of psychological research data for reanalysis*. „American Psychologist” 61 (7), s. 726–728.
- Wierzbicka A. (2007). *Słowa klucze. Różne języki – różne kultury*. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Wilson M. (2002). *Six views of embodied cognition*. „Psychonomic Bulletin and Review” 9 (4), s. 625–636.
- Yeh W., Barsalou L.W. (2006). *The situated nature of concepts*. „American Journal of Psychology” 119 (3), s. 349–384.
- Załaźnińska A. (2001). *Schematy myśli wyrażane w gestach. Gesty metaforyczne obrazujące wybrane abstrakcyjne relacje i zasoby podmiotu mówiącego*. Kraków: Universitas.

ANEKS 1. Przykładowe bodźce z Eksperymentu 1

Negatywne				Pozytywne			
bezcelowy	chory	gorzki	krytykować	aktywny	delikatny	grzeczny	miłość
bluźnierczy	daremnym	klęska	martwy	ambitny	dojrzały	hojny	miły
brutalny	diabeł	klamca	mierny	bohater	dowcipny	lojalny	niebo
brzydki	drażliwy	klótnia	niebezpieczeństwo	cukierek	etyczny	marzenie	odważny
chciwy	głupi	kradzież	nowotwór	czysty	geniusz	mądry	pocałunek

ANEKS 2. Bodźce z Eksperymentu 2

POZA	POZP	NEGA	NEGP
energia	duma	agresja	apatia
entuzjazm	ekstaza	desperacja	bezsilność
euforia	estyma	furia	depresja
fascynacja	satisfakcja	gniew	melancholia
miłość	spełnienie	napięcie	pogarda
podekscytowanie	szacunek	pragnienie	przygnębienie
radość	ukontentowanie	wściekłość	rezygnacja
rozradowanie	ulga	zazdrość	rozpacz
szczęście	zadowolenie	złość	smutek
zakochanie	zrelaksowanie	żądza	strapienie

Rozpoznawanie stanów emocjonalnych jako biologiczne źródło zmysłu moralności

The recognition of emotional states as a biological source of the moral sense

Abstract: People possess and use instinctive moral evaluations, while the rational thinking is of a minor importance. According to some hypotheses human reactions, which lead to ethical results, were shaped by the natural selection, not by the influence of a culture. The practice of moral judgments generating consists of few repeatable stages and it is characterised by some automatisms and innate reactions based on biological determinants. A sense of the morality appears to be correlated unmediatedly with the emotional states, with the recognition of such states and with instinctive demeanors. Subsequently the main source of moral intuition seems to be the instinctive behaviours and emotions. It is confirmed by the list of universal moral moduls, which are simple reactions to some stimulus, from where the society derives complex moral virtues. Perhaps the natural sense of the morality is a final effect of the functioning of different emotions, which are based on so-called emotional contagion and empathy, in other words – some forms of emotional states recognition. Additionally, neurotransmitters appears to play an essential role in the process of moral and social demeanors.

Key words: emotional contagion, moral sense, moral moduls, empathy, instinct

1. Narodziny idei

Na przełomie XVII i XVIII wieku w filozofii brytyjskiej pojawiło się pojęcie zmysłu moralnego, którego autorami byli lord Shaftesbury i kontynuator jego myśli – Francis Hutcheson. Według Shaftesbury'ego wartość moralną – harmonię naszych skłonności i proporcję czynów oraz praw – możemy rozpoznawać bądź realizować dzięki wrodzonej dyspozycji zwanej przez niego entuzjazmem. Dyspozycja ta urzeczywistniała się jako postawa wyjątkowej wrażliwości na harmonię, dobro i piękno. Refleksję Shaftesbury'ego rozwinął w 1725 roku Hutcheson¹. Zdolność rozpoznawania i tworzenia tego, co dobre lub złe Hutcheson określał terminem «zmysł moralny» (*The Sense*

¹ W 1725 roku opublikowano pierwsze wydanie *Inquiry into the Original of Our Ideas of Beauty and Virtue* Hutchesona. Patrz: wstęp do elektronicznej wersji *Inquiry into the Original... – Note on the Text*, URL: <http://facultypages.morris.umn.edu/~mcollier/Moral%20Sentimentalism/hutcheson%20ebook%20inquiries.pdf> (data pobrania: 19.07.2014).

of *Virtue*), który to zmysł wyraża się w formie wrodzonego poczucia sympatii wobec innych jednostek. Tak ujęty zmysł moralny miał charakter uczuciowy lub intuicyjny, w swej istocie zaś funkcjonował na podobieństwo zmysłu estetycznego. Zdaniem Shaftesbury'ego i Hutchesona rodzimy się wyposażeni w umiejętność automatycznego określania, co jest piękne, a co brzydkie; co moralnie poprawne, a co naganne i złe. W 1751 roku Hume odnosił się w swym etycznym dziele² już bezpośrednio do nauk przyrodniczych. Twierdził, jakoby uczucie sympatii stanowiło naturalnie zakorzeniony, pierwotny odruch emocjonalnego reagowania na świat społeczny. Codzienne wybory pomiędzy czynami ocenianymi pozytywnie bądź negatywnie stanowią – zdaniem Hume'a – bezpośrednią konsekwencję przeżywanych przez nas emocji moralnych. Z kolei owe emocje moralne – jak poczucie sprawiedliwości czy szacunku – są złożonym wynikiem procesów kojarzenia faktów i doświadczeń, podczas których pierwotnym i bazowym uczuciem jest sympatia. To sympatia stanowi naturalną, gdyż biologicznie zakorzenioną, podstawową reakcję solidarnej współpracy jednostek. Człowiek definiowany za Arystotelesem jako ζῷον πολιτικόν w sposób naturalny kieruje się ku zbiorowości, ku własnemu stadu. Stąd społeczeństwo jest naszym naturalnym środowiskiem życia. Ewentualne odruchy egoistyczne czy aspołeczne są co najwyżej wtórne względem pierwotnych odruchów sympatii oraz kooperacji. Teorię moralności Hume'a opartą na zasadzie sympatii kontynuował Adam Smith, publikując w 1759 roku *The Theory of Moral Sentiment*.

Osiemnastowieczni angielscy empiryści oraz francuscy racjoniści krytykowali oparte na emocjach i wrodzonej moralności stanowisko Shaftesbury'ego i Hutchesona. Popularne były poglądy w duchu Helvetiusa [1976], według którego cała moralność jest nabyta wyłącznie drogą sukcesywnego warunkowania. Kluczową funkcję w procesie uczenia się i przyswajania nawyków moralnych odgrywają: wychowanie domowe, obowiązujące prawodawstwo oraz całościowy ustrój polityczny. W drugiej połowie XVIII wieku i w XIX stuleciu myślenie o moralności zdominowali: najpierw Kant, a potem jego kontynuatorzy. Kant wykluczył emocje jako poznawczo bezwartościowe, a jednocześnie zaakcentował szczególną rolę rozumu dla podstaw etyki³. Mimo krytycznego podejścia do roli emocji w kształtowaniu się refleksji etycznej, już na początku XX wieku powróciła idea naturalnego zmysłu moralnego. W *Zapiskach z Piaszczystej Krainy* Aldo Leopold dowodził, że systemy etyczne powstały przede wszystkim na podstawie naszego wrodzonego poczucia moralności. Owo wrodzone poczucie moralności zostało ukształtowane – według Leopolda – drogą doboru naturalnego. Z kolei zdaniem Konrada Z. Lorenza ludzkie akty wartościowania ufundowane są na warstwie emocjonalnej, która jest ewolucyjnie starsza względem warstwy racjonalnej⁴. Wśród polskich filozofów natomiast m.in. Henryk Skolimowski [1991] jest myślicielem stojącym na stanowisku, iż posiadamy wewnętrzny zmysł moralny,

² D. Hume, *An Enquiry Concerning the Principles of Morals*. The Project Gutenberg Ebook, URL: <http://www.gutenberg.org/files/4320/4320-h/4320-h.htm> (data pobrania: 13.04.2014).

³ W *Uzasadnieniu metafizyki moralności* czytamy, że działanie moralne „nie polega wcale na uczuciach, popędach i skłonnościach, lecz tylko na stosunku istot rozumnych do siebie” [Kant 1984, s. 70–71].

⁴ „Oczywiste, że aby zamienić czysto rozumowe poznanie w nakaz czy zakaz, niezbędny jest czynnik uczuciowy”. Pozbawiając „aparatu naszego rozumu” aspektu emocjonalnego, nie byłibyśmy w stanie ani wyznaczać, ani też osiągać jakichkolwiek celów, a zatem nie byłibyśmy zdolni do jakiegokolwiek działania [Lorenz 1972, s. 325 i nast.].

czyli warstwę aksjologiczną naszej świadomości. Na szczególną uwagę zasługuje fakt, że współczesne odkrycia ze wszech miar empirycznych neuronauk mogą rzucić nowe światło na odległą refleksję Shaftesbury'ego, która w nowej neuroperspektywie przedstawia być jedynie czysto idealistyczną spekulacją.

2. Współczesne stanowiska

Shaftesbury, Hutcheson i Hume w swej idei naturalnego, wrodzonego zmysłu moralnego zrywali z racjonalistyczną tradycją, upatrującą zdolność poznania moralnego w trzeźwym rozumie czy też w intelekcie jako probierzu dobra i zła. Warto przypomnieć, że radykalny racjonalizm etyczny doszedł do głosu w 1788 roku pod postacią apriorycznego prawa moralnego, które zostało sformułowane przez Kanta w *Krytyce praktycznego rozumu*. Kant postulował, jakoby wszelkie ludzkie działanie posiadało moralną kwalifikację, którą zawdzięcza jedynie rozumowi, a konkretnie – rozumowej właściwości woli. Obowiązki czy prawa moralne, którym wola oraz działanie muszą się podporządkować, wynikają – zdaniem Kanta – wyłącznie z rozumu. Moralność jest rozumnym działaniem z obowiązku. Jako spełnianie obowiązku, w moralności rozum podporządkowuje się samemu sobie – pozostaje samowystarczalny i autonomiczny względem emocji bądź praw natury. Dokonując wyborów czy też reagując na bodźce, podlegamy rozumowi, który całkowicie determinuje naszą wolę działania [Kant 1974].

Gdy po dwustu latach neuronauki zabrały głos w sprawie szeroko ujętych władz poznawczych, racjonalne stanowisko Kanta odnośnie do etyki zostało poddane ostrej krytyce. Sztandarowym przykładem była popularna publikacja *Błąd Kartezjusza*, w której António Damasio – jako jeden z wielu neuronaukowców – zadał ogólne pytanie, dlaczego jesteśmy istotami moralnymi. Pisał on w podsumowaniu, że „chłodna strategia rozumowania, za którą obstawał między innymi Kant, ma znacznie więcej wspólnego z tym, co czynią [upośledzeni] pacjenci niż z tym, jak zachowują się zazwyczaj normalni ludzie” [2011, s. 195]. Innym, równie krytycznym głosem w literaturze popularyzującej neuronauki była wypowiedź Patricii Churchland: „przekonanie Kanta, że oderwanie się od emocji jest kluczowe w charakteryzowaniu zobowiązania moralnego, jest rażąco sprzeczne z tym, co wiemy o naszej biologicznej naturze” [Churchland 2013, s. 283]. Również Michael Gazzaniga w książce *Istota człowieczeństwa* przekonuje – na podstawie analizy wyników badań Jonathana Haidta i Marca Hausera – że jesteśmy wyposażeni we wbudowany program etyczny, który powstał drogą doboru naturalnego, w reakcji na naciski ewolucyjnej selekcji. Ten naturalnie wrodzony zmysł moralności jest dodatkowo wtórnie kształtowany przez wpływ społeczny: kultura wzmacnia lub hamuje pewne elementy naturalnych reakcji, określając, co jest cnotą, a co należy potępić.

2.1. Schemat instynktownej oceny moralnej

Biorąc pod uwagę fakt, że moralność jest pojęciem niezwykle wieloznacznym, warto uściślić, jak w tym wypadku rozumiemy tytułowy zmysł moralności. Otóż na potrzeby pracy będziemy definiować go w trojaki sposób:

- a) zdolność podejmowania jednostkowych decyzji z uwzględnieniem interesów wspólnoty;
- b) „kompas wyborów życiowych”, jak ujął to metaforycznie Frans de Waal;
- c) umiejętność oceniania, rozpoznawania wartości, porównywania zysków, wysiłku i dóbr.

Zmysł moralności zatem to przede wszystkim – w sferze kognitywnej – naturalna zdolność wartościowania zdarzeń, obiektów czy zjawisk, ale również – w sferze konatywnej – umiejętność preferowania interesów wspólnoty ponad interesami indywidualnymi. Nie jest to jedynie umiejętność rozpoznawania wartości, ale także hierarchizowania ich ze względu na kontekst sytuacji.

Skupmy się na samej hipotezie zmysłu moralnego oraz na biologicznym źródle tego zmysłu. W 2001 roku ukazał się artykuł Jonathana Haidta *The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment*, zaś dwa lata później *The Moral Emotions* [2003]. W obydwu artykułach Haidt dowodzi faktu posiadania przez ludzi intuicyjnej oceny moralnej, która to intuicja działa na zasadzie automatycznego, a raczej instynktownego decydowania. W pewnych stresujących sytuacjach, dzięki momentalnej reakcji instynktownej, dorosły człowiek (niezależnie od pochodzenia czy wykształcenia) zawsze wydaje się postępować według jednakowego schematu: (1) cieleśnie-emocjonalna reakcja na sytuację → (2) błyskawiczne reprezentacje możliwych rozwiązań tejże sytuacji → (3) ocena tych rozwiązań → (4) podjęcie decyzji → (5) racjonalne uzasadnienie. Istotny jest tutaj etap pierwszy. Jako wrażenie na bodziec/stresor występuje niekontrolowana, silna reakcja emocjonalna wraz ze znacznikiem somatycznym. Odpowiedzią na czynnik stresogenny jest emocja posiadająca funkcję wstępnie wartościującą, która mobilizuje do działania. Adekwatne odczytanie emocji ułatwia skorelowany z nią marker somatyczny. Proces wartościowania zdarzeń bądź zjawisk przebiega – zdaniem Haidta – według następującego, uniwersalnego schematu: (1) zewnętrzny bodziec wywołuje reakcję emocjonalno-cieleśną → (2) występuje reprezentacja oraz antycypacja możliwych reakcji → (3) ocena tych możliwości (pojawienie się emocji moralnej) → (4) preferowanie i decydowanie → (5) autointerpretacja. Kluczowe w tym wypadku emocje moralne Haidt wyjaśnia słowami: „moral emotions as those emotions that are linked to the interests or welfare either of society as a whole or at least of persons other than the judge or agent” [Haidt 2003, s. 853].

Bezpośrednim dowodem istnienia instynktownej, automatycznej reakcji emocjonalnej o moralnym, wartościującym zabarwieniu jest, zdaniem Haidta, odruchowo negatywna ocena zjawiska kazirodztwa mimo braku racjonalnego uzasadnienia tejże oceny [Gazzaniga 2011, s. 123–126]. Powołując się na tezy Westermarcka, można uznać tę negatywną ocenę za ponadczasową i ponadkulturową, nie licząc rzadkich wyjątków zachowania czystości krwi arystokratycznej klasy egipskich faraonów czy władców imperium Inków. Westermarck uznał kazirodztwo za kulturowe uniwersale, uzasadniając awersję kazirodczą psychologicznie – jako brak pociągu seksualnego do osób znanych nam z dzieciństwa. Haidt uzupełnił uzasadnienie Westermarcka,

dodając wrodzoną niechęć o podłożu genetycznym, na co dowody można znaleźć w obserwacjach dotyczących zachowań szympansov [Goodall 1997], ale także ptaków [Eibl-Eibesfeldt 1987]. To intuicja – manifestująca się automatycznym odruchem wypracowanym przez dobór naturalny – mówi nam, że kazirodztwo jest złe. Mamy tu do czynienia z determinizmem biologicznym, który jest dodatkowo wzmacniany (w tym akurat przypadku) przez oddziaływanie kulturowe prowadzące do społecznego tabu. Przykład Haidta został również poparty doświadczeniem Lieberman, Tooby'ego i Cosmides [2003]. Badacze ci dowiedli, że im dłużej mieszkamy w dzieciństwie z jakimkolwiek rówieśnikiem, tym bardziej złe wydaje się uprawianie z nim seksu. Racjonalnie wyuczony zakaz społeczny jest drugorzędny – tabu ukształtowało się pierwotnie w doborze naturalnym.

Kontynuując wątek badany przez Haidta, Gazzaniga podjął się głębszej analizy schematu instynktownej oceny moralnej. Powołując się na eksperymenty własne oraz badania między innymi Damasio, dotyczące uszkodzeń płata czołowego, uzasadnił drugorzędną rolę racjonalnego myślenia w kwestiach wyborów, które posiadają moralne implikacje. Krótko mówiąc, mechaniczne uszkodzenie mózgu prowadzi do spłaszczenia afektu, a więc do upośledzenia emocji pierwotnych i wtórnych, co powoduje niezdolność decydowania, planowania, organizowania czasu oraz wartościowania i oceniania. Uszkodzenie takie nie narusza jednak zdolności intelektualnych związanych z logicznym myśleniem czy wnioskowaniem. Jak wygląda wzorcowy przebieg automatycznej reakcji wartościowania zdaniem Gazzanigi? Każda decyzja moralna jest uwarunkowana podstawową emocją pozytywną „akceptacji i zbliżenia” lub negatywną „odrzućcia i unikania”. Akceptacja na podłożu emocjonalno-cieleśnym identyfikowana jest jako zjawisko pierwotnie dobre, z kolei emocjonalno-somatyczna reakcja chęci wycofania się identyfikuje zjawisko jako niepożądane, a więc złe. Jest to automatyzm, występuje on podczas działania zmysłu moralnego, a Gazzaniga [2011, s. 128–129] określił go mianem torowania afektywnego (*affective priming*). Powtórzmy: przed podjęciem decyzji występuje (1) bezwiedna reakcja emocjonalna na bodziec – albo negatywna (odrzućcie), albo pozytywna (akceptacja) – która stanowi wstępne wartościowanie zjawiska. Następnie pojawia się (2) reprezentowanie możliwych skutków działania wraz z emocją moralną i dopiero (3) decyzja. Przetwarzając dane automatycznie, poniżej progu świadomości, zabarwiamy je emocją negatywną lub pozytywną, co Gazzaniga określił właśnie jako torowanie afektywne. Zabarwienie emocjonalne wraz z wtórnym rozpoznaniem i uświadomieniem sobie owych emocji zdaje się poprzedzać podjęcie każdej decyzji o moralnych implikacjach.

Gdyby rozpisać za Gazzanigą każdy krok szczegółowo, wybór działania (czy w codziennej, banalnej sytuacji, czy też życiowo istotnej) przebiegałby następująco: (1) bodziec → (2) przetwarzanie danych percepcyjnych → (3) torowanie afektywne + marker somatyczny (wstępna aprobata lub unikanie) → (4) reprezentacje możliwości → (5) pełny stan emocjonalny + marker somatyczny → (6) intuicja moralna (ocena, wartościowanie możliwości) → (7) decyzja → (8) działanie + racjonalne uzasadnienie. O ile pierwsze etapy do (5) pełnego stanu emocjonalnego trwają ułamki sekundy, o tyle kolejne etapy wartościowania, decyzji i racjonalnej analizy znacznie się wydłużają. Ostateczne działanie i racjonalna analiza mogą przeciągać się w czasie, a decyzja bywa wielokrotnie zmieniana. Jednak wstępne etapy przebiegają błyskawicznie i niemal niezauważalnie dla świadomości.

Gazzaniga powyższy schemat uzupełnia o jeszcze jeden automatyzm, zwany przez niego „inklinacją negatywną” [Gazzaniga 2011, s. 130–131]. Otóż na skutek doboru naturalnego przejawiamy dyspozycję do szybszego i łatwiejszego, automatycznego rozpoznawania bodźców nieprzyjemnych, a więc potencjalnych zagrożeń (np. czyichś negatywnych emocji i czynów). Taka inklinacja posiada ogromną wartość adaptacyjną także do środowiska kulturowego, gdyż umożliwia błyskawiczną postawę ucieczki bądź walki, również w metaforycznym sensie. Nieprzyjemne bodźce wywołują bardziej intensywne reakcje (markery) somatyczne, dzięki czemu szybciej je rozpoznajemy; lepiej przyciągają naszą uwagę; silniej zapadają nam w pamięć; a przede wszystkim stanowią emocjonalny warunek decyzji, ocen i sądów moralnych. Zdaniem Gazzanigi można pokusić się o wskazanie hipotetycznego miejsca, w którym mieści się nasz zmysł moralny. Powołuje się on na pracę Joshui Greene’a, który ustalił metodą rezonansu magnetycznego, że osobiście dotykające nas dylematy moralne aktywizują stałe reakcje emocjonalne, a równocześnie fizycznie pobudzają tylną część bruzdy skroniowej górnej [Greene, Sommerville i Nystrom 2001]. Gazzaniga sugeruje, że być może tam znajduje się ośrodek ludzkiego zmysłu moralnego.

Wątpliwości co do istnienia takiego ośrodka budzi fakt, że jeżeli nawet faktycznie posiadamy naturalnie wbudowane reakcje, to mimo wszystko są to reakcje na bardzo konkretne bodźce, tylko subiektywnie problematyczne z etycznie indywidualnej perspektywy. Z przywołanych badań Greene’a wynika, że abstrakcyjne, czysto teoretyczne dylematy moralne nie wyzwalają w nas automatycznej reakcji emocjonalnej. Tutaj musimy posiłkować się logicznie kalkulującą czy racjonalną analizą. Mimo wszystko akcentowane przez Gazzanigę reakcje typu „zbliżanie-unikanie” (*affective priming*) oraz inklinacja negatywna wydają się być reakcjami wrodzonymi, a pewna klasa sądów moralnych bywa formułowana automatycznie czy instynktownie (np. zasadność unikania kazirodztwa). Dopiero wtórnie dążymy do logicznego wyjaśnienia formułowanych sądów.

2.2. Od podstawowych intuicji do wrodzonych zasad moralnych

Dotychczas mowa była o dowolnych, przypadkowych sytuacjach, gdzie ogólny proces wartościowania uruchamiany jest przez reakcję emocjonalno-somatyczną, niezależnie od okoliczności. Tymczasem Jonathan Haidt i Craig Joseph wykazali, że pewne wybrane sytuacje działają niczym bodziec aktywizujący kolejne automatyczne intuicje o silnie moralnych konsekwencjach [Haidt, Joseph 2004, s. 55–66]. Wszelka hierarchia wartości, fundująca dowolny system etyczny, wręcz wywodzi się bezpośrednio z pięciu podstawowych intuicji, które z kolei mają swe korzenie w zarysowanym już schemacie reakcji emocjonalnej. Ściśle rzecz biorąc, każda intuicja jest konsekwencją pewnej adaptacji do środowiska naturalnego, a następnie przeradza się w konkretną emocję moralną, ta zaś w etyczną cnotę. Według Haidta ewolucja faworyzowała jednostki posiadające wszystkie pięć intuicji, co umożliwiło wyprowadzenie szerokiego wachlarza społecznych cnót.

Owe pięć uniwersalnych intuicji można określić naturalnymi algorytmami czy też modułami moralnymi, a więc automatycznymi reakcjami typu wejście-wyjście na konkretne, nieprzypadkowe bodźce zewnętrzne. Do modułów tych należą: wzajemność,

rozpoznawanie cierpienia, hierarchizacja, tworzenie koalicji, czystość. Z owej piątki reakcji wywodzi się zdecydowana większość ponadkulturowych cnót moralnych. Dalsze społeczne różnicowanie się odmiennych hierarchii wartości wynika z faktu, że każda kultura kładzie nacisk na inne aspekty wymienionych pięciu modułów. Stanowi to źródło kulturowego relatywizmu względem moralności, gdzie społeczeństwo wtórnie wpływa na naturalnie uformowaną moralność. Każdy z modułów jest aktywowany przez bodźce środowiskowe (dane wejściowe); każdy również wywołuje – w rezultacie – złożone emocje moralne oraz cnoty (dane wyjściowe). Pewne emocje nie mieszczą się w obrębie jednego modułu, tym bardziej zaś cnoty mogą wywodzić się równocześnie z kilku automatycznych reakcji. Oto omawiane moduły/instynktowne odruchy, obserwowane głównie w świecie *Hominidae*, choć zauważane także pośród innych ssaków wyższych:

- a) „Wzajemność” albo „odwzajemnianie przysług”, czyli wymiana bazująca na wstępnej emocji akceptacji oraz altruizmie odwzajemnionym. U niektórych nietoperzy, ale też ryb z gatunku *Poecilia reticulata*, zaobserwowano akty wzajemności na poziomie indywidualnym. Podstawowym warunkiem zaawansowanej formy wzajemności jest zawarcie i przestrzeganie animalnej umowy społecznej, a także zdolność porównywania wysiłków, zysków i strat. To znaczy: «jeśli ja zrobię teraz coś dla ciebie, ty w przyszłości zrobisz coś dla mnie». Wzajemność generuje – na poziomie hominidów – następujące stany emocjonalne o moralnym zabarwieniu: zaufanie, współczucie, wdzięczność, poczucie winy i wstyd (tylko *Homo sapiens*), gniew, pogardę. Wrodzona potrzeba wzajemności prowadzi – na poziomie *Homo sapiens* – do cnoty sprawiedliwości, uczciwości, rzetelności i cierpliwości.
- b) Instynktowny odruch „wrażliwości na cierpienie” to umiejętność wykrywania cierpienia, która – jako adaptacja – zwiększa szanse przetrwania potomstwa rodziców. Główny warunek zaistnienia tego modułu to konieczność naturalnej zależności dziecka od swego rodzica. Ewolucja automatycznego rozpoznawania cierpienia – na poziomie *Homo sapiens* – generuje współczucie, empatię i życzliwość, które można uznać za abstrakcyjne cnoty, a zarazem emocje moralne.
- c) Wrodzona potrzeba „hierarchii”, uznanie dominacji, a także nierówności wynikającej z poziomu sprawności bądź atrakcyjności. Potrzeba ta, będąc adaptacją, umożliwia bezpieczne unikanie konfliktów z jednostkami posiadającymi przewagę nad nami. Warunek, jaki musi zostać spełniony, by pojawił się ten automatyzm, to życie w zbiorowości. Moduł hierarchii generuje – na poziomie hominidów – takie moralne stany emocjonalne, jak pewność siebie oraz (tylko pośród *Homo sapiens*) poczucie winy, zakłopotanie i wstyd. W ludzkim środowisku kulturowym stany emocjonalne ewoluują w abstrakcyjne cnoty: autorytetu, szacunku, lojalności, honoru i subordynacji.
- d) Naturalna potrzeba „koalicji”, sojuszu, afiliacji i kooperacji, a więc automatyczny podział na zbiorowość własną i obcą. Warunkiem koniecznym do zaistnienia tej potrzeby jest (1) życie społeczne, jak szympanse lub delfiny, oraz (2) instynktowne rozpoznawanie przyjaciela i wroga dzięki naśladowaniu, markerom somatycznym oraz prymarnej emocji unikania (wspomnianej inklinacji negatywnej). Zdolność zawierania i rozpoznawania sojuszu generuje złożone

stany emocjonalne: współczucie, winę, zakłopotanie, wdzięczność, pogardę czy gniew. Wywodzące się z nich moralne cnoty ludzkie to: zaufanie, lojalność, honor, patriotyzm, poświęcenie, bohaterstwo, rywalizacja.

e) Wrodzony moduł „czystości”, a więc instynktowne unikanie niebezpieczeństwa w postaci chorób, trucizn i pasożytów. Podstawowym warunkiem umożliwiającym zaistnienie modułu jest rozpoznawanie niebezpieczeństwa jako np. intensywne barwy. Z potrzeby czystości ewoluują złożone stany emocjonalne, takie jak wstręt, obrzydzenie (tylko *Homo sapiens*) oraz mdłości. Przy tym uczucie wstrętu pojawia się w trzech społecznych obszarach:

- cielesnym (może dotyczyć seksu, śmierci, płynów fizjologicznych czy deformacji ciała);
- w obszarze tzw. skażenia interpersonalnego (dotyczy przedmiotów i odzieży innych ludzi, zwłaszcza nielubianych);
- w obszarze moralności (dotyczy łamania prawa, zasad, norm).

Poczucie wstrętu wydaje się pozostawać pod wpływem wyjątkowo silnego warunkowania kulturowego. To, co budzi w nas jakikolwiek rodzaj wstrętu (na bazie pierwotnej emocji unikania), automatycznie kojarzone jest ze złem. To, co uznajemy za szeroko pojętą czystość duchową lub materialną (pierwotna emocja akceptacji), natychmiast kojarzymy z dobrem. Powyższe stany emocjonalne generują abstrakcyjne cnoty higieny, godności, honoru, a przede wszystkim religijności (jako kult najszerzej pojętej czystości)⁵.

Marc Hauser – wzorując się na natywistycznej teorii Chomsky’ego odnośnie do genetycznych kompetencji do nauki języka oraz tworzenia i rozumienia zdań – zaproponował własną koncepcję uniwersalnej gramatyki moralności [Hauser 2006]. Analogicznie do Haidta i Josepha, Hauser dowodził na podstawie własnych obserwacji, jakoby człowiek posiadał wrodzony zestaw narzędzi do budowania konkretnych systemów moralnych. Ponownie Hauser wnioskował, że rodzimy się z biologicznie wrodzonymi regułami, które pobudzone prowadzą do powstania pewnych zasad moralnych. Do tych naturalnych reguł Hauser zaliczał: terytorializm, hierarchiczność, koalicje (my – obcy), wzajemność jako odwzajemnianie przysług. Dodatkowo posiadamy gotowość do przyswojenia sobie następnych reguł moralnych na podobieństwo języka. Powyższe reguły stanowią kolejne uniwersalia kulturowe, które wprawny badacz zaobserwuje u osób rozwiązujących konkretne problemy moralne. Zdaniem Hausera nieświadomy zmysł moralny istnieje kolektywnie – można go określić mianem uniwersalnego sumienia. Owo sumienie, wzorem teorii Chomsky’ego, kształtuje moralne intuicje odnośnie do dobra i zła w poszczególnych sytuacjach, pod wpływem nacisku społecznego i procesu wychowania.

Z wrodzonych zasad moralnych, składających się na zmysł rozpoznający wartości oraz hierarchizujący je, najistotniejszą według Hausera jest zasada „odwzajemniania

⁵ Jak twierdzi Gazzaniga [2011], niezwykle złożona cnota religijności prawdopodobnie wynika po części z wszystkich wrodzonych modułów po kolei. Odnajdujemy w niej wzajemność (jeśli teraz czcisz, po śmierci uzyskasz gratyfikację); wrażliwość na cierpienie (dogmaty koją cierpienie lub je podkreślają); hierarchię (autorytet Kościoła to zaufanie, szacunek, lojalność); koalicję (moja grupa wyznaniowa jest najlepsza); czystość (skażone ciało i umysł to grzech). Choć nie sposób wyodrębnić jednego, uniwersalnego źródła religii, to najsilniej zdaje się ona korelować z wrodzoną potrzebą czystości oraz emocją wstrętu.

przysług” – opiera się na niej przedłużona wymiana społeczna. Jednak ponownie musi zostać spełniony pewien warunek konieczny, aby pojawiło się tak istotne dla społeczeństwa zachowanie, jak wzajemność. Mianowicie musimy być zdolni do rozpoznawania nie tylko własnych stanów emocjonalnych, ale również cudzych. Rozpoznawanie stanów emocjonalnych u siebie i u innych jednostek umożliwia pojawienie się kolejnej zdolności konstytuującej zmysł moralny – zdolności do porównywania wysiłków i zysków; do preferowania czyichś interesów ponad własne, a więc ostatecznie umiejętności odrzucania własnej gratyfikacji. Pojawia się pytanie, co stanowi naturalną bazę do rozpoznawania – przede wszystkim cudzych – stanów emocjonalnych, które są podstawą nie tylko w procesie wartościowania, ale też wynikania abstrakcyjnych cnót moralnych. Poniżej postaramy się znaleźć na to odpowiedź, odnosząc się do koncepcji Fransa de Waala i Patricii Churchland.

2.3. Krótkie podsumowanie

Podsumowując rozważania Hausera, Haidta i Josepha, można stwierdzić, że każdą jednostkę ludzką cechuje wrodzona gotowość do reagowania na bodźce w konkretny, choć ograniczony sposób. Świat społeczny jako środowisko kulturowe różnicuje i wzbogaca te uniwersalne formy reakcji. Przykładowo atawistyczny strach przed wężami może ewoluować w strach przed potworami czy smokami, jednakże nigdy nie będziemy potrafili wzbudzić w sobie naturalnego poczucia strachu na widok choćby kwiatów. Jak dowodzą wymienieni trzej badacze, konsekwencją społecznej modyfikacji instynktownych zachowań są moralne cnoty, stanowiące szczególne atrybuty zachowań – chodzi tylko o te zachowania, które stają się pożądane i naśladowane w danej zbiorowości. Wybrana kultura może odmiennie rozumieć i interpretować cnotę wywodzącą się ze wspólnej, ponadkulturowej reakcji instynktownej. Kultura może również kreować złożone cnoty wywodzące się z kilku automatycznych potrzeb, jak to ma miejsce w przypadku litości, honoru czy patriotyzmu. Zmysł moralności jako zdolność dostrzegania wartości, jak również same abstrakcyjne wartości moralne wydają się bezpośrednio powiązane z emocjami, rozpoznawaniem tych emocji oraz z zachowaniami instynktownymi.

Wnioski, do jakich skłania analiza refleksji samego Jonathana Haidta, to przede wszystkim hipoteza posiadania przez zwierzęta zmysłu protomoralnego. Wiele bodźców aktywizuje zarówno u ludzi, jak i u zwierząt automatyczny odruch pierwotnych emocji aprobaty (do...) lub unikania (od...), który przeistacza się w złożony stan emocjonalny. Z kolei stan emocjonalny posiada walor wartościującej intuicji moralnej, która mobilizuje do działania i preferowania, w tym przedkładania dobra wspólnoty ponad własne. Bezpośrednim źródłem takiej emocji, czy lepiej intuicji moralnej, są powszechne wśród zwierząt, instynktowne *modi* zachowań: reakcje terytorialne, odwzajemnianie przysług, tworzenie koalicji, opieka nad potomstwem, subordynacja wobec lidera. Ciężko jest zaprzeczyć, że na pewne bodźce reagujemy emocjonalnie tak samo, jak inne zwierzęta, lecz nazywamy te reakcje moralnością. Przykładowo ludzie, szympansy i psy jednakowo łączy doświadczenie gniewu za naruszenie własności bądź doświadczenie napaści na koalicję. W tym sensie zwierzęta poza-ludzkie posiadają mniej lub bardziej rozwinięty zmysł moralny oparty na przynależności do

gatunku, zbiorowości i hierarchii społecznej. Zmysł ten dodatkowo jest warunkowany przez dane emocje. Różnica między *Homo sapiens* a innymi zwierzętami jest taka, że ludzie posiadają nieporównywalnie szerszy zakres i większą złożoność doświadczanych emocji moralnych, jak choćby długotrwały altruizm odwzajemniony czy wyłącznie ludzkie wstęty i wstyd. Wielbiona od czasów oświecenia racjonalność pojawia się nie przed podjęciem (moralnie istotnej) decyzji, ale *post hoc*. Owa racjonalizacja zdaje się być kolejną automatyczną potrzebą tzw. logotaksji – potrzebą interpretacji i wyjaśnienia swego zachowania pod względem semantycznym i aksjotycznym [Wilson 1979]. Wówczas ponownie wartościujemy swe działania, porównując, czy ocena czynu pokrywa się z towarzyszącą mu emocją. Wstępny osąd został jednak wydany wcześniej na podstawie szybkich, równie automatycznych intuicji. Można wysnuć wnioski, że poczucie odrazy, jaką budzą w nas zabijanie, kradzież, kazirodztwo czy inne występki, stanowi wytwór mechanizmów biologicznych. Z tych naturalnych algorytmów wywodzą się tysiące norm, obyczajów, rytuałów i zakazów, powstałych na przestrzeni dziejów społecznych interakcji. Jednym zdaniem, moralność i jej zmysł są ewolucyjnie zakorzenione w naturalnych tendencjach stadnych czy – lepiej – społecznych. Stanowią rezultat uzupełniania nieświadomego systemu emocjonalnego przez domieszkę świadomego rozumu.

3. Rozpoznawanie stanów emocjonalnych

Gdy wiemy już, że jest wielce prawdopodobne istnienie zmysłu moralnego, którego korzenie tkwią w naturalnych reakcjach emocjonalno-somatycznych, należy zastanowić się nad zdolnością rozpoznawania tych stanów. Pamiętając, że zmysł moralny to wykraczanie w procesie wartościowania poza egoistyczne interesy i preferowanie dobra innych jednostek, szczególnie istotne wydaje się rozpoznawanie cudzych stanów emocjonalnych. Kwestię tę badał szczegółowo Frans de Waal [2013]. Wychodzi on z (wspólnego nam) założenia, że zmysł moralny w makroskali całych populacji, biocenoz, ekosystemów i biomów to rezultat działania doboru naturalnego. W mikroskali pojedynczych organizmów zarówno zmysł moralny, jak i rozpoznawanie stanów emocjonalnych są skutkiem tzw. zakażenia emocjonalnego oraz zdolności empatii. Te zaś mają swe źródło w mechanizmie neuronów lustrzanych uzupełnionym o markery somatyczne. Z perspektywy de Waala proces podejmowania moralnych wyborów ma następujący schemat: (1) bodziec → (2) aktywność neuronów lustrzanych + znaczniki somatyczne → (3) zakażenie emocjonalne (cielesno-neuronalne reprezentacje) → (4) empatia (pełne rozpoznanie stanów emocjonalnych) → (5) emocje społeczne/emocje moralne (wartościująca intuicja moralna, np. współczucie bądź inne manifestacje moralności). Dalej następują znane nam etapy podjęcia decyzji oraz (pośród *Homo sapiens*) racjonalnej obróbki i wtórnej oceny dokonanego wyboru, a ostatecznie konkretne działanie. Każde przejście „→” jest wynikiem coraz wyższych zdolności kognitywnych wśród zwierząt, gdzie szczególną rolę (na etapie emocji społecznych i emocji moralnych) odgrywają: pamięć; «teoria umysłu» jako umiejętność przyjmowania cudzej perspektywy; empatia kognitywna jako zdolność oceny cudzej sytuacji i celowej pomocy (charakteryzuje człowiekowate). Ponadto na etapie empatii uru-

chamiają się, wskazane przez Haidta i Hausera, automatyczne moduły wzajemności, kooperacji czy wspólnoty/koalicji. Przeanalizujemy dokładnie pierwszą, najważniejszą dla nas część schematu między aktywnością lustrzanych neuronów a empatią.

3.1. Neurony i zakażenie

Neurony lustrzane aktywują się najsilniej podczas wykonywania czynności manualnej oraz w czasie obserwacji czyjejś czynności⁶. W kwestii zmysłu moralnego neurony lustrzane są przede wszystkim mechanizmem odpowiedzialnym za pojawienie się „zakażenia emocjonalnego”, a także empatii i dalszych czynności kognitywnych – naśladowania oraz odczytywania intencji. Równocześnie mózg aktywuje stany somatyczne, czyli markery, które informują podmiot o adekwatnym lub błędnym sposobie jego działania. Informacja ta wykorzystywana jest nieświadomie w rozpoczynaniu czynności, a następnie podczas świadomej racjonalizacji *ex post factum*, jak twierdził de Waal. Akcentowana przez Haidta pierwotna emocja „do...” lub „od...” pojawia się momentalnie już na etapie aktywności neuronów lustrzanych. Aktywność ta to neuronalno-somatyczne mechanizmy inicjujące – według de Waala – najistotniejszą reakcję do rozpoznania emocji oraz samego zmysłu moralnego. Chodzi o reakcję zakażenia emocjonalnego (*emotional contagion*), jak np. doświadczanie niepokoju na widok czyjegoś cierpienia, która to reakcja stanowi wręcz podstawę wszelkich działań moralnych [de Waal 2013, s. 49–50]. Dalszymi warunkami moralności są konsekwencje zakażenia emocjonalnego, czyli empatia i wzajemność.

De Waal, mówiąc o zjawisku zakażenia emocjonalnego, posługuje się pojęciem zaproponowanym przez Hatfield, Cacioppo i Rapsona [1993]. Zakażenie polega na wywołaniu reprezentacji stanu emocjonalnego u drugiego osobnika przez stan emocjonalny pierwszego osobnika, przykładowo niepokoju, radości, złości. Mechanizm emocjonalnego zakażenia zaobserwował już w XIX wieku Robert Vischer [1994], a następnie Theodor Lipps, według którego była to automatyczna, wewnętrzna mimikra uczuć współmieszkańców. Jako zwolennik freudowskiej idei nieświadomego, Lipps umieszczał mimikrę uczuć poniżej progu świadomości, na płaszczyźnie czysto fizjologicznej. Zdaniem de Waala zakażenie emocjonalne transformuje się w empatię wtedy, gdy zwiększają się zdolności poznawcze danego osobnika. Chodzi tu głównie o pojawienie się następujących zdolności: odróżniania siebie od innych, przyjmowania perspektywy innych, rozpoznawania bodźców, które wywołały bieżący stan emocjonalny, autorefleksji. Dalszą konsekwencją zakażenia emocjonalnego są bardziej złożone formy emocji moralno-społecznych, przykładowo współczucie (tendencja do pocieszania, np. obejmowanie i całowanie się szympanсів), bezinteresowna pomoc (wśród szympanсів niespokrewnionych). Wspomniałem wcześniej, że równoległe z uczuciem empatii zakażenie emocjonalne aktywizuje wrodzone moduły moralne.

⁶ „Komórki nerwowe, które realizują określone programy we własnym organizmie i które stają się aktywne także wtedy, gdy obserwujemy lub w inny sposób współodczuwamy, jak osobnik realizuje ów program, nazywane są neuronami lustrzanymi. (...) nie tylko obserwacja, ale także każdego rodzaju spostrzeżenie dotyczące określonej czynności wykonywanej przez innych może doprowadzić do tego, że lustrzane neurony w mózgu obserwatora zaczną wysyłać impulsy nerwowe” [Bauer 2008, s. 21 i nast.].

Dla de Waala nadrzędnym modułem spośród wszystkich wymienionych przez Haidta i Hausera jest moduł wzajemności, czyli wrodzona reakcja odwzajemniania przysług. Reakcja wzajemności najmocniej ze wszystkich modułów wspomaga uczucie empatii, wyzwalaając dalsze złożone emocje moralno-społeczne, aż do pojawienia się bądź moralnych decyzji, bądź abstrakcyjnych cnót.

3.2. Empatia

Jak zostało powiedziane, podstawową reakcją emocjonalną umożliwiającą rozpoznanie cudzych potrzeb i emocjonalnych stanów jest empatia – wynik zakażenia emocjonalnego oraz wrodzonego modułu wzajemności. Zdaniem de Waala to empatia generuje emocje społeczno-moralne prowadzące nie tylko do abstrakcyjnych cnót, ale przede wszystkim do realnych zachowań, między innymi współczucia, altruizmu, poczucia sprawiedliwości, troski zorientowanej na innych, celowego pomagania czy podejmowania ryzyka (baza cnoty „bohaterstwa”) dla czyjegoś dobra. Wszystkie te zachowania były obserwowane wśród zwierząt, zwłaszcza w rządzie naczelnych, w rodzinie *Hominidae*. Otóż empatia to skutek bardziej pierwotnego rozpoznawania obcych emocji za pomocą cielesnych i neuronalnych reprezentacji stanów przedmiotu. Im bliższe sobie (genetycznie lub społecznie) rozpoznający podmiot i rozpoznawany przedmiot, tym łatwiejsza reakcja reprezentowania u podmiotu. Reakcja ta umożliwia podmiotowi «wejście w skórę» i umysł przedmiotu. Empatia jest więc wynikiem mechanizmu percepcja–działanie na poziomie komórkowym (neuronów lustrzanych) oraz somatycznym (markery). Wyłania się w trakcie komunikacji cudzych stanów emocjonalnych oraz automatycznego odwzajemniania przysług. Bywa też pobudzana w procesie uczenia się przez naśladowanie.

Aby w wyniku współdziałania empatii oraz wrodzonych modułów moralnych wygenerować emocje społeczno-moralne, w tym złożoną empatię kognitywną, potrzebny jest adekwatnie wysoki rozwój centralnego układu nerwowego (zwłaszcza kresomózgowia). Jak stwierdza de Waal, dopiero występująca w rządzie prymatów korelacja empatii i odpowiednio rozwiniętych struktur pamięci (krótko- oraz długotrwałej) jest warunkiem możliwości dla: odroczonej wzajemności (wymiany przysług u kapucynek i szympanów), wdzięczności (dzielenia się pożywieniem), pozytywnego nastawienia (pamięci o przysłudze), poczucia sprawiedliwości (porównywania wysiłków i zysków oraz interesów własnych i wspólnych), oczekiwania zysków przy podziale dóbr (baza dla powinnościowego wymiaru moralności). Nie należy przy tym zapominać o ewolucyjnej perspektywie w makroskali. Przetrwanie populacji jest uzależnione od współpracy, od lojalności wobec grupy czy też od skłonności do pomocy innym. Zarówno wśród gryzoni, słoni, wilków, jak i delfinów czy małych automatyczny impuls do pomagania innym zwiększa szanse przetrwania i jest faworyzowany przez dobór naturalny. Innymi słowy, gatunki żyjące stadnie muszą posiadać automatyczne reakcje prospołeczne jako rezultat adaptacji do naturalnego środowiska.

3.3. Krótkie podsumowanie

Pierwszy nasuwający się wniosek jest taki, że pojawienie się dobra moralnego wymaga uprzedniego rozpoznawania stanów emocjonalnych oraz potrzeb innych. Zdolność do czynienia dobra ma swoje korzenie ewolucyjne w doświadczanych emocjach jako mimowolnych reakcjach na bodźce. Rdzeń emocji moralnych i społecznych znajduje się w uczuciu empatii. Zmysł moralny oraz zjawisko moralności są zakorzenione w naszej zwierzęcej naturze. Znaczy to, że ludzie prawdopodobnie są z natury dobrzy, dążąc do zrównoważonej kooperacji, a ową «dobrą» naturę odziedziczyliśmy jako społeczne tendencje oraz wrodzone moduły występujące pośród szympanсів, bonobo czy kapucynek. Innymi słowy, protomoralność nie-ludzi stanowi fundament moralności *Homo sapiens*. Zachowania moralne nie mogą być jedynie zasłoną dla egoistycznych motywacji, nie da się bowiem wskazać źródła takiej ewentualnej dobroci fasadowej. Teoria naturalnego „egoisty aspołecznego”, jak chciałby Hobbes, jest sprzeczna z obserwacjami i danymi dotyczącymi roli emocji w podejmowaniu moralnie istotnych decyzji.

Ponadto można uznać, że de Waal zgodziłby się z *lex continui* Leibniza: zachowania moralne posiadają rozległą skalę stopni, od postaci tak znikomej, że trudno ją nazwać nawet protomoralnością (instynkt macierzyński wśród ryb), poprzez sądy oparte na emocjach oraz zarys perspektywy moralnej u dwu-, trzyletnich dzieci i ssaków społecznych (np. intencyjne pocieszanie się małp), aż po celowe pomaganie, świadome osądy moralne z działającą teorią umysłu (reprezentacje treści cudzego umysłu jako różne od własnego u szympanсів) i ostatecznie ludzkie kodeksy etyczne jako najwyższy stopień skali. Pośród tradycji filozoficznej de Waalowi bliżej jest do Arystotelesowskiej idei abstrakcyjnej moralności niż do koncepcji Epikura czy Hobbesa. Pewne obszary kultury są bowiem odległym rezultatem naturalnych tendencji kooperatywnych czy społecznych instynktów starszych niż gatunek *Homo sapiens*. Zmysł moralny wyewoluował wewnątrz wspólnot jako suma reakcji emocjonalnych, współpracy, wzajemności, dzielenia się, wspólnych interesów w utrzymaniu harmonii i równowagi czy troski o wspólnotę. Poczucie wartości wspólnoty, a więc równoczesna wrogość wobec obcych, wymusiło współpracę i solidarność. Paradoksalnie zjawisko moralności w formie preferowania dobra zbiorowości jest powiązane ze zjawiskiem wojny jako naturalną niechęcią do obcych.

4. Krytyka przedstawionych stanowisk

Poglądy Haidta i Hausera, na których opiera się część niniejszych rozważań, mocno skrytykowała Patricia Churchland [2013, s. 173–195]. Główne zarzuty zostały wymierzone w hipotezę wrodzonych modułów moralności – zdaniem Churchland takie stanowisko nie jest możliwe do obrony z kilku powodów. Otóż sama idea natywizmu jest dość problematyczna i rodzi wiele interpretacji. Wrodzone treści mogą odnosić się albo do genetycznego algorytmu, czyli naturalnego zaprogramowania zachowań, albo do (statystycznej) powszechności występowania. Naturalne skłonności – np. alergie czy predyspozycje do chorób – nie muszą jednak być uniwersalne,

a powszechne występowanie pewnych zachowań (np. używanie danych artefaktów) nie musi mieć genetycznych podstaw. Dość często nie wiadomo, jakie sekwencje genów są odpowiedzialne za konkretne zachowania. Stąd pojęcie wrodzoności wiąże się raczej z powszechnością. Problem w tym, że np. zjawisko masowych ludobójstw w historii kultury Zachodu wydaje się zaprzeczać istnieniu uniwersalnego sumienia czy też istnieniu «dobrej» natury ludzkiej. Ponadto uniwersalność pewnego zachowania może wynikać wyłącznie z pragmatyzmu. Powszechne zachowania są skuteczną formą rozwiązywania geograficznie powszechnie występujących problemów; taktyką łatwego dostępu do tych samych surowców naturalnych. Zatem przyczyna uniwersalnego zachowania może mieć solidne wyjaśnienie funkcjonalne i praktyczne, bez angażowania genetyki. Wartość troski, prawdomówności czy kooperacji posiada większą użyteczność i efektywność rozwiązywania zadań, jakkolwiek wyselekcjonowaną w drodze ewolucji. Kolejny wniosek, jaki się nasuwa, to zbędność posiłkowania się hipotetycznymi modułami wrodzonej moralności oraz pozorna powszechność w obliczu prostszego wytłumaczenia. Co więcej, nieświadome, automatyczne intuicje moralne nie mają żadnego zastosowania w przypadku tak skomplikowanych dylematów etycznych, jak kara śmierci, uczestnictwo w konflikcie zbrojnym, kwestia aborcji i eutanazji czy etyczne uzasadnienie praw zwierząt. Jonathan Haidta zdaje się sugerować ideę, że mechanizmy ewolucji preferują pewne intuicje i reakcje jako podstawowe, a następnie intuicje te wywołują charakterystyczną emocję. Jeżeli nawet jest to prawdą, to nadal problem polega na braku niepodważalnych dowodów empirycznych wskazujących na powszechność i genetyczną wrodzoność wybranych cnót.

4.1. Neurologiczne uzupełnienie modułów moralności

Gdy zgodzimy się z argumentami Churchland, że koncepcje Haidta oraz Hausera budzą zbyt wiele zastrzeżeń, aby uznać je za naukowo wartościowe, to nadal posiadamy alternatywną teorię de Waala. W ujęciu tym wrodzone moduły moralności zostały potraktowane marginalnie na rzecz jednego wyróżnionego modułu wzajemności. Siła stanowiska de Waala odnośnie do zmysłu moralności opiera się na mechanizmie zakażenia emocjonalnego oraz empatii jako jego skutku, a więc na rozpoznawaniu stanów emocjonalnych. Patricia Churchland dodatkowo pogłębia podejście de Waala o wskazanie neurologicznych źródeł zmysłu moralności oraz rozpoznawania emocji na poziomie neurosomatycznym. Aby przedstawić uzupełnienie dokonane przez Churchland, należy krótko przypomnieć dotychczasowe wnioski. Zmysł moralności, jako „kompas wyborów życiowych”, został ufundowany na emocjach moralno-społecznych krystalizujących się w postaci cnót i wartości. Emocje te stanowią konsekwencję zachowań empatycznych (przy współdziałaniu automatycznych reakcji prospołecznych, takich jak odwzajemnianie przysług bądź współpraca), a więc rezultat prawidłowego rozpoznawania swoich oraz cudzych stanów emocjonalnych, które z kolei aktywizują się drogą zakażenia emocjonalnego. Za owo zakażenie odpowiedzialne są markery somatyczne wraz z neuronami lustrzanymi.

Patricia Churchland zgadza się z de Waalem, Haidtem czy Hauserem, że fundamentem praktyk moralnych muszą być zachowania społeczne, wśród których najważniejszą rolę odgrywają: skłonność do opieki i troska, przynależność do grupy,

kooperacja i zaufanie. Moment poszerzenia kręgu troski o innych (najpierw potomstwo, następnie jednostki niespokrewnione) oznacza rozkwit zmysłu moralności. Jednocześnie nie mamy gwarancji, że owe zachowania społeczne są zaprogramowane genetycznie. Możemy za to postawić pytanie, czy istnieje głębszy, bardziej podstawowy poziom – poza neuronami lustrzanymi – stanowiący naturalną bazę empatii oraz zakażenia emocjonalnego. Churchland odpowiada pozytywnie na takie pytanie, podsuwając naukowe dowody dotyczące funkcji konkretnych neuroprzebieżników. Neuronalna konstytucja moralnych praktyk to zarazem podłoże całościowych zachowań społecznych. Zdaniem Churchland neuronalne podłoże cementowane jest współdziałaniem trzech układów:

- Systemu wydzielenia neurotransmiterów – główną rolę konstytucji zachowań społecznych oraz moralnych odgrywają tu oksytocyna i wazopresyna (poczucie zaufania, szczodrości oraz kooperację), dopamina (procesy emocjonalne) oraz serotonina (kontrola impulsywności, agresji).
- Układu nagrody, który odpowiada za doświadczanie bądź emocjonalnego bólu, bądź radości w zależności od decyzji i zachowań. Układ ten umożliwia internalizację standardów społecznych drogą wzmocnienia (nagroda) lub osłabienia (kara). Wzmacniane i nagradzane standardy stają się nawykami, czyli kolektywnymi reprezentacjami treści, które wspólnota uważa za dobre i słuszne. „Ból spowodowany odrzuceniem oraz przyjemność towarzysząca poczuciu przynależności i naśladowaniu zachowań tych, których podziwiamy, powodują silne przekonanie o absolutnej słuszności zachowań” [Churchland 2013, s. 309].
- Systemu neuronów lustrzanych. Jest on odpowiedzialny za rozpoznawanie stanów emocjonalnych innych osób drogą symulacji oraz ewentualnego przypisywania zamiarów i celów. Symulowanie emocji smutku, strachu, złości czy obrzydzenia jako emocjonalna mimikra – zdaniem Churchland – „poprzedza rozpoznanie uczuć innych i stanowi podstawę do empatii” [Churchland 2013, s. 309]. Rozpoznanie cudzego stanu mentalnego, jako wstęp do moralności, wymaga mimikry oraz życia społecznego umożliwiającego taką mimikrę⁷.

Istotnym dla moralności i zachowań społecznych miejscem anatomicznym jest kora przedczołowa, związana ze sprawnością ludzkich zachowań społecznych, zdolnością do przewidywania zdarzeń, odrzucania gratyfikacji i zachowania samokontroli. Jak wiadomo, dzięki osobie Phineasa Gage’a czy pacjentom Damasio, uszkodzenie tej części kory skutkuje między innymi brakiem empatii, stopieniem reakcji emocjonalnych itd.

⁷ Rola neuronów lustrzanych w rozpoznawaniu i symulacji stanów emocjonalnych jest wciąż kwestią niejasną i dyskutowaną, między innymi ze względu na ograniczenia funkcjonowania neuronów lustrzanych, wynikające z konieczności percepcji. Wydaje się, że niekiedy jesteśmy w stanie rozumieć czyjeś emocje bez udziału bezpośredniej percepcji – wystarczą dane kontekstowe. Ponadto Shaun Gallagher krytykuje tezę, jakoby neurony lustrzane cokolwiek symulowały, pojęcie symulacji zakłada bowiem proces dwuetapowy, działanie neuronów zaś wydaje się jednoetapową percepcją-działaniem [Gallagher 2005].

4.2. Krótkie podsumowanie

Zdaniem Patricii Churchland poprawnie funkcjonująca korelacja neuronalna trzech systemów prowadzi – poprzez reakcje emocjonalne – do zachowań społecznych, takich jak troska, zaufanie, odwzajemnianie przysług. Im bardziej złożona budowa centralnego układu nerwowego, tym zachowania społeczne stają się bardziej zaawansowane – opieka i troska rozszerzają się na organizmy coraz bardziej oddalone więziami biologicznymi. Te zaś zachowania przekazywane są już w formie transmisji kulturowej poprzez imitację prototypów zachowań i lokalnych praktyk wspólnotowych. Ocena działania jako dobre lub złe jest zakorzeniona (naturalnie) w emocjach oraz (kulturowo) w nawykach społecznych. Poziom naturalny pozostaje jednak fundamentalny względem wtórnego poziomu społecznego. Gdyby opisać w trzech etapach proces wyłaniania się zmysłu moralnego według Churchland, wyglądałby on następująco:

1. Etap, w którym presja selekcyjna wytworzyła troskę organizmu o samego siebie. Troska ta polega na podtrzymywaniu dobrostanu i unikaniu złego samopoczucia. Stanowi główną funkcję układu nerwowego. Może być zredukowana do wartości przeżycia oraz dobrego samopoczucia jako unikania bólu.
2. Etap rozszerzenia troski na swoje potomstwo. Dobrostan potomstwa ma znaczenie dla przetrwania jednostki. Nowo narodzone potomstwo zostaje instynktownie włączone we własny układ homeostatyczny.
3. Etap ponownego rozszerzenie troski na bliskich krewnych i przyjaciół. Zachodzi dzięki sprawnemu funkcjonowaniu reakcji empatycznych, a te stanowią rozwinięcie funkcji rozpoznawania stanów emocjonalnych oraz działania neuronów lustrzanych.

5. Zakończenie

Z powyższych rozważań wyłania się pogląd, że umysł ludzki jest raczej narzędziem praktycznego wygrywania konfliktów niż abstrakcyjnego poszukiwania dobra lub prawdy. Konflikty implikują pojęcie moralności, a zwycięstwa w konfliktach są uzależnione głównie od emocjonalnego zaangażowania oraz, być może, instynktownej oceny, w mniejszym zaś stopniu od racjonalnej kalkulacji. Gdybyśmy posługiwali się wyłącznie rozumem, wówczas poza niezdolnością do szybkiego szacowania oraz działania, dokonywalibyśmy autodestrukcyjnych decyzji, logicznie wybierając np. stronę naszych wrogów lub zabijając niewinne jednostki. Realnym przykładem są osoby dotknięte psychopatią, które – choć znają racjonalne normy społeczne i zasady etyczne – postępują niemoralnie z powodu nieprawidłowości w działaniu mózgu. Nieprawidłowości te polegają na braku doświadczania emocji moralnych, jak współczucie lub poczucie winy. Również czysto racjonalnie wywnioskowana złota reguła etyczna nie jest ani bezwarunkowa, ani powszechnie stosowana.

Jeżeli zmysł moralności nie wymagałby zdolności rozpoznawania emocji, ale np. mowy, wtedy faktycznie tylko ludzie by go posiadali. Jak wskazują obserwacje i badania, rozpoznawanie cudzych emocji na podłożu neuronalnym występuje w świecie zwierząt. Jeśli dodatkowo moralność identyfikujemy z wartościami społecznymi,

np. opieka czy kooperacja, wówczas posługują się nią również zwierzęta. „Moralność jest zjawiskiem naturalnym – ograniczonym przez siły doboru naturalnego, zakorzenionym w neurobiologii, kształtowanym przez lokalne środowisko i modyfikowanym przez wydarzenia kulturowe. (...) Moralność jest zakorzeniona w naszej biologii, w naszej zdolności do współczucia i naszej zdolności do uczenia się i rozumienia” [Churchland 2013, s. 307–320]. Można słusznie wnioskować, że – wbrew Kantowi – rudymenarne zasady i obowiązki nie wynikają bezpośrednio z czystego rozumu, sam rozum zaś warunkowany jest przez naturalne mechanizmy, których nigdy nie kontroluje.

BIBLIOGRAFIA

- Bauer J. (2008). *Empatia. Co potrafią lustrzane neurony*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Churchland P. (2011). *Braintrust: What Neuroscience Tells Us about Morality*. Princeton and Oxford: Princeton University Press.
- Churchland P. (2013). *Moralność mózgu*. Kraków: Copernicus Center Press.
- Damasio A. (2011). *Błąd Kartezjusza*. Poznań: Dom Wydawniczy Rebis.
- Waal, F., de (2006). *Primates and Philosophers. How Morality Evolved*. Princeton: Princeton University Press.
- Waal, F., de (2013). *Małpy i filozofowie*. Kraków: Copernicus Center Press.
- Eibl-Eibesfeldt I. (1987). *Miłość i nienawiść*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Gallagher S. (2005). *How the Body Shapes the Mind*. Oxford: Oxford University Press / Clarendon Press.
- Gazzaniga M. (2011). *Istota człowieczeństwa*. Sopot: Smak Słowa.
- Gazzaniga M. (2005). *The Ethical Brain*. Chicago: Dana Press.
- Goodall J. (1997). *Przez dziurkę od klucza*. Warszawa: Prószyński i S-ka.
- Greene J., Sommerville R., Nystrom L. (2001). *An fMRI investigation of emotional engagement in moral judgment*. „Science” 293 (5537), s. 2105–2108.
- Haidt J. (2003). *The Moral Emotions* [w:] R.J. Davidson, K.R. Scherer (red.), *Handbook of Affective Sciences* (s. 852–870). Oxford: Oxford University Press.
- Haidt J., Joseph C. (2004). *Intuitive ethics: how innately prepared intuitions generate culturally variable virtues*. „Daedalus” 133 (4), s. 55–66.
- Hatfield E., Cacioppo J., Rapsod R. (1993). *Emotional contagion*. „Current Directions in Psychological Sciences” 2, s. 96–99.
- Hauser M.D. (2006). *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York: HarperCollins Ecco Press.
- Helvétius C. (1976). *O człowieku, jego zdolnościach umysłowych i wychowaniu*. Wrocław: Państwowa Akademia Nauk.
- Kant I. (1974). *Krytyka praktycznego rozumu*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Kant I. (1984). *Uzasadnienie metafizyki moralności*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Lieberman D., Tooby J., Cosmides L. (2003). *Does morality have a biological basis?* „Proceedings of the Royal Society: Biological Sciences” 270 (1517), s. 819–826.
- Lorenz K. (1972). *Tak zwane zło*. Warszawa: Państwowy Instytut Wydawniczy.
- Ridley M. (2000). *O pochodzeniu cnoty*. Poznań: Dom Wydawniczy Rebis.
- Singer P. (1981). *The Expanding Circle: Ethics and Sociobiology*. Oxford: Oxford University Press.

- Skolimowski H. (1991). *Medytacje o prawdziwych wartościach człowieka*. Wrocław: Astrum.
- Vischer R. (1994). *Empathy, Form and Space: Problems in German Aesthetics 1873–1893*. Santa Monica: University of Chicago Press.
- Westoll A. (2013). *Szympany z azylu Fauna*. Wołowiec: Wydawnictwo Czarne.
- Wilson E.O. (1979). *On Human Nature*. Harvard: Harvard University Press.
- Wright R. (2004). *Moralne zwierzę*. Warszawa: Prószyński i S-ka.



Niezależność czy współpraca? Zagadnienie interakcji wzbudzeniowego i wykonawczego systemu uwagi

Independence or cooperation? On the issue of interaction between alerting and executive attention systems

Abstract: According to Posner and colleagues, mechanism of attention consists of three systems: alerting, orienting and executive. Here, we focus on relationship between two of them, namely, on the interactions between alerting and executive attention in tasks involving cognitive or motor conflicts. Previous studies suggest that the impact of alerting on the efficiency of conflict resolution may depend on various factors, including the presence of established stimulus-response associations, speed of the reaction selection process, characteristics of the allocation of spatial attention, and a stage of processing at which conflict occurs i.e., either perceptual analysis or response selection. However, the available results are to some extent contradictory and no single hypothesis is sufficient to explain all cases of the interactions between alerting and executive attention observed so far. The present paper provides a comprehensive review of these results and hypothetical mechanisms underlying the relationship between alerting and executive attention.

Key words: attention, alerting, executive attention, interaction, cognitive conflict

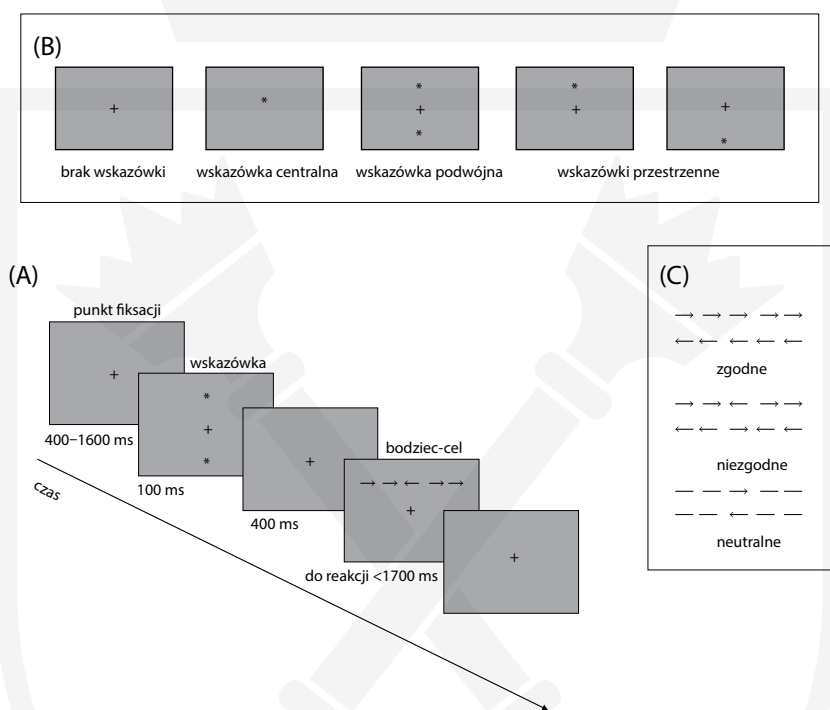
Trzy systemy uwagi

Według teorii uwagi proponowanej przez Posnera i współpracowników [Posner i Petersen 1990; Posner i Rothbart 2007], mózgowy mechanizm uwagi składa się z trzech systemów: wzbudzeniowego (*alerting*), orientacyjnego (*orienting*) oraz wykonawczego (*executive*). Pełnią one odmienne funkcje, a ich podłożem mózgowym są względnie niezależne od siebie sieci neuronalne [Posner i Petersen 1990; zob. też: Asanowicz, Siedlecka i Michalczyk 2009; Wronka 2004]. System wzbudzeniowy odpowiada za wzbudzenie i utrzymanie określonej wrażliwości systemów sensorycznych na bodźce zewnętrzne. Jest automatycznie angażowany zawsze wtedy, gdy zmysły zarejestrują nowy, zwłaszcza nieoczekiwany bodziec (np. nagły bodziec wizualny lub dźwiękowy). Zaangażowanie systemu wzbudzeniowego stanowi zatem między innymi część reakcji orientacyjnej [Sokolov 1963]. Działanie systemu wzbudzeniowego nie powinno

być jednak utożsamiane z pobudzeniem (*arousal*), gdyż nie wiąże się z pobudzeniem organizmu (np. emocjonalnym lub fizjologicznym), lecz jedynie z uaktywnieniem bądź zwiększeniem gotowości systemów sensorycznych. Neuronalne podłoże systemu wzbudzeniowego jest lokalizowane w obrębie wzgórza oraz płatów czołowych i ciemieniowych [Fan, McCandliss, Fossella, Flombaum i Posner 2005]. Funkcją systemu orientacyjnego jest selekcja informacji, w tym kierowanie ogniska uwagi w określone miejsce w przestrzeni, np. w kierunku nieoczekiwanego dźwięku lub w miejsce, w którym ma się pojawić oczekiwany bodziec wzrokowy. System orientacyjny jest angażowany w takich zadaniach, jak przeszukiwanie pola wzrokowego (*visual search*) [Treisman 1980] i wskazywanie lokalizacji (*cueing task*) [Posner 1980]. Neuronalnym podłożem orientacji uwagi są między innymi: górna kora ciemieniowa (*superior parietal lobe*), styk skroniowo-ciemieniowy (*temporo-parietal junction*) oraz przednie pole wzrokowe (*frontal eye field*) [Corbetta i Shulman 2002; Fan i in. 2005]. Z kolei system wykonawczy pełni funkcje, które ogólnie można nazwać kontrolą oraz regulacją myśli i działań bądź zarządzaniem szeroko rozumianym zachowaniem [Posner i Fan 2008]. Jest to między innymi: monitorowanie procesu wyboru poprawnej reakcji, detekcja błędu i rozwiązywanie konfliktu pomiędzy niespójnymi, konkurującymi reakcjami lub alternatywnymi decyzjami [por. Norman i Shallice 1986]. Główne ośrodki neuronalne wchodzące w skład sieci uwagi wykonawczej to przednia część zakrętu obręczy (*anterior cingulate cortex*) oraz brzuszno-boczna i grzbietowo-boczna kora przedczołowa (*ventrolateral prefrontal cortex, dorsolateral prefrontal cortex*) [Botvinick, Cohen i Carter 2004; Fan, Hof, Guise, Fossella i Posner 2008; Fan i in. 2005].

Do pomiaru efektywności funkcjonowania systemów uwagi opracowano komputerowy test sieci uwagowych (*Attention Network Test, ANT*) [Fan, McCandliss, Sommer, Raz i Posner 2002]. ANT jest połączeniem dwóch klasycznych zadań eksperymentalnych: zadania flankerów [Eriksen i Eriksen 1974] oraz zadania wskazywania lokalizacji [Posner 1980], co umożliwia pomiar trzech aspektów uwagi w jednym zadaniu. Procedura ANT (zob. Rysunek 1) składa się zwykle z kilkuset prób, w których uczestnicy badania identyfikują kierunek prezentowanej na ekranie komputera strzałki. Na strzałkę wskazującą lewą stronę należy reagować naciśnięciem np. lewego klawisza myszki, a na strzałkę w prawo – klawisza prawego. Mierzone są czas oraz poprawność reakcji. Strzałka-cel jest prezentowana w otoczeniu tzw. flankerów, czyli bodźców otaczających cel. Flankery mogą być spójne z wymaganą reakcją (np. strzałki wskazujące ten sam kierunek, co bodziec-cel), niespójne (np. strzałki wskazujące kierunek odwrotny do wskazywanego przez bodziec-cel) lub neutralne (np. proste linie otaczające strzałkę-cel). Gdy flankery są niespójne z celem, zostają uaktywnione dwa przeciwstawne wzorce reakcji, powodując konflikt. Wykonanie poprawnej reakcji wymaga zatem rozwiązania konfliktu. Miarę efektywności uwagi w rozwiązywaniu konfliktu otrzymujemy przez odjęcie reakcji (czasu lub poprawności) w warunku spójnym od reakcji w warunku niespójnym. Aby zmierzyć zaangażowanie systemów wzbudzeniowego i orientacyjnego, wyświetlenie bodźca docelowego jest poprzedzane prezentacją jednej z kilku możliwych rodzajów wskazówek (*cues*). Wskazówką może być np. znak gwiazdki („*”) wyświetlony kilkaset milisekund przed bodźcem docelowym. Wskazówka wyświetlona w lokalizacji bodźca docelowego pełni rolę wskazówki orientacyjnej (*orienting cue*). Porównanie czasu reakcji na cel poprzedzany wskazów-

ką orientacyjną z czasem reakcji na cel poprzedzany wskazówką wyświetloną w miejscu fiksacji wzroku (czyli na środku ekranu) informuje nas o tym, ile czasu zajmuje przeniesienie ogniska uwagi ze środka ekranu do lokalizacji, w której wyświetlono cel. Z kolei porównanie reakcji na cel poprzedzany wskazówką wzbudzeniową (*alerting cue*; np. wskazówką centralną wyświetloną w miejscu fiksacji wzroku lub wskazówką podwójną, tj. dwiema gwiazdkami prezentowanymi jednocześnie w miejscach, w których może pojawić się cel) [por. Fan i in. 2005] z reakcją na cel niepoprzedzony żadną wskazówką informuje nas o tym, jak duży jest zysk z zaangażowania uwagi wzbudzeniowej, tj. o ile szybsze lub bardziej poprawne są reakcje na cel poprzedzony wskazówką. Podsumowując, wskazówka wzbudzeniowa sygnalizuje czas prezentacji celu, natomiast wskazówka orientacyjna sygnalizuje zarówno czas, jak i lokalizację celu.



Rysunek 1. Procedura zadania ANT (Attention Network Test) [Fan i in. 2002]: (A) przykład procedury eksperymentalnej ze wskazówką podwójną i flankerami niezgodnymi z bodźcem docelowym (strzałką środkową), (B) rodzaje wskazówek, (C) rodzaje flankerów

Badania z zastosowaniem ANT przynoszą wyniki potwierdzające względną niezależność trzech systemów uwagi zarówno na poziomie behawioralnym [Fan i in. 2002], jak i neuronalnym [Fan i in. 2005]. Nie oznacza to jednak, że należy oczekiwać zupełnej niezależności pomiędzy nimi. Skoordynowane działanie systemu poznawczego wymaga skoordynowanego działania systemów uwagowych, które nadzorują i regu-

lują jego pracę. Taką współzależność zaobserwowano w niektórych eksperymentach. Po pierwsze, o ile w części eksperymentów z ANT nie stwierdzono korelacji pomiędzy poszczególnymi wskaźnikami efektywności systemów uwagowych [np. Fan i in. 2002], o tyle w innych pracach takie korelacje zaobserwowano [Fossella, Sommer, Fan, Wu, Swanson, Pfaff i Posner 2002; MacLeod, Lawrence, McConnell, Eskes, Klein i Shore 2010]. Po drugie, w szeregu badań z zadaniem ANT zaobserwowano związki interakcyjne pomiędzy uwagą orientacyjną i efektem wzbudzenia [Callejas, Lupiáñez i Tudela 2004; Callejas, Lupiáñez, Funes i Tudela 2005; Fuentes i Campoy 2008], uwagą orientacyjną i efektywnością uwagi wykonawczej [Callejas i in. 2004, 2005; Fan, Gu, Guise, Liu, Fossella, Wand i Posner 2009; Fan i in. 2002; Ishigami i Klein 2010; MacLeod i in. 2010] oraz pomiędzy efektem wzbudzenia i efektywnością uwagi wykonawczej [Callejas i in. 2004, 2005; Fan i in. 2009; Fan i in. 2002; Ishigami i Klein 2010]. Zaobserwowano między innymi, że: zaangażowanie systemu wzbudzeniowego przyspiesza proces reorientacji ogniska uwagi [Callejas i in. 2004, 2005], poprawne ukierunkowanie systemu orientacyjnego zwiększa efektywność systemu wykonawczego w rozwiązywaniu konfliktu [Callejas i in. 2004, 2005; Fan i in. 2009], zaangażowanie systemu wzbudzeniowego zmniejsza efektywność systemu wykonawczego w rozwiązywaniu konfliktu [Callejas i in. 2004, 2005; Fan i in. 2002; Fan i in. 2009].

Wpływ wzbudzenia na efektywność uwagi wykonawczej

O ile relacje pomiędzy wzbudzeniem i orientacją uwagi oraz pomiędzy orientacją uwagi i efektywnością rozwiązywania konfliktu są konsekwentnie potwierdzane w kolejnych eksperymentach i teoretycznie zrozumiałe [por. Callejas i in. 2005], o tyle mechanizm wpływu wzbudzenia na uwagę wykonawczą wciąż pozostaje niejasny. Co więcej, nie są znane warunki konieczne i wystarczające do tego, aby taki interakcyjny efekt zaobserwować. Z jednej strony w standardowym zadaniu ANT [Fan i in. 2002] regularnie obserwuje się negatywny wpływ wzbudzenia uwagi na efektywność rozwiązywania konfliktu, tj. obserwowany koszt konfliktu jest większy w próbach ze wskazówką wzbudzeniową niż w próbach bez wskazówki [zob. metaanalizę w: MacLeod i in. 2010]. Efekt ten jest obserwowany zarówno w pomiarze czasu reakcji, jak i poprawności, z tym że pomiar poprawności w zadaniu ANT nie zawsze przynosi wyniki istotne statystycznie, ze względu na bardzo wysoką, przekraczającą 95%, ogólną poprawność wykonania zadania. Podobną zależność pomiędzy efektem wzbudzenia a rozwiązywaniem konfliktu zaobserwowano, używając zmodyfikowanej wersji zadania (tzw. ANTI), w której rolę wskazówki wzbudzeniowej pełnił sygnał dźwiękowy prezentowany w miejsce wskazówki wizualnej [Callejas i in. 2004, 2005; Chica, de Schotten, Toba, Malhotra, Lupiáñez i Bartolomeo 2012]. Jednak z drugiej strony późniejsze prace sugerują, że ten negatywny efekt wzbudzenia może być ograniczony do zadania flankerów, ponieważ wydaje się nie występować w innych zadaniach z konfliktem, takich jak zadanie Stroopa [Weinbach i Henik 2012]. Co więcej, stwierdzono również, że gdy trudność zadania ANT się zwiększa (tzn. ogólna poprawność spada poniżej 90%), to zaangażowanie systemu wzbudzeniowego zwiększa, a nie zmniejsza efektywność rozwiązywania konfliktu w pomiarze poprawności, przy jednocześnie

obserwowanym braku interakcji pomiędzy wzbudzeniem i konfliktem w pomiarze czasu reakcji [Asanowicz, Marzecová, Jaśkowski i Wolski 2012; Marzecová, Asanowicz, Krivá i Wodniecka 2013].

Wzrost kosztu konfliktu wskutek zaangażowania systemu wzbudzeniowego był początkowo wyjaśniany na dwa sposoby (raczej uzupełniające się wzajemnie niż wykluczające). Po pierwsze, wskazywano, że efekt wzbudzenia wiąże się z automatycznym przyspieszeniem procesów przetwarzania informacji i selekcji reakcji, co w sytuacji konfliktu reakcji nie pozostawia wystarczającej ilości czasu na jego efektywne rozwiązanie [Fossella i in. 2002; Posner 2008]. (Skądinąd wiadomo, że błędy w warunku konfliktu popełniane są głównie wtedy, gdy reakcje są szybkie, czyli gdy aktywizacja błędnego programu motorycznego jest na tyle silna, że jest uruchamiana automatycznie i nie można jej zahamować. Gdy reakcje są wolniejsze, są też na ogół poprawne, pomimo konfliktu, ponieważ dłuższy czas przygotowania reakcji ułatwia uniknięcie błędu; zob. np. Ridderinkhof, van den Wildenberg i Wylie 2012). Po drugie, przywoływano tzw. hipotezę czyszczenia świadomości (*clearing of consciousness*) [Posner 1994]. Zgodnie z tą hipotezą nagle pojawienie się w otoczeniu niezidentyfikowanego jeszcze bodźca hamuje bieżącą, wewnętrzną aktywność systemu wykonawczego i kieruje „zasoby” uwagi na „zewnątrz” systemu, między innymi poprzez przyspieszenie działania systemu orientacyjnego [por. Callejas i in. 2005]. Innymi słowy, aktualna treść procesów umysłowych zostaje „wyczyszczona”, aby zrobić miejsce informacji o nowym, potencjalnie ważnym bodźcu. Na poziomie neuronalnym efekt ten ma być związany między innymi z obniżeniem aktywności struktur składających się na sieć uwagi wykonawczej, w tym przedniej części zakrętu obręczy i bocznej kory przedczołowej [Posner 1994]. Niemniej na razie brakuje potwierdzonych dowodów wspierających tę predykcję.

Alternatywne wyjaśnienie relacji pomiędzy wzbudzeniem i rozwiązywaniem konfliktu zaproponowali niedawno Fischer, Plessow i Kiesel [2010, 2011]. Badacze do wywołania konfliktu reakcji użyli zadania Simona [Simon 1990], a efekt wzbudzenia był wywoływany prezentacją krótkiego bodźca dźwiękowego. Konflikt w badaniu Simona powstaje na skutek niezgodności przestrzennej pomiędzy bodźcem i reakcją. Przykładowo bodziec prezentowany po lewej stronie wykonującej zadanie osoby badanej automatycznie uaktywnia najbardziej adekwatny program działania, czyli reakcję lewą ręką [zob. np. Hommel 2004]. Jeśli jednak jest to bodziec, na który zgodnie z instrukcją zadania należy zareagować prawą ręką, pojawia się podobny do obserwowanego w zadaniu flankerów konflikt pomiędzy dwoma alternatywnymi planami działania. Fischer i współpracownicy [2010, 2011] w swoich badaniach, podobnie jak inni badacze w standardowym zadaniu ANT, zaobserwowali zwiększenie kosztu konfliktu w próbach angażujących system wzbudzenia. Jednak tłumaczą ten efekt w inny sposób, odwołując się do modelu tzw. podwójnego procesu (*dual-process* lub *dual-route*) [De Jong, Liang i Lauber 1994]. Zgodnie z tym modelem plan działania (reakcji) jest aktywowany na dwa sposoby: automatycznie, drogą oddolną (wstępującą), poprzez automatyczne odczytanie takich cech, jak położenie bodźca lub kierunek strzałki, oraz w sposób kontrolowany, drogą odgórną (zstępującą), poprzez odczytanie zgodnego z instrukcją zadania znaczenia bodźca. W warunku konfliktowym automatycznie uaktywniony wzorzec jest niezgodny ze wzorcem aktywowanym w sposób

kontrolowany. Fischer i współpracownicy [2010, 2011] dowodzą, że efekt wzbudzenia oddziałuje na oba procesy, przyspieszając tym samym proces selekcji i uruchomienia reakcji (hipoteza *facylitacji aktywacji reakcji*), co w konsekwencji zwiększa zarówno szanse popełnienia błędu, jak i koszt konfliktu w reakcjach poprawnych.

Zgodnie z powyższą hipotezą *facylitacji aktywacji reakcji*, wskazówka wzbudzeniowa zwiększa dostępność lepiej utrwalonych bądź bardziej naturalnych związków bodziec-reakcja (takich jak zgodność przestrzenna bodźca i reakcji w zadaniu Simona lub kierunek strzałki i reakcja w zadaniu flankerów). Są to związki oparte na prostym skojarzeniu reprezentacji bodźca z reakcją motoryczną, z pominięciem głębszego przetwarzania (np. na poziomie semantycznym). W konsekwencji wzbudzenie systemu powinno mieć negatywny wpływ na konflikt jedynie wtedy, gdy takie związki są silne. Przypuszczenia te potwierdzono w badaniu z zastosowaniem zadania kategoryzacji emocjonalnej walencji bodźców z użyciem flankerów werbalnych. Wykorzystano trzy rodzaje flankerów. W wariancie pierwszym flankerami były czasowniki pozytywne lub negatywne, które pojawiały się w innych próbach jako bodźce docelowe. W kolejnym wariancie flankery były nazwami kategorii (pozytywne/negatywne). Natomiast w wersji trzeciej flankerami były słowa należące do tej samej kategorii semantycznej co target, jednakże bodźce użyte jako flankery nigdy nie pojawiały się na pozycji targetu. Mimo że każdy typ flankerów w warunkach niezgodnym doprowadził do powstania podobnej wielkości efektu konfliktu (między dwoma alternatywnymi planami działania), to tylko konflikt powstały na bazie silnych związków reprezentacji wzrokowej bodźca z reakcją motoryczną (wariant pierwszy) został nasilony przez działanie wskazówki wzbudzeniowej. Wydaje się więc, że wpływ wskazówki wzbudzeniowej jedynie na specyficzny rodzaj konfliktu przeczy opisanej wcześniej hipotezie osłabiania kontroli wykonawczej [Callejas 2004, 2005; Posner 1994, 2008], ponieważ w myśl tej ostatniej wywołane prezentacją wskazówki hamowanie uwagi wykonawczej powinno wywierać negatywny wpływ na efektywność rozwiązywania konfliktu niezależnie od jego specyfiki.

Inne wyjaśnienie relacji pomiędzy wzbudzeniem uwagi i rozwiązywaniem konfliktu oferuje model zaproponowany przez Nieuwenhuisa i de Kleijna [2013]. Model ten bazuje na dwóch podstawowych założeniach: 1. wzbudzenie uwagi skraca czas kodowania bodźca, co jest powodem przyspieszenia procesu selekcji reakcji [Seifried, Ulrich, Bausenhardt, Rolke i Osman 2010; por. Posner 2008]; 2. uruchomienie i efektywne wdrożenie procesu kontrolnego wymaga zwykle kilkuset milisekund od momentu prezentacji bodźca [Cohen, Servan-Schreiber i McClelland 1992; Gratton, Coles, Sirevaag, Eriksen i Donchin 1988]. Przyczyną negatywnego wpływu wzbudzenia na efektywność rozwiązywania konfliktu byłoby więc po prostu skrócenie czasu reakcji, co skutkuje skróceniem czasu na korektę błędnej reakcji w warunkach konfliktu. Symulacja komputerowa opisywanego modelu z użyciem zadania flankerów była zgodna z przewidywaniami badaczy, ujawniając dobrze dopasowany do wyników behawioralnych efekt skrócenia czasów reakcji po prezentacji wskazówki wzbudzeniowej, przy jednoczesnym zmniejszeniu ogólnej poprawności i wzroście kosztu konfliktu w próbach poprawnych [Nieuwenhuis i de Kleijn 2013]. Główną różnicą między proponowanym modelem a pozostałymi hipotezami jest to, że w tym modelu wzbudzenie uwagi nie modyfikuje efektywności funkcjonowania kontroli poznaw-

czej bezpośrednio, lecz jedynie prowadzi do wyboru reakcji, zanim procesy kontroli poznawczej zostaną w pełni rozwinięte i wdrożone. A zatem odpowiednie manipulacje strategią podejmowania decyzji, prowadzące do preferowania szybkości reakcji nad jej poprawnością, lub odwrotnie, poprawności reakcji nad jej szybkością, powinny systematycznie modulować efekt interakcyjny pomiędzy wzbudzeniem a rozwiązywaniem konfliktu. Ta hipoteza nie była jednak dotychczas badana empirycznie.

Alternatywne ujęcie mechanizmu leżącego u podłoża interakcji pomiędzy systemami wzbudzeniowym i wykonawczym zaproponowali Weinbach i Henik [2011, 2012, 2013]. Autorzy wysunęli tzw. hipotezę modulacji alokacji uwagi, zgodnie z którą zaangażowanie systemu wzbudzenia priorytetyzuje przetwarzanie informacji przestrzennych w polu widzenia. Jeśli więc bodziec-cel jest otoczony przez aktywujące przeciwny wzór reakcji flankery, pod wpływem wskazówki wzbudzeniowej są one przetwarzane skuteczniej, nasilając tym samym konflikt reakcji. Warunkiem koniecznym do wystąpienia negatywnego efektu wzbudzenia na efektywność rozwiązywania konfliktu jest tutaj przestrzenne rozdzielanie cechy bodźca wywołującej błędną automatyczną reakcję i cechy, zgodnie z którą należy reagować. Hipotezę tę autorzy potwierdzili w szeregu badań, między innymi w eksperymencie z modyfikacją klasycznego zadania Navona [1997], w której wywołany wskazówką wzbudzeniową wzrost konfliktu (efekt podobny jak w standardowym zadaniu ANT) był obserwowany w przypadku konieczności identyfikacji bodźca lokalnego przy jednoczesnym ignorowaniu informacji globalnej, ale już nie w sytuacji odwrotnej, tj. konieczności identyfikacji bodźca globalnego przy ignorowaniu informacji lokalnej [Weinbach i Henik 2011]. Zgodne z tą hipotezą są również wyniki ukazujące brak wpływu wskazówki wzbudzeniowej na wielkość konfliktu w klasycznym zadaniu Stroopa [Weinbach i Henik 2012]. W tym zadaniu cecha wymagająca reakcji (kolor czcionki) oraz aspekt dystrakcyjny bodźca (znaczenie słowa np. „czarny” – warunek zgodny, lub „czerwony” – warunek konfliktowy) są zintegrowane w jeden obiekt, a więc obecność wskazówki mającej modulować przetwarzanie jedynie informacji przestrzennych nie ma tutaj znaczenia. Podobny efekt uzyskano, zastępując słowo pojedynczą kolorową strzałką (będącą swego rodzaju hybrydą zadania Stroopa i zadania flankerów) [Weinbach i Henik 2012]. W tej procedurze instrukcja wymagała reakcji zależnej od koloru bodźca, a kierunek strzałki pełnił funkcję dystraktora. Również tutaj nie zaobserwowano wpływu wskazówki na koszt konfliktu. Wyniki tego eksperymentu nie tylko pozostają w zgodzie z hipotezą modulacji alokacji uwagi [Weinbach i Henik 2011, 2012], ale też przeczą wspomnianej wcześniej koncepcji facylitacji aktywacji reakcji [Fischer i in. 2010, 2011]. Mimo że konflikt wynika z istnienia silnych związków pomiędzy bodźcem i reakcją, zaangażowanie systemu wzbudzenia uwagi nie wywiera wpływu na wielkość kosztu interferencji.

Podsumowując, wyniki badań Weinbacha i Henika [2011, 2012] podkreślają rolę przetwarzania informacji przestrzennych dla powstania interakcji między wzbudzeniowym oraz wykonawczym systemem uwagi, upatrując w przestrzennym rozdzielaniu celu oraz dystraktorów warunku koniecznego i wystarczającego do jej wystąpienia. Stanowią zarazem wyzwanie dla wyjaśnień proponowanych przez Posnera [1994], Callejas i in. [2004] oraz Fischera i in. [2011]. Niestety, żadna z zaproponowanych dotychczas koncepcji teoretycznych nie wyjaśnia odwrócenia efektu interakcyjnego

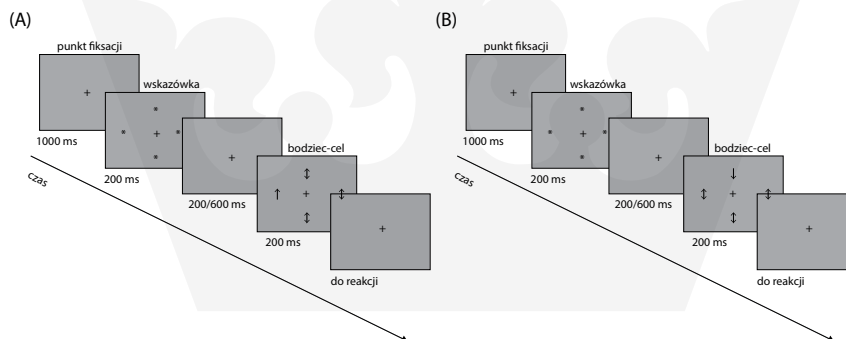
pomiędzy wzbudzeniem uwagi i rozwiązywaniem konfliktu, obserwowanego przy zwiększonym poziomie trudności zadania ANT [Asanowicz i in. 2012; Marzecová i in. 2013].

Poziomy konfliktu a wpływ wzbudzenia na jego rozwiązywanie

Większość dotychczasowych prac dotyczących interakcji pomiędzy wzbudzeniowym i wykonawczym systemem uwagi koncentrowała się na efektach związanych z konfliktem reakcji. Nie jest to jednak jedyny rodzaj konfliktu, którego rozwiązanie wymaga zaangażowania uwagi wykonawczej. Konflikt może być również wywołany przez niespójność w ramach percepcyjnej reprezentacji bodźca, np. wtedy, gdy dwa aspekty bodźca zawierają informacje sugerujące bądź automatycznie aktywujące dwie przeciwstawne reakcje. Taka sytuacja ma miejsce w tzw. przestrzennej wersji zadania Stroopa (*spatial Stroop*) [Liu, Banich, Jacobson i Tanabe 2004]. W tym zadaniu konflikt ma miejsce wtedy, gdy dwa atrybuty tego samego bodźca interferują między sobą. Może to być np. znaczenie słowa lub kierunek strzałki, które interferuje z jej położeniem: słowo „góra” (lub strzałka wskazująca górę) jest umiejscowione u góry (warunek zgodny) lub u dołu (warunek konfliktowy) ekranu monitora. Konflikt na poziomie reakcji, np. w zadaniu Simona, jest często określany jako konflikt S-R (*stimulus-response conflict*), natomiast konflikt na poziomie bodźca – jako konflikt S-S (*stimulus-stimulus conflict*). Neuroobrazowanie czynności mózgu za pomocą fMRI potwierdziło specyfikę konfliktu na poziomie selekcji reakcji oraz konfliktu na poziomie percepcyjnym, pokazując, że te dwa rodzaje konfliktu angażują częściowo inne struktury mózgu [Liu i in. 2004; Schulte, Müller-Oehring, Vinco, Hoeft, Pfefferbaum i Sullivan 2009]. Podczas gdy konflikt reakcji (S-R) angażował głównie przednie części mózgu, między innymi przedni zakręt obręczy i dodatkowe pole ruchowe (*supplementary motor area*), to konflikt S-S angażował także obszary ciemieniowe, w tym okolice dolnej kory ciemieniowej (*inferior parietal cortex*), będące podłożem systemu związanego między innymi z selekcją ważnych aspektów bodźców. Wspólny obszar aktywacji obejmował natomiast regiony klasycznie wiązane z kontrolą uwagową, między innymi grzbietowo-boczną korę przedczołową.

Opierając się na postulowanej specyfice konfliktów na poziomie reakcji i na poziomie percepcyjnym, postanowiliśmy sprawdzić, czy wzbudzenie uwagi wywiera odmienny wpływ na rozwiązanie konfliktu zależnie od jego poziomu. Zgodnie z hipotezą „czyszczenia świadomości” [Posner 1994] nagłe wzbudzenie systemu uwagi ma hamować bieżącą aktywność uwagi wykonawczej w zakresie kontroli reakcji i kierować uwagę na przetwarzanie informacji percepcyjnych. A zatem jeśli konflikt pojawi się w obrębie reprezentacji percepcyjnej, to efekt wzbudzenia uwagi nie powinien już utrudniać jego rozwiązania, a wręcz je ułatwiać. Analogiczny wpływ na te dwa rodzaje konfliktu zaobserwowano wcześniej w zakresie czasowej orientacji uwagi (*temporal orienting*) [Correa i in. 2010]. Powyższą hipotezę testowaliśmy w eksperymencie z połączonymi zadaniami Simona i przestrzennego Stroopa [Correa i in. 2010; Liu i in. 2004]. Aby uzyskać efekt wzbudzenia, w połowie prób

bodźce docelowe poprzedzane były prezentacją wskazówki w postaci gwiazdek (*) pojawiających się jednocześnie we wszystkich możliwych lokalizacjach bodźca-celu (zob. Rysunek 2).



Rysunek 2. Zadanie do pomiaru konfliktu powstałego na etapie selekcji reakcji (zadanie Simona) i na etapie przetwarzania percepcyjnego (przestrzenny Stroop). Przykłady procedur eksperymentalnych z wizualną wskazówką wzbudzeniową: (A) sekwencja bodźców z zadaniem Simona w warunkach niespójnym (bodziec pojawiający się po lewej stronie wymagający reakcji prawą ręką), (B) sekwencja bodźców w przestrzennym zadaniu Stroopa w warunkach niespójnym (strzałka w dół pojawiająca się powyżej punktu fiksacji)

Wyniki eksperymentu ukazały wyraźny i istotny statystycznie efekt konfliktu w obu zadaniach. Jednakże efektywność rozwiązywania konfliktu była w obu przypadkach niezależna od zaangażowania systemu wzbudzeniowego. Brak interakcji pozostaje w zgodzie z tezą, że interakcje pomiędzy tymi systemami zależą od specyficznych warunków zadania [np. Weinbach i Henik 2012], potwierdza założenie o względnej niezależności systemów uwagi [Posner i Rothbart 2007] i jednocześnie przeczy hipotezie „czyszczenia świadomości” [Callejas 2004, 2005; Posner 1994]. Otrzymane wyniki uzupełniają też wnioski Fischera i współpracowników [2010, 2011], pokazując, że obecność silnych związków bodziec-reakcja, chociaż być może konieczna, nie jest jednak wystarczająca do ujawnienia się negatywnego wpływu wzbudzenia na efektywność rozwiązywania konfliktu. Wyniki są również spójne z hipotezą Weinbacha i Henika [2012], zakładającą konieczność przestrzennego rozdzielania celu i dystraktów do powstania interakcji, ponieważ takie rozdzielanie nie miało miejsca ani w zadaniu Simona, ani w przestrzennym Stroopie. Jednak brak pozytywnych efektów interakcyjnych pomiędzy efektem wzbudzenia i efektem konfliktu ogranicza możliwości interpretacyjne tych badań w kwestii potencjalnego mechanizmu odpowiadającego za obserwowane w innych eksperymentach zarówno negatywne, jak i pozytywne efekty wpływu wzbudzenia uwagi na rozwiązywanie konfliktu.

Podsumowanie

Badania związków pomiędzy funkcjonowaniem systemu wzbudzenia uwagi oraz systemu uwagi wykonawczej potwierdzają ich względną niezależność, wskazując jednocześnie na możliwość zachodzenia między nimi interakcji i współpracy. Mimo że kolejne badania przynoszą coraz jaśniejszy obraz obserwowanych zależności, czynniki warunkujące pojawienie się interakcji wciąż nie są jasne, co oczywiście utrudnia poznanie jej mechanizmu. Przede wszystkim wciąż nie wiadomo, dlaczego niektóre eksperymenty przynoszą efekt interakcji negatywny, inne zaś pozytywny, a w części prac w ogóle nie obserwuje się takiej interakcji. W dalszych badaniach powinno się zatem podjąć próbę ustalenia warunków koniecznych oraz wystarczających do ujawnienia się efektu wpływu wzbudzenia uwagi na efektywność uwagi wykonawczej.

BIBLIOGRAFIA

- Asanowicz D., Marzecová A., Jaśkowski P., Wolski P. (2012). *Hemispheric asymmetry in the efficiency of attentional networks*. „Brain and Cognition” 79 (2), s. 117–28.
- Asanowicz D., Siedlecka M., Michalczyk Ł. (2009). *Zintegrowane podejście do badań nad uwagą jako model badawczy dla kognitywistyki* [w:] L. Hess, T. Konik, A. Pohl, M. Siedlecka, J. Sitnicka (red.), *Rocznik Kognitywistyczny*, t. II, s. 9–15.
- Botvinick M.M., Cohen J.D., Carter C.S. (2004). *Conflict monitoring and anterior cingulate cortex: An update*. „Trends in Cognitive Sciences” 8 (12), s. 539–546.
- Callejas A., Lupiáñez J., Funes M.J., Tudela P. (2005). *Modulations among the alerting, orienting and executive control networks*. „Experimental Brain Research” 167 (1), s. 27–37.
- Callejas A., Lupiáñez J., Tudela P. (2004). *The three attentional networks: on their independence and interactions*. „Brain and Cognition” 54 (3), s. 225–227.
- Chica A.B., Thiebaut de Schotten M., Toba M., Malhotra P., Lupiáñez J., Bartolomeo P. (2012). *Attention networks and their interactions after right-hemisphere damage*. „Cortex. A Journal Devoted to the Study of the Nervous System and Behavior” 48 (6), s. 654–663.
- Cohen J.D., Servan-Schreiber D., McClelland J.L. (1992). *A parallel distributed processing approach to automaticity*. „The American Journal of Psychology” 105 (2), s. 239–269.
- Corbetta M., Shulman G.L. (2002). *Control of goal-directed and stimulus-driven attention in the brain*. „Nature Reviews Neuroscience” 3 (3), s. 201–215.
- Correa A., Cappucci P., Nobre A.C., Lupiáñez J. (2010). *The two sides of temporal orienting: facilitating perceptual selection, disrupting response selection*. „Experimental Psychology” 57 (2), s. 142–148.
- Eriksen B., Eriksen C. (1974). *The impact of noise letters upon the identification of a target letter in a nonsearch task*. „Perception & Psychophysics” 16 (1), s. 143–149.
- Fan J., Gu X., Guise K.G., Liu X., Fossella J., Wang H., Posner M.I. (2009). *Testing the behavioral interaction and integration of attentional networks*. „Brain and Cognition” 70 (2), s. 209–220.
- Fan J., Hof P., Guise K., Fossella J., Posner M. (2008). *The functional integration of the anterior cingulate cortex during conflict processing*. „Cerebral Cortex” 18 (4), s. 796–805.
- Fan J., McCandliss B.D., Fossella J., Flombaum J.I., Posner M.I. (2005). *The activation of attentional networks*. „NeuroImage” 26 (2), s. 471–479.
- Fan J., McCandliss B.D., Sommer T., Raz A., Posner M.I. (2002). *Testing the efficiency and independence of attentional networks*. „Journal of Cognitive Neuroscience” 14 (3), s. 340–347.

- Fischer R., Plessow F., Kiesel A. (2010). *Auditory warning signals affect mechanisms of response selection: evidence from a Simon task*. „Experimental Psychology” 57 (2), s. 89–97.
- Fischer R., Plessow F., Kiesel A. (2011). *The effects of alerting signals in action control: Activation of S-R associations or inhibition of executive control processes?* „Psychological Research” 76 (3), s. 317–328.
- Fossella J., Sommer T., Fan J., Wu Y., Swanson J.M., Pfaff D.W., Posner M.I. (2002). *Assessing the molecular genetics of attention networks*. „BMC Neuroscience” 11, s. 1–12.
- Fuentes L.J., Campoy G. (2008). *The time course of alerting effect over orienting in the attention network test*. „Experimental Brain Research” 185 (4), s. 667–672.
- Gratton G., Coles M.G., Sirevaag E.J., Eriksen C.W., Donchin E. (1988). *Pre- and poststimulus activation of response channels: a psychophysiological analysis*. „Journal of Experimental Psychology. Human Perception and Performance” 14 (3), s. 331–344.
- Hommel B. (2004). *Event files: Feature binding in and across perception and action*. „Trends in Cognitive Sciences” 8 (11), s. 494–500.
- Ishigami Y., Klein R.M. (2010). *Repeated measurement of the components of attention using two versions of the Attention Network Test (ANT): stability, isolability, robustness, and reliability*. „Journal of Neuroscience Methods” 190 (1), s. 117–128.
- Jonathan D., Schreiber D.S., McClelland J.L. (1992). *A parallel distributed processing approach to automaticity*. „American Journal of Psychology” 105 (2), s. 239–269.
- Jong, R., De, Liang C., Lauber E. (1994). *Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus-response correspondence*. „Journal of Experimental Psychology. Human Perception and Performance” 20 (4), s. 731–750.
- Liu X., Banich M.T., Jacobson B.L., Tanabe J.L. (2004). *Common and distinct neural substrates of attentional control in an integrated Simon and spatial Stroop task as assessed by event-related fMRI*. „NeuroImage” 22 (3), s. 1097–1106.
- Macleod J.W., Lawrence M. a, McConnell M.M., Eskes G. a, Klein R.M., Shore D.I. (2010). *Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test*. „Neuropsychology” 24 (5), s. 637–651.
- Marzecová A., Asanowicz D., Krivá L., Wodniecka Z. (2013). *The effects of bilingualism on efficiency and lateralization of attentional networks*. „Bilingualism: Language and Cognition” 16 (3), s. 608–623.
- Navon D. (1977). *Forest before trees: The precedence of global features in visual perception*. „Cognitive Psychology” 9 (3), s. 353–383.
- Nieuwenhuis S., Kleijn R. de (2013). *The impact of alertness on cognitive control*. „Journal of Experimental Psychology: Human Perception and Performance” 39 (6), s. 1797–1801.
- Norman D.A., Shallice T. (1986). *Attention to Action* [w:] R.J. Davidson, G.E. Schwartz, D. Shapiro (red.), *Consciousness and Self-Regulation* (s. 1–18). New York: Springer US.
- Posner M. (1994). *Attention: the mechanisms of consciousness*. „Proceedings of the National Academy of Sciences USA” 91 (August), s. 7398–7403.
- Posner M.I. (2008). *Measuring alertness*. „Annals of the New York Academy of Sciences” 1129, s. 193–199.
- Posner M.I., Fan, J. (2008). *Attention as an organ system* [w:] *Topics in Integrative Neuroscience: From Cells to Cognition* (s. 31–61). New York: Cambridge University Press.
- Posner M.I., Petersen S.E. (1990). *The attention system of the human brain*. „Annual Review of Neuroscience” 13, s. 25–42.
- Posner M.I., Rothbart M.K. (2007). *Research on attention networks as a model for the integration of psychological science*. „Annual Review of Psychology” 58, s. 1–23.
- Ridderinkhof R.K., van den Wildenberg W.P.M., Wylie S.A. (2012). *Action control in times of conflict. Analysis of reaction time distributions in healthy and clinical populations* [w:] M.I. Posner (red.), *Cognitive Neuroscience of Attention* (2nd ed.) (s. 409–420). New York, London: Guilford Press.

- Schulte T., Müller-Oehring E.M., Vinco S., Hoefft F., Pfefferbaum A., Sullivan E.V. (2009). *Double dissociation between action-driven and perception-driven conflict resolution invoking anterior versus posterior brain systems*. „NeuroImage” 48 (2), s. 381–390.
- Seifried T., Ulrich R., Bausenhart K.M., Rolke B., Osman A. (2010). *Temporal preparation decreases perceptual latency: Evidence from a clock paradigm*. „The Quarterly Journal of Experimental Psychology” 63 (12), s. 2432–2451.
- Simon J.R. (1990). *The effects of an irrelevant directional cue on human information processing* [w:] R.W. Proctor, T.G. Reeve (red.), *Stimulus-response Compatibility* (s. 31–86). Amsterdam: Elsevier.
- Sokolov E.N. (1963). *Higher nervous functions: the orienting reflex*. „Annual Review of Physiology” 25 (1), s. 545–580.
- Spencer K.M., Coles M.G. (1999). *The lateralized readiness potential: relationship between human data and response activation in a connectionist model*. „Psychophysiology” 36 (3), s. 364–370.
- Treisman A.M., Gelade G. (1980). *A feature-integration theory of attention*. „Cognitive Psychology” 12 (1), s. 97–136.
- Weinbach N., Henik A. (2011). *Phasic alertness can modulate executive control by enhancing global processing of visual stimuli*. „Cognition” 121 (3), s. 454–458.
- Weinbach N., Henik A. (2012). *The relationship between alertness and executive control*. „Journal of Experimental Psychology: Human Perception and Performance” 38 (6), s. 1530–1540.
- Weinbach N., Henik A. (2013). *The interaction between alerting and executive control: dissociating phasic arousal and temporal expectancy*. „Attention, Perception & Psychophysics” 75 (7), s. 1374–1381.
- Wronka E. (2004). *Uwaga! – mózg w działaniu. O neuronalnych podstawach mechanizmu uwagi*. „Studia Psychologiczne” 4 (1), s. 11–23.