

MRT w roli wykrywacza kłamstw i czytelnika myśli? Krytyka i refleksje

M. Ruchsow,¹ L. Hermle,¹ M. Kober²

Nervenarzt 2010, 81:1085–1091

W ostatnich latach intensywnie badano, czy do wykrywania kłamstw można zastosować tomografię czynnościową rezonansu magnetycznego (fMRT). W artykule zanalizowano i przedyskutowano argumenty za i przeciw takiemu zastosowaniu fMRT. Rozważania naukowo-teoretyczne jak również metodologiczne pokazały, że na podstawie wyniku fMRT nie można wysnuwać żadnych wniosków dotyczących stanu umysłu. Tym samym na podstawie badania fMRT nie da się ustalić tego, co dokładnie jakaś osoba myśli, ani też – czy kłamie, czy mówi prawdę.

Słowa kluczowe

czynnościowa tomografia rezonansu magnetycznego, wykrywanie kłamstw, radykalna interpretacja, niedookreśloność przekładu/interpretacji wyniku fMRT

W ostatnich latach intensywnie badano, czy nowoczesne techniki obrazowe mogą być wykorzystane w roli wykrywacza kłamstw i czy możliwe jest odczytanie myśli na podstawie skanów mózgu względnie czynnościowej tomografii rezonansu magnetycznego (fMRT).¹⁻⁹ W wielu badaniach z zastosowaniem fMRT zidentyfikowano sieci neuronalne, których aktywacja wiązała się ze szczerymi odpowiedziami badanych. Można było w ten sposób z dużym prawdopodobieństwem odróżnić kłamliwych uczestników badania od prawdomównych.

Coraz więcej krytycznych opinii dotyczy nie tylko kwestii etycznych,¹⁰ ale również podstawowych problemów metodologicznych.² Część z tych zarzutów odnosi się do tego, że – podobnie jak w technikach diagnostycznych – w przypadku wykrywania kłamstw fMRT może oddziaływać¹¹ na stan psychiczny badanego.

W artykule przedstawiono zarówno metodologiczne, jak również teoretyczne zarzuty, sięgające do teorii dwóch amerykańskich filozofów Willarda Van Ormana Quine'a (1908-2000) i Donalda Davidsona (1917-2003). W szczególności Quine, na podstawie swego przykładu „radykalnego przekładu” wykazał (np. 12), że żadna z teorii naukowych nie daje się do końca udowodnić empirycznie i że „wychodzą” one poza doświadczenia dostępne za pomocą pięciu zmysłów. Każdy język używany do próby opisu czy ujęcia danych pochodzących z naszych

zmysłów, zawiera uogólnienia np. w formie ogólnikowych terminów, które przekraczają granice konkretnych doświadczeń i tym samym stają się interpretacjami (podlegającymi rewizji). W przypadku wykrywania kłamstw na podstawie badania fMRT lub w ogóle możliwości odczytywania myśli oznacza to, że wszystkie wnioski dotyczące stanów psychicznych takich, jak przekonania lub zamiary są one *de facto* zawsze tylko przejściowymi interpretacjami, podlegającymi za każdym razem rewizji, gdyż z zasady nigdy nie można jednoznacznie ustalić co dana osoba dokładnie myśli.^{12,13} Należy brać to pod uwagę w rozważaniach dotyczących pewnej i wiarygodnej możliwości oceny, czy jakaś osoba kłamie czy nie, za pomocą badania fMRT. Nigdy w związku z tym nie można dokładnie określić, z powodu jakich przekonań lub zamiarów konkretna osoba w danej chwili kłamie albo mówi prawdę.

Czytanie myśli

Od czasu wydarzeń 11 września i początku wojny z terroryzmem wielu życzyłoby sobie, a w szczególności siły bezpieczeństwa i służby specjalne, by móc czytać myśli innych. Jonathan D. Moreno z Uniwersytetu Wirginii w swojej książce pod tytułem *Mind Wars. Brain Research and National Defence*¹⁵ opowiada bardzo szczegółowo o staraniach, które w Stanach Zjednoczonych podejmowane są w ramach tak zwanej wojny

¹Fachkrankenhaus Christophsbad, Göppingen

²Philosophisches Seminar, Albert-Ludwigs-Universität Freiburg

Prof. Dr. M. Ruchsow Fachkrankenhaus Christophsbad Faurndauer Straße 6-28, 73035 Göppingen martin.ruchsow@christophsbad.de

Autor korespondujący nie zgłasza żadnego konfliktu interesów.

poznawczej. Szczególne oczekiwania związane są z zastosowaniem nowoczesnych technik obrazowych, a z wykrywaniem kłamstw za pomocą fMRT w szczególności, które to z zasady opiera się na założeniach „klasycznej” techniki wykrywania kłamstw (poligrafii).⁵

Wykrywacz kłamstw (poligraf)

Idea czytania myśli to nie tylko produkt współczesnych fantazji i lęków. Już wkrótce po odkryciu elektroencefalografii (EEG) przez Hansa Bergera z Uniwersytetu w Jenie pojawiła się nadzieja, że sygnały elektrofizjologiczne pochodzące od badanego pozwolą na wysnucie wniosków na temat treści jego myśli, a nawet że w przyszłości uda się rozszyfrować pojedyncze myśli.¹⁶ W 1930 roku tak temat EEG pisał rozentuzjasmowany dziennikarz gazety Duesseldorfer Stadt-Anzeiger:

„Dzisiaj są to jeszcze tajemne znaki, być może jutro na tej podstawie rozpoznawać się będzie choroby psychiczne i mózgu, a pojutrze można nawet będzie pisać listy językiem (pismem) mózgu.” (cytat za: 17).

Źródła powstania wykrywacza kłamstw (poligrafu) dają się prześledzić od początków XX wieku. Prace między innymi V. Benussiego, L. Keelera, W.M. Marstona, C.G. Junga i Maxa Wertheimera wyprzedzają ideę wykorzystania badań psychofizjologicznych w kryminologii. Przy czym główną ideą nie były pomiary samego kłamstwa, ale takich parametrów fizjologicznych towarzyszących kłamaniu, jak: zmiana częstości oddechu, czynności serca, ciśnienia tętniczego i reakcji psychogalwanicznej skóry.¹⁶

Główne zarzuty wobec stosowania wykrywacza kłamstw dotyczą przede wszystkim braku wiarygodności i podatności na manipulację. Jak stwierdzają Rill i Vossel, w badaniu wykrywaczem kłamstw na podstawie testu składającego się z pytań z kontrolnych przynajmniej 15% „niewinnych” zostaje zakwalifikowanych jako „winni”.⁷ Ponadto, napięcie i lęk występujące podczas badania poligraficznego niekoniecznie musi być wywołane kłamaniem. Przyczyną może być sama procedura badania.⁶ Do podobnego wniosku doszedł niemiecki Bundesgerichtshof (Sąd Federalny) w orzeczeniu z 17 grudnia 1998 roku, popartym dowodami naukowymi, w którym między innymi stwierdza, że u osoby niebędącej sprawcą czynu lęk przed karą może prowadzić do nieprawidłowych wyników badania poligraficznego.³ Markowitsch i Siefer podsumowują, że w badaniu poligraficznym decydującą rolę odgrywa efekt placebo i strachu. Celem wykrywacza kłamstw jest tym samym stworzenie wrażenia, że maszyna jest nieomylna: „Zwykle ten, kto w to uwierzy, już przegrał”.¹⁸ W wielu badaniach wykazano ponadto, że parametry fizjologiczne w trakcie badania wykrywaczem kłamstw są podatne na manipulację ze strony badanych, jak na przykład przez celowe ugryzienie się w język lub przywoływanie wspomnień lub obrazów wywołujących lęk (prześląd w [1]).

Wykrywanie kłamstw przy użyciu fMRT

Neurolog z Uniwersytetu Pensylwanii Daniel Langleben wraz z kolegami należą do pionierów badań nad wykorzystaniem fMRT do wykrywania kłamstw. W jednym z pierwszych badań przeprowadzili z udziałem 23 osób doświadczenie z zastosowaniem MRT i kart do gry. Karty były pokazywane na ekranie komputera w określonej kolejności.¹⁹ Na początku badania uczestnicy otrzymali 20 USD i kartę do gry (piątkę trefl) w zamkniętej kopercie. Wszystkich poinstruowano, aby w dalszej części badania zaprzeczali, jakoby mieli w posiadaniu taką kartę, w przeciwnym razie będą musieli oddać otrzymane pieniądze. Następnie na ekranie komputera pokazywano badanym kolejno różne karty. Mieli oni, naciskając odpowiedni klawisz, udzielić odpowiedzi na pytanie, czy dana karta jest w ich posiadaniu. Podczas 22 minutowej sesji pokazywano w sumie 16 razy piątkę trefl (warunek „kłamstwo”) i 16 razy inną kartę (dwójkę kier – warunek „prawda”), dzięki czemu można było bezpośrednio porównać aktywację kory mózgowej w przypadku kłamania i mówienia prawdy.

Langleben i wsp. odkryli, że w trakcie kłamania zwiększona aktywacja występowała między innymi w obrębie przedniej części zakrętu obręczy (ACC) i zakrętu czołowego górnego (GFS), dwu obszarów mózgu, które związane są z tak zwaną kontrolą funkcji wykonawczych, jak na przykład hamowaniem reakcji wykonawczych. Langleben wnioskuje, że w świadomości w czasie kłamania informacja o prawdziwej odpowiedzi musi zostać zahamowana, co aktywuje odpowiednie obszary mózgu (ACC i GFS). Mówienie prawdy jest w pewnym sensie podstawowym stanem poznawczym (...suggesting that truth is the baseline cognitive state”).¹⁹ Wynik ten jest zgodny z definicją kłamstwa jako zamierzoną (intencjonalną) negacją prawdy, pochodzącą od Aureliusza Augustyna (354-430), Doktora Kościoła. W prostych słowach: kłamstwo to „denying what is”, lub dokładniej: „intentionally denying what one knows what actually is”.²⁰

Langleben i wsp. tak podsumowują uzyskane wyniki:¹⁹

Odpowiedzi zgodne z prawdą i kłamstwo mają różne korelaty neuronalne.

- Kłamstwo jest zamierzoną negacją prawdy.
- ACC i GFS tworzą podstawową sieć neuronalną kłamstwa („deception”).

Tymczasem badania Langlebena były wielokrotnie powtarzane (np. 21,22). Andrew Kozelowi i wsp. z Medycznego Uniwersytetu Karoliny Południowej udało się na podstawie badania fMRT określić z około 90% trafnością, czy badany w trakcie pozorowanej kradzieży ukradł zegarek czy pierścionek. W ten sposób Kozel stworzył neuronalny model kłamstwa, który obejmuje między innymi znane już obszary ACC i GFS. Model został następnie sprawdzony na kolejnej próbie testowej. W badaniu tym ponownie udało się z 90% dokładnością ustalić, który z dwu przedmiotów został skradziony.²²

Wobec zastosowania techniki fMRT sformułowano wiele zarzutów etycznych. Postulowano na przykład, by między

innymi wprowadzić odpowiednie procedury etyczne postępowania w przypadku przypadkowych diagnoz (np. wykrycia guza mózgu) i pewnych informacji z wywiadu (np. dotyczących choroby psychicznej, cech charakteru, aktywnego lub utajonego uzależnienia od substancji psychoaktywnych itp.).¹⁰ Sednem tej pracy nie będą jednak etyczne zarzuty wobec procedury wykrywania kłamstw, lecz zarzuty metodologiczne i naukowe, których dyskusja znajduje się w dalszej części artykułu.

Zarzuty metodologiczne i naukowe

PSYCHOLOGIZM

Podczas wykorzystywania poligrafu należy brać pod uwagę, że za jego pomocą ustala się raczej szczerść osoby wypowiadającej się niż to, czy wypowiedzi są prawdziwe. Podmiotem wykrywania kłamstw jest więc nie tyle prawdziwość jakiejś wypowiedzi (sąd w sensie logicznym, ang. proposition), ale tak zwane nastawienie podmiotowe wyrażone poprzez sądy logiczne – postawa osądzania (propositional attitude), przy czym poznawcza relacja danej osoby do odpowiedniego sądu jest zrozumiała.²³ Postawa osądzania prawdomównego autora wypowiedzi (A) da się opisać w następujący sposób: A wierzy, że p, i mówi, że p (gdzie p to wartość logiczna dowolnego zdania, np. „pada deszcz”). Przeciwnie, nieszczera, kłamiąca osoba B wierzy, że p, mówi jednak, że nie-p.

W przypadku przekonania osób A i B, z postawy osądzania nie wynika co jest prawdą: p czy nie-p, to znaczy nie wynika, która z wypowiedzi: wypowiedź p czy wypowiedź nie-p jest prawdziwa. Co więcej, A i B mogą się oczywiście również mylić, czyli szczerze wierzyć, że p jest prawdą, mimo iż w rzeczywistości prawdą jest nie-p.

Wpływ postawy osądzania również udało się wykazać w badaniach fMRT. Okazało się, że subiektywne przekonania i założenia wywierają silny wpływ na aktywność mózgu.²⁴ U badanych aktywacja mózgu zależała, między innymi, od tego czy wiedzieli, że komunikują się z komputerem czy z inną osobą, mimo że bodziec w obu przypadkach był dokładnie taki sam.^{25,26} Sip i wsp.²⁴ są zdania, że w takich przypadkach decydujące znaczenie mają nastawienia poznawcze badanego i należy po przeprowadzeniu badania o nie spytać.

Poligraf mierzy więc raczej stan psychiczny badanych (np. stan umysłowy związany z przekonaniem o czymś lub stan psychiczny, względnie postawę związaną z tak zwaną uczciwością), nie mówi jednak nic na temat prawdy obiektywnej na temat przypuszczeń, myśli lub przekonań. Dla policji jest ważniejsze, by dowiedzieć się, czy osoba A w plecaku ma bombę czy też nie, niż odpowiedź na pytanie: czy A jest o tym przekonana (wierzy w to), ponieważ A również bez własnej wiedzy mogła być użyta przez innych jako kurier. Policja powinna więc raczej wykorzystywać techniki MTR do przeszukiwania plecaków niż głów czy myśli ich właścicieli.

Bezkrytyczne podejście do wykrywania kłamstw prowadzi więc do psychologizacji argumentacji, która myli określenie prawdy o p na poziomie logicznym lub naukowym z psychologizacją płaszczyzną dotyczącą przekonań, wiary i myślenia. Ogólne założenia tej niezaprzeczalnie rozumiejącej psychologizacji argumentacji są następujące:

- Logika jest teorią właściwego wnioskowania.
- Właściwe wnioskowanie jest stanem umysłowym (lub procesem umysłowym).
- Psychologia zajmuje się wszelkimi stanami umysłowymi.
- Ergo: Logika jest częścią psychologii.

Błąd takiej argumentacji ujawnia się w poglądzie wielu autorów (przegląd w [27]), że psychologia może być określona jako w najlepszym wypadku nauka deskryptywna dotycząca tego, co i jak ludzie właściwie myślą (odnośnie granic takiej możliwości patrz poniżej). Logiką jako teorią właściwego wnioskowania jest jednak teorią normatywną lub preskryptywną, gdyż wskazuje, jak należy myśleć i formułować wnioski, jeżeli oczekuje się, by były one prawidłowe. Z teorii deskryptywnej nie udaje się z reguły wysnuć żadnych normatywnych twierdzeń, w szczególności w sytuacji, gdy – co można udowodnić doświadczalnie – wielu ludzi de facto dochodzi do fałszywych wniosków, czyli w rzeczy samej nie myślą oni tak, jak powinni myśleć.

Powołując się na nasz przykład dotyczący wykrywacza kłamstw wspomaganego fMRT, stwierdzenie zwiększonej aktywacji podczas kłamania (patrz powyżej) nie warunkuje w żadnym wypadku tego, że w plecaku podróżnego A jednak rzeczywiście znajduje się bomba, choć on sam zapewnia, że jej tam nie ma.

Postawy osądzania

Kolejną trudnością związaną z wykrywaniem kłamstw z użyciem fMRT jest to, że jak dotychczas nie ma żadnego uzgodnienia dotyczącego operacjonalizacji definicji kłamstwa. Kłamstwo (i oszustwo) są jednak bardzo heterogenne, co czyni mało prawdopodobnym, aby stosując jeden doświadczalny paradygmat fMRT udało się zrozumieć problem.² Na przykład pozostaje kwestią otwartą, w jakim stopniu wyniki badań Langlebena i wsp.^{19,20} oraz Kozela i wsp.²² odnoszą się do wspomnień dotyczących wcześniejszych kłamstw, względnie kłamstwa pożądanego społecznie („Tak, bardzo dziękuję, jedzenie było wyśmienite”).⁵ Także Ganis i wsp. znaleźli istotne różnice między różnymi „rodzajami” kłamstwa. I tak, inaczej wyglądała aktywacja zarejestrowana przez fMRT w przypadku kłamstwa w kontekście autobiograficznym (memorized-scenario lies) niż kłamstwa całkowicie niezwiązanego z biografią danej osoby (spontaneous-isolated lies).⁴

Także filozoficzne definicje kłamstwa sprawiają często trudności. Według Mahona²⁸ kłamstwo składa się z czterech istotnych elementów:

- pewnego twierdzenia (statement condition),
- znanej nadawcy fałszywości tego twierdzenia (untruthfulness condition),

- jakiegoś adresata (addressee condition),
- zamiaru nadawcy, aby adresat uwierzył w fałszywe twierdzenie (intention to deceive addressee condition).

Czwarte kryterium Mahona oznacza, że adresat nie może wiedzieć, że został okłamany. Oznacza to na przykład, że w badaniach Langlebena i wsp.^{19,20} i Kozela²² w rzeczywistości nie badano żadnego kłamstwa.

Wadą definicji Mahona jest niezbyt dokładne rozróżnienie między kłamstwem a pomyłką (patrz drugie kryterium: „znanej fałszywości”). I tak, osoba A może być wszak szczerą, a mimo to mówić nieprawdę (A wierzy, że p i mówi, że p, jednak p nie jest prawdą). Nie mówimy wtedy, że A skłamał, ale że się pomylił. Przykładowo, A zapytany może uczciwie twierdzić, że jest zdrowy (A wierzy, że p i mówi, że p), ale lekarz zakładowy diagnozuje u niego napady padaczkowe, co wiąże się z tym, że absolutnie nie nadaje się do pracy przy montażu rusztowań (prawdą jest nie-p). Podobnie inna osoba B może uważać, że nie jest zdrowa (że ciemna plama na jej twarzy to czerniak złośliwy), jednak po przeprowadzeniu szczegółowego badania dermatolog stwierdza, że zmiana skórna jest nieszkodliwą plamą wątrobową (B wierzy, że nie-p i mówi, że nie-p, prawdą jest jednak p). Również w tym przypadku można powiedzieć o B, że się pomylił, a nie, że skłamał.

Z drugiej strony można sobie wyobrazić sytuacje, w których osoby nie są szczerze pod tym względem, że chcą skłamać, choć w rzeczywistości nie kłamią. Na przykład niefortunny symulant: C wierzy, że jest zdrowy (p), twierdzi jednak, że jest chory (nie-p), w trakcie badania lekarz rodzinny stwierdza guz (C wierzy, że p, ale mówi, że nie-p, i prawdą jest nie-p). Można też wyobrazić sobie czwartą ewentualność: pracownik D uważa, że nie jest zdrowy, ale mimo to pełen poświęcenia idzie do pracy i (z łęką przed utratą posady) twierdzi oficjalnie, że jest zdrowy; w trakcie rutynowej kontroli u lekarza rodzinnego nie zostaje zdiagnozowana żadna choroba (D wierzy, że nie-p, a mówi, że p i prawdą jest p). Także w tym przypadku nie mamy do czynienia z kłamstwem.

W przypadku połączeń różnych rodzajów postaw osądzania (wierzy, że p, względnie wierzy, że nie-p), wypowiedzi (mówi, że p, względnie mówi, że nie-p) i możliwości dotyczących prawdy rzeczywistości (wypowiedź p jest prawdziwa albo fałszywa) daje to w sumie 23 możliwości, z których tylko dwie, w ograniczonym sensie, można uznać za właściwe oszustwo lub kłamstwo:

- przypadek typowego symulanta: E wierzy lub wręcz wie, że jest zdrowy, mówi jednak, że jest chory mimo, iż w rzeczywistości nie cierpi na żadne dolegliwości (E wierzy, że p i mówi, że nie-p, a prawdą jest p). Do tej grupy można zaliczyć również przypadek z badania Langlebena i wsp.:¹⁹ badany E* wie, że posiada piątkę trefl, kłamie jednak, że nie ma tej karty. Prawdą jest natomiast, że badany E* ma w swoim posiadaniu piątkę trefl.
- przypadek sumiennego pracownika: F wierzy lub wręcz wie, że jest chory, chciałby jednak koniecznie pracować i dla-

tego mówi, że jest zdrowy. Jakiś czas później lekarz zakładowy rozpoznaje u niego ciężkie zapalenie płuc (F wierzy, że nie-p, a mówi, że p, ale prawdą jest nie-p).

Oba warianty szczerych wypowiedzi – a) osoba G wierzy, że p i mówi, że p, gdy p jest prawdą oraz b) osoba H wierzy, że nie-p i mówi, że nie-p, gdy prawdą jest nie-p – wspomniane zostają tylko celem uzupełnienia.

Z reguły wykrywacz kłamstw prawidłowo wskazuje na kłamstwo w dwóch przypadkach (osoby E i F) i w dwóch przypadkach (osoby G i H) prawidłowo na wypowiedzi prawdziwe, podczas gdy w dwóch przypadkach (osoba C i D) daje fałszywie dodatni wynik i w kolejnych dwóch przypadkach (osoba A i B) wynik fałszywie ujemny.

Okazuje się tym samym, że precyzja wykrywacza kłamstw zależy także od założonej definicji kłamstwa. U Langlebena operacjonalizacja kłamstwa jako „zaprzeczanie temu, co jest”¹⁹ zakłada, że wypowiadający się wie, jaka jest prawda. Podobnie zakłada Mahon – autorowi wypowiedzi wiadomo, że jego twierdzenie jest fałszywe (porównaj drugie kryterium kłamstwa wg Mahona²⁸). W przypadku błędu – jak przy naszej operacjonalizacji definicji kłamstwa, czyli możliwości, że autor wypowiedzi myli się co do prawdziwości lub fałszu swojego twierdzenia, powoduje to, że wykrywacz kłamstw podaje wyniki przypadkowe (w 50% prawidłowe, w 50% fałszywe).

Sytuacja jeszcze bardziej komplikuje się, gdyż w przypadku korelacji wymienionych ośmiu możliwości i wyników badania fMRT chodzi nie o prosty, lecz o nakładający się związek, czyli przynajmniej w pewnym procencie wynik fMRI może być sprzeczny na przykład z badaniem psychologicznym lub nastawieniem poznawczym. W zasadzie w ten sposób zaistnieć może przypadek, że osoba K wierzy, a także mówi, że p, gdzie również prawdą jest p. Badanie rezonansu magnetycznego wskazuje natomiast, że dana osoba kłamie, w wyniku czego liczba możliwych kombinacji wzrasta w sumie do 16 (24).

Ważniejsze jest jednak to, czy wykryte w badaniu fMRT „sieci kłamstwa” (patrz powyżej) są wystarczająco selektywne i swoiste (w sensie statystycznym), by jednoznacznie odróżnić kłamstwo lub oszustwo od innych procesów poznawczych. Wiedzieć, kiedy ktoś kłamie, znaczy również wiedzieć, kiedy ta osoba nie kłamie, przy czym liczba możliwości interpretacji wyniku fMRT po raz kolejny wzrasta geometrycznie. W szczególności należy odróżnić kłamstwo lub oszustwo od procesów poznawczych aktywujących „sieć kłamstwa”, jak na przykład hamowanie odpowiedzi 29, wykrywanie błędów 30 i inne procesy związane z kontrolą funkcji wykonawczych (porównaj [2]). Wykorzystując odpowiednio coraz bardziej wyrafinowane metody, można dalej rozróżniać kolejne stany poznawcze (można na przykład różnicować między posiadaną wiedzą, ugruntowanym przeświadczeniem, założeniem, przypuszczeniem) i dlatego dopuszczalne jest tylko stwierdzenie: „[...] rzeczywiste czytanie w myślach wymagałoby nieprzewidywalnie wiele zestawień oznaczonych danych.”² Bles i Haynes² wnioskuje z tego dalej:

Dlatego, w dającej się przewidzieć przyszłości hipotetyczne, uniwersalne maszyny do czytania w myślach pozostaną science fiction.

Neuroplastyczność a swoistość

Dotychczasowe rozważania można podsumować w ten sposób, że dane pochodzące z badania fMRT to nie wydruk z urządzenia technicznego lub poziom rtęci w termometrze, a stanów umysłowych nie da się odczytać podstawie ze skanu fMRT. Działanie mózgu jest tak samo niejednoznaczne jak działania osoby i dlatego poddawane są interpretacji (podniesienie rąk może być zinterpretowane jako zgłoszenie się, jako pozdrowienie znajomego itp.). W przypadku wykrywania kłamstw szczególne problemy dotyczą wyciągania wniosków odnośnie treści myślowych danego probanta, gdyż jak dotąd nie ma żadnych systematycznych badań na temat ewentualnych źródeł błędów. Tak więc w przypadku badania wykrywania błędów z zastosowaniem techniki fMRT nie wiadomo więc, jaki wpływ na probanta mają wiek, płeć, ręczność lub pora dnia. Jeżeli chodzi o inne paradygmaty poznawcze to wiadomo, że aktywacja typu BOLD (blood oxygen level dependent) zmienia się wraz z wiekiem,^{31,32} ale również w związku z chorobami sercowo-naczyniowymi³³ oraz w przypadku nadużywania narkotyków³⁴ i alkoholu.³⁵ Co więcej BOLD reaguje także na zmiany częstości oddechu, ruchy głową, polykanie i zaburzenia pola magnetycznego.² Opisano ponadto, że badani dzięki ćwiczeniom mogą wpływać na aktywność BOLD w konkretnych rejonach mózgu.^{36,37} W niektórych przypadkach już samo zróżnicowanie prawdziwych wspomnień autobiograficznych i wyimaginowanych „pseudowspomnień” może być trudne, gdyż aktywność odpowiednich rejonów mózgu częściowo się pokrywa.³⁸

Kolejnym dużym problemem są indywidualne różnice morfologiczne i histologiczne mózgu. Na podstawie badań *post mortem* stwierdzono u badanych znaczące różnice cyto- i mieloarchitektury niektórych rejonów mózgu. Wynika z tego również, że przypisywanie zmian strukturalnych (morfologicznych) i czynnościowych (rezonansu magnetycznego) na podstawie standardowego atlasu Talairacha jest bardzo niedokładne^{39,40} i dlatego na przykład aktywacja w rejonie kory przedczołowej u osoby A może oznaczać coś zupełnie innego niż aktywacja w rejonie kory przedczołowej u osoby B. Różnice w cyto- i mieloarchitekturze mózgu są między innymi odzwierciedleniem indywidualnych historii procesów uczenia. Z tego i podobnych danych wynika, że wnioskowanie na temat stanu umysłu (content decoding) na podstawie fMRT wymaga najwyższej ostrożności a przyporządkowanie treści myślowych i aktywności mózgu opiera się na indywidualnych różnicach badanych osób.² Znaczyłyby to, że nie istnieje ogólny (międzyosobniczy) wzór aktywności kłamstwa lub oszustwa w fMRT, ale za każdym razem (wewnątrz)osobniczy. Bles i Haynes² dochodzą do wniosku, że głównym problemem jest w mniejszym stopniu rozpoznawanie wzoru niż jego interpretacja.

Przedstawione poniżej przemyślenia Quine'a i Davidsona pokażą, że wyniki każdego fMRT dopuszczają wiele interpretacji i są niedokładne z zasady.

Niedokładność przekładu

Obaj amerykańscy filozofowie W.V.O. Quine i D. Davidson wykazali, że przypadek pragmatycznej interpretacji jakiejś wypowiedzi jest określony przez triangulację, która jest następstwem trójkątnej zależności między autorem wypowiedzi, interpretatorem (lub słuchaczem) i pozajęzykowym elementem lub wydarzeniem, będącym tematem wypowiedzi.⁴¹ Obaj wychodzą od podstawowych zasad argumentacji i metodologii oraz zilustrowania konstelacji tak zwanej „radykalnej interpretacji”, która powstaje w sytuacji pierwszego spotkania dwóch dotychczas zupełnie nieznanymi sobie postaci, autora wypowiedzi i słuchacza, próbujących zrozumieć swoje werbalne i niewerbalne działania. Przedstawia to następujący przykład: w dżungli podczas wyprawy niemieckojęzyczny etnolog napotyka pewne plemię i chciałby zrozumieć jego kulturę oraz język. Nie ma tam żadnego słownika ani tłumacza, który wyjaśniłby etnologowi wypowiedzi, działania, obyczaje i zwyczaje tego nieznanego plemienia (porównaj [41]). Aby zinterpretować wypowiedzi tubylców staje się on metodycznym behawiorystą, co znaczy, że nie ma innej możliwości, jak potraktować werbalne i niewerbalne zachowania tych osób oraz określony kontekst jako dane i oprzeć na nich swoje badanie.

Quine wyjaśnia tę koncepcję na następującym znanym przykładzie „gavagai”: W pobliżu przemyka królik, tubylec mówi „gavagai”, a etnolog lub badacz języka notuje „królik” (lub „to jest królik”) jako robocze tłumaczenie, które następnie będzie testowane.¹²

Głównym celem etnologa stworzenie pewnego rodzaju podręcznika przekładu języka tubylców. W tym miejscu pojawia się jednak problem, który Quine określa jako „niedookreślenie” lub „niedokładność tłumaczenia”. W ramach metodycznego behawioryzmu, który należy przyjąć, wkład sensoryczny „gavagai” nie wystarcza, by biorąc pod uwagę kontekst zdecydować „czy przetłumaczyliśmy [„gavagai”, uzupełnienie M.R.] jako

- „królik” w sensie „nieulegająca podziałowi całościowa postać królika” czy
- „stan, w którym tymczasowo znajduje się królik”.¹³

Oznacza to, że również z powtarzanych i wskazujących gestów tubylca (wskazanie bezpośrednio) nie udaje się jednoznacznie wyekstrahować, do czego on się odnosi.

Różne systemy hipotez na temat tłumaczenia mogą prowadzić do całkowicie odmiennych i wzajemnie nieporównywalnych zasad przekładu. Podsumowując to, tezę o niedokładności przekładu można podzielić na dwie części:

„Po pierwsze można przypuszczać, że mogą powstawać słowniki, które stoją ze sobą w sprzeczności, i po drugie, że nie można stwierdzić, iż któryś z nich w stosunku do drugiego jest tym jedynym,

właściwym. Quine nie chce jednak samym przez to powiedzieć, że żaden przekład nie jest poprawny, ani że przekład w ogóle nie jest możliwy. Raczej mówi, że istnieje więcej poprawnych przekładów niż jeden.^{39,42}

Według Quine'a niedookreśloność danych nie jest ograniczona do przekładu, ale dotyczy każdej dziedziny nauki, która chce tworzyć teorie na podstawie obserwacji. Będzie to omówione w następnym rozdziale.

Niedokładność teorii empirycznych

Podstawowym celem filozoficznej teorii Quine'a jest pokazanie, jak względnie „mizerny” wkład doznań zmysłowych, na które jesteśmy wystawieni prowadzi do powstania naszego rozległego obrazu świata.⁴³ Jedyne wiarygodne źródło informacji o świecie zewnętrznym powstaje „podczas natknięcia się fal światła i cząsteczek na nasze powłoki sensoryczne”.^{44,45}

Jednak kiedy, ściśle biorąc dysponujemy tylko wzorem pobudzenia naszych receptorów zmysłów (i na przykład nie ma żadnego obiektu fizycznego), już wtedy każda obserwacja wyrażona słowami czyli każdy „samowystarczalny język opisujący dane zmysłowe”,⁴² którym próbujemy ująć odpowiednie dane, zawiera uogólnienia językowe pod postacią określeń (jak „królik”), i które z uwagi na związane z nimi ontologiczne założenia wykraczają poza konkretną obserwowaną sytuację. Dlatego też Quine mówi o niedookreśloności wszystkich naszych teorii empirycznych.¹²

Według Quine'a w rzeczywistości jest więc tak, że „również zdania prawdziwe na temat zwyczajnych rzeczy [...] wykraczają daleko poza dostępne dane”.¹² W konsekwencji istnieje wiele alternatywnych hipotez lub teorii naukowych wyjaśniających wzór pobudzenia naszych receptorów zmysłowych, albo krótko za Quine'em: teorie naukowe „mogą być logicznie nie do przyjęcia, natomiast empirycznie równoważne”.¹⁴

Praktyczne podsumowanie

W artykule wykazano, że, z uwagi na leżące u podstaw stany umysłowe, wszystkie wnioski pochodzące z danych uzyskanych na podstawie badania fMRT są empirycznie niedokładne. Z przemyśleń amerykańskich filozofów W.V.O. Quine i D. Davidsona wynika, że dane z fMRT z założenia muszą podlegać interpretacji, oraz że z reguły dla jednego wyniku istnieje wiele możliwych, lecz różnych interpretacji. W konsekwencji stany umysłowe w żadnym wypadku nie mogą być bezpośrednio odczytywane z wyniku fMRT, jak to bywa sugerowane lub oczekiwane.

Poza tymi naukowo-teoretycznymi zarzutami Bles i Hayes² w szczególności wskazują na wiele metodologicznych problemów związanych z wykrywaniem kłamstw przy użyciu fMRT. Główną komplikacją okazuje się wysoka międzypersonalna różnorodność w zakresie mapy korowej, która prowadzi do bardzo różnych wyników fMRT u osób badanych, co skrajnie utrudnia porównania ze zbiorem danych grupy kontrolnej. Bles i Hay-

nes² stwierdzają, że wykrywanie kłamstw przy zastosowaniu fMRT jest możliwe – jeżeli w ogóle, tylko w sposób intraindywidualny czyli na podstawie „single-subject-basis”. Aby rzetelnie twierdzić, czy jakaś osoba kłamie czy nie, należałoby jednak wcześniej określić i skategoryzować trudną do przewidzenia dużą liczbę możliwych stanów umysłowych oraz odpowiadających im wyników fMRT. Wydaje się, że w tej sytuacji wykrywanie kłamstw przy zastosowaniu techniki fMRT w przewidywalnej przyszłości należy raczej do sfery science fiction.²

© Springer Medizin Verlag 2010. This article MRT als Lügendetektor und Gedankenleser? Kritische Bestandsaufnahme und Reflexion by M. Ruchsov, L. Hermle, i M. Kober is translated and reproduced with permission from Springer.

Piśmiennictwo

1. Ben-Shakhar G, Elaad E (2003) The validity of psychophysiological detection of information with the Guilty Knowledge Test: a meta-analytic review. *J Appl Psychol* 88:131–151
2. Bles M, Haynes J-D (2008) Detecting concealed information using brain-imaging technology. *Neurocase* 14:82–92
3. Bundesgerichtshof (BGHSt 44, 308 ff.) (1998) Beweiserhebung mit Hilfe eines „Lügendetektors“ (Polygraphen). *NJW* 27:1938–1939
4. Ganis G, Kosslyn SM, Stose S et al (2003) Neural correlates of different types of deception: an fMRI investigation. *Cereb Cortex* 13:830–836
5. Greely HT, Illes J (2007) Neuroscience-based lie detection: the urgent need for regulation. *Am J Law Med* 33:377–431
6. Peters K (1975) Eine Antwort auf Undeutsch: Die Verwertbarkeit unwillkürlicher Ausdruckserscheinungen bei der Aussagenwürdigung. *ZStW* 87:663–687
7. Rill H-G, Vossel G (1998) Psychophysiologische Täterschaftsbeurteilung („Lügendetektion“, „Polygraphie“): Eine kritische Analyse aus psychophysiologischer und psychodiagnostischer Sicht. *NStZ* 18:481–486
8. Schleim S, Walter H (2007) Gedankenlesen mit dem Hirnscanner? *Nervenheilkunde* 26:505–510
9. Schleim S, Spranger TM, Walter H (2007) Von der Neuroethik zum Neurorecht? Der Beginn einer neuen Debatte. *Nervenheilkunde* 26:813–818
10. Wolpe PR, Foster KR, Langleben DD (2005) Emerging neurotechnologies for lie-detection: promises and perils. *Am J Bioeth* 5:39–49
11. Haynes JD (2008) Detecting deception from neuroimaging signals – a data-driven perspective. *Trends Cogn Sci* 12:126–127
12. Quine WVO (1980) Wort und Gegenstand. Reclam, Stuttgart
13. Lauener H (1982) Willard Van Orman Quine. Beck, München
14. Quine WVO (1970) On the reasons for indeterminacy of translation. *J Phil* 67:178–183
15. Moreno JD (2006) Mind Wars. Brain research and national defense. Dana Press, New York
16. Schleim S (2008) Gedankenlesen. Pionierarbeit der Hirnforschung. Heise, Hannover
17. Borck C (2005) Hirnströme. Eine Kulturgeschichte der Elektroenzephalographie. Wälstein, Göttingen
18. Markowitsch H, Siefer W (2007) Tatort Gehirn. Auf der Suche nach dem Ursprung des Verbrechens. Campus, Frankfurt/Main
19. Langleben DD, Schroeder L, Maldjian JA et al (2002) Brain activity during simulated deception: an event-related functional magnetic resonance study. *Neuroimage* 15:727–732
20. Langleben DD, Loughhead JW, Bilker WB et al (2005) Telling truth from lie in individual subjects with fast event-related fMRI. *Hum Brain Mapp* 26:262–272
21. Davatzikos C, Ruparel K, Fan Y et al (2005) Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *Neuroimage* 28:663–668
22. Kozel FA, Johnson KA, Mu Q et al (2005) Detecting deception using functional magnetic resonance imaging. *Biol Psychiatry* 58:605–613
23. McKay T, Nelson M (2008) Propositional attitude reports. In: Zalta EN (Hrsg) The stanford encyclopedia of philosophy, forthcoming. <http://plato.stanford.edu/archives/fall2008/entries/prop-attitude-reports/>
24. Sip KE, Roepstorff A, McGregor W, Frith CD (2008) Response to Haynes: There's more to deception than brain activity. *Trends Cogn Sci* 12:127–128
25. Gallagher HL, Jack AI, Roepstorff A, Frith CD (2002) Imaging the intentional stance in a competitive game. *Neuroimage* 3:814–821

26. Rilling JK, Sanfey AG, Aronson JA et al (2004) The neural correlates of theory of mind within interpersonal interactions. *Neuroimage* 22:1694–1703
27. Kusch M (2008) Psychologism. In: Zalta EN (Hrsg) *The stanford encyclopedia of philosophy*, forthcoming. <http://plato.stanford.edu/archives/fall2008/entries/psychologism/>
28. Mahon JE (2008) The definition of lying and deception. In: Zalta EN (Hrsg) *The stanford encyclopedia of philosophy*, forthcoming. <http://plato.stanford.edu/archives/fall2008/entries/lying-definition/>
29. Botvinick M, Nystrom LE, Fissell K et al (1999) Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature* 402:179–181
30. Carter CS, Braver TS, Barch DM et al (1998) Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280:747–749
31. Buckner RL, Snyder AZ, Sanders AL et al (2000) Functional brain imaging of young, nondemented, and demented older adults. *J Cogn Neurosci* 12 [Suppl 2]: 24–34
32. D'Esposito M, Deouell LY, Gazzaley A (2003) Alterations in the BOLD fMRI signal with ageing and disease: a challenge for neuroimaging. *Nat Rev Neurosci* 4:863–872
33. Pineiro R, Pendlebury S, Johansen-Berg H, Matthews PM (2002) Altered hemodynamic responses in patients after subcortical stroke measured by functional MRI. *Stroke* 33:103–109
34. Bruhn H, Kleinschmidt A, Boecker H et al (1994) The effect of acetazolamide on regional cerebral blood oxygenation at rest and under stimulation as assessed by MRI. *J Cereb Blood Flow Metab* 14:742–748
35. Levin JM, Ross MH, Mendelson JH et al (1998) Reduction in BOLD fMRI response to primary visual stimulation following alcohol ingestion. *Psychiatry Res* 82:135–146
36. Rota G, Sitaram R, Veit R et al (2009) Self-regulation of regional cortical activity using real-time fMRI: the right inferior frontal gyrus and linguistic processing. *Hum Brain Mapp* 30:1605–1614
37. Weiskopf N, Veit R, Erb M et al (2003) Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fMRI): methodology and exemplary data. *Neuroimage* 19:577–586
38. Markowitsch HJ, Thiel A, Reinkemeier M et al (2000) Right amygdalar and temporofrontal activation during autobiographic, but not during fictitious memory retrieval. *Behav Neurol* 12:181–190
39. Rademacher J, Bürgel U, Geyer S et al (2001) Variability and asymmetry in the human precentral motor system. A cytoarchitectonic and myeloarchitectonic brain mapping study. *Brain* 124:2232–2258
40. Rademacher J, Bürgel U, Zilles K (2002) Stereotaxic localization, intersubject variability, and interhemispheric differences of the human auditory thalamocortical system. *Neuroimage* 17:142–160
41. Kober M (2002) *Bedeutung und Verstehen*. Mentis, Paderborn
42. Keil G (2002) *Quine zur Einführung*. Junius, Hamburg
43. Quine WVO (1975) *Ontologische Relativität und andere Schriften*. Reclam, Stuttgart
44. Quine WVO (1987) *Die Natur natürlicher Erkenntnis*. In: Bieri P (Hrsg) *Die analytische Philosophie der Erkenntnis*. Athenäum, Frankfurt/Main, S 422–435
45. Quine WVO (2000) *Naturalismus – oder: Nicht über seine Verhältnisse leben*. In: Keil G, Schnädelbach H (Hrsg) *Naturalismus*. Philosophische Beiträge. Suhrkamp, Frankfurt/Main, S 113–127

KOMENTARZ

Dr n. med. Grzegorz Witkowski

I Klinika Neurologiczna, Instytut Psychiatrii i Neurologii w Warszawie

Aparat, kolokwialnie nazywany wykrywaczem kłamstw czy maszyną prawdy, to urządzenie, którego celem jest wykrycie kłamstwa w wypowiedzi osoby badanej. Taka maszyna działająca ze skutecznością 100-procentową, choć pożądana i poszukiwana, pozostaje wciąż domeną fantastyki naukowej. W praktyce techniką od około 75 lat stosowaną podczas przesłuchań jest badanie poligraficzne (wariograficzne). Naukową podstawą wariografii jest teoria zakładająca, że wypowiedzane kłamstwo wiąże się z aktywacją układu współczulnego, której różne elementy można w sposób łatwy zarejestrować. Tymi elementami są: zmiana częstości pracy serca, oddechów, wzrost ciśnienia tętniczego, zwiększona produkcja potu powodująca, że rośnie przewodnictwo elektryczne skóry. Założenie, że każda osoba kłamiąca odczuwa pewien niepokój, który przekłada się na stymulację układu autonomicznego, stanowi jednocześnie podstawową słabość tej metody badawczej. Nieswoista aktywacja układu współczulnego związana z pobudzeniem emocjonalnym wywołanym innymi przyczynami, stresem podczas testu, silnym bólem itp., może powodować uzyskanie wyników fałszywie dodatnich, fałszywie ujemnych lub wręcz uniemożliwić badanie. Z drugiej strony, odpowiednie przeszkolenie (np. słynne historie szpiegów, których ewidentne kłamstwa nie

były wykrywane badaniem wariograficznym), zaburzenia osobowości bądź zaburzenia psychiczne, na przykład związane z występowaniem urojeń, mogą ograniczać aktywację układu współczulnego i tym samym przyczynić się do uzyskiwania wyników fałszywie ujemnych. Z tych powodów skuteczność badania wariograficznego oceniania jest na ok. 70% i w postępowaniach sądowych może ono pełnić jedynie rolę pomocniczą. Alternatywą jest opisane w prezentowanym artykule zastosowanie obrazowania czynnościowego magnetycznego rezonansu jądrowego (functional magnetic resonance imaging, fMRI) w testach prawdomówności. Waznym argumentem pomysłodawców wykorzystania fMRI jako wykrywacza kłamstw jest całkowicie inna metodologia badania, pomijająca aktywację układu autonomicznego. Ta metodologia zakłada istnienie „neuronalnego obwodu kłamstwa” obejmującego obszary kory przedczołowej i przednią część zakrętu obręczy, którego aktywacja może być uchwyciona badaniem czynnościowym magnetycznego rezonansu jądrowego. Celem działania tego obwodu ma być supresja naturalnej tendencji do udzielenia odpowiedzi prawdziwej, i wybór odpowiedzi alternatywnej, będącej kłamstwem. Dokładność (accuracy) badania fMRI jest przez zwolenników tej metody oceniana na powyżej 90%. W Stanach Zjednoczonych pojawiły się już firmy, które komercyjnie oferują badanie prawdomówności tym sposobem. W kilku przypadkach (m.in. przy podejrzeniu o molestowanie dziecka) był on wykorzystany pomocniczo w postępowaniu procesowym. Zastosowanie fMRI jako wykrywacza kłamstw, choć eli-

minuje problemy związane z nieswoistą aktywacją układu autonomicznego, tworzy inne, które obecnie każą wątpić w możliwość szerszego wykorzystania tej techniki podczas śledztw i procesów sądowych. Poza kwestiami moralnymi i filozoficznymi w prezentowanym artykule słusznie podkreślono wiele technicznych i metodologicznych wątpliwości. Badania nad czułością i swoistością fMRI prowadzone były z udziałem relatywnie niewielkich (20-30 osób) grup wyselekcjonowanych zdrowych uczestników, z wykorzystaniem dość prostych paradygmatów „doświadczalnego kłamstwa”. Nie jest więc jasne, czy opisany wzorec aktywacji kory mózgu jest właściwy dla szerszej grupy osób. Do tej pory nie ma danych na temat tego, czy w podobny sposób można oceniać wynik fMRI w odniesieniu do np. różnych grup wiekowych, osób cierpiących na zaburzenia poznawcze, po urazach głowy, z zaburzeniami psychicznymi i zaburzeniami osobowości. Korowa reprezentacja wyższych czynności poznawczych cechuje się bardzo dużą osobniczą zmiennością, wydaje się więc wątpliwe, czy u każdej badanej osoby możliwa jest prawidłowa identyfikacja obszarów aktywujących się w czasie udzielania nieprawidłowych odpowiedzi. Niektórzy sceptyczni badacze wskazują, że słabością fMRI jest rozpoznawanie kłamstwa na podstawie wykrywania obszaru kory zaangażowanego w supresję jednej odpowiedzi (w założeniu odpowiedzi prawdziwej) w celu udzielenia odpowiedzi innej (kłamstwa). W ich opinii supresja może dotyczyć też alternatywnych odpowiedzi, które wcale nie muszą być prawdziwe, lub stanów emocjonalnych, które badana osoba próbuje w danej chwili opanować. Skutkiem może być uży-

skanie wyników fałszywie dodatnich. Kolejnym problemem jest duże ryzyko uzyskania wyników fałszywie ujemnych lub po prostu oszustwa w czasie badania. W trakcie dotychczasowych doświadczeń badani otrzymywali polecenie udzielenia prostej odpowiedzi nieprawdziwej. Czy na przykład u przestępcy, który przez długi czas konsekwentnie trzyma się nieprawdziwej wersji wydarzeń, „przyzwyczajając się” niejako do niej, dojdzie do podobnie jednoznacznej i silnej aktywacji obszarów odpowiadających za supresję odpowiedzi prawdziwej? Nie ma jednoznacznej odpowiedzi na tego typu pytania. Bardzo prawdopodobna wydaje się możliwość opracowania prostych ćwiczeń myślowych, które „oswajając” umysł z kłamstwem, pozwoliłyby zafałszować wynik badania fMRI. Wreszcie, fMRI w większym stopniu niż konwencjonalna poligrafia wymaga ścisłej współpracy ze strony badanego. Minimalny ruch głową czy nieadekwatna odpowiedź na pytanie czyni badanie niemożliwym do interpretacji.

Podsumowując, na obecnym etapie zarówno wiedzy na temat wyższych czynności poznawczych, jak i rozwiązań technicznych pozwalających na ich ocenę, wydaje się, że praktyczne znaczenie fMRI jako wykrywacza kłamstw jest mocno ograniczone. Prawdopodobnie w wybranych przypadkach metoda ta może pełnić funkcję pomocniczą, brak jednak jednoznacznych przesłanek, aby uznać, że jest ona czulsza i bardziej swoista od konwencjonalnej poligrafii. Nie można natomiast nie doceniać znaczenia poznawczego badania fMRI. Ta technika zdecydowanie przyczyniła się do poszerzenia wiedzy na temat neuronalnego podłoża procesów poznawczych związanych z mówieniem prawdy i kłamstwa.