

M i c h a ł B i a ł e k

Przegląd badań współczesnej kognitywistyki nad efektem przekonań

Słowa kluczowe: *racjonalność, teoria dwusystemowa, efekt przekonań, rozumowanie, kryteria racjonalności*

Logika i rachunek prawdopodobieństwa a rozumowanie

Obiektywność nabywanej wiedzy jest tematem dyskusji filozofów od lat. Na przestrzeni epok kształtowano kryteria, na podstawie których oceniać można poprawność myślenia, np. logikę czy rachunek prawdopodobieństwa. Normatywne podejście do procesów poznawczych doprowadziło do znakomitego usprawnienia sposobu nabywania wiedzy, dostarczyło narzędzi jej obiektywizacji i kryteriów weryfikacji. Jednak od czasu, kiedy temat myślenia podejmują kognitywiści, rozwija się podejście deskryptywne. Prace z tego nurtu nie zajmują się tym, jak powinno wyglądać poprawne myślenie, a opisem tego, jakie ono jest na co dzień. I tak np. w nurcie finansów behawioralnych prace skupiają się na opisaniu tego, jak zachowują się rzeczywiście inwestorzy (Kubińska, Markiewicz, Tyszka 2012; Markiewicz, Weber 2013) w porównaniu z tym, co zakładają teorie normatywne (Fama 1970). Psychologia moralności rozważa nie to, jak powinni zachowywać się moralni ludzie, ale jak faktycznie oceniają sytuacje pod kątem moralności czy jakich dokonują zniekształceń w tych ocenach (Terbeck i in. 2012; Unger 1996).

Poszukiwanie kryteriów i metod usprawniających i obiektywizujących myślenie jest nadal jednym z najważniejszych zadań nauk, szczególnie aplikacyjnych. Logika czy matematyka, ale także usprawnianie analizy ryzyka, kryteriów decyzyjnych w finansach, analizy metodologiczne i statystyczne dla

badaczy eksperymentalnych to przykłady nauk o podejściu normatywnym, które dynamicznie się rozwijają.

Podejście normatywne nie jest jednak powszechnie akceptowane jako kryterium oceny poprawności codziennego funkcjonowania poznawczego, ponieważ jest ono zbyt restrykcyjne (dyskusji tej poświęcono numer czasopisma *Behavioral and Brain Sciences* 2011, vol. 5; por. Elqayam, Evans 2011; Sun, Wang 2011). Niektóre działania ludzi, normatywnie uznawane za niepoprawne, wydają się jednak ekologicznie racjonalne. Dlatego nawet w tak kluczowych aspektach, jak wydawanie sądów prawniczych, nie przymusza się decydentów do stosowania normatywnych reguł myślenia (Markiewicz, Markiewicz-Żuchowska 2012).

W świetle badań wydaje się, że ludzie raczej nie posługują się logiką dwuwartościową w swoich wnioskowaniach, gdyż często mają pewien poziom niepewności dla każdego rodzaju wniosku. To znaczy, że wnioski zarówno prawdziwe, jak i nieprawdziwe są takie w ocenie ludzi jedynie do pewnego stopnia. Logika klasyczna, ze swoim twierdzeniem, że jeden kontrprzykład falsyfikuje regułę, zdaje się nie być w rzeczywistości akceptowana przez ludzi, gdyż badani, znajdując kontrprzykład, nierzadko twierdzą, że zdanie jest nadal prawdziwe, tyle że w mniejszym stopniu (Evans, Handley, Harper 2001).

Od lat 1980. powstał szereg koncepcji starających się wyjaśnić rozumowanie niezależnie od jego formalnej klasyfikacji, przyjmując za prawdopodobne, że mają one jednakową realizację umysłową. Pierwsza historycznie koncepcja, według której rozumowanie jest zinternalizowaną realizacją rachunku zdań, została już praktycznie wyparta z literatury tematu i uznana za niefalsyfikowalną. W jej miejsce rozwijają się dwie konkurencyjne koncepcje: dwusystemowa (Evans 2009, 2011; Kahneman 2011; Stanovich 2010) oraz modeli umysłowych (Johnson-Laird 2009). Koncepcja dwusystemowa ma wiele różnych wersji, które błędnie sprowadzane są do jednej i atakowane zbiorczo (np. Kruglanski 2013; Osman 2013; Thompson 2013; por. odpowiedź na zarzuty: Evans, Stanovich 2013b).

Na co zwracamy uwagę oceniając rozumowanie?

Kluczowym odkryciem dla współczesnej wiedzy o rozumowaniu jest badanie Evansa, Barstona i Pollarda (1983), w którym stwierdzono, że ludzie wyjątkowo mocno przywiązują się do wiarygodności wniosków (rozumianej jako zgodność z uprzednio nabytą wiedzą) przedkładając ją nad ocenę logiczną rozumowania, które ten wniosek uzasadnia. Zjawisko to nosi nazwę efektu przekonania (*belief bias*). Weźmy za przykład klasyczne zadanie ze wspomnianego eksperymentu:

*Żadne papierosy nie są tanie
Niektóre uzależniające rzeczy są tanie*

Niektóre uzależniające rzeczy nie są papierosami.

Ludzie mają tendencję, by uznawać takie rozumowanie za poprawne i przyjmować wnioski jako prawdziwe. Gdy jednak wymienić terminy w przesłankach miejscami, otrzymujemy kolejne wnioskowanie, również powszechnie uznawane za poprawne:

*Żadne uzależniające rzeczy nie są tanie
Niektóre papierosy są tanie*

Niektóre uzależniające rzeczy nie są papierosami.

To drugie rozumowanie nie jest jednak logicznie poprawne, ponieważ przyjmując obie przesłanki za prawdziwe, nie mamy gwarancji prawdziwości wniosku. Jak widać na powyższym przykładzie, niezależnie od struktury logicznej ludzie uznają oba rozumowania za poprawne, gdyż prowadzą one do tego samego, wiarygodnego wniosku.

Efekt przekonania występuje, mimo że autorzy badań podkreślają w instrukcji eksperymentów, aby badani koncentrowali się na poprawności logicznej (Trippas, Handley, Verde 2013; Evans, Allen, Newstead, Pollard 1994) lub by ignorowali swoją wiedzę (Bialek 2008, 2010). Mimo tych prób, efekt treści przekonania nadal występuje, co pozwala przyjąć, że ludzie nie są w stanie świadomie ignorować wiarygodności wniosku jako kryterium oceny całego rozumowania.

Istnieją badania, które pokazują, że naturalny sposób prezentacji danych ułatwia ich poprawne przetwarzanie (Goldstein, Gigerenzer 2002). Na przykład rozumienie informacji probabilistycznych może być uzależnione od formatu komunikowania prawdopodobieństwa (Tyszka, Sawicki 2011). Także efekt przekonania znika, gdy poprzedzony jest serią wyborów, w których osoba musi wybierać wniosek poprawny logicznie z pary niewiarygodnych (Trippas, Handley, Verde 2014).

Efekt przekonania nie jest skutkiem ignorowania praw logiki

Ludzie, analizując przedstawione im do oceny rozumowanie, czasem mają odczucie, że wnioskowanie jest poprawne. Odczucie to jest osadzone nie tylko w ocenie wiarygodności wniosku (Thompson, Evans 2012), ale także w poprawności logicznej. Badani, proszeni o ocenę, jak bardzo lubią dane

wnioskowanie, wykazywali pozytywniejszy afekt w kierunku wnioskowań poprawnych logicznie niż niepoprawnych (Morsanyi, Handley 2012, ale por. też dyskusję Klauer, Singman 2012).

Yates i Carlson (1986), Buciarelli i Johnson-Laird (1999), Newstead i in. (2004) oraz Białek (2012) pokazali, że badani proszeni o ewaluację wnioskowań o wyjątkowo prostej strukturze logicznej dokonują jej szybciej, niż gdy jest ona bardziej złożona, ale wiarygodna. Zatem można postulować, że gdy problem jest wyjątkowo prosty, posługujemy się kryterium formalnym i nie angażujemy zbędnych już wtedy porównań z wiedzą. Handley, Newstead i Trippas (2011) pokazali ponadto, że prawomocność logiczna wpływa silniej na ocenę wiarygodności wniosku niż wiarygodność na ocenę prawomocności. Zatem wpływ procesów określających status logiczny nie jest jedynie okazjonalny, ale zachodzi także wtedy, gdy nie jest niezbędny do wykonania zadania.

Pozostaje jednak pytanie, czy efekt przekonania jest następstwem niewykrucia konfliktu między logiką a przekonaniem? Czy wybór pojemnika A, z 99 wygrywającymi kulami na 1000, zamiast pojemnika B, z 1 wygrywającą kulą na 10, jest wynikiem niezdawania sobie sprawy z konfliktu między prawdopodobieństwem wygranej a bezwzględną liczbą wygrywających kul?

Wspomniane zadanie z kulami jest znakomitym materiałem do badania procesu decyzyjnego. Dla oceny, czy badani zdają sobie sprawę ze wspomnianego konfliktu, zadaniom rodzącym konflikt między intuicją a oceną formalną przeciwstawia się problemy analogiczne, tyle że nietworzące konfliktu. Na przykład zwiększenie pojemności drugiego pojemnika do 20 kul powoduje, że zarówno rachunek prawdopodobieństwa jak i bezwzględna liczba kul wygrywających wspierają wybór pojemnika A. Liczne badania porównawcze między konfliktowymi i niekonfliktowymi wersjami tych samych zadań wskazują, że badani potrzebują więcej czasu na te pierwsze (Bonner, Newell 2010; De Neys, Glumicic 2008; Stuppel, Ball 2008; Thompson, Striemer, Reikoff, Gunter, Campbell 2003).

Także badania eytrackerem wykazały, że dla zadań konfliktowych, w których wiarygodność i logiczna poprawność wnioskowań nie są zgodne, badani mają tendencję do reinspekcji przesłanek. Mimo że odpowiedzi z reguły pokrywają się z oceną wiarygodności, to badani zachowują się jednak znacząco inaczej w obu rodzajach zadań (Ball, Philips, Wade, Quayle 2006). Podobnych danych dostarczają wyniki dotyczące neuobrazowania. Pokazują one, że podczas rozwiązywania zadań konfliktowych większą aktywność wykazuje przednia część zakrętu obręczy (*anterior cingulate cortex* – ACC), która pełni rolę monitorującą bieżącą aktywność, hamującą niewłaściwe i inicjującą właściwe reakcje (Bush, Luu, Posner 2000).

Innym wskaźnikiem, który świadczyć ma o wczesnym wykrywaniu konfliktu między poprawnością formalną a intuicją, są miary pewności opinii

(*confidence ratings*). Badanych prosi się nie tylko o ocenę poprawności rozumowania, ale także o podanie pewności co do swojej decyzji. Trippas, Handley, Verde (2014) przedstawiali badanym pary wnioskowań z prośbą, by wskazali oni wnioskowanie poprawne z pary. Badacze wykazali, że gdy badani mogli kierować się wiarygodnością wniosku, średnia pewność własnej opinii była znacznie wyższa, niż gdy nie mieli takiej możliwości. W grupie par o różnej wiarygodności wniosku badani wyrażali niższą pewność wyboru, gdy zadania zawierały konflikt między przekonaniem a prawomocnością wnioskowania, nawet wtedy, gdy kierowali się wiarygodnością. Jednak brak możliwości korzystania z przekonania i konieczność polegania na ocenie formalnej powodowały wyraźny spadek ufności we własne sądy, co po raz kolejny podkreśla ich kluczowe znaczenie dla przebiegu myślenia. Część badań wskazuje na niższą pewność dla zadań konfliktowych niż niekonfliktowych (De Neys, Cromheeke, Osman 2011), ale tylko w grupie osób, które polegały na wiarygodności jako kryterium oceny.

Mimo, że badani proszeni o werbalizowanie swoich myśli (Evans i Over, 1996; Wason i Evans, 1975) praktycznie nigdy nie odnoszą się do formalnych aspektów problemów, to przytoczone tu dane wskazują, że nieświadomie doświadczają oni konfliktu między przekonaniem a procesami formalnymi. Pozwala to przyjąć z dużym prawdopodobieństwem, że wystąpienie efektu przekonania wiąże się z równoczesnym działaniem zarówno procesów heurystycznych (ocena wiarygodności), jak i algorytmicznych (ocena prawomocności).

Wczesne sposoby wyjaśniania efektu przekonania

Opisany powyżej efekt przekonania wywołał poruszenie wśród badaczy i doczekał się setek badań z nim związanych. We wstępnej fazie badań pojawiły się konkurujące ze sobą wyjaśnienia tego zjawiska, dość często powtarzane zresztą po dziś dzień, mianowicie: (1) wybiórcza analiza, (2) źle pojmowana konieczność oraz (3) teoria modeli umysłowych, oparta jednak na innym paradygmacie badawczym¹. Omówię pokrótce każdy z tych sposobów wyjaśniania efektu przekonania.

¹ Evans i inni preferują badanie wspak, czyli ocenę wiarygodności wnioskowania zawierającego już wniosek, podczas gdy Johnson-Laird praktykuje badania, w których ludzie samodzielnie wyciągają wnioski z podanych przesłanek.

Wybiórcza kontrola

Evans, postulując koncepcję wybiórczej analizy (*selective scrutiny*, por. Evans i in. 1983), dowodził, że ludzie w pierwszej kolejności analizują wiarygodność wniosku, a dopiero wtedy, gdy jest on niewiarygodny, przyglądają się jego logicznemu uzasadnieniu. Mówiąc prościej, ludzie nie wykazują żadnej dodatkowej działalności umysłowej (ponad ocenę wiarygodności), jeśli analizują wnioskowanie z konkluzją zgodną z ich przekonaniem. Gdy wniosek jest niewiarygodny, pojawia się motywacja do jego analizy pod kątem prawomocności jego wyprowadzenia i poszukiwania metod jego obalenia.

Istnieją liczne dane, które popierają te przypuszczenia (Ball i in. 2006; Morley i in. 2004), ale istnieje też szereg badań, które pokazują, że prawomocność wnioskowania uwzględniana jest także w problemach, których wniosek jest wiarygodny (Handley, Newstead, Trippas 2010; przegląd innych badań w: Klauer i in. 2000). Teoria ta zatem nie pozostała jedyna i doczekała się alternatywnych wobec siebie propozycji.

Źle pojmowana konieczność wynikania

Aby dobrze zrozumieć koncepcję źle pojmowanej konieczności wynikania (*misinterpreted necessity*), przeanalizujmy następujący problem:

Niektóre uzależniające rzeczy są tanie.

Żadne papierosy nie są tanie.

Zatem niektóre uzależniające rzeczy nie są papierosami.

Rozumowanie to różni się nieco od poprzednio omawianych, gdyż jego nieprawomocność logiczna jest niejednoznaczna (*indeterminant*), tzn. jego wniosek jest spójny z przesłankami, ale nie jest w sposób konieczny prawdziwy. Ludzie są zakłopotani takim rodzajem wniosku, i szczególnie wtedy, aby ocenić jego zasadność, poszukują wsparcia w źródłach innych niż logika. Ponieważ analiza formalna wnioskowania wydaje się ludziom niekonkluzywna (tzn. wniosek może, ale nie musi być prawdziwy), badani ocenę wnioskowania uzależniają od wiarygodności wniosku. Niekonkluzywność oceny formalnej z kolei wynika z nieprzyjmowania kryteriów logiki formalnej i dwuwartościowej tabeli prawdziwości, o czym wspominałem we wstępie. Evans i jego współpracownicy (Evans, Handley, Harper, Johnson-Laird 1999; Evans, Handley, Harper 2001) wykazali, że poprzez prośzenie badanych, by określili, czy wniosek mógłby być prawdziwy (zamiast czy jest koniecznie prawdziwy), można podnieść wskaźniki akceptacji sylogizmów niezależnie od ich struktury logicznej.

Według niektórych badaczy (Dickstein 1981; Markovits, Nantel 1989; Newstead i in. 1992) to właśnie źle rozumiana konieczność wynikania jest powodem występowania efektu przekonań. W ich ujęciu to analiza logiczna jest wcześniejsza, a dopiero jej niekonkluzywność prowadzi do kierowania się wiarygodnością wniosku jako wskazówką do jego oceny.

Mimo różnic między tym wyjaśnieniem a wybiórczą analizą, ich wspólnym mianownikiem, a zarazem słabością w świetle danych empirycznych, jest to, że przewidują występowanie efektu przekonań tylko dla wnioskowań błędnych formalnie.

Teoria modeli umysłowych

Trzecią koncepcją wyjaśniającą efekt przekonań jest teoria modeli umysłowych. Zyskała ona sobie niezwykłą popularność, rozwijając się dynamicznie od połowy lat 1980., głównie dzięki jej twórcy Johnson-Lairdowi i jego współpracownikom (Johnson-Laird 1983; Johnson-Laird, Bara 1984). Według tej teorii rozumowanie przebiega w trzech etapach: (1) tworzenia maksymalnie oszczędnego modelu reprezentującego przesłanki, (2) porównania wyprowadzonego na jego podstawie wniosku z wiedzą (poszukiwanie kontrprzykładów), oraz, jeśli jakieś zostaną znalezione, (3) poszukiwanie alternatywnych modeli umysłowych reprezentujących przesłanki (*flashing-out*, por. Johnson-Laird 2006; Mackiewicz 2000). Jeśli nie znaleziono w 2. etapie kontrprzykładu, czyli wniosek jest zgodny z wiedzą, nie poszukuje się alternatywnych modeli i całe wnioskowanie uznaje za poprawne. Znalezienie kontrprzykładu powoduje poszukiwanie alternatywnych modeli, i jedynie gdy przy rozpatrzeniu kolejnych modeli wniosek jest ciągle ten sam, zostaje uznany za poprawny.

Teoria modeli umysłowych przewiduje zatem, że interakcja między prawomocnością a wiarygodnością zostanie ujawniona jedynie wtedy, gdy możliwe jest stworzenie większej liczby alternatywnych modeli umysłowych. Kolejnym wnioskiem wypływającym z teorii modeli umysłowych jest fakt, że na wystąpienie efektu przekonań nie powinien mieć wpływu rodzaj nieprawomocności wniosku (jednoznaczna bądź nie), który kluczowy jest dla koncepcji źle pojmowanej konieczności. Danych potwierdzających tą koncepcję dostarczył mimowolnie choćby Newstead (1992), który w swojej serii badań, mających potwierdzić koncepcję źle pojmowanej konieczności, użył zadań jednomodowych, niejednoznacznie błędnych, i nie znalazł interakcji między wiarygodnością a prawomocnością wnioskowań.

Podstawowa różnica między tą koncepcją a poprzednio omawianymi jest taka, że posługuje się ona inną metodologią badawczą. Johnson-Laird prezentuje badanym przesłanki prosząc ich o wygenerowanie wniosku, podczas

gdy Evans i in. przedstawiają badanym pełne sylogizmy i proszą o ich ocenę. Morley i in. (2004) dostarczają dowodów, że wnioskowanie do przodu (od przesłanek do wniosku) jest procesem innym niż wsteczne (od wniosku do przesłanek), a teoria modeli umysłowych nie wyjaśnia, jak wniosek może wpływać na konstrukcję modelu umysłowego.

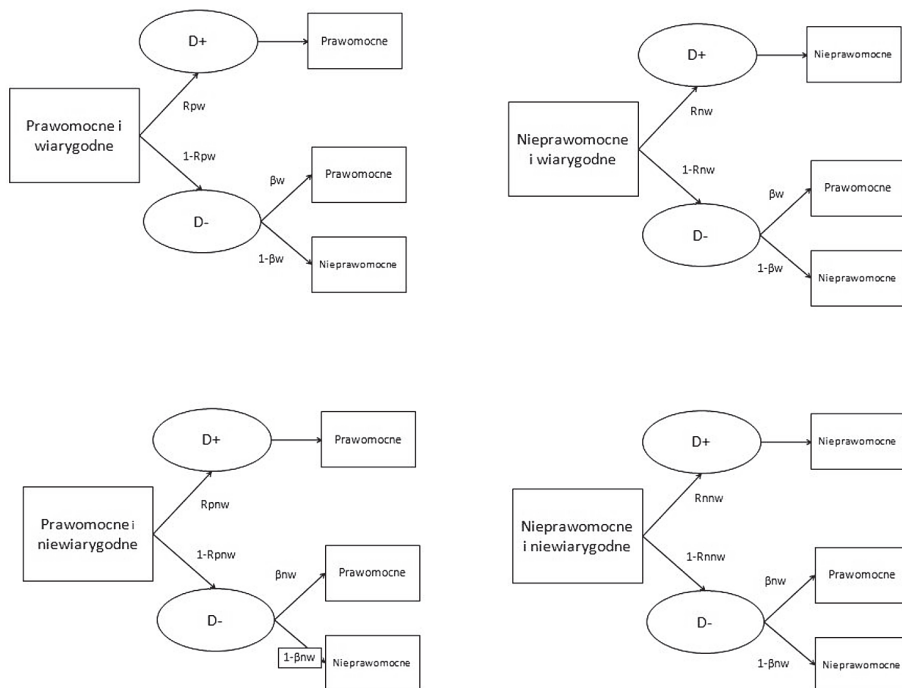
Słabością teorii modeli umysłowych jest to, że efekt przekonań odnajdywany jest także dla wnioskowań poprawnych oraz jednomodelowych, dla których w świetle omawianej teorii nie powinien występować. Dodatkowo Ball, Philips, Wade i Quayle (2006) oraz Thompson i in. (2003) pokazali, że badani więcej czasu poświęcają na analizę wiarygodnych rozumowań. Także tu wyniki nie pokrywają się z przypuszczeniami teorii modeli umysłowych, w świetle której więcej czasu powinno poświęcić się na analizę wniosków niewiarygodnych, gdyż wymagają one generowania większej liczby modeli umysłowych.

Quayle i Ball (2000) rozwinęli teorię modeli umysłowych, postulując, że efekt przekonań jest wynikiem przeciążenia pamięci roboczej przez nadmiar alternatywnych modeli dla nieprawomocnych wniosków, które należy utrzymywać w pamięci. Jeśli zatem alternatyw jest zbyt wiele, aby ocenić prawomocność wniosku, badani posługują się ich wiarygodnością. Wspomniani autorzy dostarczyli danych wskazujących, że niska pojemność pamięci roboczej sprzyja występowaniu interakcji między wiarygodnością a prawomocnością wniosków. Także ta wersja nie jest w pełni potwierdzana przez dane empiryczne, gdyż na przykład nie potwierdzono, by ludzie generowali więcej alternatywnych modeli dla nieprawomocnych wnioskowań.

Współczesne koncepcje efektu przekonań

Wybiórcze przetwarzanie

Za rewizję opisanych wyżej wyjaśnień efektu przekonań zabrali się Klauer, Musch i Naumer (2000). Rozpisali oni drzewa decyzyjne dla rozumowania w zależności od jego prawomocności oraz wiarygodności i ustalili dla nich dwa parametry – „ r ” oraz „ β ”. Parametr r jest wskaźnikiem szansy określenia przez podmiot prawomocności logicznej w analizowanym wnioskowaniu (stan $D+$), a „ $1-r$ ” odpowiada niemożności jej ustalenia i konieczności zgadywania (stan $D-$). Jeżeli analiza wnioskowania nie kończy się powodzeniem w pierwszym etapie, ludzie zmuszeni są kierować się wiarygodnością wniosku (przesłanek). Parametr β odpowiada szansie, że zgadujący uzna rozumowanie za „poprawne”, podczas gdy $1-\beta$ odpowiada wartości niepoprawnej. Autorzy stworzyli 4 takie drzewa, dla wniosków poprawnych i niepoprawnych formalnie oraz wiarygodnych i niewiarygodnych, testując dla nich te same, jak i różne wartości współczynników r oraz β . Modele takie przedstawia rys. 1.



Rys. 1. Drzewa decyzyjne w modelu Klauera i in. (2000)

Opisana wcześniej koncepcja wybiórczej kontroli zakłada, że ludzie nie są w stanie wykrywać prawomocności logicznej dla zdań wiarygodnych, a zatem $R_{(\text{prawomocny, wiarygodny})} = R_{(\text{nieprawomocny, wiarygodny})} = 0$. Zarówno tego założenia, jak i założenia o równości wartości R w różnych modelach nie udało się utrzymać w świetle przeprowadzonej przez wspomnianych autorów serii 8 eksperymentów. Także zrównanie parametrów β dla wniosków wiarygodnych i niewiarygodnych nie poprawiło dopasowania modelu do danych empirycznych.

Po tych ustaleniach Klauer i in. (2000) oraz Evans i jego współpracownicy (Evans, Handley, Harper 2001; Thompson, Evans 2011) zaczęli postulować model wybiórczego przetwarzania (*selective processing*), w którym proces tworzenia modeli umysłowych uzależniony jest od wiarygodności wniosku. Ludzie konstruują inny model umysłowy dla wniosków, które uznają za wiarygodne, a inny dla wniosków niewiarygodnych. Te pierwsze ludzie starają się potwierdzić, konstruując model, w którym wniosek wynika z przesłanek. W drugim, motywacją ludzi jest odrzucenie wniosku przez konstrukcję takiego modelu, który zaprzecza wnioskowi. Następstwem tego jest występowanie dwóch rodzajów efektu przekonań: (1) pozytywnego, w którym niepoprawne wiarygodne wnioskowania akceptowane są częściej niż takie same formalnie, ale przed-

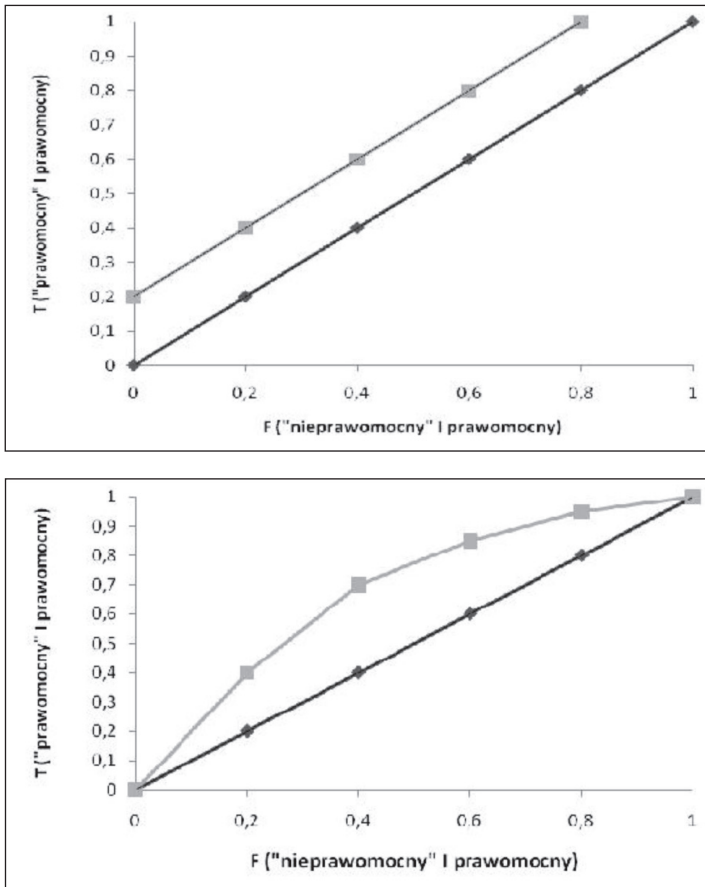
stawione na materiale abstrakcyjnym, oraz (2) negatywnego, który polega na rzadszym akceptowaniu wniosków niewiarygodnych w porównaniu do wniosków abstrakcyjnych. Ludzie nie wpadają zatem w pułapki związane ze skomplikowaną formą logiczną rozumowania (*fallacies*) dla niewiarygodnych wniosków nie dlatego, że poszukują kontrprzykładów dla narzucającego się pierwszego modelu problemu (Oakhill i in. 1989), ale ponieważ od samego początku konstruują modele, które je obalają. Obserwowana wyższa trafność rozumowania dla przesłanek niewiarygodnych jest powodowana tym, że nie da się znaleźć kontrprzykładów dla wnioskowań prawomocnych i niewiarygodnych, zatem ich akceptacja jest wyższa niż niewiarygodnych nieprawomocnych. Gdy wniosek jest wiarygodny, łatwiej jest skonstruować model, w którym jest on prawdziwy nawet dla nieprawomocnych wnioskowań, stąd prawomocne i nieprawomocne wiarygodne wnioskowania są akceptowane praktycznie w tym samym stopniu.

Efekt przekonania jest efektem potakiwania

We wszystkich dotychczasowych badaniach wyniki ujmuje się jako procent akceptacji przedstawianych wnioskowań, czyli analizuje się poprawność rozumowania określaną jako stosunek między trafieniami (H; stwierdzenie „poprawny” dla wnioskowań prawomocnych) a liczbą fałszywych alarmów (F; stwierdzenie „fałszywy” dla wnioskowań prawomocnych). Znajdywana w analizie wariancji interakcja między prawomocnością a wiarygodnością wniosku stanowiła podstawę dywagacji, który z tych aspektów jest oceniany jako pierwszy i modyfikowany przez drugi. Stwierdzony m.in. przez Evansa i in. (2001) wzrost akceptowalności zarówno poprawnych, jak i niepoprawnych, ale wiarygodnych wnioskowań generuje stały przyrost zarówno H, jak i F, dając jednakową poprawność mierzoną wzorem $H-F=k$. Stała wielkość k nie znajduje jednak potwierdzenia w danych empirycznych, dostarczonych przez Dube’go, Rotello i Heita (2010, 2011). Badacze ci dowodzą, że zdolność rozróżniania między poprawnymi a niepoprawnymi wnioskami daje się mierzyć metodologią znaną z teorii detekcji sygnałów, w której wykreśla się krzywe ROC (*receiver operation characteristic*), a pole pod tymi krzywymi odpowiada poprawności oceny prawomocności. Osiami dla krzywych są wartości H i F, a przekątna odpowiada zgadywaniu – człowiek, który osiągałby wyniki leżące na przekątnej, z równym prawdopodobieństwem orzekałby „poprawne” w reakcji na każdy wniosek, niezależnie od jego wiarygodności.

Wszystkie dotychczasowe koncepcje efektu przekonania zakładają, że linie te mają jednakowy, liniowy charakter dla wniosków wiarygodnych oraz niewiarygodnych. Miejsce, jakie zajmuje konkretny badany na wykresie wartości K , odpowiada różnemu natężeniu potakiwania (*responce bias*), czyli chęci odpo-

wiadania ogólnie „poprawny” na każde przedstawiane do oceny zdanie. Jednak znajdowanie się dwóch osób na tej samej linii prostej oznacza, że badani charakteryzują się taką samą poprawnością, a różnią jedynie poziomem potakiwania. Wykres dla stałej wartości różnicy między H a F przedstawia rys. 2.

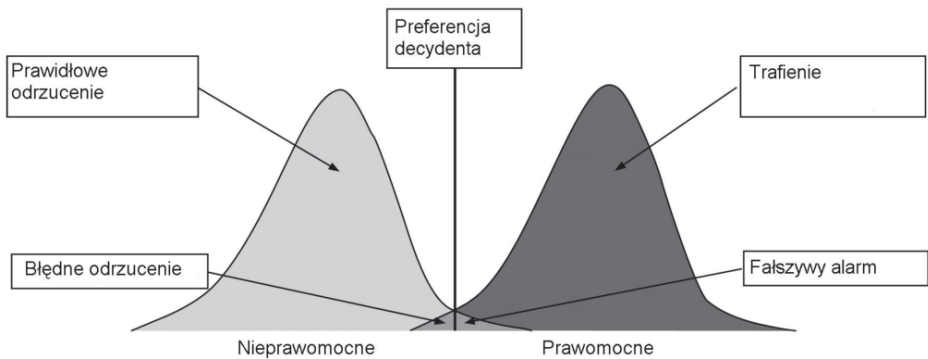


Rys. 2. Linie ROC w wersji prostoliniowej oraz krzywoliniowej

W prawym górnym rogu wykresu znajdują się zatem osoby, które na wszystkie problemy odpowiedzą „poprawne”, a w dolnym lewym „niepoprawne”, każdorazowo notując jednak poprawność na poziomie zgadywania. Im dany punkt znajduje się bliżej lewego górnego rogu, tym większa szansa na trafienie i mniejsza na fałszywy alarm, a zdolność określania prawomocności rozumowania jest wyższa.

Dube, Rotello i Heit (2010, 2011) stwierdzają, że linie ROC dla oceny prawomocności syllogizmów mają charakter krzywoliniowy, a obserwowana

interakcja między prawomocnością a wiarygodnością jest błędem I rodzaju, wynikającym z użycia analizy wariancji. W rzeczywistości argumenty mają niejednakową średnią siłę przekonywania – wyższą dla wniosków wiarygodnych niż niewiarygodnych, oraz mają jednakową wartość odchylenia standardowego, a ich rozkłady częściowo się pokrywają. Ludzie samodzielnie ustalają wartość dyskryminującą między porównywanymi argumentami, określającą akceptowany poziom niepewności. Im bardziej są restrykcyjni, tym mniejsza szansa na błąd I rodzaju (uznanie niepoprawnego wniosku za poprawny), a większa szansa na błąd II rodzaju (uznanie prawomocnego wniosku za błędny). Zależność tę przedstawia rys. 3.



Rys. 3. Rozkłady argumentów w świetle teorii detekcji sygnałów

Na wykresie przedstawiono dystrybucje dwóch argumentów, a pionowa linia oznacza ustaloną przez podmiot wartość dyskryminacyjną. Im większa restrykcyjność decydenta, tym mniejsza skłonność do potakiwania. Mimo zmiany skłonności do potakiwania poprawność rozróżniania między wnioskami prawomocnymi a nieprawomocnymi jest stała.

Konkludując, wspomniani autorzy dowodzą, że postulowana interakcja między wiarygodnością a prawomocnością w ocenie wnioskowań jest artefaktem (sztucznym wytworem) używanej metodologii, a wiarygodność argumentu i wniosku ma przełożenie jedynie na skłonność do potakiwania, a nie wpływa na zdolność określania statusu logicznego rozumowania. Poprawność wnioskowania mierzona jako pole pod krzywą ROC dla argumentów wiarygodnych i niewiarygodnych jest jednakowa, a obserwowane różnice wynikają z błędnego założenia o liniowym charakterze krzywej i dokonywaniu pomiaru w różnych jej miejscach w zależności od spostrzeganej wiarygodności argumentu.

Oczywiście koncepcja ta spotkała się z dużym odzewem ze strony kognitywistów. Niektórzy, jak Klauer i Kellen (2011) twierdzą, że ich koncepcja drzew decyzyjnych (*multinomial trees*) lepiej pasuje do danych empirycznych

zebranych przez Dube'go i in. oraz że ROC ma kształt linii prostej dla dychotomicznych decyzji „prawomocny” vs „nieprawomocny”.

Inni, jak Trippas, Verde i Handley (2011) przyznają rację Dube'mu co do wpływu wiarygodności na skłonność do bardziej lub mniej restrykcyjnego ustalania granicy ewaluacji prawomocności, ale starają się wykazać, że efekt przekonania nie polega jedynie na zwiększaniu lub zmniejszaniu chęci do potakiwania. Wykazali oni, że istnieją różnice w wielkości pola pod krzywą ROC dla wniosków wiarygodnych i niewiarygodnych. Różnic tych nie obserwowano jednak zawsze, nie było ich w przypadku bardzo prostych formalnie problemów, u osób o niskim poziomie zdolności poznawczych oraz przy presji czasowej. Wyniki te potwierdzają krzywoliniowość ROC dla rozumowania sylogistycznego, ale wskazują, że postawiona hipoteza o nieistnieniu interakcji między wiarygodnością a zdolnością do różnicowania między prawomocnymi a nieprawomocnymi wnioskowaniami jest przedwczesna i w sposób nieuzasadniony upraszcza opis efektu przekonania.

Podsumowanie

Efekt przekonania jest przykładem na przetwarzanie odgórne, gdzie posiadane przekonania kształtują proces tworzenia nowej wiedzy oraz podejmowania decyzji. Przetwarzanie odgórne jest procesem szeroko analizowanym w wielu aspektach ludzkiego funkcjonowania poznawczego, począwszy od percepcji, przez pamięć, po podejmowanie decyzji. Badanie efektu przekonania ma więc znaczenie dla wielu subdyscyplin kognitywistyki, a toczący się spór powinien znaleźć swoje odzwierciedlenie w prowadzonych badaniach.

Na przestrzeni ostatnich 10 lat dokonano wielu zmian w opisie efektu przekonania, który jest kluczowym tematem w rozumieniu procesu wnioskowania, ale i szerzej – procesów poznawczych. Od 2010 roku i tekstu Dube'go i in. trwa poważna dyskusja nad tym, czy efekt przekonania jest faktycznie efektem wpływu procesów heurystycznych na algorytmiczne, czy też jedynie pochodną większej chęci potakiwania dla wniosków wiarygodnych.

Niezbędna jest aktualizacja kognitywistyki w różnych jej subdziedzinach o te istotne zmiany w rozumieniu efektu przekonania, a także pogłębianie wiedzy na jego temat, gdyż jest to obszar niezwykle dynamicznej debaty odbywającej się na łamach najbardziej prestiżowych czasopism poświęconych naukom poznawczym. Świat nauki czeka na relewantne dane empiryczne, a mam nadzieję, że ten artykuł pomoże włączyć się w dyskusję także polskim naukowcom zainteresowanym tym tematem.

Bibliografia

- Białek M. (2010), „Związek cech analizowanych treści z pojawiającymi się błędami rozumowania”, w: „*IV Rocznik Kognitywistyczny. W cieniu zakwitających anomalii*”, red. A. Pohl, M. Siedlecka, Kraków, Uniwersytet Jagielloński, s. 210–225
- Białek M. (2012). *Wiem, więc nie myślę. Znaczenie entymematów w wyjaśnianiu tendencyjności rozumowania*. Niepublikowana praca doktorska.
- Białek M. (2008). Język, logika rozumowanie. *Acta Universitatis Lodziensis, Series Folia Psychologica*, (1), 27–44.
- Bonner C., Newell B.R. (2010). In conflict with ourselves? An investigation of heuristic and analytic processes in decision making. *Memory cognition*, 38(2), 186–196.
- Bucciarelli M., Johnson-Laird P.N. (1999). Strategies in syllogistic reasoning. *Cognitive Science*, 23(3), 247–303.
- Bush G., Luu P., Posner M.I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in cognitive sciences*, 4(6), 215–222.
- De Neys W., Glumicic T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, 106, 1248–1299;
- De Neys W., Cromheeke S., Osman M. (2011). Biased but in doubt: Conflict and decision confidence. *PloS one*, 6(1), e15954.
- Dickstein L.S. (1981). Conversion and possibility in syllogistic reasoning. *Bulletin of the Psychonomic Society*, 18, 229–232.
- Dube C., Rotello C.M., Heit E. (2010). Assessing the belief bias effect with ROCs: It's a response bias effect. *Psychological review*, 117(3), 831.
- Dube C., Rotello C.M., Heit E. (2011). The belief bias effect is aptly named: A reply to Klauer and Kellen (2011). *Psychological review*, 118(1), 155–163.
- Elqayam S., Evans J.S.B. (2011). Subtracting „ought” from „is”: Descriptivism versus normativism in the study of human thinking. *Behavioral and Brain Sciences*, 34(5), 233.
- Evans J.S.B. (2011). Dual-process theories of reasoning: Contemporary issues and developmental applications. *Developmental Review*, 31(2), 86–102.
- Evans J.S.B., Over D.E. (1996). *Rationality and reasoning*. Psychology Pr.
- Evans J.S.B., Stanovich K.E. (2013a). Dual-Process Theories of Higher Cognition Advancing the Debate. *Perspectives on Psychological Science*, 8(3), 223–241.
- Evans J.S.B., Stanovich K.E. (2013b). Theory and Metatheory in the Study of Dual Processing. Reply to Comments. *Perspectives on Psychological Science*, 8(3), 263–271.
- Evans J.S.B., Barston J.L., Pollard P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory cognition*, 11(3), 295–306.

- Evans J.S.B., Handley S.J., Harper C.N., Johnson-Laird P.N. (1999). Reasoning about necessity and possibility: A test of the mental model theory of deduction. *Journal Of Experimental Psychology Learning Memory And Cognition*, 25, 1495–1513.
- Evans J.S.B., Newstead S.E., Allen J.L., Pollard P. (1994). Debiasing by instruction: The case of belief bias. *European Journal of Cognitive Psychology*, 6(3), 263–285.
- Evans J.S.B., Handley S.J., Harper C.N.J. (2001). Necessity, possibility and belief: A study of syllogistic reasoning. *Quarterly Journal of Experimental Psychology A*, 54, 935–958.
- Evans J., Frankish K. (2009). *In two minds: Dual processes and beyond*. Oxford University Press.
- Fama E. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 383–417.
- Goldstein D.G., Gigerenzer G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109(1), 75–90
- Handley S.J., Newstead S.E., Trippas D. (2011). Logic, beliefs, and instruction: A test of the default interventionist account of belief bias. *Journal of experimental psychology. Learning, memory, and cognition*, 37(1), 28.
- Handley S.J., Newstead S.E., Trippas D. (2011). Logic, beliefs, and instruction: A test of the default interventionist account of belief bias. *Journal of experimental psychology. Learning, memory, and cognition*, 37(1), 28.
- Johnson-Laird P.N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness* (No. 6). Harvard University Press.
- Johnson-Laird P.N. (2006). *How we reason*. Oxford University Press, USA.
- Johnson-Laird P.N., Bara B.G. (1984). Syllogistic inference. *Cognition*, 16(1), 1–61.
- Kahneman D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Klauer K.C., Kellen D. (2011). Assessing the belief bias effect with ROCs: reply to Dube, Rotello, and Heit (2010). *Psychological review*, 118(1), 164–173.
- Klauer K.C., Singmann H. (2012). *Does Logic Feel Good? Testing for Intuitive Detection of Logicality in Syllogistic Reasoning*.
- Klauer K.C., Musch J., Naumer B. (2000). On belief bias in syllogistic reasoning. *Psychological Review*, 107(4), 852.
- Kruglanski A.W. (2013). Only one? The default interventionist perspective as a unimodel – Commentary on Evans Stanovich (2013). *Perspectives on Psychological Science*, 8(3), 242–247.
- Kubińska E., Markiewicz Ł., Tyszka T. (2012). Disposition Effect Among Contrarian and Momentum Investors. *Journal of Behavioral Finance*, 13(3), 214–225.
- Mackiewicz R. (2000). *Rozumowanie warunkowe w interpretacji teorii modeli umysłowych: psychologiczne badania eksperymentalne* (Vol. 60). Towarzystwo Naukowe Katolickiego Uniwersytetu Lubelskiego.

- Markiewicz Ł., Markiewicz-Żuchowska A. (2012). Skłonności poznawcze sędziego wpływające na wysokość wymierzonej kary. *Decyzje*, 18, 49–81.
- Markiewicz Ł., Weber E.U. (2013). DOSPERT's Gambling Risk-Taking Propensity Scale Predicts Excessive Stock Trading. *Journal of Behavioral Finance*, 14(1), 65–78.
- Markovits H., Nantel G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory Cognition*, 17(1), 11–17.
- Morley N.J., Evans J.S.B., Handley S.J. (2004). Belief bias and figural bias in syllogistic reasoning. *Quarterly Journal of Experimental Psychology Section A*, 57(4), 666–692.
- Morsanyi K., Handley S.J. (2012). Logic feels so good – I like it! Evidence for intuitive detection of logicity in syllogistic reasoning. *Journal of Experimental Psychology–Learning Memory and Cognition*, 38(3), 596.
- Newstead S.E., Handley S.J., Harley C., Wright H., Farrelly D. (2004). Individual differences in deductive reasoning. *Quarterly Journal of Experimental Psychology Section A*, 57(1), 33–60.
- Oakhill J., Johnson-Laird P.N., Garnham A. (1989). Believability and syllogistic reasoning. *Cognition*, 31(2), 117–140.
- Osman M. (2013). A Case Study Dual-Process Theories of Higher Cognition –Commentary on Evans Stanovich (2013). *Perspectives on Psychological Science*, 8(3), 248–252.
- Phillips P., Ball L., Wade C.N., Quayle J.D. (2006). Effects of belief and logic on syllogistic reasoning: Eye-movement evidence for selective processing model. *Experimental Psychology* 53 (1), 77–86
- Quayle J.D., Ball L.J. (2000). Working memory, metacognitive uncertainty, and belief bias in syllogistic reasoning. *The Quarterly Journal of Experimental Psychology: Section A*, 53(4), 1202–1223.
- Stanovich K.E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory. *In two minds: Dual processes and beyond*, 55–88.
- Stuppelle E.J., Ball L.J. (2008). Belief-logic conflict resolution in syllogistic reasoning: Inspection-time evidence for a parallel-process model. *Thinking Reasoning*, 14(2), 168–181.
- Sun Y., Wang H. (2011). Probability theory and perception of randomness: Bridging „ought” and „is”. *Behavioral and Brain Sciences*, 34(05), 271–272.
- Thompson V.A. (2013). Why It Matters The Implications of Autonomous Processes for Dual Process Theories – Commentary on Evans Stanovich (2013). *Perspectives on Psychological Science*, 8(3), 253–256.
- Thompson V.A., Striemer C.L., Reikoff R., Gunter R.W., Campbell J.I. (2003). Syllogistic reasoning time: Disconfirmation disconfirmed. *Psychonomic Bulletin Review*, 10(1), 184–189.

- Thompson, V., Evans, J.S.B. (2012). Belief bias in informal reasoning. *Thinking Reasoning*, 18(3), 278–310.
- Trippas D., Handley S.J., Verde M.F. (2013) The SDT Model of Belief Bias: Complexity, Time and Cognitive Ability Mediate the Effects of Believability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Trippas D., Verde M.F., Handley S.J. (2014). Using forced choice to test belief bias in syllogistic reasoning. *Cognition*, 133(3), 586–600.
- Tyszka T., Sawicki P. (2011). Affective and Cognitive Factors Influencing Sensitivity to Probabilistic Information. *Risk Analysis*, 31(11), 1832–1845. doi: 10.1111/j.1539-6924.2011.01644.x
- Wason P.C., Evans J.S.B. (1975). Dual processes in reasoning. *Cognition*, 3(2), 141–154.
- Yates J.F., Carlson B.W. (1986). Conjunction errors: evidence for multiple judgment procedures, including „signed summation”. *Organizational Behavior and Human Decision Processes*, 37(2), 230–253.

Streszczenie

Już od lat 1980. efekt *belief bias*, czyli preferencja wniosków wiarygodnych niż poprawnych logicznie, stanowi ważny temat dla psychologii rozumowania. Badacze starają się wyjaśnić obserwowaną interakcję między wiarygodnością a prawomocnością wniosku na różne sposoby. W artykule przedstawiam kilka klasycznych modeli teoretycznych: teorię wybiórczej kontroli, modeli umysłowych oraz źle rozumianej konieczności, i omawiam różnice między nimi a ostatnio zaproponowanym modelem wyбіórczego przetwarzania. Wspominam także o debacie nad ideą intuicji logicznej oraz o zastosowaniu teorii detekcji sygnałów do badania rozumowania. Ta druga sprowadza *belief bias* do efektu potakiwania (*response bias*) – tendencji do przyjmowania wszystkich wniosków wiarygodnych i negowania niewiarygodnych. *Response bias* nie wpływa jednak na poprawność oceny prawomocności wniosków.