

Iluzoryczne wnioskowania – czyli o zaskakujących konsekwencjach „teorii modeli umysłowych”

Robert Mackiewicz*

Katedra Psychologii Eksperymentalnej, Katolicki Uniwersytet Lubelski, Lublin

ILLUSORY INFERENCES, OR THE SURPRISING CONSEQUENCES OF THE MENTAL MODEL THEORY OF REASONING

The aim of this paper is to present two competing theories of reasoning: the mental logic theory and the mental model theory. According to the mental logic theory, mind is a syntactic device, and mental processes are governed by a set of rules which are equivalent to those known from classical logic. On the other hand, the mental model theory claims that mind works on semantic principles, building mental models of the premises and drawing inferences based on what is represented in such models. One of the results of this process is the existence of specific reasoning „illusions”. This paper presents such „illusory inferences” as well as the situations in which they are suppressed. Finally, the place of illusory inferences in building the algorithmic theory of mind and the rationality debate is dealt with.

WSTĘP

Wnioskowanie to przechodzenie od zdań uznanych za prawdziwe do innych zdań połączonych z nimi stosunkiem wynikania (Łukasiewicz, Mała Encyklopedia Logiki). Życie codzienne pełne jest różnorodnych wnioskowań i trudno w zasadzie wyobrazić sobie świat w którym wyciąganie nowych wniosków na podstawie danych przesłanek nie byłoby możliwe. Jaki jednak jest psychologiczny mechanizm wnioskowania lub szerzej rozumowania?¹ Zanim przejdę do przedstawienia ostatnio zaprezentowanej teorii rozumowania, teorii modeli umysłowych Johnson-Lairda (Johnson-Laird, 1983, Johnson-Laird i Byrne, 1991), oraz zaskakujących konsekwencji jakie z tej teorii wynikają niech mi będzie wolno poprosić czytelników o rozwiązanie następującego zadania:

Załóżmy, że grasz w karty z Piotrem i masz dwie wskazówki odnośnie kart jakie ma on w ręku. Wiesz, że jedna z tych wskazówek jest prawdziwa a jedna fałszywa, ale, niestety, nie wiesz, która jest prawdziwa, a która jest fałszywa:

Jeżeli w ręku Piotra jest król, to jest w jego ręku as.

Jeżeli w ręku Piotra nie ma króla, to jest w jego ręku as.

Równocześnie wiesz na pewno, że:

W ręku Piotra jest król.

Co można na pewno powiedzieć o kartach Piotra:

¹ Tradycyjnie rozumowanie jest terminem szerszym od wnioskowania, gdyż ten drugi termin dotyczy raczej dedukcji, podczas gdy w ramach rozumowania umieszcza się takie formy aktywności jak myślenie przez analogię czy rozumowania probabilistyczne. Dokładne rozróżnienia terminologiczne są szczególnie charakterystyczne dla prac polskich logików zwłaszcza z tak zwanej szkoły lwowsko-warszawskiej (por. Woleński, 1985). W piśmiennictwie angielskim rozumowanie (*reasoning*) także jest terminem szerszym niż wnioskowanie (*inference*), jednak często terminy te są stosowane dużo bardziej swobodnie niż w języku polskim (por. np. Johnson-Laird, 1993; Garnham i Oakhill, 1994).

* Korespondencję dotyczącą artykułu można kierować na adres: Robert Mackiewicz, Katedra Psychologii Eksperymentalnej, Katolicki Uniwersytet Lubelski, ul. Aleje Racławickie 14, 20-950 Lublin. e-mail: robert.mackiewicz@kul.lublin.pl

- a) Piotr ma w ręku asa,
- b) Piotr nie ma w ręku asa,
- c) Piotr może mieć, ale może też nie mieć asa.

Problem ten jest pewną wersją zadania zaprezentowanego w pracy Johnson-Lairda i Savary'ego (1999). Jeżeli to możliwe, proszę zapisać odpowiedź, aby móc się do niej później odwołać.

Sposób rozwiązywania przedstawionego wyżej problemu (nazwijmy go Problemem Kart w ręku Piotra) jest jednym z ostatnich argumentów teorii modeli umysłowych przeciwko konkurencyjnej do niej teorii logiki umysłu. Poniżej postaram się krótko przedstawić obie rywalizujące teorie kompetencji dedukcji (tak teorie te określa Evans, 1991). W części następnej powrócę do przedstawionego wyżej problemu i innych jemu podobnych oraz do tego jakie jest ich znaczenie w dyskusji pomiędzy teorią logiki umysłu i teorią modeli umysłowych.

LOGIKA UMYŚŁU

Zwolennicy teorii logiki umysłu powołują się na Piageta (1953), według którego myślenie na poziomie operacji formalnych to nic innego jak praktyczna realizacja reguł rachunku zdań. Braine (1994, s. 243) twierdzi, że panuje zgoda wśród zwolenników tej teorii co do tego, że „logika umysłu musi składać się ze zbioru schematów wnioskowania razem z pewnego rodzaju podstawowym programem ich aplikacji”.

Najpełniej teorię logiki umysłu przedstawiają Braine i Romain (1983), Braine i O'Brien (1991) oraz Rips (1983; 1994). Rips (1994) twierdzi, że pewne schematy wnioskowania wydają się oczywiste dla wszystkich co oznacza, że każdemu z tych schematów powinna odpowiadać reguła w jakiś sposób zapisana w umyśle. Weźmy jako przykład następujące rozumowanie:

Jeżeli Piotr wrzuci monetę do automatu, to dostanie on ciastko
 Piotr wrzucił monetę do automatu
 ∴ Piotr dostał ciastko (znak ∴ oznacza „więc”)

Jeżeli pierwsze dwa zdania są prawdziwe to trzecie musi być prawdziwe, gdyż wnioskowanie to opiera się o poprawny logicznie schemat Modus Ponens (MP)². Według Ripsa w umyśle zapisana

jest reguła odpowiadająca właśnie temu schematowi. Oznacza to, że „włącza się” ona ilekroć dany problem zostanie zidentyfikowany właśnie jako Modus Ponens.

Teoria logiki umysłu nazywana jest także „teorią reguł” (Evans, Newstead i Byrne, 1993), ponieważ twierdzi ona, że umysł jest „urządzeniem” dokonującym operacji na symbolach w oparciu o zapisane w nim reguły. Nie wszystkie jednak poprawne reguły logiczne są regułami umysłu. Nie jest regułą umysłu inny schemat znany w logice jako Modus Tollens (MT):

JEŻELI p TO q
 NIE q
 ∴ NIE p

O ile niemal wszyscy zgadzają się, że rozumowanie Modus Ponens jest zawsze poprawne, to czasami zaledwie 17% badanych uważa za poprawne rozumowanie oparte o schemat Modus Tollens (Evans, Newstead i Byrne, 1993). Zgodnie z teorią logiki umysłu, wnioskowanie Modus Tollens jest trudniejsze, ponieważ wymaga ono zbudowania całego dowodu umysłowego wymagającego więcej niż jednego kroku i wykorzystania większej liczby reguł niż Modus Ponens. Oto jak przebiega umysłowy dowód typu Modus Tollens:

1. JEŻELI p TO q	(założenie)
2. NIE q	(założenie)
3. p	(założenie tymczasowe)
4. ∴ q	(Modus Ponens zastosowany do wierszy 1 i 3)
5. q i NIE q	(konjunkcja wierszy 2 i 4)
6. NIE p	(<i>reductio ad absurdum</i> dla wierszy 3 – 5)

Rips opiera swój system logiki umysłu o teorię tak zwanej dedukcji naturalnej pochodzącą od polskich logików ze szkoły lwowsko-warszawskiej (Jaśkowski, 1934)³. Jak widać również na przykładzie prostego wnioskowania typu MT system dedukcji naturalnej korzysta z tymczasowych założeń, które są wprowadzane pomocniczo do dowodu. Wszystkie kroki dowodu, które odwołują się do takiego tymczasowego założenia nazywane są subdomeną dowodu. Jeżeli w wyniku stosowania innych reguł okaże się, że w danej subdomenie występuje sprzeczność (albo wewnątrz niej, albo sprzeczność z domeną nadrzędną), wtedy należy

² Dokładniej jest to schemat Modus Ponendo Ponens.

³ Warto przy tym zwrócić uwagę na to, że sam Rips *explicite* powołuje się na polskich logików jako twórców teorii dedukcji naturalnej.

odrzuć tymczasowe założenie, a przyjąć jego zaprzeczenie. Taka reguła znana jest w logice pod nazwą „doprowadzenia do absurdu” (łac. *reductio ad absurdum*) i jest integralną częścią logiki umysłu. Ta właśnie zasada znajduje swoje zastosowanie w przedstawionym wyżej rozumowaniu MT. Podobna reguła jest również obecna w innym systemie logiki umysłu – systemie Braine i O’Brien (1991), gdzie nazywana jest dowodem warunkowym.

Dowód umysłowy składa się z trzech kroków. Po pierwsze: odkrycie logicznej formy przesłanek. Etap drugi to przeprowadzenie kolejnych kroków rozumowania na symbolach umysłowych odpowiadających poszczególnym zdaniom zapisanym w przesłankach. Krok ostatni to zamiana konstrukcji z języka symboli na język przesłanek. W dowodzie umysłowym operacje są dokonywane na symbolach niezależnych od kontekstu, a ich wynik jest dopiero potem „tłumaczony” na konkretną treść zawartą w przesłankach. Stąd też określenie systemów opartych na dedukcji naturalnej jako syntaktycznych lub niezależnych od treści.

Każda teoria kompetencji dedukcji musi równocześnie wyjaśnić skąd się biorą błędy w rozumowaniu, które jak wiadomo nie tylko zdarzają się często, ale są także popełniane w sposób systematyczny (por. Lewicka, 1993). W teoriach logiki umysłu mówi się zasadniczo o trzech źródłach błędów. Wynikają one albo z konieczności zastosowania większej liczby reguł, albo z konieczności budowania dłuższego dowodu umysłowego (w obu przypadkach ograniczenia pamięci pracującej nie pozwalają na poradzenie sobie ze zbyt złożonym zadaniem), albo, wreszcie, czynniki pragmatyczne mogą mieć wpływ na tłumaczenie przesłanek na symbole umysłowe.

MODELE UMYŚLOWE

Wróćmy do Problemu Kart w ręku Piotra. Poprawne rozwiązanie tego problemu wymaga wzięcia pod uwagę faktu, że obydwie wskazówki („Jeżeli król, to as” i „Jeżeli nie ma króla, to as”) połączone są dysjunkcją co oznacza, że gdy jedna jest prawdziwa, to druga musi być fałszywa. Taka sytuacja jest możliwa tylko w jednym przypadku – gdy Piotr nie ma asa. Pierwsze zdanie jest fałszywe, gdy Piotr ma króla i nie ma asa. Wtedy drugie

zdanie jest prawdziwe między innymi wtedy, gdy poprzednik i następnik są fałszywe, a więc gdy Piotr ma króla i nie ma asa. I odwrotnie, drugie zdanie jest fałszywe, gdy Piotr nie ma króla i nie ma asa, a wtedy pierwsze zdanie jest prawdziwe między innymi wtedy, gdy Piotr nie ma króla i nie ma asa. Analiza całej tablicy prawdziwości (por. Borkowski, 1991) pokazuje, że są to jedyne niewykluczające się sytuacje. Tak więc bez względu na to czy przesłanka kategorialna (trzecia) twierdzi, że Piotr ma czy też nie ma króla, w żadnym wypadku Piotr nie może mieć asa.

Czemu jednak większość osób badanych nie jest w stanie podać poprawnej odpowiedzi? Na przykład na około 1000 kandydatów na studia w Uniwersytecie w Padwie tylko 2% z nich stwierdziło, że Piotr nie ma asa (Vittorio Girotto, informacja własna). Teoria reguł umysłowych nie jest w stanie podać żadnej zadowalającej odpowiedzi (Johnson-Laird i Savary, 1999). Poprawne rozumowanie wymagałoby zastosowania wielu reguł logiki umysłu oraz zbudowania dosyć długiego dowodu umysłowego. Zgodnie z przewidywaniem teorii reguł, gdy sytuacja jest zbyt trudna, człowiek wycofuje się z podawania jakiegokolwiek konkluzji (Rips, 1994; Braine i O’Brien, 1991). Jeżeli tak, to najczęstszą odpowiedzią w Problemie Kart powinno być: „Piotr może mieć asa albo nie”. Tak jednak nie jest.

Odpowiedź jaka najczęściej pada w Problemie Kart wynika bezpośrednio z teorii modeli umysłowych. Teoria ta, oparta na tak zwanej semantyce modelowo – teoretycznej (por. np. Barwise, Etchemendy, 1989; Manktelow i Over, 1990), twierdzi, że rozumowanie polega na budowaniu umysłowych modeli reprezentujących tylko prawdziwe „obiekty” obecne w danej sytuacji problemowej. Tak więc, na przykład zdanie:

„Jeżeli na tablicy narysowane jest koło,
to jest także narysowany trójkąt”

reprezentowane jest przez następujący model umysłowy:

O Δ

Symbole te są w jakiś sposób reprezentowane w umyśle, nie są one jednak dostępne świadomości inaczej jak tylko poprzez swoją treść (Johnson-Laird, informacja ustna). W praktyce ludzie raczej nie reprezentują wszystkich modeli możliwych do utworzenia w danej sytuacji, a poprzestają je-

dynie na takim modelu, który zawiera niezbędne minimum danych. Model ten nosi nazwę modelu początkowego (Johnson-Laird, Byrne, 1991; Johnson-Laird, Byrne i Schaeken, 1992). W przypadku powyższej implikacji model zawierający jedynie koło i trójkąt to właśnie model początkowy. Jest on zbudowany w oparciu o tak zwaną zasadę prawdziwości (Johnson-Laird, 1999; Johnson-Laird, Legrenzi, Girotto, Legrenzi, Caverni, 1999). W celu zminimalizowania obciążenia pamięci pracującej w modelu początkowym reprezentowane jest tylko to co prawdziwe w danej sytuacji. Zasada ta realizowana jest na dwóch poziomach. Po pierwsze w modelu reprezentowana jest jedynie sytuacja bezpośrednio nazwana w przesłance. Tak więc zdaniu „Jeżeli p to q ” odpowiada jeden model zawierający jedynie p i q . I choć zdanie „Jeżeli koło to trójkąt” jest prawdziwe nie tylko wtedy, gdy prawdziwe są zarówno poprzednik jak i następnik implikacji, ale także wtedy, gdy poprzednik jest fałszywy a następnik prawdziwy i gdy oba człony implikacji są fałszywe, to jednak te inne sytuacje nie są początkowo reprezentowane, o ile to nie jest to konieczne. Bardzo często jednak jesteśmy świadomi tego, że te inne sytuacje istnieją. W teorii modeli fakt ten zaznaczany jest z reguły przez wielokropek (. . .), który nazywany jest modelem *implicit*. Pełny zestaw modeli dla powyższej implikacji wygląda następująco:

O	Δ
$\neg O$	Δ
$\neg O$	$\neg \Delta$

Modele te zwane są modelami *explicite* i odpowiadają poszczególnym wierszom logicznej tablicy prawdziwości. Symbol $\neg$ jest w tym wypadku operatorem negacji i oznacza odpowiednio „nie ma koła” oraz „nie ma trójkąta”.

Drugi poziom realizacji zasady prawdziwości dotyczy właśnie reprezentacji negacji. W modelu początkowym obecne są tylko symbole odpowiadające temu co prawdziwe w danej sytuacji. Tak więc reprezentacja zdania „Jeżeli p to q ” zawiera jedynie symbole odpowiadające p i q . Gdyby jednak przesłanka brzmiała „Jeżeli nie p to q ”, wtedy model początkowy dla tego zdania byłby następujący:

$\neg p$	q
----------	-----

Zgodnie z zasadą prawdziwości jeżeli element jest zanegowany w przesłance, to jest w taki właśnie sposób reprezentowany w modelu początkowym.

Zgodnie z teorią modeli dwie wskazówki na temat kart trzymanyh przez Piotra są reprezentowane przez następujące dwa modele początkowe:

Król	As
$\neg$ Król	As

Dodanie trzeciej przesłanki „Piotr ma króla” powoduje wyeliminowanie drugiego z modeli. Tak więc jedyna odpowiedź na pytanie co można powiedzieć o kartach Piotra, zgodna z pozostawionym modelem brzmi: Piotr ma asa. Tak też zapewne odpowiedziała większość czytelników tego tekstu.⁴

WNISKOWANIA ILUZORYCZNE

Problem Kart jest jednym z kilku problemów prowadzących do tak zwanych iluzorycznych wnioskowań (Johnson-Laird, Goldvarg, 1997; Johnson-Laird, Savary, 1996; 1999; Mackiewicz, Schaeken, Schroyens, d'Ydewalle, 1998). Wspólną cechą wszystkich tych problemów jest zależność od zasady prawdziwości. Model początkowy dla każdej z takich sytuacji budowany jest właśnie w oparciu o tę zasadę, co z kolei prowadzi do systematycznych błędów we wnioskowaniu.

Teoria modeli umysłowych nie jest jedynie teorią wyjaśniającą tak zwane rozumowania warunkowe typu „Jeżeli ... to ...”. Również rezultaty badań nad wnioskowaniami w oparciu o inne funktory logiczne otrzymują w tej teorii swoje wyjaśnienie. Najpełniejszą prezentację rozumowań w ramach rachunku zdań zawiera praca Johnson-Lairda, Byrne i Schaekena, 1992. Inne funktory logiczne w równie dobrym stopniu jak implikacja mogą stać się podstawą iluzorycznych wniosków. Przykładem może być poniższy problem będący dysjunkcją dwóch alternatyw (Johnson-Laird i Savary, 1999):

⁴ Warto przy okazji zauważyć, że sam Problem Kart jest dziełem programu komputerowego symulującego teorię modeli umysłowych w języku LISP. Sami autorzy programu (Philip Johnson-Laird i Fabien Savary) początkowo uznali tę odpowiedź za błędną i rozpoczęli poszukiwania błędów w programie.

Tylko jedno z poniższych zdań jest prawdziwe:

Albert jest tutaj lub Betty jest tutaj lub oboje są tutaj
Charlie jest tutaj lub Betty jest tutaj lub oboje są tutaj.

Równocześnie wiadomo na pewno, że:

Alberta tu nie ma i Charliego tu nie ma.

Co z tego wynika?

Jeżeli oznaczymy wszystkie trzy osoby przez ich imiona, to problemowi temu odpowiadają następujące modele początkowe:

Albert

Betty

Albert

Betty

Charlie

Charlie Betty

Pierwsze trzy wiersze są modelami pierwszej z alternatyw, dwa ostatnie oraz wiersz drugi są modelami drugiej z alternatyw. Zgodnie z przesłanką kategoriową (trzecią) wyeliminowane zostają wszystkie modele reprezentujące Alberta i Charliego. Tak więc jedyny model jaki pozostaje to:

Betty

Model ten jest sprzeczny z obydwoma przesłankami, podczas gdy jedna z nich musi być prawdziwa, a druga fałszywa. Zatem, z logicznego punktu widzenia, przedstawiona sytuacja zawiera sprzeczność. Ze sprzeczności zaś może wynikać każde dowolne zdanie. Fakt ten jednak nie przeszkodził badanym przez Johnson-Lairda i Savary'ego studentom Princeton w wyciągnięciu iluzorycznego (czyli błędnego) wniosku: „Betty jest tutaj”.

Trzeci typ iluzji wynikających z teorii modeli dotyczy rozumowań na temat relatywnego prawdopodobieństwa zdarzeń. Przykładem może być następujący problem (Johnson-Laird i Savary, 1996):

Tylko jedno z poniższych twierdzeń, dotyczących kart w ręku pewnego gracza, jest prawdziwe:

Ma on króla lub asa lub obie te karty na raz

Ma on damę lub asa lub obie te karty na raz

Co jest bardziej prawdopodobne: gracz ma króla czy asa?

Choć pytanie dotyczy prawdopodobieństwa, jest to jednak problem dedukcyjny, gdyż posiada on odpowiedź, która musi być prawdziwa, jeżeli przesłanki są prawdziwe (tak definiują dedukcję Johnson-Laird i Byrne, 1991). Utworzenie samych

tylko modeli początkowych nie wystarcza, by tę odpowiedź podać. Oto te modele (k, a, d – oznaczają odpowiednio gracz ma w ręku króla, asa, damę):

k		a
k		a
	d	
		a
	d	a

Odpowiedź jaka najczęściej pada w tym zadaniu zależy od interpretacji prawdopodobieństwa (Johnson-Laird, 1994). W teorii modeli prawdopodobieństwo określane jest przez jedną z dwóch zasad. Pierwsza z nich to tak zwana zasada podzbioru: „Jeżeli A występuje w tych samych modelach co B, a B występuje także w innych modelach, to B jest bardziej prawdopodobne od A”. Zgodnie z tą zasadą nie można jednak określić co jest bardziej prawdopodobne – król czy as. Zasada druga jest łagodniejsza i wynika z milcząco przyjmowanego założenia o jednakowym prawdopodobieństwie modeli. Jeżeli wszystkie modele są jednakowo prawdopodobne, to A jest bardziej prawdopodobne niż B, jeżeli występuje w większej liczbie modeli. Okazuje się, że osoby badane przez Johnson-Lairda i Savary'ego kierują się raczej tą drugą zasadą: 75% z nich stwierdziło, że as jest bardziej prawdopodobny niż król, 4% badanych odpowiedziało w oparciu o pierwszą bardziej konserwatywną zasadę (as i król są jednakowo prawdopodobne), 25 % badanych podało odpowiedź prawidłową – król jest bardziej prawdopodobny niż as. Prawidłowa odpowiedź – ponownie – ukryta jest za zasadą prawdziwości. Wystarczy rozważyć fałszywe przypadki. Jeżeli pierwsze zdanie jest fałszywe, wtedy gracz nie ma ani króla, ani asa. Jeżeli drugie zdanie jest fałszywe, wtedy gracz nie ma ani damy, ani asa. Tak więc w żadnym wypadku as nie może się pojawić w jego ręku. Król zaś jest jak najbardziej możliwy, gdy pierwsze zdanie jest prawdziwe a drugie fałszywe.

Czy można wyjaśnić przedstawione błędy w rozumowaniu dedukcyjnym inaczej niż poprzez teorię modeli umysłowych i zasadę prawdziwości? Szereg argumentów przeciwko temu wyjaśnieniu przedstawiają i po kolei odrzucają Johnson-Laird i Savary (1999). Są to między innymi następujące twierdzenia:

Problemy iluzoryczne są zbyt skomplikowane i odbiegają od codziennej praktyki: gdyby tak jednak było, to osoby badane powinny wycofywać się z podawania odpowiedzi a nie dawać odpowiedzi błędnej.

Osoby badane mają problemy z dostrzeżeniem dysjunktywnego charakteru prezentowanych im przesłanek. Rzeczywiście wszystkie problemy iluzoryczne zaczynają się od: „Tylko jedno z poniższych twierdzeń jest prawdziwe”. Argument ten jednak łatwo odrzucić, gdy weźmie się pod uwagę inne problemy – również dysjunkcje, które do iluzji nie prowadzą (problemy takie były wykorzystywane przez Johnson-Lairda i Savary’ego, 1996 jako tak zwane problemy kontrolne).

Zdania warunkowe mają inną interpretację niż to zakładają. Argument ten odwołuje się do tak zwanej „defektywnej tablicy prawdziwości” (Watson, 1966). Większość osób badanych (por. Evans, Newstead i Byrne, 1993, rozdz. 2) uważa że zdanie „nie p i q ” nie ma żadnego związku z implikacją „Jeżeli p to q ”. Argument ten może dotyczyć jedynie problemu Kart w rękę Piotra, nie ma jednak znaczenia w przypadku Alberta, Betty i Charliego (oraz innych tego typu), gdyż implikacja nie jest obecna w tym modelu. Co więcej, odwołanie się do defektywnej tablicy prawdziwości powinno powodować raczej zawieszanie konkluzji niż podawania wniosków błędnych.

Czy można wyobrazić sobie taką sytuację, w której osoby badane nie popełnią błędu, rozwiązując przedstawione wyżej problemy? Okazuje się, że tak, jeżeli uznamy, że odpowiedź niezdeterminowana (a więc np. „Piotr może mieć, ale może też nie mieć asa”) nie jest odpowiedzią błędną. Tak właśnie jest w obu prezentowanych tutaj teoriach. Według teorii reguł osoby badane wycofują się z podania jakiegokolwiek odpowiedzi, gdy dowód umysłowy wymaga zastosowania zbyt wielu reguł lub gdy wymaga zbyt dużej liczby kroków. Podobnie w teorii modeli zbyt duża liczba modeli i zbyt ich skomplikowanie również powodują wycofywanie się z podania jakiegokolwiek konkluzji. Tabossi, Bell i Johnson-Laird (w druku) pokazali, że wprowadzenie zdania: „Tylko jedno z poniższych zdań jest fałszywe” powoduje istotny spadek liczby błędnych wnioskowań. Podobnie w badaniach Newsome i Johnson-Lairda (1996) wcześniejsza prezentacja fałszywych przypadków dla każdej z reguł również powoduje spadek liczby osób, które wyciągają wniosek błędny. Okazuje

się jednak, że żadna z tych manipulacji nie daje badanym większego „wglądu” w zadanie, gdyż odsetek tych, którzy podają prawidłową odpowiedź ciągle pozostaje niewielki.

Zjawisko zawieszania iluzorycznych wnioskowań zostało poddane dokładniejszym testom w pracy Mackiewicza i współpracowników (1998). Zadaniem osób badanych było podanie odpowiedzi w zadaniach logicznie tożsamy z podanym na początku problemem kart w rękę Piotra. Jedno z tych zadań (nazwane Problemem Wuja Leona) brzmiało tak:

Grasz w karty ze swoim wujem Leonem. W rodzinie słynie on z bawienia się w różnego rodzaju gry i z zadawania szarad. Niektórzy mówią, że ma bzika na tym punkcie ale zawsze gra uczciwie. Zanim rozpoczniecie grę wuj mówi:

Chcę Ci ułatwić grę dlatego dam Ci wskazówkę:

Tylko jedna z poniższych reguł jest prawdziwa, ale nie powiem która:

Jeżeli podrapię się w nos, to mam asa

Jeżeli nie podrapię się w nos, to mam asa

Po czym widzisz, że wuj drapie się w nos.

Zadaniem osób badanych było zakreślenie jednej z podanych odpowiedzi: „Wuj ma asa” (odpowiedź iluzoryczna), „Wuj nie ma asa” (odpowiedź poprawna), „Wuj może mieć ale może też nie mieć asa” (odpowiedź niezdeterminowana). Problem Wuja Leona różni się od problemu Piotra zarówno co do treści jaka jest zawarta w przesłankach, jak i co do kontekstu w jakim zaprezentowana jest cała „zagadka”. W przypadku Piotra przesłanki dotyczyły kart w jego rękę, w przypadku Wuja dotyczą one łatwego do wyobrażenia zachowania się wuja Leona. Również sama osoba wuja może skłaniać do większej ostrożności w przyjmowaniu jego wskazówek (pomimo stwierdzenia, że zawsze gra on uczciwie). W eksperymentach Mackiewicza i innych systematycznie manipulowano treścią przesłanek i kontekstem w jakim zostały one zaprezentowane. Okazało się, że liczba iluzorycznych wnioskowań istotnie spada, gdy kontekst staje się coraz bardziej realny i gdy przesłanki dotyczą czynności osób grających w karty, a nie bezpośrednio kart trzymanyh w rękę gracza. W jednym z eksperymentów zastosowano technikę głośnego myślenia, nagrywając wypowiedzi osób badanych, które padały podczas rozwiązywania Problemu Piotra i Problemu Wuja Leona. Z wypowiedzi osób badanych można wnioskować, że traktują oni obie wskazówki (drapie się w nos to

as / nie drapie się w nos to as) jako równie prawdziwe. Zarówno w przypadku problemu Piotra jak i w problemie wuja eliminują oni tę, która nie jest zgodna z przesłanką kategorialną („Piotr ma asa”, „Wuj drapie się w nos”). Różnica jednak polega na odmiennym traktowaniu drugiej przesłanki. Kontekst i treść w Problemie Wuja powodują, że osoby badane ciągle pamiętają, że przesłanka która pozostała może równie dobrze być prawdziwa i fałszywa. Tak więc nie pozwala to na wyciągnięcie iluzorycznego wniosku „wuj ma asa”. Interpretacja ta ciągle jednak jest zgodna z teorią modeli umysłowych. Można więc podejrzewać, że oprócz mechanizmu budowania modeli istnieje także umysłowa procedura oceny prawdopodobieństwa tychże modeli – pewien swoisty system umysłowych adnotacji „zapisujący” szanse na to, aby odpowiednie sytuacje zaistniały w rzeczywistości (Johnson-Laird, 1994). Modelowa reprezentacja Problemu Wuja może więc wyglądać następująco:

Wuj drapie się w nos	As	(może być prawdziwe albo fałszywe)
Wuj nie drapie się w nos	As	(może być prawdziwe albo fałszywe)

Interpretacja taka byłaby zgodna z modelową teorią rozumowań probabilistycznych (Giroto, Gonzalez, w druku; Johnson-Laird i inni, 1999). Zgodnie z tą teorią modeli osoby badane nie tylko budują modele, ale także są w stanie ocenić prawdopodobieństwo ich występowania.

REGUŁY CZY MODELE?

Zasadnicza różnica pomiędzy teorią reguł umysłu a teorią modeli umysłowych wynika z przyjęcia innej definicji logicznego wynikania. Najogólniej należy wyróżnić wynikanie typu inferencyjnego i wynikanie typu semantycznego. (Manktelow i Over, 1990, Barwise i Etchemendy, 1989)⁵. Pomiedzy pewnym zbiorem zdań *A* oraz zdaniem *z* zachodzi relacja wynikania inferencyjnego wtedy, gdy zdanie *z* można uzyskać przez jedno lub wielokrotne zastosowanie reguł obowiązujących w systemie do zbioru zdań *A*. Z kolei, zdanie *z* wynika semantycznie ze zbioru zdań *A* wtedy i tylko wtedy, gdy nie istnieje żadna taka sytuacja,

w której zdania w *A* są prawdziwe, a zdanie *z* jest fałszywe. O ile więc wynikanie inferencyjne opiera się na możliwości wyprowadzenia jednego zdania z innych przy zastosowaniu pewnych reguł w systemie, to wynikanie semantyczne nie może obyć się bez pojęcia prawdy i warunków prawdziwości.

Wynikanie inferencyjne jest więc wynikiem syntaktycznym ponieważ odwołuje się do istnienia reguł w systemie pozwalających na przejście od przesłanek do konkluzji. Teorie logiki umysłu są właśnie oparte o taką syntaktyczną teorię wynikania. Na przykład w teorii Braine’a i O’Brien’a (1991) w umyśle zapisany jest pewien zbiór reguł (w przypadku rozumowania warunkowego są to dwie reguły: Modus Ponens i tak zwany dowód warunkowy odpowiadający stawianiu czasowych założeń w teorii Ripsa) oraz program sterujący stosowaniem tych reguł. Program ten przeszukuje przesłanki, stara się dopasować do nich którąś z reguł, a ewentualne konkluzje dodaje do systemu. System syntaktyczny nie musi więc odwoływać się do tego czy sytuacje, które analizuje są prawdziwe czy też nie. Zupełnie inaczej jest w systemie semantycznym, którego przykładem jest właśnie teoria modeli umysłowych Johnson-Lairda. W tym przypadku poprawne schematy wnioskowania to takie, które gwarantują dojście do prawdziwej konkluzji wtedy, gdy przesłanki są prawdziwe. Nie da się jednak ocenić prawdziwości przesłanek bez odwołania się do pojęcia modelu. Manktelow i Over (1989) definiują jako model dla zdania logicznego nic innego, jak tylko przypisanie prawdy lub fałszu dla poszczególnych składowych tego zdania. W tym sensie teoria modeli umysłowych jest teorią ekstensjonalną, jako że odwołuje się do znaczenia zdań w danym modelu.

Każda teoria funkcjonowania umysłu musi przede wszystkim adekwatnie wyjaśniać fakty eksperymentalne. Dokładna analiza danych eksperymentalnych z różnych dziedzin psychologii rozumowania nie pozwala na jednoznaczne rozstrzygnięcie pomiędzy teorią reguł a teorią modeli. O ile teoria modeli lepiej radzi sobie z wyjaśnianiem charakteru błędów w rozumowaniu, to teoria reguł wydaje się oferować bardziej adekwatne wyjaśnienie kompetencji rozumowania (Evans, 1991, Falmagne, Gonsalves, 1995). Dokładną analizę danych eksperymentalnych oraz sposób ich wyjaśnienia w obu teoriach przedstawiają Evans, Newstead i Byrne (1993). Zarówno teoria reguł jak i teoria modeli nie ograniczają

⁵ *Mała Encyklopedia Logiki* wymienia aż cztery jego rodzaje: wynikanie inferencyjne, implikacyjne, logiczne i semantyczne

się jedynie do rozumowań w ramach rachunku zdań (Braine i Romain, 1983; Rips, 1989; 1994, Johnson-Laird, Byrne, Schaeken, 1992). Oba podejścia dostarczają też wyjaśnienia szerszej gamy zjawisk określanych ogólnie jako rozumowania. Oprócz zaprezentowanych wyżej eksperymentów nad rozumowaniami w oparciu o rachunek zdań, prace prowadzone w obu paradygmatach dotyczą między innymi: sylogizmów kategoryalnych (Johnson-Laird, 1983; Yang, Braine, i O'Brien, 1998), rozumowań na temat relacji (Byrne, Johnson-Laird, 1989; Schaeken, Johnson-Laird, d'Ydewalle, 1996; Rips, 1994,), czy wnioskowań statystycznych (Giroto, Gonzalez, w druku).

Istnienie iluzorycznych wnioskowań może okazać się jednym z mocniejszych argumentów na rzecz teorii modeli umysłowych. Argument ten jest istotny nie tylko dlatego, że pokazuje wartość predykcyjną tej teorii, ale także dlatego, że wskazuje na semantyczny charakter wnioskowania. Trudno bowiem inaczej, jak przez odwołanie się do treści reprezentacji wyjaśnić powszechność iluzji w cytowanym wyżej Problemie Kart w ręku Piotra oraz równie powszechny fakt zawieszania iluzji, gdy zadanie dotyczy „rubasznego” wuja Leona.

Umiejętność myślenia dedukcyjnego jest jedną z najważniejszych zdolności w funkcjonowaniu poznawczym (Rips, 1994; Johnson-Laird, Byrne, 1991). Nic więc dziwnego, że każdy spór o kompetencję dedukcji (Evans, 1991) jest równocześnie sporem o racjonalność człowieka. Jak do tej pory podano chyba wszystkie możliwe odpowiedzi: od zanegowania w ogóle możliwości znalezienia odpowiedzi (Cohen, 1981) do zupełnej wiary w racjonalność człowieka (Henle, 1962). Istnieje też wiele teorii poszukujących meta – kryterium racjonalności. Przykładami mogą tu być: teoria heurystyk i odchyłeń od racjonalności Evansa (1989), wprowadzenie perspektywy osoby rozumującej – aktor lub obserwator (por. Lewicka, 1993); wyodrębnienie racjonalności-1, która sprowadza się do osiągania celów i racjonalności-2, którą jest rozumowanie zgodne z regułami logiki (Evans, Over, Manktelow, 1993); poszukiwanie ewolucyjnych uzasadnień racjonalności – teoria kontraktów społecznych (Cosmides, 1989); czy wreszcie twierdzenie, że rozumowanie jest tylko reagowaniem na współwystępowanie zjawisk w realnym świecie – teoria frekwencyjna (Gigerenzer, 1994, 1997). Trudno polemizować z faktami: nasze rozumowanie jest w ogromnym stopniu

uzależnione od treści przesłanek i od kontekstu w jakim one są prezentowane. Nie przeszkadza nam to jednak w codziennym funkcjonowaniu. Zadania, takie jak przedstawione wyżej wnioskowania iluzoryczne, prowadzą nas do systematycznych błędów, ale przecież umiejętność rozwiązywania takich problemów jak „Karty w ręku Piotra” nie jest konieczna nawet dla zawodowego brydżysty. Można więc powiedzieć, cytując słynne zdanie Kahnemana i Tversky'ego (1996), że „jesteśmy racjonalni co do zasady, ale błędzimy w szczegółach”.

LITRRATURA

- Barwise, J. Etchemendy, J. (1989) Model-theoretic semantics. W: M. Posner (red.), *Foundations of cognitive science* (s. 207-243). Cambridge, MA: MIT Press.
- Borkowski, L. (1991). *Wprowadzenie do logiki i teorii mnogości*. Lublin: RW KUL.
- Braine, M.D. (1994). Mental logic and how to discover it. W: J. Macnamara, E.R. Gonzalo (red.), *The logical foundations of cognition* (s. 241-263). New York: Oxford University Press
- Braine, M.D., O'Brien, D.P. (1991). A theory of if: A lexical entry, reasoning program and pragmatic principles. *Psychological Review*, 98, 182-203.
- Braine, M.D.S., Romain, B. (1981). Development of comprehension of „of”. *Journal of Experimental Child Psychology*, 31, 46-70.
- Byrne, R.M., Johnson Laird, P.N. (1989). Spatial reasoning. *Journal of Memory and Language*, 28, 564-575.
- Cohen, L.J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and brain sciences*, 4, 317-370.
- Cosmides, L. (1989). The logic of social exchange. Has natural selection shaped how human reason? Studies with the Wason selection task. *Cognition*, 31, 187-276.
- Evans, J. St. (1989). *Bias in human reasoning: Causes and consequences*. Hove, UK: Erlbaum.
- Evans, J.S., Over, D.E., Manktelow, K.I. (1993). Reasoning, decision making and rationality. *Cognition*, 49, 165-187.
- Evans, J.St. (1991). Theories of human reasoning: the fragmented state of the art. *Theory and psychology*, 1, 83-105.
- Evans, J.St., Newstead, S.E., Byrne, R.M. (1993). *Human reasoning: The psychology of deduction*. Hove: Erlbaum.
- Falmagne, R.J., Gonsalves, J. (1995). Deductive inference. *Annual Review of Psychology*, 46, 525-559.
- Gigerenzer, G. (1994) Why the distinction between single event probabilities and frequencies is important for psychology (and vice versa). W: G. Wright, P. Ayton (red.), *Subjective probability* (s. 129-161). New York: Wiley.
- Gigerenzer, G. (1998). Ecological intelligence: an adaptation for frequencies. W: D. Cummins, C. Allen (red.) *The evolution of mind* (s. 9-29). New York: Oxford University Press.
- Giroto, V., Gonzalez, M. (w druku). Strategies and models in statistical reasoning. W: W. Schaeken (red.), *Strategies in deductive reasoning*. Mahwah, N.J.: Erlbaum.
- Henle, M. (1962). On the relation between logic and thinking. *Psychological Review*, 69, 366-378.

- Jaśkowski, S. (1934). On the rules of suppositions in formal logic. *Studia Logica*, 1, 5-32.
- Johnson-Laird, P.N. (1983). *Mental models*. Cambridge: Cambridge University Press.
- Johnson-Laird, P.N. (1994). Mental models and probabilistic thinking. *Cognition*, 50, 189-209.
- Johnson-Laird, P.N. (1999). Deductive reasoning. *Annual Review of Psychology*, 50, 109-135.
- Johnson-Laird, P.N., Byrne, R.M. (1991). *Deduction*. Hove: Erlbaum.
- Johnson-Laird, P.N., Byrne, R.M., Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, 99, 418-439.
- Johnson-Laird, P.N., Goldvarg, Y. (1997). *How to make the impossible seem probable*. *Proceedings of the Nineteenth Annual Conference of the Cognitive science Society*. Mahwah, NJ: Erlbaum.
- Johnson-Laird, P.N., Legrenzi, P., Girotto, V., Legrenzi, M.S., Caverni, J.P. (1999). Naive probability: A mental model theory of extensional reasoning. *Psychological Review*, 106, 62-88.
- Johnson-Laird, P.N., Savary, F. (1996). Illusory inferences about probabilities. *Acta Psychologica*, 93, 69-90.
- Johnson-Laird, P.N., Savary, F. (1999). Illusory inferences: a novel class of erroneous deductions. *Cognition*, 71, 191-229.
- Kahneman, D., Tversky A. (1996). On the reality of cognitive illusions. *Psychological Review*, 103, 582-591.
- Lewicka, M. (1993). *Aktor czy obserwator. Psychologiczne mechanizmy odchylenia od racjonalności w myśleniu potocznym*. Warszawa: PTP Pracowania Wydawnicza.
- Mackiewicz, R., Schaeken, W., Schroynes, W., d'Ydewalle, J. (1998). *Suppression of the illusory inferences*. University of Leuven, Laboratory of Experimental psychology: Research report.
- Manktelow, K.I., Over, D.E. (1990). *Inference and understanding*. London: Rutlege.
- Marciszewski, W. (red.) (1988). *Mała Encyklopedia Logiki*. Wrocław: Ossolineum.
- Newsome, M.R., Johnson-Laird, P.N. (1996). An antidote to illusory inferences? W: G.W. Cottrell (red.), *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* (s. 820). Mahwah: Erlbaum.
- Piaget, J. (1953). *Logic and psychology*. Manchester: Manchester University Press.
- Rips, L.J. (1983). Cognitive processes in propositional reasoning. *Psychologica Review*, 90, 38-71.
- Rips, L.J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA: MIT Press.
- Schaeken, W., Johnson-Laird, P.N., d'Ydewalle, G. (1996). Tense, aspect and temporal reasoning. *Thinking and reasoning*, 2, 309-327.
- Tabossi, P., Bell, V.A., Johnson-Laird, P.N. (w druku). Mental models in deductive, modal, and probabilistic reasoning. W: C. Habel, G. Rickheit (red.), *Mental models in discourse processing and reasoning*. Amsterdam: North-Holland.
- Wason, P.C. (1966). Reasoning. W: B.M. Foss (red.), *New horizons in psychology* (s. 135-151). Harmondsworth: Penguin.
- Woleński, J. (1985). *Filozoficzna szkoła lwowsko-warszawska*. Warszawa: PWN.
- Yang, Y., Braine, M.D.S., O'Brien, D.P. (1998). Some empirical justifications of the mental predicate logic model. W: M.D. Braine, D.P. O'Brien (red.), *Mental Logic* (s. 335-365). Mahwah, NJ: Erlbaum.