

# Mechanika moralności. Dylematy moralne i intuicyjne rozumienie dobra i zła

Michał Białek\*

Centrum Psychologii Ekonomicznej i Badań Decyzji, Akademia Leona Koźmińskiego

MECHANICS OF MORALITY. MORAL DILEMMAS AND INTUITIVE UNDERSTANDING OF RIGHT AND WRONG

The paper presents the state of the art knowledge on moral judgments, focusing mainly on the cognitive processes responsible for producing them. The author discusses the controversies surrounding the most widely accepted two-system model of Joshua Greene, pointing out the simplifications it involves. A few ground-breaking discoveries in this area from recent years are discussed: the application of the process dissociation method and non-conflicting task analysis, which ultimately led to the development of two newer moral judgment models: the hybrid model and the moral myopia model. Similarities and differences between the three aforementioned models are also discussed, and emerging new research directions are indicated.

**Key words:** morality mechanics, moral dilemma, right, wrong, Greene's two-system model, hybrid model, Moral Myopia Model

## WPROWADZENIE

Podczas II Wojny Światowej trwał pontyfikat Piusa XII. Związana z nim jest poważna kontrowersja, gdyż papieżowi zarzucano, że nie potępił nigdy Hitlera i jego zbrodni, trwając w milczeniu przez cały okres II wojny światowej. obrońcy papieża twierdzili, że sam akt potępienia nie dałby nic, może poza silniejszym prześladowaniem Kościoła w Rzeszy. W zamian papież nakazał nuncjaturom, by starały się chronić Żydów przed hitlerowcami, dzięki czemu ocalało wiele ludzkich istnień.

Pytania o moralny wymiar opisanego postępowania ma dwa wymiary – zgodność z pewnymi uniwersalnymi i abstrakcyjnymi zasadami moralnymi oraz konsekwencje praktyczne. Z jednej strony przykazanie „nie zabijaj” jasno sugeruje, należy przeciwstawiać się zabijaniu niewinnych, a jak pisała Zofia Kossak-Szczucka: „Kto milczy w obliczu mordy – staje się współnikiem mordercy. Kto nie potępia – ten przyzwala” (Żukowski, 2015). Z drugiej strony, liczba żyć ludzkich, które ocalono poprzez zachowanie neutralnej postawy Kościoła jest poważnym argumentem za obroną linią postępowania Piusa XII. Osoba, która dokonując oceny moralnej kieruje się kategorycznymi zasadami takimi jak „nie zabijaj”, „nie krzywdź innych”, nazywana jest deontologiem; osoba taka uzna postępowanie Piusa za naganne. Inne osoby – tak zwani utylitarysty – uznają za etyczne takie działania, które przynoszą jak najwięcej korzyści jak największej liczbie osób (nawet kosztem naruszenia pewnych innych zasad). Utylitarysta doceni korzyść, jaką przyniosły działania

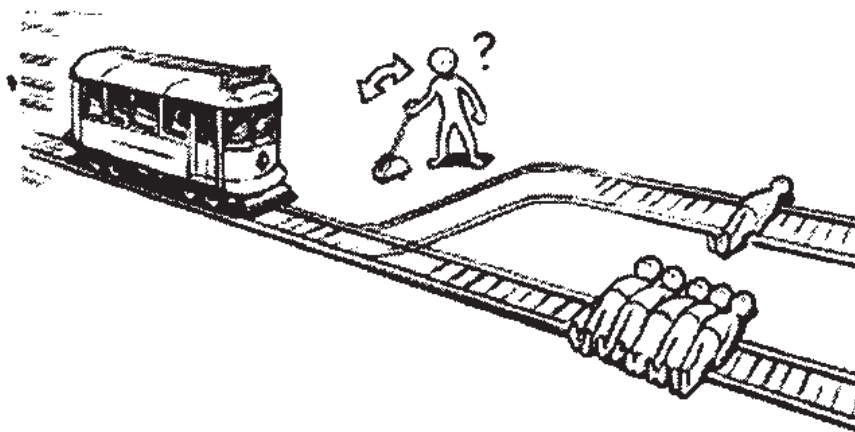
papieża, i uzna jego postępowanie za jak najbardziej moralne. Psychologia ostatnich dwudziestu lat intensywnie analizuje, jak działa umysł deontologa oraz utylitarysty, czym się od siebie różnią i czy któryś z nich wydaje się być „lepszy” (Greene, 2015b).

Aby lepiej zrozumieć moralność w jej potocznym rozumieniu, psychologowie odeszli od normatywnych sporów etyków, dotyczących na przykład niuansów utylitarysty (które z rozwiązań w powyższym dylemacie ma faktycznie wyższą użyteczność?, czy deontologia faktycznie wymaga jawnego potępienia Hitlera?). To, czym się zajęli w zamian, to deskryptywne badania ocen moralnych takich, jakie one są na co dzień (Greene, 2015a; Parzuchowski, Bocian i Wojciszke, 2016). Badacze przedstawiają ludziom serie problemów moralnych, o znacznie uproszczonym w stosunku do wspomnianego wcześniej problemu, z którym zmierzyć musiał się Pius XII. Najpopularniejszym z nich jest problem zwrotnicy (Foot, 1967, 1978), przedstawiony na **Rycinie 1**.

Dylemat wagonika ma liczne zalety: opcje wyboru są jednoznacznie określone, a ich skutki pewne i nieodwracalne<sup>1</sup>, dzięki czemu decyzja w takim dylemacie doskonale różnicuje osoby utylitarystyczne od deontologicznych. Najczęstszą odpowiedzią w tego typu dylematach jest utylitarystyczna (procent utylitarystów waha się pomiędzy 60-80% (Baron, Gürçay, Moore i Starcke, 2012; Bartels, 2008; Greene, 2007). Psychologia w ostatniej dekadzie skupiła się na analizowaniu, kiedy i jacy ludzie przyjmują poszczególne ze stanowisk (Greene, 2015b). Kluczowym ustaleniem jest to, że rzadko kto jest

\* Korespondencję dotyczącą artykułu można kierować na adres: Michał Białek, Akademia Leona Koźmińskiego, ul. Jagiellońska 59, 03-311 Warszawa. mbialek@kozminski.edu.pl

<sup>1</sup> Taka sztuczność sytuacji wiąże się z dyskusją nad trafnością ekologiczną uzyskanych w ten sposób danych (Bauman, McGraw, Bartels, i Warren, 2014).



Wagonik kolejki wyrwał się spod kontroli i pędzi w dół po torach. Na jego drodze znajduje się pięciu ludzi przywiązanych do torów przez szalonego filozofa. Ale możesz przestawić zwrotnicę i w ten sposób skierować wagonik na drugi tor, do którego przywiązany jest jeden człowiek. Co powinieneś zrobić?

Ryc. 1. Dylemat wagonika i zwrotnicy. Stworzony przez filozofów problem obrazuje spór pomiędzy deontologami, których zdaniem przełączenie zwrotnicy jest niemoralne, a utilitarystami, którzy twierdzą, że ocalenia życia pięć osób niesie za sobą większą korzyść, a zatem jest moralne

konsekwentny odpowiadając na szereg dylematów o konstrukcji podobnej do dylematu wagonika – co wskazuje na znaczący wpływ czynników innych niż wyznawana filozofia moralna. Jak inaczej bowiem wytłumaczyć niekonsekwencję w wyborach?

Przedstawiona niekonsekwencja ludzi świetnie obrazuje problem między sytuacją a jej interpretacją (Tomaszewski, 1975). Badani, w takich samych obiektywnie warunkach (wybór między poświęceniem jednej osoby dla ocalenia piętki innych), zdają się dokonywać innych interpretacji sytuacji w zależności od drugorzędnych jej elementów, takich jak rodzaj akcji, którą należy wykonać (Royzman i Baron, 2002), albo to, czy osoba jest celowo poświęcana, czy też ginie „przy okazji”, jako skutek uboczny podjętego działania (Waldmann i Dieterich, 2007). Wpływ na oceny moralne ma na przykład także fakt, czy decydent jest u władzy, czy nie (Lammers, Galinsky, Dubois, i Rucker, 2015; Lammers i Stapel, 2009; Wojciszke, 2011) oraz język, w jakim przedstawiono problem (Hayakawa, Costa, Foucart, i Keysar, 2016). Wszystkie te elementy nie stanowią osi problemu, który należy rozważyć, ale zdają się mieć istotne znaczenie dla ostatecznych decyzji. W efekcie, jak już wspomniałem, niewielu badanych jest konsekwentnym deontologiem czy utilitarystą, gdy ocen mają dokonać dla czterech problemów moralnych (Baron i in., 2012). Skoro dokonywane oceny nie są prostym przełożeniem poglądów decydenta na moralność, to trzeba dobrze zrozumieć, co za nie odpowiada: zrozumieć mechanikę moralności (*moral cognition*).

W kolejnych sekcjach omówię aktualny stan wiedzy i debaty na temat sądów moralnych, ze szczególnym zwróceniem uwagi na procesy poznawcze, które za nie odpowiadają. Omówię poszczególne elementy: teorię

dwusystemową, jej krytykę, oraz szereg najnowszych propozycji alternatywnych do panującego (a niedoskonałego w moim odczuciu) modelu dwusystemowego Greena: propozycję Conwaya i Gawrońskiego, by użyć dysocjacji procesu (Conway i Gawronski, 2013), model moralnej krótkowzroczności Landyego i Royzmana (Landy i Royzman, 2017; Royzman, Landy, i Leeman, 2015) oraz własną propozycję hybrydowego modelu dwusystemowego (Białek i De Neys, 2016, 2017; De Neys i Białek, 2017).

### TEORIA DWUSYSTEMOWA (DUAL-PROCESS THEORY)

W ostatnich latach doszło do ciekawej zmiany ról w nauce. Zwykle ton dyskusji w nauce nadawała filozofia, szukająca odpowiedzi na najbardziej fundamentalne pytania na temat świata i ludzi, takich jak wolna wola i moralność, a w ślad za filozofią podążały nauki empiryczne. Od niedawna jest odwrotnie: to filozofia wywodzi swoje tezy z danych empirycznych, a jej trzon stanowi jej nowy nurt – filozofia eksperymentalna (Knobe i in., 2012; Knobe i Nichols, 2008). Filozofowie eksperymentalni, aby odpowiedzieć na pytania typowe dotąd dla rozważań „w bujanym fotelu”, używają metod psychologii eksperymentalnej, a nawet ekonomii eksperymentalnej czy neuroobrazowania (Kahane, 2013). Podobieństwa między psychologią a filozofią eksperymentalną stają się tak ściśle, że coraz trudniej jest rozróżnić jednych od drugich. Pionierami badań nad mechaniką moralnością są Amerykanie: Jonathan Haidt oraz Joshua Greene. Ich rozumienie moralności bierze swoje korzenie w wszechobecnym w nauce dążeniu do prostych dychotomii – tu: intuicja kontra rozum (Bedford, 1997; Haman, 2008). Greene upatruje źródeł utilitaryzmu w rozumie, podczas

gdy deontologia korzenie swoje mieć ma w intuicji; utilitarysta swoje przekonania i decyzje wywodzi z umiejętności opanowania wrodzonej i automatycznej chęci do unikania czynienia krzywdy innym, kluczowej dla deontologa (Greene, 2014).

Być może nawet lepszą niż wspomniany powyżej problem typu wagonika ilustracją różnicy między deontologią i utilitaryzmem jest seria problemów stworzonych przez Haidta (Haidt, Koller, i Dias, 1993). Jedna z takich historii opisuje, jak pewien mężczyzna kupuje co weekend mrożonego kurczaka, na osobności kopuluje z nim, a następnie gotuje go i zjada. Nikt tego nie widzi, ani o tym nie wie.

Większość z nas czuje, że zachowanie to jest odrażające. Nic zatem dziwnego, że duża część z nas, w oparciu o to uczucie, uzna zachowanie opisywanego mężczyzny za niemoralne (Eskine, Kacinek, i Prinz, 2011; Prinz, 2006; Wheatley i Haidt, 2005). Są jednak ludzie, którzy zaczynają się zastanawiać, komu szkodzi to, że mężczyzna robi to, co robi. W końcu nikt nie widzi jego zachowania, nie grozi ono żadną chorobą, a nawet jedzenie się nie marnuje. Taka refleksja prowadzi do wniosku, że zachowanie mężczyzny – mimo że odrażające – nie jest niemoralne. Nie ma szkody – nie ma podstaw, by potępiać moralnie. To stanowisko ilustruje pozycje utilitarystyczną, która moralność zachowania ocenia przez pryzmat ich następstw.

Widzimy wyraźnie, jak namysł i analiza pozwala przezwyciężyć odruch oceniania moralności w oparciu o wrodzone emocje, jak choćby wstręt. Haidt (2001, 2007) sugeruje, że zdecydowana większość naszych sądów bierze się właśnie z intuicji, rzucając tym samym wyzwanie wcześniejszym, racjonalistycznym ujęciom moralności (Kohlberg, 1981; Piaget, 1997; Turiel, 1998). W rozumieniu Haidta człowiek, a raczej jego sądy, są jak słoń, który idzie własną drogą, podczas gdy rozum jest jedynie pozornie kierowcą. Tak naprawdę rozum ma za zadanie utrzymywać słońca w przekonaniu, że idzie we właściwym kierunku, oraz dostarczać argumentów, by przekonać do tego innych (Haidt, 2012), por. też (Białek, 2015a; Jost, 2012). Także Greene, mimo że ma silniejszą wiarę w rozum ludzki, przyznaje, że większość naszych ocen moralnych bierze się z intuicji i emocji (Greene, 2014, 2015a).

Model dwusystemowy wspiera szereg badań, prowadzonych głównie przy zastosowaniu dylematów moralnych będących wariacją wspomnianego już dylematu wagonika. Badania wskazują, że osoby, które zmuszone są do odpowiadania pod presją czasu, dokonują znacząco mniej utilitarystycznych wyborów (Suter i Hertwig, 2011) niż osoby mające czas do namysłu; podobny efekt daje zajęcie zasobów poznawczych innym zadaniem (Conway i Gawronski, 2013; Trémolière i Bonnefon, 2014; Białek i De Neys, 2017). W efekcie ludziom najwidoczniej brakuje zasobów poznawczych, by skupić się na konsekwencjach wyboru, a w zamian bardziej polegają na swoich emocjach. Więcej utilitarystycznych wyborów obserwujemy, gdy prosimy osoby, by rozważyły problem

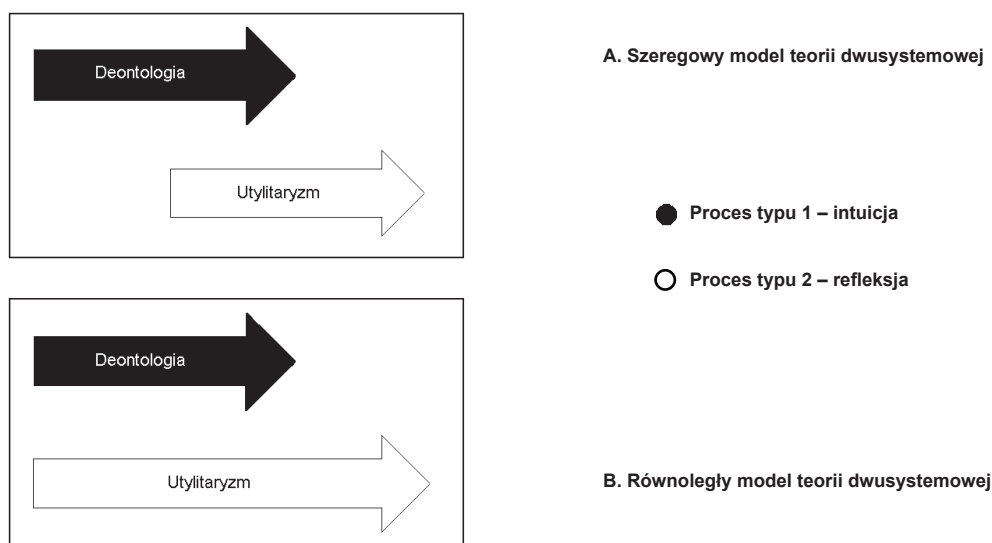
moralny, posługując się językiem obcym zamiast ojczystym (Cipolletti, McFarlane, i Weissglass, 2015; Corey i Costa, 2015; Geipel, Hadjichristidis, i Surian, 2015b). Ten ostatni efekt autorzy tłumaczą niższą aktywnością emocji, gdyż język obcy nie ma tak silnych afektywnych konotacji, jakie ma język ojczysty, w efekcie słabiej odczuwamy emocje, takie jak oburzenie czy odrazę – stanowiące podstawę deontologii. Gdy rozważamy problemy moralne w języku obcym, mniejsze znaczenie mają też intencje (Geipel, Hadjichristidis, i Surian, 2015a) oraz normy społeczne (Gawinkowska, Paradowski, i Bilewicz, 2013). Te z kolei stanowią podstawę dla sądów utilitarystycznych: użyteczność decyzji pojawia się w miejsce brakujących afektywnych komponentów oceny moralnej. Istnieje też cała seria danych z neuroobrazowania, które łączą utilitaryzm z ośrodkami odpowiedzialnymi za funkcje wykonawcze czy myślenie abstrakcyjne (Greene, Nystrom, Engell, Darley, i Cohen, 2004), a deontologię z ośrodkami odpowiedzialnymi za emocje (Moll, Zahn, de Oliveira-Souza, Krueger, i Grafman, 2005).

#### WARIANTY TEORII DWUSYSTEMOWEJ

Istotą teorii dwusystemowej jest założenie, że utilitaryzm opiera się na procesach refleksyjnych, podczas gdy deontologia ma swoje korzenie w emocjach i intuicji, i jest ślepa na argumenty. Teoria dwusystemowa nie jest jednak jednolita, istnieje raczej cała grupa modeli o tym wspólnym założeniu (por. Evans i Stanovich, 2013). Tym, co je różni najbardziej, jest sposób, w jaki współpracują ze sobą procesy typu 1 – te szybkie, odpowiedzialne za intuicję, i procesy typu 2 – te wolne, odpowiedzialne za namysł (Evans, 2008; Evans i Stanovich, 2013)<sup>2</sup>. Istnieją dwa podstawowe typy modeli dwusystemowych: szeregowy i równoległy (**Ryc 2**, panele A i B).

Modele szeregowy zakładają, że ludzie uruchamiają procesy typu 1 zawsze, gdy natrafiają na odpowiedni bodziec ze środowiska. Nie mają nad tym żadnej kontroli. Od decyzji człowieka zależy za to, czy uruchomi też procesy typu 2. Taka wersja (model *default-interventionist*) jest najpopularniejsza, np. w obszarze rozumowania, gdzie zakłada się, że człowiek dokonuje pobieżnej analizy swojej intuicyjnej odpowiedzi, i gdy czuje, że nie jest ona doskonała, to stara się skorygować ją o procesy refleksyjne (Evans, 2008; Evans i Stanovich, 2013). Na przykład, gdy intuicja podpowiada nam, że dany wniosek jest prawdziwy, przyjmujemy go jako prawomocny,

<sup>2</sup> Warto zwrócić uwagę na to, że w polskiej literaturze przyjęło się mówić o teorii „dwusystemowej”, podczas gdy w anglojęzycznych tekstach mówi się o teorii dwóch procesów (*dual-process theory*). W ostatnich latach sami teoretycy z tego nurtu jasno mówią, że nie istnieje spójny „system”, który odpowiadałby za jakieś konkretne procesy poznawcze; mamy raczej do czynienia z interakcją szeregu procesów, z których niektóre mają charakterystykę typu 1, a inne typu 2 (Evans i Frankish, 2009). Na przykład w rozumowaniu przekonania o prawdziwości wniosku mają źródło w procesach typu 1, podczas gdy przekonania o prawomocności wniosku w procesach typu 2. Stąd konsekwentnie w reszcie tekstu będę odwoływał się do procesów, a nie systemów.



**Ryc. 2. Szeregowy i równoległy ujęcie teorii dwusystemowej. Procesy typu 1 są szybkie, automatyczne, niekontrolowalne wolicjonalnie i nie wymagają zasobów pamięci roboczej. Procesy typu 2 są wolicjonalnie uruchamiane, powolne i uzależnione od pojemności pamięci roboczej. Wielkość strzałek odpowiada sile intuicji, a ich długość i miejsce, gdzie się zaczynają oraz kończą – charakterystyce czasowej**

ale jeśli sądzimy że jest on nieprawdziwy, czujemy opór przed jego zaakceptowaniem i szukamy kontrargumentów, nierzadko używając do tego procesów typu 2 (Białek, 2015b; Howarth, Handley, i Walsh, 2016; Newstead, Pollard, Evans, i Allen, 1992)<sup>3</sup>.

Równoległa wersja teorii dwusystemowej zakłada, że procesy typu 1 i 2 startują równocześnie, gdy w otoczeniu pojawi się stosowny bodziec, ale procesy typu 1 są szybsze i z reguły przejmują kontrolę nad działaniem. Gdy decydent pozna odpowiedź intuicyjną, to w zależności od jej subiektywnej jakości, procesy typu 2 są wygaszane albo kontynuowane. W tym drugim przypadku mają szansę na przejęcie kontroli nad zachowaniem (Stanovich, 2009; Stupple i Ball, 2008). Model ten narusza jednak zasadę skapstwa poznawczego, w świetle której umysł skonstruowany jest tak, by pracować jak najefektywniej, przy jak najmniejszym użyciu zasobów poznawczych (Fiske i Taylor, 2013). Mało tego, ostatnie dane wskazują też, że procesy typu 1 mogą być wolniejsze niż typu 2 (Bago i De Neys, 2017; Trippas, Thompson i Handley, 2017), co znacznie komplikuje klasyfikację danego procesu jako typu 1 lub 2 w oparciu o czas reakcji, tak jak postuluje choćby noblista, Daniel Kahneman (2011).

Greene nie dookreśla, którą wersją teorii dwusystemowej jest jego model sądów moralnych, szczęśliwie uni-

kając konieczności przyjęcia wszystkich jej założeń (De Neys i Białek, 2017; Gürçay i Baron, 2017). Nie wiemy zatem, czy utylityzm pojawia się wtedy, gdy intuicja deontologiczna daje niesatysfakcjonującą odpowiedź, czy też równoległe prowadzona jest refleksja utylitarystyczna. Na to pytanie starają się odpowiedzieć inni badacze, gdyż jeśli model Greena nie daje się osadzić spójnie w ramach żadnego z modeli – szeregowego bądź równoległego – to oznacza to, że teoria ta jest nadmiernie uproszczona, albo nawet całkiem błędna. Oczekujemy przecież w nauce pewnej spójności w zakresie tworzonych modeli, i nieracjonalne wydaje się tworzyć jakościowo odmiennego modelu wydawania sądów, a odmiennego modelu dla wydawania sądów moralnych (Białek i Terbeck, 2016; Osman i Wiegmann, 2017).

#### REFLEKSJA I REFLEKSYJNY STYL POZNAWCZY

Jak wynika z powyższych rozważań, dla zrozumienia teorii dwusystemowej kluczowa jest odpowiedź na pytanie: pod jakim warunkiem dochodzi do wygaszenia procesów typu 2 (wersja równoległa) lub ich uruchomienia (wersja szeregową). Aby to wyjaśnić, musi istnieć jakiś metapoznawczy proces, śledzący jakość odpowiedzi intuicyjnych i decydujący, czy są one satysfakcjonujące (Thompson i Johnson, 2014; Thompson, Turner, i Pennycook, 2011; Thompson i in., 2013). W tym zakresie istnieją różnice indywidualne: niektórzy łatwiej zadowolają się pierwszym rozwiązaniem, które pojawia się na horyzoncie, a inni raczej szukają dalej, zatem ci pierwsi częściej polegają na procesach typu 1, a ci drudzy skłonni są częściej

<sup>3</sup> Warto zauważyć, że mechanika moralności czerpie wiele z psychologii rozumowania – zarówno na poziomie metodologicznym (Tremoliere, De Neys, i Bonnefon, 2016), jak i konceptualnym (Białek i Terbeck, 2016). Już sama teoria dwusystemowa początki ma w badaniach na temat rozumowania sylogistycznego (Newstead i in., 1992).

angażować procesy typu 2. Kwantyfikacją tych różnic jest refleksja poznawcza, mierzona przy pomocy niezwykle popularnego 3-pytaniowego testu CRT (*cognitive reflection test*, Frederick, 2005). Przykładowe zadanie tego testu brzmi:

Jeziro porastają lilie wodne, które każdego dnia podwajają swoją powierzchnię. Zarosną one całe jezioro w 48 dni. Ile czasu potrzeba, by porosły one połowę jeziora?

Intuicyjna i najczęstsza odpowiedź to 24 dni (połowa czasu na połowę jeziora), podczas gdy prawidłowa odpowiedź to 47 dni (dzień przed zarośnięciem całej tafli lilie były dwa razy mniejsze, czyli pokrywały dokładnie pół jeziora!). Większość ludzi jest jednak zadowolona z błędnie udzielonej odpowiedzi, twierdzą przy tym, że zadanie to było łatwe, a oni rozwiązali je poprawnie (Thompson i in., 2011; Thompson i in., 2013). Osoby, które lepiej radzą sobie z zadaniami w teście CRT, znacznie lepiej radzą sobie w zadaniach dotyczących różnorodnych błędów poznawczych (Noori, 2016), a refleksja poznawcza jest lepszym predyktorem wyniku w tych zadaniach niż iloraz inteligencji (Toplak, West, i Stanovich, 2011) czy zdolności poznawcze (Białek i Domurat, 2017).

Badacze wiążą refleksyjność poznawczą ze sposobem rozstrzygania dylematów moralnych, pokazując że bardziej refleksyjne osoby częściej skłonne są łamać zasady, jeśli przynosi to więcej korzyści dla społeczeństwa (Patton, Ungar, i Greene, 2012; Pennycook, Cheyne, Barr, Koehler, i Fugelsang, 2014). Okazuje się jednak, że podejście to jest wątpliwe. Na przykład korelacja CRT i utilitaryzmu jest, zdaniem Barona, sztuczna, a wynik CRT jest silnie skorelowany z innymi zmiennymi psychologicznymi, na przykład skłonnością do aktywnego myślenia (*Active Openminded Thinking*, AOT, Baron, 1993), czy religijnością, które w dalszej kolejności powiązane są z przekonaniami moralnymi danych osób (Baron, Scott, Fincher, i Emlen Metz, 2015). Refleksja poznawcza nie jest zatem realnym predyktorem sądów moralnych, a jedynie tych predyktorów korelatem.

Innym problemem, z którym boryka się model dwusystemowy Greena, jest to, że model ów przewiduje, że sądy utilitarystyczne mające swoje korzenie w refleksji, powinny wymagać dłuższego czasu niż automatycznie dostępne oceny deontologiczne. Baron (Baron i in., 2012; Gürçay i Baron, 2016) pokazał jednak, że najdłuższe czasy reakcji badani mają nie wtedy, gdy najczęstszą decyzją jest utilitarystyczna, ale wtedy, gdy odpowiedzi dzielą się po równo pomiędzy deontologów i utilitarystów. Predyktorem czasu potrzebnego na podjęcie decyzji jest zatem siła wewnętrznego konfliktu pomiędzy deontologią a utilitaryzmem, a nie prawdopodobieństwo podjęcia decyzji utilitarystycznej. Jest to ewidentna sprzeczność z szeregowym modelem teorii dwusystemowej.

Ostatni kamyk do ogródka teorii Greena pochodzi z prac Guya Kahane z Oksfordu, który pokazał że w niektórych dylematach pierwszą, intuicyjnie dostępną odpowiedzią jest ta utilitarystyczna. Na przykład, gdy wiem, że nasz znajomy przechowuje zakazane książki na strychu, wolimy

skłamać pytającym nas o to policjantom, aby uchronić go przed karą. Namysłu wymaga w ogóle uświadomienie sobie fakt złamania deontologicznego nakazu mówienia prawdy, niezależnie od jego konsekwencji (Kahane i in., 2012). W tym kontekście Kahane mówi, że dylematy „wagoniopodobne” są po prostu przedstawicielami grupy problemów, w których utilitaryzm jest intuicyjny, podkreślając jednocześnie, że nie jest to grupa wyczerpująca dla wszystkich problemów moralnych (Kahane, 2015).

## NOWA MECHANIKA MORALNOŚCI

W świetle wcześniejszych rozważań kluczowe dla zrozumienia mechaniki moralności staje się pytanie o to, kiedy zaczynamy głębiej analizować problem oraz co z tej analizy wynika. W świetle badań nad rozumowaniem oraz moralnością wiemy, że za refleksję odpowiada subiektywne poczucie, że odpowiedź, która pojawiła się w umyśle, nie jest tą właściwą, że warto poszukać dalej. W świetle teorii dwusystemowej odczuwany konflikt nie jest to „rozdarcie” między dwiema opcjami, bo przecież wg modelu Greena, aby „odkryć” jedną z opcji potrzebujemy przecież pomyśleć. Po prostu coś podpowiada nam, że warto szukać dalej, bo gdzieś tam widnieje szansa znalezienia lepszego rozwiązania.

Badania nad rozumowaniem, nie tylko moralnym, pokazują, że im większy konflikt wewnętrzny i czym mniejsza subiektywna pewność co do poprawności pierwszej, nasuwającej się odpowiedzi – tym dłuższy czas namysłu, a co za tym idzie, większy udział procesów typu 2. Metapoznanie wydaje się więc kluczowym konstruktem w rozumieniu tego, jak działają procesy poznawcze odpowiedzialne za tworzenie sądów moralnych. Dobrym wskaźnikiem metapoznania jest pewność siebie – osoby które doświadczyły poczucia, że ich intuicyjna odpowiedź nie jest zadowalająca, wskazują mniejszą pewność siebie po dokonaniu wyboru (Botvinick, Braver, Barch, Carter, i Cohen, 2001; Evans, 2007), potrzebują na to więcej czasu (De Neys i Glumicic, 2008; Franssens i De Neys, 2009), a ich wzrok częściej wraca w miejsca wyświetlania konkurencyjnej opcji (Stuppel i Ball, 2008). Monitorowanie własnych procesów poznawczych daje się zatem śledzić behawioralnie, a w oparciu o te wskaźniki przewidywać, czy do rozwiązania problemu użyte zostaną procesy typu 2, czy nie (Handley i Trippas, 2015; Pennycook, Fugelsang, i Koehler, 2015). Analiza pewności siebie i czasów reakcji, a nie tylko samych wyborów, stanowi ważny element badania mechaniki moralności, brakujący element do spójnego modelu teoretycznego.

## DEONTOLOGIA I UTILITYRYZM TO NIE DWA KRAŃCE KONTINUIUM, ALE DWIE ODRĘBNE INKLINACJE

Podstawowym założeniem przyświecającym badaniu moralności jest to, że podczas wydawania ocen moralnych większa skłonność do polegania na uniwersalnych zasadach moralnych (deontologia) wiąże się z mniejszym dążeniem do maksymalizacji użyteczności (utilitaryzm).

Idąc tym tropem, pytając o to, czy ktoś akceptuje poświęcenie jednej osoby dla ocalenia większości, badacze nie tylko wnioskujeją wprost o identyfikacji danej osoby z utilitaryzmem, ale także pośrednio o jej identyfikacji z deontologią. Gdy na przykład znajdujemy się pod presją czasu i rzadziej decydujemy się na dokonanie poświęcenia, lub gdy jesteśmy bardziej surowi dla osoby naruszającej normy społeczne, to wnioskuje się, że spadła nasza inklinacja utilitarystyczna, a także wzrosła deontologiczna. Dopiero niedawno Conway i Gawronski (2013) zasugerowali, że obie inklinacje moralne mogą być dostępne równocześnie i tworzyć konflikt typu dążenie – dążenie. Faktycznie, autorzy zaaplikowali metodę dysocjacji procesu – prosty matematyczny zabieg, który pozwala wyabstrahować komponenty bardziej złożonych mechanizmów decyzyjnych. Autorzy pokazali, że na przykład obciążenie poznawcze wpływa na inklinację utilitarystyczną, ale nie deontologiczną, a empatia koreluje z inklinacją deontologiczną, a nie koreluje z inklinacją utilitarystyczną. W innym badaniu Friesdorf, Conway, i Gawronski (2015) pokazali, że kobiety są bardziej deontologiczne niż mężczyźni, nie różniąc się jednak od nich inklinacją utilitarystyczną.

Ta zmiana w myśleniu o moralności pozwala dokonać nowych odkryć, o jakie trudno byłoby w klasycznym podejściu, bo np. gdy rosną obie inklinacje równocześnie to – ponieważ są one w konflikcie – na poziomie behawioralnym nie widać zmian. Na przykład tożsamość moralna (czyli stopień, w jakim osoba sama siebie uznaje za moralną, *moral identity*; Hardy i Carlo, 2011) skorelowana jest z każdą z inklinacji moralnych, co znaczy, że im bardziej osoba uznaje się za moralną, tym bardziej identyfikuje się równocześnie zarówno z deontologią, jak i utilitaryzmem. W efekcie tożsamość moralna nie koreluje ze sposobem rozstrzygania dylematów moralnych, gdyż w dylemacie wagonika wyższa chęć ratowania pięciu osób równoważona jest większą chęcią przestrzegania norm i zasad (Conway i Gawronski, 2013). Mimo to, w świetle nowego modelu Conwaya i Gawrońskiego, obie inklinacje mają inne podłoże poznawcze – deontologia wywodzi się z procesów typu 1, podczas gdy utilitaryzm z procesów typu 2. Nie odeszliśmy zatem tak daleko od modelu dwusystemowego, jak by się mogło wydawać.

## MODELE ALTERNATYWNE DO MODELU GREENA

### MODEL HYBRYDOWY SĄDÓW MORALNYCH

Pomyślmy chwilę nad przykładowym dylematem użytym w badaniu Conwaya i Gawrońskiego. Możemy cofnąć się w czasie i zadusić bezkarnie a) młodego Hitlera, zanim dojdzie do władzy i przyczyni się do śmierci milionów, lub b) młodego człowieka, który później porwie, a po otrzymaniu okupu wypuści dziecko. Gdyby osoby deontologiczne podejmowały decyzje intuicyjnie, w oparciu o subiektywne poczucie niewłaściwości dokonania morderstwa, powinny oceniać oba problemy jednakowo (De Neys i Białek, 2017). Okazuje się jednak, że deontologowie potrzebują więcej czasu, by podjąć decyzję w

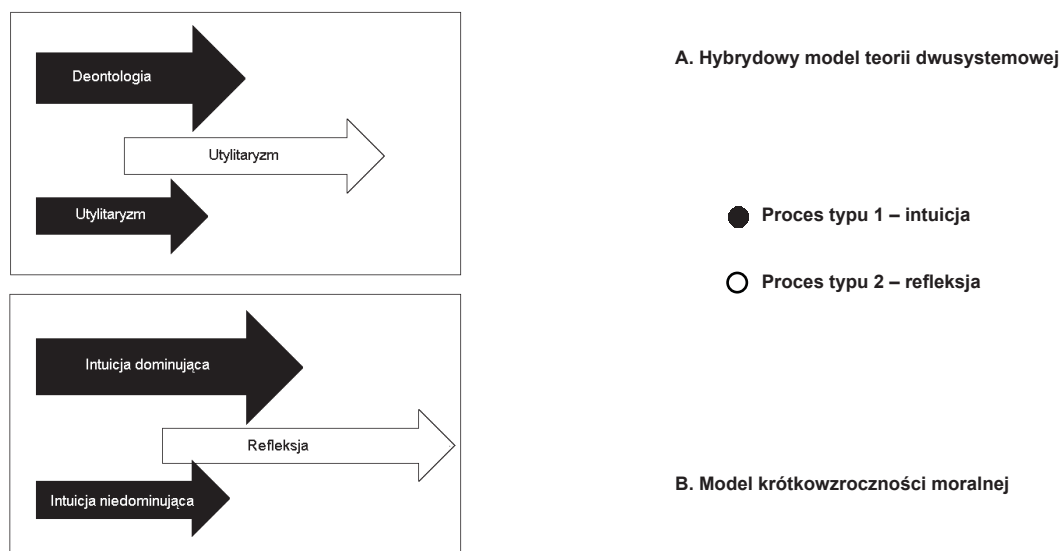
problemie B niż A, a po jej podjęciu są mniej pewni jej słuszności, co sugeruje że są oni mniej pewni słuszności swojej decyzji, gdy nieuzyskana korzyść była duża, niż gdy korzyść ta była mała (Białek i De Neys, 2016). Deontologowie, mimo że nie kierują się nimi w ostatecznym wyborze, to jednak biorą pod uwagę konsekwencje swojego wyboru. Skoro zatem decyzja o odmowie dokonania poświęcenia mimo znacznych korzyści dla innych – która w świetle modelu Greena miałyby być intuicyjna – nosi ślady konfliktu z inklinacją utilitarystyczną. Sugeruje to, że szeregowy model sądów moralnych jest mało prawdopodobny. Jak bowiem później uruchamiana refleksja miałyby wpływać na deontologiczną intuicję?

Gdyby tak było, że niektórzy ludzie refleksyjnie analizują argumenty za i przeciw dokonaniu poświęcenia, i mimo to decydują się pozostać przy deontologii, to utrudnienie im refleksyjnej analizy problemu powinno osłabić siłę wewnętrznego konfliktu. Wynikałoby to z tego, że źródło konfliktowej inklinacji utilitarystycznej byłoby stłumione, a co za tym idzie inklinacja deontologiczna nie napotykałaby silnej kontr-inklinacji. Podążając tym tropem, w serii badań sprawdzaliśmy, co się stanie, gdy badani podczas rozwiązywania problemu będą musieli utrzymywać w pamięci większą lub mniejszą ilość informacji (Białek i De Neys, 2017). Wywołane w ten sposób obciążenie poznawcze powoduje, że nie mają oni dość zasobów, by refleksyjnie analizować szczegóły problemu (dokładną analizę metody obciążenia poznawczego można znaleźć w De Neys, 2006, 2012; De Neys i Glumicic, 2008; De Neys, Rossi, i Houdé, 2013). Oczekujemy zatem, że jeżeli utilitaryzm ma swoje źródło w refleksji, to w tej sytuacji badani dokonają mniej wyborów utilitarystycznych (co już zostało wykazane, por. Greene, Morelli, Lowenberg, Nystrom, i Cohen, 2008; Trémolière i Bonnefon, 2014), ale także ci, którzy dokonali wyboru deontologicznego, będą odczuwali mniejszy wewnętrzny konflikt.

Wyniki pokazały, że o ile obciążenie poznawcze faktycznie zmniejszyło liczbę wyborów utilitarystycznych, to subiektywna różnica w pewności siebie badanych pozostała niezmienną. Niezależnie od dostępności zasobów poznawczych – niezbędnych by realizować procesy typu 2, ale zbędnych dla procesów typu 1 – badani doświadczały jednakowo silnego konfliktu wewnętrznego. Podczas oceny opcji utilitarystycznej o dużej korzyści deontologowie byli mniej pewni podjętej decyzji (badanie 1), czy oceny postępowania innej osoby (badanie 2), niż podczas oceniania opcji utilitarystycznej o małej korzyści. Dodatkowo deklarowali, że czuli silniejszy konflikt wewnętrzny oraz więcej myśleli nad odmiennym niż wybrane rozwiązaniem (badanie 3). Różnice w sile wewnętrznego konfliktu (między zadaniami o dużych i małych korzyściach) pozostały bez zmian wraz ze wzrastającym obciążeniem poznawczym. W ten sposób pokazaliśmy, że każdy decydent doświadcza konfliktu pomiędzy obiema inklinacjami moralnymi, oraz że konflikt ten ma charakter automatyczny i jego zaistnienie nie wymaga dostępu do zasobów poznawczych. Deontologia, jak i utilitaryzm mają zatem charakter procesów

typu 1. Refleksja jednak także przyczynia się do wzrostu utylitarnej inklinacji, gdyż zajęcie zasobów poznawczych zmniejsza szanse, że osoba dokona poświęcenia jednostki dla dobra większości.

Przytoczone dane wymagają rewizji modelu Greena, żeby uwzględnić zarówno równoległe przetwarzanie inklinacji moralnych na poziomie intuicyjnym, jak i refleksję, która promuje utylitaryzm. Wraz z Wimem De Neyssem zaproponowaliśmy hybrydowy model dwusystemowy (Białek i De Neys, 2017; De Neys i Białek, 2017). **Rycina 3** przedstawia strukturę tego modelu. Zakłada się w nim, że kiedy jednostka zda sobie sprawę z tego, iż stoi przed dylematem moralnym, to uruchamiają się w niej dwie intuicje: utylitarystyczna i deontologiczna. Obie te intuicje konkurują o kontrolę nad decyzją, a siła tych intuicji zależy od preferencji decydenta oraz charakteru zadania.



**Ryc 3. Propozycje hybrydowego modelu teorii dwusystemowej oraz modelu krótkowzroczności moralnej.** Refleksja, podobnie jak w innych modelach szeregowych uruchamia się, gdy procesy intuicyjne nie dają satysfakcjonującego rozwiązania (np. deontologia i utylitaryzm są o podobnej sile i tworzą silny wewnętrzny konflikt). Wielkość strzałek odpowiada sile intuicji, a ich długość i miejsce, gdzie się zaczynają oraz kończą – charakterystyce czasowej

O wpływie abstrakcyjnych przekonań moralnych na jakość i rodzaj intuicji świadczą badania wskazujące na przykład, że psychopaci zdają się nie odczuwać odrazy związanej z wyrządzeniem krzywdy innym (Koenigs, Kruepke, Zeier, i Newman, 2012; Patil, 2015; Young, Koenigs, Kruepke, i Newman, 2012), tak samo jak z resztą aleksytymicy (Patil i Silani, 2014a, 2014b) czy osoby będące pod wpływem alkoholu (Duke i Bègue, 2015).

Także charakter zadania kształtuje decyzje moralne. Dokonuje się to poprzez większą lub mniejszą dostępność myśli związanych z robieniem krzywdy innym: łatwiej

jest podjąć decyzję utylitarystyczną, gdy w celu ocalenia kilku osób, poświęcenia dokonujemy pośrednio (np. przez pociągnięcie dźwigni), niż gdy musimy działać bezpośrednio (np. popychając ciężką osobę pod koła wagonika). Podobnie, duża liczba osób, które można ocalić powoduje, że bardziej dostępne są intuicje utylitarystyczne (Trémolière i Bonnefon, 2014). Zwróćmy uwagę, jak takie ujęcie rozwiązuje zarówno problemy szeregowego, jak i równoległego modelu teorii dwusystemowej: 1) obie inklinacje są intuicyjnie przetwarzane, więc zarzut o marnowaniu zasobów poznawczych odnoszący się do modelu równoległego nie ma zastosowania; 2) utylitaryzm ma swoje źródło w intuicji, zatem nie dotyczy go zarzut wobec modelu szeregowego o nieuwzględnianiu dostępnych intuicyjnie rozwiązań utylitarystycznych.

#### MODEL MORALNEJ KRÓTKOWZROczNOŚCI (MORAL MYOPIA MODEL)

Równoległe do naszego modelu hybrydowego rozwijał się alternatywny, niezwykle podobny model – moralnej krótkowzroczności. Model moralnej krótkowzroczności stworzony został przez Amerykanów: Justina Landyego (Landy i Royzman, 2017) z University of Chicago oraz Eda Royzmana z University of Pennsylvania. Zakłada on, że osoby, które są bardziej skłonne do refleksji, częściej myślą zarówno o deontologii, jak i o utylitaryzmie. A dokładniej przyjmuje się, że osoby, które silniej pole-

gają na swojej intuicji, bazują na jednej z dwu inklinacji moralnych: czy to deontologii, czy to utylityzmu. Źródłem inspiracji autorów jest obserwacja, że osoby, które mają ograniczone możliwości myślenia (np. mają zajęty umysł innymi myślami lub są pod wpływem alkoholu), mają w efekcie zawężoną uwagę, i koncentrują się na najsilniejszej stymulacji ze środowiska, ignorując te mniej silne (Mann i Ward, 2007). Tak jak nałogowi alkoholicy koncentrują się na silniejszej pokusie, niż na mniej wyrazistej myśli o możliwych szkodach picia, tak ludzie dokonujący ocen moralnych ulegają najbardziej wyrazistej ocenie, a nie analizują wszystkich dostępnych argumentów.

Podobnie jak wcześniej robił to Kahane, tak Landy i Royzman twierdzą, że dane uzyskane w badaniach z zastosowaniem dylematów należy uznać za trudne do uogólniania na całą domenę moralnego poznania (moral cognition). W dylematach mamy do czynienia z dominującą intuicją deontologiczną, podczas gdy w innych dylematach intuicyjnie przemawia do nas wymiar utylitystyczny, a deontologiczna zasada wymaga przemyślenia (np. gdy trzeba by donieść przyjacielowi o zdradzie jego żony). W tym kontekście istotnym elementem postulowanego modelu jest odwołanie się do wcześniej omawianych prac wspierających model hybrydowy, i sugerujących, że każda inklinacja moralna jest niezależną intuicją. Dzięki uwzględnieniu takich założeń model krótkowzroczności moralnej może przewidywać, że osoba bazująca na procesach typu 1 może być zarówno deontologiem, jak i utylitystą.

Ciekawym dowodem na słuszność modelu Landyego i Royzmana są ich wcześniejsze badania (Royzman, Landy, i Goodwin, 2014), w których ludzie oceniali naruszenie szeregu zachowań, gdy 1) poinformowani zostali, że zachowanie to jest powszechnie akceptowane w danym społeczeństwie – na przykład w pewnym kraju przyjętym jest, że rodzeństwo okazuje sobie czułość namiętnie się całując, a ludzie przychodzą w piżamie do pracy, oraz 2) jawnie niemoralnych czynów, takich jak kradzież czy oszustwo. Związek między refleksyjnością a ocenami moralnymi zaobserwowano tylko w tej pierwszej grupie problemów. Co z tego wynika? Ano to, że refleksyjne osoby umieją lepiej niż osoby niskorefleksyjne odróżnić naruszenie norm moralnych od naruszenia konwencji społecznych. Jaka bowiem jest szkoda, którą wyrządza sobie bądź innym pracownik w piżamie? Rozszerzając te wyniki Landy (2016) pokazuje, że niskorefleksyjni badani są bardziej skłonni potępiać złamanie tych z fundamentalnych zasad moralnych (podział za: Haidt i Graham, 2007), które opierają się głównie na konwencji społecznej (autorytet, lojalność, świętość), ale nie różnią się od osób refleksyjnych w surowości ocen, gdy chodzi o naruszenie fundamentów, takich jak troska i uczciwość; grupa fundamentów konstytuuje moralność w odezwaniu od konwencji społecznych.

Model krótkowzroczności moralnej jest więc w jednym zakresie różny od modelu zaproponowanego przez mnie i De Neysa – mianowicie refleksja w tym modelu ma za

zadanie uwypuklać mniej wyrazista intuicję, i w efekcie zwiększyć udział obu w podejmowanej decyzji. Osoby mało refleksyjne będą dla odmiany polegały jedynie na silniejszej intuicji, i stosownie dokonywać ocen moralnych. Bliżej im w tym modelu do postulatów Guya Kahane, który w refleksji widział siłę tworzącą kontrintuicyjne sądy moralne: jeśli intuicja była utylitystyczna, to refleksja deontologiczna, jeśli z kolei intuicja była deontologiczna, to refleksja utylitystyczna (Kahane, 2012, 2015; Kahane i in., 2012).

## PODSUMOWANIE

Omówione trzy modele myślenia moralnego niewątpliwie wymagają dalszych badań i analiz. Skoro bowiem na „wejściu” mamy dostępne obie intuicje moralne, to wiemy, że nie generuje ona nowej opcji rozwiązania dylematu, a jedynie pozwala wybrać tę z nich, która wydaje się lepiej odpowiadać uniwersalnym przekonaniom decydenta oraz kontekstowi problemu. Na chwile obecną stanowisko, które mi wydaje się lepiej osadzone w danych, jest takie, że refleksja pozwala lepiej oszacować zyski i straty związane z konkretnymi ocenami moralnymi, a zatem promuje utylityzm. Brak definitywnych dowodów pozwalających rozstrzygnąć, który model lepiej oddaje rzeczywistość: czy model hybrydowy czy moralnej krótkowzroczności.

Myślę, że dzięki temu tekstowi uważnemu Czytelnikowi łatwiej będzie wejść w tematykę sądów moralnych; szczególnie Czytelnikowi, który dopiero buduje swoją wiedzę w tej dziedzinie. Jak widać choćby z rozległości cytowanej literatury – jest to „gorący” temat, z wieloma różnymi opiniami i świeżymi modelami, a który domaga się aż nowych ujęć teoretycznych integrujących dostępne dane empiryczne. Z całych sił zachęcam Czytelników tego tekstu do podjęcia takich wysiłków.

## LITERATURA

- Bago, B., De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition*, 158, 90-109.
- Baron, J. (1993). Why teach thinking? –An essay. *Applied Psychology*, 42, 191-214.
- Baron, J., Gürçay, B., Moore, A. B., Starcke, K. (2012). Use of a Rasch model to predict response times to utilitarian moral dilemmas. *Synthese*, 189, 107-117.
- Baron, J., Scott, S., Fincher, K., Emlen Metz, S. (2015). Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition*, 4, 265-284.
- Bartels, D.M. (2008). Principled moral sentiment and the flexibility of moral judgment and decision making. *Cognition*, 108, 381-417.
- Bauman, C.W., McGraw, A.P., Bartels, D.M., Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass*, 8, 536-554.



- Bedford, F.L. (1997). False categories in cognition: The not-the-liver fallacy. *Cognition*, 64, 231-248.
- Bialek, M. (2015a). Jonathan Haidt, Prawy Umysł – Recenzja i dyskusja. *Psychologia Ekonomiczna*, 8, 77-79.
- Bialek, M. (2015b). Przegląd badań współczesnej kognitywistyki nad efektem przekonania. *Przegląd Filozoficzny. Nowa Seria.*, 95, 91-107.
- Bialek, M., De Neys, W. (2016). Conflict detection during moral decision-making: evidence for deontic reasoners' utilitarian sensitivity. *Journal of Cognitive Psychology*, 28, 631-639.
- Bialek, M., De Neys, W. (2017). Dual processes and moral conflict: evidence for deontological reasoners' intuitive utilitarian sensitivity. *Judgment and Decision Making*, 12, 148-167.
- Bialek, M., Domurat, A. (2017). Cognitive abilities, analytic cognitive style and overconfidence: a commentary on Duttler (2015). *Bulletin of Economic Research*, w druku.
- Bialek, M., Terbeck, S. (2016). Can cognitive psychological research on reasoning enhance the discussion around moral judgments? *Cognitive Processing*, 17, 329-335.
- Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108, 624.
- Cipolletti, H., McFarlane, S., Weissglass, C. (2015). The moral foreign-language effect. *Philosophical Psychology*, 29, 1-18.
- Conway, P., Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology*, 104, 216-235.
- Corey, J.D., Costa, A. (2015). The foreign language effect on moral decisions. *Ciencia Cognitiva*, 9, 3-57.
- De Neys, W. (2006). Dual processing in reasoning two systems but one reasoner. *Psychological Science*, 17, 428-433.
- De Neys, W. (2012). Bias and conflict a case for logical intuitions. *Perspectives on Psychological Science*, 7, 28-38.
- De Neys, W., Bialek, M. (2017). Intuition and reflection in moral judgment: A case for utilitarian intuitions? W: J.-F. Bonnefon, B. Trémolière (red.), *Moral inferences* (s. 123-136). Hove: Psychology Press.
- De Neys, W., Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, 106, 1248-1299.
- De Neys, W., Rossi, S., Houdé, O. (2013). Bats, balls, and substitution sensitivity: cognitive misers are no happy fools. *Psychonomic Bulletin & Review*, 20, 269-273.
- Duke, A.A., Bègue, L. (2015). The drunk utilitarian: Blood alcohol concentration predicts utilitarian responses in moral dilemmas. *Cognition*, 134, 121-127.
- Eskine, K.J., Kacinik, N.A., Prinz, J.J. (2011). A bad taste in the mouth gustatory disgust influences moral judgment. *Psychological Science*, 22, 295-299.
- Evans, J.S.B. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning*, 13, 321-339.
- Evans, J.S.B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Rev. Psychol.*, 59, 255-278.
- Evans, J.S.B., Frankish, K.E. (2009). In two minds: Dual processes and beyond. New York: Oxford University Press.
- Evans, J.S.B., Stanovich, K.E. (2013). Dual-process theories of higher cognition advancing the debate. *Perspectives on Psychological Science*, 8, 223-241.
- Fiske, S.T., Taylor, S.E. (2013). *Social cognition: From brains to culture*. Los Angeles: Sage.
- Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review*, 5, 1-7.
- Foot, P. (1978). The problem of abortion and negative and positive duty: A reply to James LeRoy Smith. *Journal of Medicine and Philosophy*, 3, 253-255.
- Franssens, S., De Neys, W. (2009). The effortless nature of conflict detection during thinking. *Thinking & Reasoning*, 15, 105-128.
- Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, 19, 25-42.
- Friesdorf, R., Conway, P., Gawronski, B. (2015). Gender differences in responses to moral dilemmas: A process dissociation analysis. *Personality and Social Psychology Bulletin*, 41, 696-713.
- Gawinkowska, M., Paradowski, M.B., Bilewicz, M. (2013). Second language as an exemptor from sociocultural norms. Emotion-related language choice revisited. *PloS one*, 8, e81225.
- Geipel, J., Hadjichristidis, C., Surian, L. (2015a). The foreign language effect on moral judgment: The role of emotions and norms. *PloS one*, 10, e0131529.
- Geipel, J., Hadjichristidis, C., Surian, L. (2015b). How foreign language shapes moral judgment. *Journal of Experimental Social Psychology*, 59, 8-17.
- Greene, J.D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends in Cognitive Sciences*, 11, 322-323.
- Greene, J.D. (2014). *Moral tribes: emotion, reason and the gap between us and them*. Bloomsbury: Atlantic Books Ltd.
- Greene, J.D. (2015a). Beyond point-and-shoot morality: Why cognitive (neuro) science matters for ethics. *The Law & Ethics of Human Rights*, 9, 141-172.
- Greene, J.D. (2015b). The rise of moral cognition. *Cognition*, 135, 39-42.
- Greene, J.D., Morelli, S.A., Lowenberg, K., Nystrom, L.E., Cohen, J.D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, 107, 1144-1154.
- Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., Cohen, J.D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44, 389-400.
- Gürçay, B., Baron, J. (2017). Challenges for the sequential two-system model of moral judgement. *Thinking & Reasoning*, 23, 49-80.
- Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, 108, 814-834.
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316, 998-1002.
- Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York: Pantheon.
- Haidt, J., Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, 20, 98-116.
- Haidt, J., Koller, S.H., Dias, M.G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, 65, 613.

- Haman, M. (2008). Racjonalne powody akceptacji koncepcji pseudonaukowych: dążenie do prostoty, społeczny podział pracy w nauce i fałszywe dychotomie pojęciowe. *Psychologia Społeczna*, 3, 336-339.
- Handley, S.J., Trippas, D. (2015). Dual processes and the interplay between knowledge and structure: A new parallel processing model. W: B. Ross (red.), *Psychology of learning and motivation* (s. 33-58). Burlington: Academic Press.
- Hardy, S.A., Carlo, G. (2011). Moral identity: What is it, how does it develop, and is it linked to moral action? *Child Development Perspectives*, 5, 212-218.
- Hayakawa, S., Costa, A., Foucart, A., Keysar, B. (2016). Using a foreign language changes our choices. *Trends in Cognitive Sciences*, 20, 791-793.
- Howarth, S., Handley, S., Walsh, C. (2016). The logic-bias effect: The role of effortful processing in the resolution of belief–logic conflict. *Memory & Cognition*, 44, 330-349.
- Jost, J.T. (2012). Left and right, right and wrong. *Science*, 337, 525-526.
- Kahane, G. (2012). On the wrong track: Process and content in moral psychology. *Mind & Language*, 27, 519-545.
- Kahane, G. (2013). The armchair and the trolley: an argument for experimental ethics. *Philosophical Studies*, 162, 421-445.
- Kahane, G. (2015). Sidetracked by trolleys: Why sacrificial moral dilemmas tell us little (or nothing) about utilitarian judgment. *Social Neuroscience*, 10, 551-560.
- Kahane, G., Wiech, K., Shackel, N., Farias, M., Savulescu, J., Tracey, I. (2012). The neural basis of intuitive and counterintuitive moral judgment. *Social Cognitive and Affective Neuroscience*, 7, 393-402.
- Kahneman, D. (2011). *Thinking fast and slow*. New York: Farrar, Strauss and Giroux.
- Knobe, J., Buckwalter, W., Nichols, S., Robbins, P., Sarkisian, H., Sommers, T. (2012). Experimental philosophy. *Annual Review of Psychology*, 63, 81-99.
- Knobe, J., Nichols, S. (2008). An experimental philosophy manifesto. *Experimental Philosophy*, 1, 3-14.
- Koenigs, M., Kruepke, M., Zeier, J., Newman, J.P. (2012). Utilitarian moral judgment in psychopathy. *Social Cognitive and Affective Neuroscience*, 7, 708-714.
- Kohlberg, L. (1981). *The philosophy of moral development moral stages and the idea of justice*. San Francisco: Harper & Row.
- Lammers, J., Galinsky, A.D., Dubois, D., Rucker, D.D. (2015). Power and morality. *Current Opinion in Psychology*, 6, 15-19.
- Lammers, J., Stapel, D.A. (2009). How power influences moral thinking. *Journal of Personality and Social Psychology*, 97, 279.
- Landy, J.F. (2016). Representations of moral violations: Category members and associated features. *Judgment and Decision Making*, 11, 496-508.
- Landy, J.F., Rozman, E. (2017). The moral myopia model: Why and how reasoning matters in moral judgment. W: G. Pennycook (red.), *The new reflectionism in cognitive psychology: Why reason matters* (w druku). New York: Psychology Press.
- Mann, T., Ward, A. (2007). Attention, self-control, and health behaviors. *Current Directions in Psychological Science*, 16, 280-283.
- Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., Grafman, J. (2005). The neural basis of human moral cognition. *Nature Reviews Neuroscience*, 6, 799-809.
- Newstead, S.E., Pollard, P., Evans, J.S.B., Allen, J.L. (1992). The source of belief bias effects in syllogistic reasoning. *Cognition*, 45, 257-284.
- Noori, M. (2016). Cognitive reflection as a predictor of susceptibility to behavioral anomalies. *Judgment and Decision Making*, 11, 114-120.
- Osman, M., Wiegmann, A. (2017). Explaining moral behavior: A minimal moral model. *Experimental Psychology*, 64, 68-81.
- Patil, I. (2015). Trait psychopathy and utilitarian moral judgement: The mediating role of action aversion. *Journal of Cognitive Psychology*, 27, 349-366.
- Patil, I., Silani, G. (2014a). Alexithymia increases moral acceptability of accidental harms. *Journal of Cognitive Psychology*, 26, 597-614.
- Patil, I., Silani, G. (2014b). Reduced empathic concern leads to utilitarian moral judgments in trait alexithymia. *Frontiers in Psychology*, 5, 501.
- Paxton, J.M., Ungar, L., Greene, J.D. (2012). Reflection and reasoning in moral judgment. *Cognitive Science*, 36, 163-177.
- Parzuchowski, M., Bocian, K., Wojciszke, B. (2016). Od skrajności do codzienności: współczesna psychologia ocen moralnych. *Psychologia Społeczna*, 39, 1-12.
- Pennycook, G., Cheyne, J.A., Barr, N., Koehler, D.J., Fugelsang, J.A. (2014). The role of analytic thinking in moral judgements and values. *Thinking & Reasoning*, 20, 188-214.
- Pennycook, G., Fugelsang, J.A., Koehler, D.J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, 80, 34-72.
- Piaget, J. (1997). *The moral judgement of the child*. New York: Simon and Schuster.
- Prinz, J. (2006). The emotional basis of moral judgments. *Philosophical Explorations*, 9, 29-43.
- Rozman, E.B., Baron, J. (2002). The preference for indirect harm. *Social Justice Research*, 15, 165-184.
- Rozman, E.B., Landy, J.F., Goodwin, G.P. (2014). Are good reasoners more incest-friendly? Trait cognitive reflection predicts selective moralization in a sample of American adults. *Judgment and Decision Making*, 9, 175-190.
- Rozman, E.B., Landy, J.F., Leeman, R.F. (2015). Are thoughtful people more utilitarian? CRT as a unique predictor of moral minimalism in the dilemmatic context. *Cognitive Science*, 39, 325-352.
- Stanovich, K.E. (2009). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory. W: J.S.B. Evans, K.E. Frankish (red.), *In two minds: Dual processes and beyond* (s. 55-88). New York: Oxford University Press.
- Stuppel, E.J., Ball, L.J. (2008). Belief–logic conflict resolution in syllogistic reasoning: Inspection-time evidence for a parallel-process model. *Thinking & Reasoning*, 14, 168-181.
- Suter, R.S., Hertwig, R. (2011). Time and moral judgment. *Cognition*, 119, 454-458.
- Thompson, V.A., Johnson, S.C. (2014). Conflict, metacognition, and analytic thinking. *Thinking & Reasoning*, 20, 215-244.

- Thompson, V.A., Turner, J.A.P., Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, 63, 107-140.
- Thompson, V.A., Turner, J.A.P., Pennycook, G., Ball, L.J., Brack, H., Ophir, Y., i in. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition*, 128, 237-251.
- Tomaszewski, T. (1975). *Psychologia*. Warszawa: PWN.
- Toplak, M.E., West, R.F., Stanovich, K.E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, 39, 1275-1289.
- Trémolière, B., Bonnefon, J.-F. (2014). Efficient kill-save ratios ease up the cognitive demands on counterintuitive moral utilitarianism. *Personality and Social Psychology Bulletin*, 40, 923-930.
- Tremoliere, B., De Neys, W., Bonnefon, J.-F. (2016). Reasoning and moral judgment: A common experimental toolbox. W: L.J. Ball, V.A. Thompson (red.), *International handbook of thinking and reasoning* (w druku). Hove: Psychology Press.
- Trippas, D., Thompson, V.A., Handley, S.J. (2017). When fast logic meets slow belief: Evidence for a parallel-processing model of belief bias. *Memory & Cognition*, 45, 539-552.
- Turiel, E. (1998). The development of morality. W: N. Eisenberg (red.), *Handbook of Child Psychology, Vol. 3 Social, emotional, personality, and development* (s. 863-932). Cambridge: Cambridge University Press.
- Waldmann, M.R., Dieterich, J.H. (2007). Throwing a bomb on a person versus throwing a person on a bomb intervention myopia in moral intuitions. *Psychological Science*, 18, 247-253.
- Wheatley, T., Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, 16, 780-784.
- Wojciszke, B. (2011). Psychologia władzy. *Nauka*, 2, 51-69.
- Young, L., Koenigs, M., Kruepke, M., Newman, J.P. (2012). Psychopathy increases perceived moral permissibility of accidents. *Journal of Abnormal Psychology*, 121, 659.
- Żukowski, T. (2015). Autowizerunek po katastrofie. Zofia Kossak-Szczucka i Jerzy Andrzejewski: dwa polskie świadectwa Zagłady z lat 40. *Poznańskie Studia Polonistyczne. Seria Literacka*, 25, 165-186.

