

JANUSZ REYKOWSKI *

Teoria ograniczonej etyczności a wiedza o systemie motywacyjnym człowieka **

Wprowadzenie

Przez dość długi czas w naukach społecznych, a przede wszystkim w ekonomii i socjologii, dużą popularnością cieszyła się w istocie psychologiczna teoria racjonalnego wyboru (*Rational Choice Theory*, Almond 1991). Zgodne z nią ludzie podejmujący decyzje kierują się własnym interesem, a więc racjonalną analizą swoich potencjalnych zysków i strat. Ten sposób myślenia był rozpowszechniony i w innych naukach społecznych. Tak np. w psychologii przybierał on formę koncepcji człowieka jako intuicyjnego ekonomisty, a więc decydenta, którego głównym celem jest maksymalizacja użyteczności. Stara się on dokonywać optymalnych wyborów, czyli jak najmniejszym kosztem osiągać maksymalne korzyści.

Teorię tę zakwestionowali dwaj nobliści Herbert Simon (ekonomista i socjolog) oraz Daniel Kahneman (psycholog). Według nich racjonalność człowieka jest ograniczona przez czynniki sytuacyjne i przez ograniczoną moc obliczeniową ludzkiego umysłu. Dowodzili, że w rzeczywistości racjonalna analiza zastępowana jest przez heurystyki, czyli uproszczone reguły wnioskowania. Była to teoria ograniczonej racjonalności (*Bounded Rationality*), (Kahneman, 2003).

Teoria ograniczonej etyczności

Do teorii ograniczonej racjonalności nawiązali Dolly Chugh, Max Bazerman i Mahzarin Banaji z Uniwersytetu Harvarda, formułując teorię ograniczonej etyczności (*Bounded Ethicality*, (2005)). Badali oni zachowania ludzi w sytuacji konfliktu interesów. Doszli do wniosku, że ludzie posługują się kryteriami moralnymi, gdy abstrakcyjnie analizują jakiś problem moralny lub gdy oceniają zachowania cudze. Kiedy sami znajdują się w sytuacji konfliktu interesów posługują się heurystykami, które automatycznie przyznają słuszność tym wyborom, które są zgodne z ich interesami. Postrzegając sa-

* Prof. dr hab. Janusz Reykowski (januszre@warman.com.pl), członek korespondent PAN, Instytut Psychologii PAN, Warszawa

** Podstawą niniejszego artykułu był wygłoszony przez autora referat na XVII Zjeździe Polskiego Stowarzyszenia Psychologii Społecznej w Gdańsku, 15–18 września 2022 r.

mych siebie jako osobę moralną, kompetentną i zasługującą na korzyści, które powinni uzyskać, nie dostrzegają, czy nie uznają interesów lub perspektyw cudzych. Dla teorii tej można znaleźć potwierdzenia w wielu badaniach psychologicznych. W trakcie tego wykładu chciałbym skonfrontować jej twierdzenia z tym, co wiemy o działaniu systemu motywacyjnego człowieka.

Jako punkt wyjścia mojej analizy przedstawię trzy różne programy badawcze wykonane przez trzy zespoły – programy te rzucają różne światło na naturę ludzkiej motywacji.

Program 1: Badania zespołu Bogdana Wojciszki nad egocentrycznymi zniekształceniami ocen moralnych

Bogdan Wojciszke w swoich licznych publikacjach przedstawił przekonujące dowody, że ludzie dokonując oceny innych, a także siebie, posługują się głównie dwoma podstawowymi wymiarami – wymiarem kompetencji i wymiarem moralności, przy czym, jak się okazuje, wymiar moralności ma kluczowe znaczenie przy formułowaniu oceny drugiego człowieka (Wojciszke, 2010).

W zespole Wojciszki podjęto wiele badań dotyczących ocen moralności innych ludzi. Stwierdzono m.in. że:

- Oceny moralności zwierzchników przez badanych pracowników banku były wyższe, jeśli działania zwierzchników były korzystne dla pracowników (Wojciszke i Baryła, 2006).
- Badani emeryci postrzegali polityka opisanego jako popierającego politykę nastawioną na interesy emerytów jako moralnego, a przeciwnika tej polityki – jako niemoralnego (Cisłak i Wojciszke, 2006).
- Badani studenci oceniali sytuacje naruszania zasad sprawiedliwości w interesie określonej osoby (np. przyznanie stypendium słabemu studentowi wbrew obowiązującemu kryterium) jako niemoralne, ale studenci, którzy mieli wyobrazić sobie, że to oni korzystają na naruszeniu zasad oceniali ten fakt jako moralnie neutralny (Cisłak i Wojciszke, 2006).

Badania te nasuwają wniosek, że na ocenę moralną wpływać może własny interes oceniającego. Wniosek ten Bocian i Wojciszke potwierdzili eksperymentalnie. Bocian i Wojciszke (2014) przeprowadzili badania, w których osoba badana dokonywała oceny moralnej zachowania pomocnika eksperymentatorów:

- Bibliotekarki, która wbrew zasadom obowiązującym w bibliotece zwalniała studentkę z karnej opłaty za spóźniony zwrot książki (Badanie 1), albo
- Uczestnika badania, który oszukiwał przy wykonywaniu zadania eksperymentalnego, aby dzięki temu uzyskać zapowiadaną nagrodę finansową – w połowie przypadków nagrodą dzielił się z osobą badaną (Badanie 2).

Stwierdzono, że zachowania pomocników eksperymentatora, którzy postępowali wbrew obowiązującym normom oceniane były jako bardziej moralne, kiedy osoba oceniająca osiągała z nich korzyść. Okazało się też, że wywołują oni pozytywne uczucia – lubienie. Uczucia te odgrywały istotną (pośredniczącą) rolę przy dokonywaniu ocen moralności. Wszakże Bocian, Baryła i Wojciszke (2021) stwierdzili, że oceny te mogą ulegać zmianie, jeśli osobom zapowiadano, że będą musiały uzasadnić te oceny, co będzie rejestrowane, a następnie będą musiały je podpisać. Okazało się, że procedura ta redukowała egocentryczne zniekształcenia ocen moralnych. Tak więc opisane badania:

- Zademonstrowały zjawisko systematycznego egocentrycznego zniekształcenia ocen moralnych – jako bardziej moralne oceniane są te osoby, które działają na korzyść oceniającego.
- Pokazały, że oceny moralności osób, których działania nie mają wpływu na dobrostan oceniającego, dokonane są zgodnie ze społecznie przyjętymi kryteriami moralnego zachowania.
- Wykazały, że ten, kto działa na korzyść osoby oceniającej, wywołuje u niej pozytywną reakcję w formie reakcji „lubienia” (efekt ten ma pewne ograniczenia). Reakcja ta „pośredniczy” w powstawaniu pozytywnych ocen zachowania danej osoby.

Mechanizm: doświadczenie korzyści – pozytywna reakcja (lubienie) – pozytywna ocena działa automatycznie bez pośrednictwa świadomej refleksji. Uaktywnienie refleksji hamuje mechanizm egocentrycznego zniekształcenia.

Przedstawiony układ wyników jest w zasadzie zgodny z teorią ograniczonej etyczności.

W tekście podsumowującym różne własne i opisywane w literaturze psychologicznej badania dotyczące wpływu egocentryzmu na oceny moralne Bocian, Baryła i Wojciszke (2020) stwierdzają: oceny egocentryczne są szybkie, automatyczne, afektywne i motywowane strategicznie. Ponieważ nie wymagają wysiłku i zasobów, mogą służyć jako *defaultowa* podstawa sądów moralnych. Tak więc egocentryczne oceny subiektywnie wydają się obiektywną i wierną percepcją społecznego świata

Program 2: Badania zespołu Dariusza Dolińskiego dotyczące posłuszeństwa wobec autorytetu

Program ten zainspirowany przez słynne badania Stanleya Milgrama (1963) został przedstawiony w monografii pt. *Posłuszni do bólu* (2017), której autorami są profesoro- wie Dariusz Doliński i Tomasz Grzyb.

Celem badań było poszukiwanie „psychologicznych mechanizmów leżących u podłoża posłuszeństwa” (str. 77). Chodziło tu o szczególną jego formę – o posłuszne wykonywanie działań, które zadają ból innemu człowiekowi – „w służbie nauki”.

Schemat badań był zmodyfikowaną (łagodniejszą) wersją schematu Milgrama. Osoba badana na polecenie eksperymentatora musiała „karać” uderzeniami prądu elektrycznego o wzrastającym napięciu siedzącego w sąsiednim pokoju „ucznia” (pomocnika eksperymentatora). Uczeń, wykonując kolejne zadania testu pamięciowego, zgodnie ze scenariuszem „mylił się”, za co „nauczyciel” (osoba badana) miał go na polecenie Eksperymentatora „karać” impulsami elektrycznymi o wzrastającej sile. Wprawdzie przyciski aparatu, który obsługiwała badana osoba nie wysyłały elektrycznych impulsów, ale „uczeń” zachowywał się tak, jakby doświadczał coraz większych cierpień.

Kryterium pełnego posłuszeństwa było wykonanie wszystkich poleceń eksperymentatora, tj. użycie wszystkich 10 przycisków – od najsłabszego do najsilniejszego. Część osób badanych wyrażała niechęć do użycia przycisków wywołujących skargi i „objawy bólu” „ucznia”, próbując przerwać badanie. W tych wypadkach eksperymentator zgodnie ze scenariuszem wypowiadał określone „ponaglenia” mające skłonić badaną osobę do ich kontynuacji. Na jej stanowcze żądanie eksperyment przerywano, a następnie udzielano jej szczegółowych wyjaśnień o tym, co się tu naprawdę działo i po co.

W kolejnych seriach badań autorzy sprawdzali sformułowane przez siebie hipotezy dotyczące czynnika (czy czynników), które mogły mieć wpływ na poziom posłuszeństwa osób badanych. Innymi słowy badanie to dotyczyło motywacji do wykonywania tego rodzaju zadań. Sprawdzane hipotezy:

H1. Efekt posłuszeństwa był rezultatem zastosowania techniki wpływu społecznego określonej jako „stopa w drzwiach; – uzyskanie zgody na wykonane łatwego zadania (wysyłanie najpierw słabego bodźca elektrycznego) mogło ułatwiać akceptację zadania trudniejszego – byłaby to więc „wielokrotna stopa w drzwiach”. Sprawdzano, czy zmiana warunków – rozpoczęcie od bardzo silnego bodźca – zniweczy efekt posłuszeństwa.

Porównano dwie grupy: K, w której siła bodźca zwiększana była stopniowo i E, w której rozpoczynano od bardzo silnego bodźca. Wynik był zaskakujący: Ukończyło zadanie w grupie K 80% a w grupie E 100%.

H2. Posłuszeństwo może się zmniejszyć, jeśli osoba badana może spodziewać się odwetu ze strony karanej osoby, a więc we własnym interesie powinna sprzeciwić się poleceniom zadawania silnego bólu. W tym celu poinformowano osobę badaną („nauczyciela”) i „ucznia”, że w drugiej części badania zamienią się rolami. Wynik: badanie ukończyło 85% grupy (oczywiście żadnej zamiany nie było).

H3. Na poziom posłuszeństwa może mieć wpływ płeć osoby, która pełniła rolę ucznia. Wynik: w grupie o równej proporcji płci liczącej 80 zbadanych osób 72 osoby (90%) doszło do ostatniego przycisku. Tak więc statystycznie znaczących różnic związanych z płcią „ucznia” nie udało się zaobserwować.

H4. Posłuszeństwo mogło być wynikiem poczucia zobowiązania wynikającego z faktu, że osoby badane otrzymywały z góry 50 złotych za udział w badaniu. Wprowadzono zasadę płacenia z dołu. Wynik: badanie ukończyło w grupie K (standardowej) 84%, a w grupie E (opłacanej z dołu) – 87,5%.

Opisany program zawierał inne jeszcze elementy, które tu pominię, ponieważ nie mają one wpływu na końcową konkluzję autorów.

Autorzy z pewnym żalem konstatują, że nie udało im się potwierdzić żadnej ze sformułowanych hipotez. Uznali więc, że najlepsze wyjaśnienie otrzymanych wyników sformułował sam Milgram, wskazując na „stan pośrednictwa w działaniu” jako „podstawowy mechanizm tłumaczący zachowania badanych osób” (s. 140).

Sądzę jednak, że przedstawiony przez Dolińskiego i Grzyba program badawczy i jego rezultaty nasuwają pewne ważne ustalenia. Zwrócę uwagę na trzy z nich.

Po pierwsze, uderzający jest fakt, że wprowadzone modyfikacje warunków badania nie wywoływały żadnych zmian dominującego wzoru reakcji osób badanych. Modyfikacje odnosiły się:

- do czynników wpływu społecznego,
- do pewnej formy własnego interesu (zabezpieczenia przed odwetem),
- do cech osoby, która jest obiektem oddziaływania (płeć),
- do interpersonalnej relacji zobowiązania.

Może to świadczyć, że reprezentacja sytuacji badania w umysłach większości osób badanych była wysoce „zglobalizowana” i być może niewielka część jej ważnych szczegółów docierała do ich świadomości.

Po drugie, jak konstatują Doliński i Grzyb, w ich badaniach, w których wzięło udział 400 osób, 85-90% członków badanych grup wykonało wszystkie polecenia eksperymentatora (str. 177). Był to wynik wyraźnie wyższy, niż uzyskany przez Milgrama i w wielu innych powtórzeniach jego eksperymentu (ok. 65% grup). Ale w badaniach polskich było 10 przycisków – ostatni z najwyższą oznaczoną wartością 150 V, a w badaniach Milgrama 30 przycisków z maksymalną wartością 450 V. Nasuwa się przypuszczenie, że wielkość motywacji do przeciwstawienia się naciskom „autorytetu” może zależeć od postrzeganej wielkości krzywdy, którą na jego polecenie badany wyrządza innej osobie. Innymi słowy sytuację badawczą należałoby interpretować w kategorii dynamiki dwóch przeciwnych tendencji motywacyjnych.

Po trzecie, uzyskane wyniki wskazują na potrzebę modyfikacji teorii ograniczonej etyczności. Jak się okazuje, źródłem ograniczenia etyczności może nie być własny interes, a zupełnie inny mechanizm. W tym wypadku ograniczeniem etyczności jest uległość wobec autorytetu. Mechanizm ten, o czym będzie dalej mowa, ma zupełnie inne uwarunkowania, niż zakładała teoria ograniczonej etyczności.

Program 3: Autorzy Jane Piliavin, John Dovidio, Samuel Gaertner i Russell Clark III – badania opisane w monografii pt. *Emergency intervention* (1981)

Opisano 49 badań eksperymentalnych – swoich i cudzych, w których aranżowano sytuacje nagłego wypadku i obserwowano reakcje świadków. Były to takie sytuacje, jak:

- upadek w wagonie nowojorskiego metra,
- upadek ze schodów i objawy zemdlenia,
- wypadek pracownika serwisu w sąsiednim pokoju wołającego o pomoc,
- kontuzja nogi kobiety,
- próba gwałtu (na osobie eksperymentatorki),
- przerwanie sieci elektrycznej przez pracownika serwisu – możliwość porażenia prądem,
- przewrócenie się półki z książkami i in.

Badanymi były przypadkowe osoby (np. pasażerowie metra), osoby zaproszone na jakieś inne badanie w laboratorium.

Uwzględniano: fizyczną pozycję świadków (siedzi, stoi, idzie), jednoznaczność sytuacji wypadku, znajomość ofiary (jakaś forma interakcji przed wypadkiem), cechy ofiary (płeć, kolor skóry), stan ofiary przed wypadkiem (np. poruszanie się o lasce), stopień zagrożenia dla osoby udzielającej pomocy itp.

W poszczególnych eksperymentach warunkami takimi systematycznie manipulowano.

Na podstawie tych badań autorzy wyodrębnili i opisali zjawisko, które nazwali **pomocą impulsywną** (*impulsive helping*). Pojawia się ono w określonych sytuacjach i posiada pewne charakterystyczne cechy:

- Jest to reakcja na wypadek, tj. zdarzenie, z którym wiąże się zagrożenie dla jakiejś osoby i zagrożenie to jest oczywiste (potencjalnie) dla każdego jego świadka.
- Miejsce wypadku jest bliskie (bliskość fizyczna).
- Wypadek dotyczy osoby znanej (np. krótki kontakt) lub widzianej np. dotyczy kogoś kto niedawno przechodził w wagonie i czymś zwrócił na siebie uwagę.
- Pomoc podejmuje 85–100% świadków.
- Pomoc jest natychmiastowa, podejmowana po nie więcej niż 15 sekundach.
- Jest to reakcja bez zastanowienia (automatyczna), która nie uwzględnia ważnych cech sytuacji, w tym niebezpieczeństwa dla osoby udzielającej pomocy.

Zjawiska impulsywnej pomocy nie da się wyjaśnić w ramach modelu oczekiwanej użyteczności (*expected utility*). Musi mieć ono motywacyjne uwarunkowania, które wydają się niezgodne z Modelem Ograniczonej Etyczności. Cechą wspólną wszystkich trzech przedstawianych tu sytuacji był w zasadzie bezrefleksyjny tryb działania badanych osób. Działy one niejako automatycznie realizując cele mające jakościowo różne

znaczenia adaptacyjne. Tak więc związane były realizacją interesów badanych osób (w badaniu pierwszym), z realizacją grupowych (w badaniu drugim), z realizacją interesów pozasobistych (w badaniu trzecim). Cechy tych działań wskazują, że były one kierowane przez odrębne systemy motywacyjne, których aktywizacja nie wymagała udziału świadomości. Były to zatem jakieś bardzo podstawowe mechanizmy motywacyjne, można nawet powiedzieć – trzy aspekty ludzkiej natury.

I. Pierwszy z tych systemów motywacyjnych przejawiał się w formie działań wyrażających troskę o własny interes.

Działania tego systemu nastawione są na realizację celów adaptacyjnych jednostki. Służy on przetrwaniu jednostki i zachowaniu jej dobrostanu, co wymaga uzyskiwania odpowiednich dóbr, a także ochrony czy podwyższania społecznego statusu. Integralnym składnikiem tej motywacji bywa troska o bezpieczeństwo i dobrostan osób najbliższych – rodziny. Motywację tego rodzaju można nazwać **ipsocentryczną**.

Wtedy, gdy podmiot realizuje jakiś znaczący dla siebie cel, ignorując inne ważne społeczne lub fizyczne cechy sytuacji, motywację tę można określić jako **egocentryczną** lub egoistyczną.

Centralnym elementem systemu ipsocentrycznej motywacji jest Struktura ja. Ma ona różne aspekty, takie jak: obraz własnej osoby, poczucie tożsamości, „schematy ja” (*self-schema*) samoocena, poczucie własnej wartości.

Struktura Ja (SJ)

Struktura ja (*self, ego*) jest definiowana jako organizacja wiedzy o sobie (Greenwald, 1980), czy reprezentacja samego siebie analogiczna do reprezentacji innych osób (Swann i Bosson 2010). Ujmowana jest jako składnik systemu poznawczego.

Odgrywa ona centralną rolę w regulacji zachowania się człowieka, przede wszystkim w regulacji jego działań intencjonalnych służących:

- Zaspokojeniu potrzeb biologicznych i emocjonalnych powstających w systemie organizmalnym; ich realizacja jest warunkiem biologicznego przetrwania jednostki i hedonistycznej satysfakcji.
- Zaspokojeniu potrzeb i emocji powstających w Strukturze ja: ich realizacja wiąże się, w znacznym stopniu, ze społecznymi interakcjami i społecznym położeniem jednostki.

Ważną rolę pośredniczącą w regulacji zachowania odgrywają przeświadczenia o możliwości osiągnięcia danego wyniku. Zależą one od „zapisanych” w Strukturze Ja przekonań dotyczących własnych kompetencji. Jest to własność, którą Albert Bandura (1995) opisał jako poczucie sprawczości (*self efficacy*).

SJ powstaje pod wpływem doświadczenia, które można zaliczyć do trzech kategorii (Baumeister 2000):

- Doświadczenia samoświadomości (*reflexive consciousness*) – myślenie o sobie, obserwowanie osobistych przeżyć, tworzenie pojęcia osobistego ja.
- Doświadczenia relacji ja – inni ludzie: informacje, jak inni nas spostrzegają i oceniają, jak się do nas odnoszą.
- Doświadczenia efektów osobistych decyzji i własnej działalności (*executive functions*).

Rozwój SJ polega na stopniowym wyodrębnianiu samego siebie z fizycznego i społecznego otoczenia i kształtowaniu własnej podmiotowości (Jarymowicz, 1994).

SJ jest źródłem specyficznych tendencji motywacyjnych. Do najczęściej wymienianych i badanych motywów generowanych przez tę strukturę należą:

- Dążenie do obrony i podwyższenia samooceny i poczucia własnej wartości lub ważności swego ja (*Self-enhancement*, Hogg, 2003).
- Do potwierdzania własnej tożsamości (Wicklund and Gollwitzer 1982, Gollwitzer i Wicklund 1985, McFarland i in.2019).

Jak dowodził Greenwald, podstawowym motywem jest utrzymanie przekonania o pozytywności własnego ja i o prawdziwości obrazu własnej osoby. Innymi słowy chodzi o zachowanie „teorii własnego ja” jako spójnej wartościowej całości. Motywacja ta jest źródłem systematycznych zniekształceń poznawczych. Greenwald (1980) opisał trzy typowe zniekształcenia poznawcze generowane przez tę strukturę:

- Egocentryzm, czyli tendencję do postrzegania siebie jako postać centralną obecnych i przeszłych (zapisanych w pamięci) zdarzeń.
- Pozytywną sprawczość (*benefactance*), czyli notoryczną skłonność do przypisywania sobie sprawstwa zdarzeń pozytywnych, a zaprzeczaniu swemu udziałowi w powstawaniu zdarzeń negatywnych.
- Poznawczy konserwatyzm, czyli tendencję do zachowania trwałości utworzonych struktur wiedzy, a więc do zniekształcania nadchodzących informacji tak, aby można je było asymilować do istniejących struktur.

Pewne formy motywacji powstającej w SJ mogą ukierunkować działanie człowieka na cele pozaosobiste. Kształtują się one w sferze indywidualnej tożsamości i w sferze samooceny

Ukształtowanie tożsamości a motywacja pozaosobista (prospołeczna)

Pojęcie tożsamość i poczucie tożsamości dotyczą tych elementów obrazu własnej osoby, które charakteryzują samego siebie jako odrębny autonomiczny podmiot. Z reguły są to cechy trwałe, zapewniające poczucie ciągłości własnego ja. Cechy te są też podstawą poczucia podobieństwa/poczucia odrębności od innych ludzi, co ma istotne znaczenie dla formowania motywacji społecznych (Jarymowicz, 1994).

Według teorii „kompletności ja” (*self completion theory*) (Wicklund i Gollwitzer, 1982) wśród przekonań, które ludzie mają o sobie, określających „kim jestem”, a więc tożsamościowych, ważną rolę odgrywa pewna szczególna ich kategoria – te, które dotyczą istotnych dla nas ról, jakie pełniimy w życiu. Przekonania te mają funkcje motywacyjne, ponieważ do utrzymania tych przekonań konieczne jest ich potwierdzenie przez doświadczenie – przez praktyki, które im odpowiadają, a także przez zachowania i postawy innych ludzi, które je potwierdzają.

II. Przejawem drugiego z systemów motywacyjnych było wykonywanie zadań wymaganych w ramach społecznego układu

Ten układ tworzyła tu „grupa minimalna”. Pojęcie to wprowadził Henry Tajfel (Tajfel i in. 1976, Tajfel 1996), pokazując, że grupa i zjawiska grupowe powstać mogą, kiedy są spełnione pewne minimalne warunki, a mianowicie, jeśli A pod względem jakiegokolwiek cechy dostrzega podobieństwo do B i różnice od C, to A i B stają się „grupą minimalną”, zaś C zostaje uznane za „outgroup”. Grupę tworzyć może także wspólne zadanie lub wspólny los. Utworzenie minigrupy uruchamia elementarny proces grupowy – uprzywilejowania „swoich” i dyskryminację „obcych” (*in-group bias, out-group discrimination*). W grupach, w zależności od warunków, uruchomione zostają różne procesy grupowe takie, jak:

- powstawanie norm grupowych i presja grupowa na ich przestrzeganie,
- konformizm wobec panujących w grupie norm,
- ukształtowanie grupowych autorytetów i uległość wobec nich,
- tendencja do zachowania spójności grupowej i do faworyzowania członków własnej grupy kosztem tych, którzy nie są jej członkami (obcych),
- tendencja do ochrony dobra grupy i jej dobrego imienia,
- gotowość do pomocy członkom grupy (solidarność),
- i inne.

Wszystkie te procesy mogą stawać się źródłem specyficznych tendencji motywacyjnych. Podstawowym wskaźnikiem powstania grupy jest pojawienie się kategorii „my”. Powstanie „my” jest przejawem utworzenia społecznej tożsamości, czyli wg. teorii samokategoryzacji, zdefiniowania samego siebie i bycie zdefiniowanym przez innych jako członka pewnej grupy (Tajfel i Turner, 1979). To „my” implikuje poczucie wspólnoty z daną grupą. Mogą to być grupy różnej skali – zarówno duże, trwałe, wewnętrznie zróżnicowane, jak naród, grupa etniczna, polityczna organizacja, zwolennicy tej samej ideologii, współwyznawcy i in. A także do grup małych krótkotrwałych, tworzonych sytuacyjnie, wytwarzających chwilowe „my”.

Sądzę, że osoby badane w sytuacji stworzonej przez Dolińskiego i Grzyba stawały się członkami minigrupy obejmującej zamkniętych w jednej przestrzeni eksperymen-

tatora i osobę badaną – wykonujących wspólne zadanie, natomiast „uczeń” znajdujący się w innym pomieszczeniu i wykonujący inne zadanie mógł być postrzegany w kategoriach *outgroup*. Milgram (1977 s. 108) określił tę sytuację jako „rodzenie się grupy” (*incipient group formation*). A Gibson (2022) uwzględnił kategorię „my” w swojej analizie jednego z eksperymentów Milgrama.

Tak więc można przyjąć, że motywacja osoby badanej wynikała z działania sił grupowych, solidarności z eksperymentatorem jako członkiem grupy i z uległości wobec autorytetu grupowego. Ale istniało także inne oddziaływanie – sygnały cierpienia „ucznia, potencjalne źródło innej, „pozagrupowej” motywacji.

W konkluzji można powiedzieć, że w omawianej tu sytuacji eksperymentalnej mamy do czynienia z zupełnie innym niż opisywany poprzednio systemem motywacyjnym.

III. Przejawem trzeciego z systemów motywacyjnych były działania ukierunkowane na realizację celów o znaczeniu pozaosobistym – prospołecznych

Ratowanie ludzi w sytuacji w zagrożenia badane przez Jane Piliavin i jej kolegów nie miało związku z adaptacyjnymi celami jednostki ani z celami adaptacyjnymi własnej grupy, a nawet było z nimi w jakimś stopniu sprzeczne. Motywację tę określa się jako altruistyczną. Motywacja altruistyczna należy do szerszej kategorii procesów motywacyjnych ukierunkowujących działania na cele pozaosobiste. Kategoria ta obejmuje takie zjawiska, jak np.:

- Empatia (Hoffman, 2000; Eisenberg i in., 2010) lub potrzeba przynależności (Baumeister i Leary, 1995), które mogą powstawać w systemie organizminalnym.
- Motywacja normatywna, która może kształtować się dzięki procesom identyfikacji (Jarymowicz, 1994, 2008) i procesom uwewnętrzniania norm. Normy te mogą być powiązane z poczuciem własnej wartości (Higgins, 1999) lub z ukształtowaniem tożsamości (Mc Farland i in., 2019; Gollwitzer 2018), czyli z mechanizmami Struktury ja.
- System autonomicznych wartości ukształtowanych w sieci poznawczej (Tetlock 2003).

Scharakteryzuję dokładniej mechanizm związany z funkcjonowaniem sieci poznawczej.

Dygresja terminologiczna

Pojęcia „sieć poznawcza” i „struktura w sieci poznawczej” miały dotychczas status „konstruktów teoretycznych”, czyli narzędzi intelektualnych służących organizowaniu pewnych obszarów wiedzy psychologicznej i tworzenia hipotez mających wyjaśniać pewne obserwowalne zjawiska.

Postępy neurokognitywistyki w ostatnich latach zdają się dowodzić, że pojęcia te mogą odnosić się do pewnej konkretnej neurofizjologicznej rzeczywistości.

Jak opisuje Ken Norman (za Somers, 2021), szef (*Chair*) Departamentu Psychologii w Princeton University i ekspert w zakresie problematyki dekodowania myśli (*thought decoding*), w mózgu powstaje zapis przedmiotów i scen w formie wzorów aktywacji. Wzory zachowują tę samą formę bez względu na to, czy osoba ogląda daną scenę, myśli o niej, czy opowiada komuś o niej.

Wzory te składają się z jednostek tzw. *voxels* o rozmiarze milimetra sześciennego – każdy z nich konsumuje tlen w ilości odpowiadającej ok. miliona neuronów.

Można przypuszczać, że struktury poznawcze, może także takie jak Struktura Ja, mają swój neurofizjologiczny odpowiednik (czy substrat) w postaci wielowymiarowych wzorów aktywacji wyższego rzędu.

Systemy wartości jako specyficzne ukształtowanie struktur w sieci poznawczej

Sieć poznawcza to obszar aparatu psychicznego, w ramach którego zachodzą procesy tworzenia reprezentacji rzeczywistości, w tym także konstrukcje różnych jej możliwych czy niemożliwych wersji. Dokonują się tam także procesy wyboru i planowania czynności. W sieci poznawczej tworzą się struktury będące reprezentacjami wartości, które w odróżnieniu od innych pojęć nie są wyłącznie pewnym zapisem wiedzy o rzeczywistości. Należą one do innego porządku.

Pojęcie wartość może mieć dwa zupełnie różne znaczenia. W potocznym języku odnosi się je na ogół do takich obiektów i relacji, które są dla jednostki mniej lub bardziej użyteczne, czy też społecznie za użyteczne uznawane. W tym sensie wartość to stopień użyteczności przedmiotu lub stanu rzeczy. Mają one wartość wymienną – można je kupić, sprzedać, zamienić na coś innego. Ale istnieją też wartości, które mają charakter absolutny – są ujmowane w kategoriach moralności i/lub świętości. Sama myśl o ich wymianie na jakąś inną wartość może budzić zgrozę (jak np. myśl, że można sprzedać własne dziecko).

Model ochrony świętych wartości (Tetlock)

Tetlock (2003) zdefiniował święte wartości, jako takie, którym moralna wspólnota przypisuje znaczenie transcendentalne, co wyklucza ich porównywanie do innych czy wymianę na inne (*taboo-trade off*, Tetlock i in. 2000). Należą tu takie jak np. honor, miłość, sprawiedliwość, życie ludzkie, integralność ludzkiego ciała (nie wolno kupować organów od żywego człowieka), dobro dziecka, lojalność wobec własnego kraju, przyjaźń i inne. Świętość może być przypisywana przedmiotom (np. świętym księgom), miejscom (takim jak ziemia ojczysta, cmentarze, Wzgórze Świątynne w Jerozolimie), wydarzeniom

historycznym (np. powstania narodowe) i in. Fiske i Tetlock (1997) dowodzą, że w większości kultur rozróżnia się sfery, w których wymiana i ocena wg. rynkowej wartości jest dopuszczalna, i takie, w których jest niedopuszczalna. Tetlock (2003) sformułował Model Ochrony Świętych Wartości (*Sacred Values Protection Model* – SVPM), który dotyczy reakcji ludzi na sytuacje pogwałcenia tych wartości. Taką reakcją jest moralne oburzenie, gniew, potępienie osoby, która takiego gwałtu się dopuszcza lub go toleruje, poddawanie jej ostracyzmowi, żądanie dla niej kary. Wartości absolutne (moralne, święte) wzbudzają bardzo silne emocje i mają bardzo duży potencjał motywacyjny. Skłaniają ludzi do działań w ich obronie bez względu na koszty osobiste, także zagrożenia dla ich zdrowia i życia.

Osoby Badane w programie Piliavin i in. reagowały na zagrożenie zdrowia, czy życia innej osoby, nie bacząc na osobiste konsekwencje. Tak rozumiane wartości stanowią jedno z głównych źródeł prospołecznej motywacji.

Konkluzja

Teoria ograniczonej etyczności stwierdza, że w sytuacjach konfliktu interesów naturalną reakcją ludzi jest tendencja, aby uznawać decyzje służące ich interesom jako moralnie słuszne. Jej autorzy uważają, że tego rodzaju sytuacje są w życiu ludzi powszechne. Tak więc zdają się sugerować, że własny interes jest dominującym motywem człowieka, natomiast moralność służy głównie do oceny czynów cudzych i usprawiedliwiania czynów własnych. Wydaje się, że pogląd taki jest dość rozpowszechniony (Miller, 1999).

Jak starałem się to uzasadnić, zgromadzona w psychologii wiedza o motywacyjnych mechanizmach człowieka wskazuje na znaczne ograniczenia tego sposobu myślenia. Dowodzi, że w toku biologicznego i społecznego rozwoju kształtują się w człowieku odrębne, działające na różnych zasadach trzy takie subsystemy. Można założyć, iż u normalnie usocjalizowanego człowieka osiągają dojrzałość wszystkie wyróżnione tu subsystemy, choć w poszczególnych kulturach pewne subsystemy mogą być bardziej rozwijane niż inne. Tak np. w kulturach liberalnych szczególny nacisk kładziony jest na rozwój Struktury ja jako autonomicznego systemu, a w kulturach tradycyjnych – na rozwój ja kolektywnego (Markus i Kitayama, 1991; Markus i Conner, 2013).

Wszakże aktywność poszczególnych podsystemów może być różna w zależności od sytuacji, w jakiej człowiek się znalazł, a także od potencjału regulacyjnego każdego z nich (w danym przedziale czasu). Ilustrują to trzy prezentowane tu programy badawcze. Każdy z nich koncentrował uwagę osób badanych na wymaganiach „zaadresowanych” do tylko jednego spośród wyróżnionych tu podsystemów motywacyjnych, utrudniając czy uniemożliwiając zaktywizowanie innych. W tych warunkach okazywało się, że „naturalną”, oczywistą bezrefleksyjną reakcją człowieka może być „ograniczona etycz-

ność”, związana z automatycznym poznawczym egocentryzmem, albo „ograniczona etyczność”, związana z automatyczną uległością wobec autorytetu, albo „nieograniczona etyczność”, związana z bezrefleksyjną reakcją na zagrożenie zdrowia lub życia innego człowieka.

Okazywało się także, iż w tych samych okolicznościach pod wpływem zewnętrznych lub wewnętrznych czynników może dochodzić do „przełączania motywacji”. Tak np. w badaniach nad egocentrycznym zniekształcaniem ocen moralnych (Bocian, Baryła i Wojciszke, 2021) okazywało się, że włączenie społecznej kontroli (odpowiedzialności za słowo) aktywowało socjocentryczną motywację – badani kierowali się przyjętymi normami społecznymi, dokonując oceny zachowania osób łamiących normę.

W badaniach nad posłuszeństwem wobec autorytetu (Doliński i Grzyb, 2017) można było zaobserwować jeszcze bardziej złożone formy relacji między różnymi systemami motywacyjnymi. Podczas gdy u ogromnej większości osób badanych mogła dominować (jak starałem się to uzasadnić) motywacja zadaniowa związana z podporządkowaniem wymaganiom grupowej sytuacji (socjocentryczna), to u części osób pojawiały się inne tendencje motywacyjne. Często były to reakcje, które można interpretować jako przejawy empatii (silne emocjonalne pobudzenie wzrastające pod wpływem skarg „drażnionej prądem” osoby) – reakcje te u niektórych osób przekształcały się w usiłowania przerwania badań. Były też osoby, których reakcje można interpretować jako przejawy społecznej odpowiedzialności (motywacji prospołecznej) – osoby takie bez objawów silnego wzburzenia emocjonalnego odmawiały dalszego udziału w badaniach.

Z przedstawionego wywodu zdaje się wynikać następująca konkluzja. Zjawisko ograniczonej etyczności i inne przejawy egocentrycznego myślenia i działania mogą pojawiać się przede wszystkim wtedy, gdy motywacja ipsocentryczna jest odpowiednio pobudzona, natomiast inne subsystemy nie są aktywne.

Trzeba zauważyć, że w zwykłych sytuacjach życiowych motywacja ipsocentryczna może być często angażowana, ponieważ człowiek musi rozwiązywać wiele różnych osobistych problemów adaptacyjnych. Musi więc jako żywy organizm stale dbać o zachowanie homeostatycznej równowagi swego organizmu, a także jako członek społeczności dbać o równowagę swoich społecznych relacji. Dotyczą one nie tylko tu i teraz, ale także bliższej i dalszej przyszłości – nie tylko swojej, ale i bliskich. Są to zadania adaptacyjne, które muszą angażować dużo uwagi i dość dużą ilość zasobów człowieka, choć różną u różnych ludzi. Zależy to w dużym stopniu od sytuacji materialnej i społecznej człowieka, ale także od ukształtowanych trwałych relacji między systemami. Tak np. u niektórych ludzi można zaobserwować stałą dominację Struktury Ja nad innymi systemami, co przejawia się w postaci takich cech jak narcyzm czy egoizm.

Ale właściwie każdy człowiek jest członkiem różnych grup – wspólnot, takich jak rodzina, społeczność miejsca pracy, organizacje i związki, grupa przyjaciół, własny kraj,

wyznawcy określonej religii itp. Każda z tych grup generuje określone wymagania, będąc źródłem różnych form motywacji socjocentrycznej.

Każdy z nas żyje w świecie przedmiotów, symboli, wyobrażeń, idei ukształtowanych w ciągu bardzo wielu pokoleń i składających się na świat wartości przechowywanych przez kulturę. Choć ma ona wiele różnych specyficznych form, to są w niej też takie obszary, z którymi identyfikują się ludzie należący do bardzo wielu wspólnot – prawie uniwersalne. Wartości kulturowe są jednym z istotnych źródeł motywacji, której w życiu codziennym możemy nie zauważać, ale noszą w sobie duży motywacyjny potencjał.

Ten tok myślenia prowadzi do wniosku, że system motywacyjny człowieka, to wielowymiarowa struktura w ramach której Struktura ja i jej cele adaptacyjne wymagają „uzgadniania” z celami generowanymi przez inne struktury systemu. Przekonanie, że stanem naturalnym jest podporządkowanie wszelkich innych celów celom własnego ja może być raczej wynikiem zaakceptowania ideologii głoszącej prymat dobra własnego nad każdym innym dobrem (np. Ayn Rand) niż opisem psychologicznej rzeczywistości.

Ma to oczywiście pewne znaczenie dla psychologicznej praktyki. Czy należy ją pojmować jako wsparcie dla samorealizacyjnych celów człowieka, ma wspomagać go w osiągnięciu celów, które maksymalizują jego własne dobro, czy celów, które sprzyjają harmonizacji własnego dobra z dobrem szeroko rozumianej wspólnoty?

To oczywiście jest pytanie, na które odpowiedź opiera się nie tylko na przesłankach profesjonalnych, ale w jakimś stopniu na przesłankach ideologicznych. Odpowiedź, która mnie wydaje się słuszna, opiera się na założeniu, że ostatecznie dobro jednostki i dobro szeroko rozumianej wspólnoty są nierozłączne (Reich 2018).

Bibliografia

- Almond G. (1991). Rational choice theory and the social sciences. [w:] K.R. Monroe (red.), *The economic approach to politics. A critical reassessment of the theory of rational action* (s. 32–52). New York: HarperCollins Publisher.
- Bandura A. (red.). (1995). *Self-efficacy in changing societies*. Cambridge University Press
<https://doi.org/10.1017/CBO9780511527692>
- Baumeister R. (2000). The nature and structure of the self: An overview. [w:] R.F. Baumeister (red.), *The self in social psychology. Key readings* (s. 1–20) Philadelphia, PA: Psychology Press.
- Baumeister R.F., Leary, M.R. (1995). The need to belong. Desire for Interpersonal attachments as fundamental human motivation. [w:] E.T. Higgins, A.W. Kruglanski (red.), *Motivational science: Social and personality perspective*. (s. 24–49). Philadelphia, PA: Psychology Press.
- Bocian K., Baryła W., Wojciszke B. (2020). Egocentrism shapes moral judgements. *Social and Personality Psychology Compass*. 1–14. 10.1111/spc3.12572.
<https://doi.org/10.1111/spc3.12572>

- Bocian K., Baryła W., Wojciszke B. (2022). Moderators of the Liking Bias in Judgments of Moral Character. *Personality and Social Psychology Bulletin*, 48(4), 596–605.
<https://doi.org/10.1177/01461672211013272>
- Bocian K., Wojciszke, B. (2014). Self-interest bias in moral judgments of others' actions. *Personality and Social Psychology Bulletin*, 40, 898–909.
<https://doi.org/10.1177/0146167214529800>
- Chugh D., Bazerman, M.H., Banaji, M.R. (2005). Bounded ethicality as a psychological barrier to recognizing conflicts of interest. [w:] D.A. Moore, D.M. Cain., G. Loewenstein, M.H. Bazerman (red.), *Conflict of interest: Challenges and solutions in business, law, medicine, and public policy* (s. 74–95). New York, NY: Cambridge University Press.
- Doliński D., Grzyb T. (2017). *Posłuszni do bólu. O uległości wobec autorytetu w 50 lat po eksperymencie Milgrama*. Sopot: Smak Słowa.
- Eisenberg N., Eggum, N.D., Di Giunta, L. (2010). Empathy-related responding: Associations with prosocial behavior, Aggression, and intergroup relations. *Social Issues and Policy Review*. Dec 1; 4(1): 143–180.
<https://doi.org/10.1111/j.1751-2409.2010.01020.x>
- Fiske A.P., Tetlock, P.E. (1997). Taboo Trade-offs: Reactions to transactions that transgress the spheres of justice. *Political Psychology*, 18 (2), 255–297.
<https://doi.org/10.1111/0162-895X.00058>
- Gibson S. (2020). 'We have a choice': Identity construction and the rhetorical enactment of resistance in the 'two peers rebel' condition of Stanley Milgram's obedience experiment. *European Journal of Social Psychology*, 52(3).
<https://doi.org/10.1002/ejsp.2734>
- Gollwitzer P.M. (2018). The goal concept: A helpful tool for theory development and testing in motivation science. *Motivation Science*, 4(3), 185–205.
<https://doi.org/10.1037/mot0000115>
- Gollwitzer P.M., Wicklund, R.A. (1985). The Pursuit of Self-Defining Goals. [w:] J. Kuhl, J. Beckmann (Red.), *Action control. From cognition to behavior*. (s. 61–85). Heidelberg: Springer-Verlag.
- Greenwald A. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist*, 35(7), 603–618.
<https://doi.org/10.1037/0003-066X.35.7.603>
- Higgins E. (1999). Self-discrepancy: A theory relating self and affect. [w:] R. Baumeister (red.), *The self in social psychology* (s. 150–175). Philadelphia, PA: Taylor & Francis.
- Hoffman M. (2000). *Empathy and moral development: Implications for caring and justice*. Cambridge, MA: Cambridge University Press.
- Hogg M.A. (2003). Social identity. [w:] M.R. Leary, J.P. Tangney (red.), *Handbook of self and identity* (s. 462–479). New York, NY: The Guilford Press.
- Jarymowicz M. (1994). W stronę indywidualnej podmiotowości i zbliżeń z innymi: podmiotowe podstawy społecznych identyfikacji. [w:] M. Jarymowicz (red.), *Poza egocentryczną perspektywą widzenia siebie i świata* (s. 11–36). Wydawnictwo Instytutu Psychologii PAN.
- Jarymowicz M. (2008). O roli myślenia w regulacji reakcji automatycznych. *Czasopismo Psychologiczne* 14(2), 187–199.
- Kahneman D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58(9), 697–720.
<https://doi.org/10.1037/0003-066X.58.9.697>

- Markus H.R., Conner A. (2013). *Clash!: 8 cultural conflicts that make us who we are*. New York, NY: Hudson Street Press.
- Markus H.R., Kitayama S. (1991). Cultural variation in the self-concept. Culture and self: Implications for cognition. *Emotion and Motivation. Psychological Review*, 98(2), 224–253. [doi:10.1037/0033-295X.98.2.224](https://doi.org/10.1037/0033-295X.98.2.224)
- McFarland S., Hackett J., Hamer K., Katzarska-Miller I., Malsch A., Reese G., Reysen S. (2019). *Global human identification and citizenship: A review of psychological studies*. *Political Psychology*, 40(Suppl. 1), 141–171. <https://doi.org/10.1111/pops.12572>
- Milgram S. (1977). *The individual in a social world: Essays and experiments*. New York: Addison-Wesley Publishing Company
- Miller D.T. (1999). The norm of self-interest. *American Psychologist*, 54(12), 1053–1060. <https://doi.org/10.1037/0003-066X.54.12.1053>
- Piliavin J., A., Dovidio J.F., Gaertner S.L. Clark R.D. (1981). *Emergency intervention*. New York, NY: Academic Press.
- Rand A. (1962). *Introducing objectivism by Ayn Rand*. <https://courses.aynrand.org/works/introducing-objectivism>
- Reich R. (2018). *The common good*. New York, NY: Knopf.
- Somers, J. (2021, December 6). The science of mind reading. *The New Yorker*.
- Swann Jr. W., Bosson J. (2010). Self and identity. [w:] S. Fiske, D. Gilbert, G. Lindzey (red.), *Handbook of social psychology* (wyd. 5, t. 1, s. 589–628). New York, NY: J. Wiley.
- Tajfel H. (1996). The minimal group studies. [w:] Banyard, P., Grayson, A. (red.), *Introducing psychological research. Sixty studies that shape psychology* (s. 69–73). London: Red Globe Press.
- Tajfel H., Turner J. (1979). An integrative theory of intergroup conflict. [w:] W.G. Austin, S. Worchel (red.), *The social psychology of intergroup relations*. (s. 33–48). Monterey, CA: Brooks/Cole Publishing Company.
- Tajfel H., Billig M.C., Bundy R.P., Flament C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149–178. <https://doi.org/10.1002/ejsp.2420010202>
- Tetlock, P. E. (2003). Thinking the unthinkable: sacred values and taboo cognitions. *Trends in Cognitive Sciences*, 7(7), 320–324. [https://doi.org/10.1016/S1364-6613\(03\)00135-9](https://doi.org/10.1016/S1364-6613(03)00135-9)
- Tetlock P.E., Kristel O.V., Elson S.B., Green M.C., Lerner J.S. (2000). The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*, 78(5) 853–870. <https://doi.org/10.1037/0022-3514.78.5.853>
- Wicklund R.A., Gollwitzer P.M. (1982). *Symbolic self – completion*. Hillsdale, NJ: L. Erlbaum.
- Wojciszke B. (2010). *Sprawczość i wspólnotowość*. Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- Wojciszke B., Baryła W. (2006). Self-interest in interpersonal perception and attitudes. *Polish Psychological Bulletin*, 37(1), 43–50.
- Wojciszke B., Cisiak A. (2006). Self interest in social and political perception. [w:] A. Golec de Zavala, K. Skarżyńska, (red.), *Understanding social change. Political psychology in Poland* (s. 51–70). New York, NY: Nova Science Pub.

The Theory of Bounded Ethicality and knowledge of human motivation systems

The Theory of Bounded Ethicality postulates that in situations involving consequences for self and/or others people make ethical decisions that bring self-oriented motivational forces to bear on decision-making. These ethical decisions are biased by a stubborn view of oneself, as moral, competent, and deserving and thus, not susceptible to conflicts of interest. People unconsciously favor this particular vision of the self being unaware of data that may contradict it. This conception of ethical decision making seems to imply that self-oriented motivation plays a dominant role in regulation of human behavior. But there are good theoretical reasons for questioning this view. In the paper, I describe three different research programs that may illustrate operation of the three different motivational systems – only one of them seems to correspond with the Bounded Ethicality model. It means that that “bounded ethicality” may appear in pure form in specific situations – when self system is activated by specific exogenous and endogenous factors while other systems are not remain in the latent state.

Key words: motivation, self, cognitive bias, group forces, sacred values

