

DYLEMAT WIĘZNI A DYLEMATY ETYCZNE

Streszczenie: *Podstawę wielu teorii ekonomicznych, w tym głównie neoklasycznej oraz teorii podejmowania decyzji (zwłaszcza szeroko stosowanej teorii oczekiwanej użyteczności), stanowi naczelną postulat racjonalności – maksymalizacji własnych korzyści. Jednak wiele zachowań ekonomicznych dotyczy interakcji między ludźmi, m.in. zawierania umów i kontraktów. Dla trwałości funkcjonowania procesów ekonomicznych, w których podstawą są właśnie umowy i kontrakty, niezbędne jest wzmocnienie ich systemem zasad moralnych. Celem niniejszego rozdziału jest wskazanie na możliwości wykorzystania dylematu więźnia w analizie i kształtowaniu etycznych zachowań w działaniach ekonomicznych wymagających interakcji między ludźmi poprzez omówienie wybranych eksperymentów opisanych w literaturze. Przebieg i wyniki tych eksperymentów pokazują, jak racjonalne postępowanie każdego z uczestników gry prowadzi do suboptymalnych wyników, zaś kierowanie się racjonalnością grupową, spójną z teoriami etycznymi pozwala na osiągnięcie wyników optymalnych, zarówno z punktu widzenia interesów indywidualnych, jak i grupowych. Motywy postępowania mogą być uzasadnione zarówno na gruncie teorii deontologicznych, jak i teleologicznych. Ćwiczenie i analiza szeregu eksperymentów wykorzystujących dylemat więźnia mogłoby przyczynić się do lepszego zrozumienia sensu racjonalnego postępowania i towarzyszących mu zasad moralnych, zwłaszcza w kontekście konfliktu interesów indywidualnych i grupowych.*

Słowa kluczowe: *dylemat więźnia, teoria gier, etyka, moralność, racjonalność.*

WSTĘP

Podstawę wielu teorii ekonomicznych, w tym głównie neoklasycznej oraz teorii podejmowania decyzji (zwłaszcza szeroko stosowanej teorii oczekiwanej użyteczności), stanowi naczelną postulat racjonalności – maksymalizacji własnych korzyści. Jednak wiele zachowań ekonomicznych dotyczy interakcji między ludźmi, m.in. zawierania umów i kontraktów. Dla trwałości funkcjonowania procesów ekonomicznych, w których podstawą są właśnie umowy i kontrakty, niezbędne jest wzmocnienie ich systemem zasad moralnych. Przejawem tego są np. kodeksy etyczne wprowadzane w korporacjach. Ważną kwestią jest, czy pracownicy korporacji rzeczywiście identyfikują zapisane w tych kodeksach zasady jako własne, czy rozumieją ich sens w kontekście interesów własnych i grupowych.

Celem niniejszego rozdziału jest wskazanie na możliwości wykorzystania *dylematu więźnia* w analizie i kształtowaniu etycznych zachowań w działaniach

* Dr Wojciech Zatoń, Wydział Ekonomiczno-Socjologiczny, Uniwersytet Łódzki.

ekonomicznych wymagających interakcji między ludźmi. *Dylemat więźnia*, jeden z najbardziej znanych problemów teorii gier, jest popularnym narzędziem do analizy zachowań ekonomicznych, społecznych i psychologicznych.

Walorem pracy jest zebranie, analiza i przedstawienie wybranych, wartościowych badań we wskazanej powyżej dziedzinie. Kryterium selekcji było wskazanie specyfiki zależności występującej między racjonalnością indywidualną i grupową przez pryzmat zasad moralnych i teorii etycznych. Przebieg i wyniki opisanych eksperymentów pokazują, jak racjonalne postępowanie każdego z uczestników gry prowadzi do suboptymalnych wyników, zaś kierowanie się racjonalnością grupową, spójną z teoriami etycznymi pozwala na osiągnięcie wyników optymalnych, zarówno z punktu widzenia interesów indywidualnych, jak i grupowych.

Pierwsza część rozdziału zawiera zarys teoretycznych problemów racjonalności i moralności, jak również motywów psychologicznych w ekonomii. W drugiej, wykorzystując opisane w literaturze eksperymenty, przedstawiono analizę problemu *dylematu więźnia* w kontekście racjonalności indywidualnej i grupowej oraz zasad moralnych i teorii etycznych.

1. RACJONALNOŚĆ I MORALNOŚĆ – ZARYS TEORII

Modele teoretyczne stosowane w finansach i wynikające z nich zachowania obserwowane na rynkach finansowych i w szeroko rozumianej działalności gospodarczej, mają swoje podstawy w neoklasycznej teorii ekonomii. Uzasadnieniem dla ludzkiego postępowania jest skupienie się na własnym interesie i dążeniu do maksymalizacji osobistych korzyści. Adam Smith, twórca ekonomii klasycznej uważał, że siłą napędową gospodarki jest interes własny. W tym kontekście często przytaczany jest fragment z jego dzieła: „Nie od przychyłności rzeźnika, piwowara, czy piekarza oczekujemy naszego obiadu, lecz od ich dbałości o własny interes. Zwracamy się nie do ich humanitarności, lecz do egoizmu i nie mówimy im o naszych własnych potrzebach, lecz o ich korzyściach”¹. Choć za główny mechanizm powstrzymujący nieograniczony egoizm i chciwość, A. Smith uważał konkurencję, to jednak w swoim drugim ważnym dziele² wskazywał na ważność tzw. etyki sympatii w postępowaniu ludzi, przez którą rozumiał zdolność współprzeżywania cudzych uczuć i ich wartościowania. Choć A. Smith nie określał normatywów dobra i zła i nie tworzył zwartego zbioru zasad etycznych, to było dla niego oczywiste, że ludzie odbierają pewne działa-

¹ A. Smith, *Badania nad naturą i przyczynami bogactwa narodów*, Wydawnictwo Naukowe PWN, Warszawa 2007, s. 21–22.

² A. Smith, *Teoria uczuć moralnych*, Wydawnictwo Naukowe PWN, Warszawa 1989; Warto zauważyć, że było to chronologicznie (1759) dzieło wcześniejsze od *Badania nad naturą i przyczynami bogactwa narodów* (1776).

nia jako dobre a inne za złe i na tej podstawie dokonują ich ocen moralnych. Poza tym zdecydowanie wyznawał zasadę priorytetu dobra ogólnospołecznego nad dobrem prywatnym. To wszystko świadczyło o złożoności motywów ekonomicznych działalności ludzi w rozumienia A. Smitha³.

W dalszym rozwoju teorii ekonomii wybiórczo wykorzystywano założenia przyjęte przez A. Smitha, przyjmując jako podstawę działanie ukierunkowane na własny interes i pomnażanie osobistego bogactwa. Kwestie moralne i psychologiczne pozostały przez wiele lat całkowicie pominięte w normatywnym opisie działania ekonomicznego ludzi.

Teoria oczekiwanej użyteczności przedstawiona najpierw przez D. Bernoulliego w 1738 r., a następnie sformalizowana przez J. von Neumanna i O. Morgensterna w 1944 r.⁴, stała się podstawą teorii racjonalnego wyboru we współczesnej ekonomii i finansach. Racjonalność ekonomiczna opiera się na pięciu aksjomatach użyteczności kardynalnej (kompletności, przechodniości, ciągłości, niezależności i wyniku zupełnie pewnego)⁵, zdefiniowanych pierwotnie przez J. von Neumanna i O. Morgensterna. Dodatkowym aksjomatem jest twierdzenie o dążeniu do maksymalizacji bogactwa (z racjonalnego, finansowego punktu widzenia lepiej jest mieć więcej pieniędzy niż mniej). Założeniem działalności *homo oeconomicus* jest też stwierdzenie, że przestrzegając powyższych aksjomatów, człowiek potrafi dokonać racjonalnych wyborów spośród wielu alternatyw⁶.

Analiza decyzji podejmowanych przez ludzi wykazała jednak, że często zachowują się oni nieracjonalnie, nie przestrzegając aksjomatów teorii oczekiwanej użyteczności⁷. Dotyczy to zwłaszcza zasady stałości i spójności preferencji. Problem niestosowania się ludzi do założeń o racjonalności działania w ekonomii dostrzegł H. Simon, który wprowadził pojęcie ograniczonej racjonalności⁸. Wynika ona z prostego faktu, że ludzie nie są w stanie przetwarzać wszystkich informacji (ze względu choćby na ich ilość, niekompletność, koszt przetwarzania) i w wielu sytuacjach posługują się prostymi regułami decyzyjnymi (heurystykami), które często prowadzą do nieoptymalnych, ale zadowalających decy-

³ Interesującą, kompleksową analizę twórczości A. Smitha przeprowadził G. Szulczewski w pracy: G. Szulczewski, *Krytyka „etyki sympatii” i „problem A. Smitha” w świetle lektury Teorii uczuć moralnych*, „Prakseologia” 2015, nr 157, t. 1.

⁴ J. von Neumann, O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton (NJ) 2007.

⁵ J. Koziński, *Psychologiczna teoria decyzji*, PWN, Warszawa 1977.

⁶ T. E. Copeland, J. F. Weston, *Financial Theory and Corporate Policy*, Addison Wesley, New York 1988, s. 80

⁷ Szereg ciekawych przykładów można znaleźć np. w pracy: T. Zaleskiewicz, *Psychologia ekonomiczna*, Wydawnictwo Naukowe PWN, Warszawa 2011.

⁸ H. A. Simon, *Models of Man: Social and Rational*, Wiley, New York 1957.

zji. G. Gigerenzer i P. M. Todd⁹ zwrócili uwagę, że ludzie podejmują decyzje w oparciu o proste heurystyki, szybkie i oszczędne, a ich wybór nie zależy od logicznej spójności, ale od realizacji praktycznych celów i adaptacyjności do danej sytuacji i warunków otoczenia (postępowanie racjonalnie ekologicznie)¹⁰. Dalsze próby wyjaśnienia braku racjonalności w podejmowaniu decyzji doprowadziły do powstania psychologicznej teorii decyzji. Do rozwoju zainteresowania wpływem czynników behawioralnych na podejmowanie decyzji przyczyniły się zwłaszcza wyniki badań D. Kahnemana i A. Tversky'ego¹¹. Na ich gruncie powstała teoria perspektywy badająca wpływ czynników psychologicznych (emocji, sposobów formułowania przewidywań i konstruowania preferencji) na podejmowanie decyzji, zwłaszcza tych o charakterze finansowym. Podstawą tej teorii jest założenie, że wybory ludzi są niestałe i zależą od kontekstu¹², zwłaszcza od tego, czy podejmujemy decyzję w warunkach ponoszenia strat, czy w sytuacji osiągania zysków¹³.

Do znaczenia roli własnego interesu i egoizmu w teorii A. Smitha, a jednocześnie znaczenia zasad etycznych w gospodarce powrócił A. Sen. Pisząc o racjonalności wyborów zauważał, że muszą istnieć przekonujące relacje między celami i zamiarami postępowania i samo pojęcie użyteczności jest niewystarczające do analizy ekonomicznych zachowań¹⁴. Argumentował, że w zachowaniach ekonomicznych ludzi (wymianie dóbr) musi istnieć pewne zaufanie. Wracając do wspomnianego cytatu z dzieła A. Smitha, A. Sen zauważał, że konsumenci kupując wyroby rzeźnika, piwowara, czy piekarza ufają, że są one odpowiedniej jakości, a surowce, z których zostały wyprodukowane nie zostały wymienione na gorsze. Z kolei producenci muszą mieć zaufanie do konsumentów, licząc na określony popyt z ich strony. Poza tym, oprócz dóbr prywatnych, w gospodarce

⁹ G. Gigerenzer, P. M. Todd, *Simple Heuristics That Make Us Smart*, Oxford University Press, New York 1999.

¹⁰ Na kwestię stosowania heurystyk uwzględniających problemy moralne zwrócił uwagę C. R. Sunstein wskazując na ich często małą wiarygodność i podatność na błędy.

¹¹ D. Kahneman, A. Tversky, *Prospect Theory: An Analysis of Decision under Risk*, „Econometrica” 1979, no. 47, s. 263–91.

¹² Interesującą z punktu widzenia wyborów moralnych, ilustracją efektu obramowania, czyli uzależnienia podjęcia decyzji od sposobu przedstawienia sytuacji decyzyjnej jest „problem choroby azjatyckiej” opisany przez A. Tversky'ego i D. Kahnemana w pracy: A. Tversky, D. Kahneman, *The Framing of Decisions and the Psychology of Choice*, „Science” 1981, no. 211(4481), s. 453–458.

¹³ Podobne zachowanie, jak w eksperymentach D. Kahnemana i A. Tversky'ego zostało zaobserwowane w kwestii etyki podatkowej np. w badaniach opisanych w pracy: O. Chang, D. Nichols, J. Schulz, *Taxpayer Attitude Toward Tax Audit Risk*, „Journal of Economic Psychology” 1987, no. 8, s. 299–309. Znacznie częściej na oszustwo podatkowe decydowali się podatnicy rozliczający się z urzędem skarbowym w sytuacji konieczności dopłaty podatku niż mający już nadpłatę. Ci pierwsi działali w obliczu strat i podejmowali ryzyko, ci drudzy znajdowali się w sytuacji osiągnięcia zysku i wykazywali awersję do ryzyka.

¹⁴ A. Sen, *Rational Behaviour*, [w:] J. Eatwell, M. Milgate (red.), *The New Palgrave: Utility and Probability*, W. W. Norton, New York 1990, s. 198–216.

i społeczeństwie mamy do czynienia z dobrami publicznymi, wspólnymi dla wszystkich, jak np. czyste środowisko. Produkcja takich dóbr nie jest motywowana egoizmem i korzyściami osobistymi. Nie sposób zatem wyobrazić sobie funkcjonowania społeczeństwa, w którym wszyscy obywatele kierowaliby się wyłącznie maksymalizacją osobistych korzyści. Ponadto zawsze istnieje pewna grupa osób motywowanych w swoim postępowaniu chęcią czynienia dobra dla innych¹⁵.

W szerszym społecznym kontekście włączenia zasad moralnych do organizacji społeczeństwa warto nawiązać do Th. Hobbesa, który w opublikowanym w 1651 r. *Lewiatanie*¹⁶ przedstawił teorię umowy społecznej. Według niej, ludzie kierują się wyłącznie własnym interesem i egoizmem, co prowadzi do nieustannego konfliktu, stanu natury będącego walką o przetrwanie. W oczywisty sposób taka sytuacja bardzo utrudnia realizację celów poprzez brak zabezpieczenia korzyści płynących ze współpracy. W interesie każdego człowieka jest opanowanie tego chaosu, aby móc bezpiecznie realizować swoje cele. Można to osiągnąć poprzez dokonanie racjonalnego wyboru, polegającego na ustanowieniu struktur władzy oraz wprowadzeniu zasad moralnych. Zasady te powinny dotyczyć: ograniczeń w dążeniu do realizacji wyłącznie własnych interesów oraz uwzględniania wzajemnych korzyści oraz zaufania. Właśnie osiąganie wzajemnych korzyści jest kluczową cechą moralności. Dla Th. Hobbesa moralność jest rodzajem kontraktu, umowy kooperacyjnej, którą strony zawierają w celu realizacji własnych interesów. Dla osiągnięcia tego celu niezbędne jest honorowanie warunków umowy przez wszystkie strony.

Opis stanu natury i umowy społecznej przedstawione przez Th. Hobbesa oraz ogólnie, możliwe współzależności występujące między racjonalnym ekonomicznym postępowaniem a zasadami moralnymi są interesującym obszarem zastosowań teorii gier. Do analizy skutków wyborów indywidualnych i społecznych w warunkach współpracy lub nie uczestników gry (stron umowy), można wykorzystać m.in. jeden z najbardziej znanych problemów teorii gier, *dylemat więźnia*.

2. DYLEMAT WIĘŹNIA JAKO PROBLEM ETYCZNY

Teoria gier umożliwia modelowanie podejmowania decyzji (wybór strategii) w sytuacjach, w których uczestnicy (gracze) zachowują się świadomie i racjonalnie, dążąc do podjęcia optymalnych decyzji. Konsekwencje działań podejmowanych w danej sytuacji zależą od działań pozostałych uczestników, dlatego podejmując decyzję dany uczestnik musi przewidywać decyzje pozostałych uczestników zakładając racjonalność ich postępowania. W teorii gier anali-

¹⁵ A. Sen, *Does Business Ethics Make Economic Sense?*, „Business Ethics Quarterly” 1993, vol. 3, issue 1.

¹⁶ Th. Hobbes, *Lewiatan, czyli materia, forma i władza państwa kościelnego i świeckiego*, Fundacja Aletheia, Warszawa 2005.

zie podlegają głównie sytuacji konfliktu interesów uczestników oraz sytuacji, w których interesy są zbieżne, ale z racjonalnego punktu widzenia niemożliwe jest uzgodnienie podejmowanych decyzji¹⁷.

Teoria gier znalazła szerokie zastosowania w ekonomii, biologii, ewolucyjnej, socjologii, naukach politycznych. Teoria gier jest wykorzystywana również w etyce do m.in. badania funkcjonowania teorii umowy społecznej¹⁸ oraz wytworzenia i uwzględniania norm moralnych w interakcjach racjonalnie działających podmiotów¹⁹.

Klasycznym, często cytowanym w odniesieniu do różnych dziedzin, problemem teorii gier jest *dylemat więźnia*²⁰. W problemie tym dostępne strategie określić można jako działanie kooperacyjne, na rzecz wspólnego dobra oraz działanie egoistyczne, wyłącznie na własną korzyść.

Historia przedstawia sytuację dwóch mężczyzn zatrzymanych przez policję, podejrzanych o współudział w popełnieniu przestępstwa. Brak jednak dostatecznych dowodów na ich winę. Zostali oni umieszczeni w osobnych pomieszczeniach. Każdego z nich poinformowano, że jeśli tylko jeden z nich się przyzna i będzie zeznawać przeciw drugiemu, wówczas ten, który się przyznał zostanie uwolniony, zaś ten drugi zostanie skazany na 20 lat więzienia. Jeśli obaj się przyznają, każdy z nich otrzyma karę 5 lat więzienia. Jeśli żaden się nie przyzna, wówczas każdy z nich otrzyma karę roku więzienia za wykroczenie. Macierz wypłat w postaci ilości lat do odsiedzenia w więzieniu przedstawia tab. 1.

Tabela 1. Macierz wypłat w dylemacie więźnia

		Decyzja podejrzanego B	
		Przyznać się	Nie przyznać się
Decyzja podejrzanego A	Przyznać się	(-5, -5)	(0, -20)
	Nie przyznać się	(-20, 0)	(-1, -1)

Źródło: opracowanie własne.

¹⁷ Szersze omówienie różnych aspektów teorii gier można znaleźć np. w pracy: M. Malawski, A. Wieczorek, H. Sosnowska, *Konkurencja i kooperacja. Teoria gier w ekonomii i naukach społecznych*, Wydawnictwo Naukowe PWN, Warszawa 2011.

¹⁸ Pionierska publikacja w tej dziedzinie szeroko wykorzystująca teorię gier to praca: D. Gauthier, *Morals by Agreement*, Clarendon Press, Oxford 1986, w której (jak również w innych pracach) D. Gauthier nawiązywał do teorii umowy społecznej Th. Hobbesa.

¹⁹ W pracy: B. Skyrms, *The Stag Hunt and the Evolution of Social Structure*, Cambridge University Press, Cambridge 2004, do analiz wykorzystana została ewolucyjna teoria gier.

²⁰ Pierwotnie problem *dylematu więźnia* został sformułowany przez dwóch pracowników RAND Corporation, Dreshera i Flooda w 1950 r. natomiast później został sformalizowany przez Alberta Tuckera.

Strategiami dominującymi tzn. najlepszymi ze wszystkich możliwych strategii, niezależnie od decyzji, jaką podejmie drugi gracz, jest strategia „Przyznać się”. Równowaga Nasha, czyli sytuacja, w której żaden z graczy nie może poprawić swojego wyniku poprzez jednostronną zmianę wybranej strategii, jest w tym przypadku osiągnięta wtedy, gdy każdy z podejrzanych przyzna się do winy. Natomiast punkt $(-1, -1)$, będący optimum Pareto, nie jest punktem równowagi, gdyż każdy z podejrzanych może wtedy poprawić swoją sytuację zmieniając jednostronnie swoją decyzję i przyznać się do winy zdradzając współtowarzysza. W tej grze każdy gracz może zyskać zdradzając drugiego (jeśli ten nie zmieni swojej strategii). Jeśli obaj zdradzają się nawzajem, wówczas obaj tracą w porównaniu z wyborem strategii kooperacyjnej, nieprzyznawania się. Dylemat więźnia ilustruje konflikt pomiędzy racjonalnością indywidualną (obaj podejrzani się przyznają) prowadzącą do równowagi Nasha, a racjonalnością grupową (współpraca, obaj podejrzani się nie przyznają) prowadzącą do optimum Pareto.

Szersze rozumienie dylematu więźnia i znaczenia poszczególnych elementów macierzy wypłat, również w kontekście etycznym, zawiera tab. 2.

Tabela 2. Macierz wypłat w uogólnionym dylemacie więźnia

		Decyzja B	
		N	T
Decyzja A	N	(KN, KN)	(PN, SBNA)
	T	(SANB, PN)	(NT, NT)

N – przyznanie się do winy, brak współpracy (zdrada), nieprzestrzeganie umowy, nieetyczne postępowanie, nadużycie zaufania drugiej osoby;

T – nieprzyznanie się do winy, współpraca, przestrzeganie umowy, etyczne postępowanie, zaufanie do drugiej osoby;

KN – konsekwencje jednoczesnego postępowania N przez obie osoby;

PN – pokusa dla postępowania N;

NT – nagroda za jednoczesne postępowanie T przez obie osoby;

SBNA (SANB) – strata dla B (A) wynika z postępowania N przez A (B), w przypadku dylematu więźnia SBNA = SANB.

Źródło: opracowanie własne.

Prostą ekonomiczną ilustracją powyższego problemu może być sytuacja, w której dwa przedsiębiorstwa działają na tym samym rynku, którego wartość oceniana jest perspektywicznie na 100 mln zł w przypadku przeprowadzenia odpowiednich działań marketingowych. Załóżmy, że podział tego rynku zależy od wydatków marketingowych każdego przedsiębiorstwa, które mogą wynieść

7 mln lub 10 mln zł (osiągnięcie sprzedaży na poziomie 100 mln zł łącznie jest możliwe zarówno, poprzez wydatek 7 mln zł, jak i 10 mln zł przez każde z przedsiębiorstw). Jeśli koszty marketingu w obu przedsiębiorstwach będą takie same, wówczas udział w rynku będzie też równy (50 mln zł). Przedsiębiorstwo, które wyda więcej będzie mogło przejąć 80% rynku. Przedsiębiorstwa mogą współpracować i zawrzeć umowę w kwestii wydatków marketingowych (zakładając, że nie skutkuje to naruszeniem prawa antymonopolowego). Rentowność sprzedaży każdego przedsiębiorstwa wynosi 30%. Macierz wypłat dla tego przykładu zawiera tab. 3.

Tabela 3. Macierz wypłat (zysk/strata w mln zł) dla przykładu dwóch przedsiębiorstw w zależności od wielkości wydatków marketingowych

		Firma B	
		Wydatek 10 mln zł	Wydatek 7 mln zł
Firma A	Wydatek 10 mln zł	(5, 5)	(14, -1)
	Wydatek 7 mln zł	(-1, 14)	(8, 8)

Źródło: opracowanie własne.

Najlepszym rozwiązaniem dla obu przedsiębiorstw jednocześnie jest przeznaczanie na marketing mniejszej kwoty 7 mln zł. Wymaga to jednak przestrzegania umowy (formalnej bądź nie) i wzajemnego zaufania. Konsekwencje nierepektowania umowy przez obie strony skutkują gorszym wynikiem, a pokusa, żeby to zrobić i zwiększyć wydatki, a tym samym udział w rynku i zysk jest duża.

Zachowanie uczestników w prostej wersji jednookresowej *dylematu więźnia* jest przewidywalne – zdecydowana większość tego typu sytuacji kończy się wyborem strategii prowadzących do równowagi Nasha. Znacznie ciekawsze są bardziej złożone sytuacje bazujące na iterowanej (wielookresowej) rozgrywce. Są to okoliczności bardziej realistyczne, uwzględniające fakt, że człowiek lub przedsiębiorstwo funkcjonuje w określonym środowisku, w którym wzajemne relacje są częste, a ich konsekwencje mają znaczenie dla kolejnych okresów. Robert Axelrod²¹ przeprowadził eksperyment komputerowy dla wyłonienia najlepszej strategii w iterowanym *dylemacie więźnia*. Poprosił ekspertów z teorii gier o nadesłanie programów komputerowych z różnymi strategiami dotyczącymi sytuacji, której macierz wypłat została przedstawiona w tab. 4.

W pierwszej rundzie eksperymentu testowano 14 strategii, w drugiej 52 (po znanych wynikach pierwszej rundy dołączono dodatkowe bardziej skompli-

²¹ R. Axelrod, *The Evolution of Cooperation*, Basic Books, New York 1984.

kowe strategie). Każda strategia została skonfrontowana ze wszystkimi pozostałymi w grach liczących 200 iteracji (okresów). W obu rundach zwyciężyła strategia wet za wet polegająca na rozpoczynaniu gry od kooperacji, a w kolejnych iteracjach na wyborze takiej strategii, którą wybrał drugi gracz w poprzedniej iteracji²².

Tabela 4. Macierz wypłat w eksperymencie R. Axelroda, iterowanym *dylemacie więźnia*

		Gracz B	
		Zdrada	Współpraca
Gracz A	Zdrada	(1\$, 1\$)	(5\$, 0\$)
	Współpraca	(0\$, 5\$)	(3\$, 3\$)

Źródło: R. Axelrod, *The Evolution of Cooperation*, Basic Books, New York 1984.

Z obserwacji R. Axelroda wynikało, że najlepsze strategie charakteryzowały się czterema cechami:

- *przyjaznością* – współpracuj, nigdy nie zdradzaj pierwszy,
- *odwzajemnieniem* – reaguj natychmiast na zachowanie drugiej strony, w przypadku zdrady stosuj odwet i zdradzaj w następnej iteracji, w przypadku współpracy – współpracuj,
- *brakiem zazdrości* – nie myśl o tym, żeby osiągnąć lepszy wynik niż druga strona, skoncentruj się na maksymalizacji swojego wyniku, a najlepszą drogą do tego celu jest założenie i promowanie współpracy,
- *przejrzystością* – nie szukaj skomplikowanych i wyrafinowanych strategii starając się przewidywać ruch drugiej strony; tylko strategia systematycznej współpracy może być lepsza od strategii wet za wet.

Wyniki eksperymentu R. Axelroda wykazały, że strategie altruistyczne miały znaczną przewagę nad strategiami egoistycznymi i że współpraca może prowadzić do równowagi w horyzoncie wielookresowym. Zasadniczo to właśnie długotrwałość relacji, a nie zaufanie jest podstawą współpracy. Zwracał na to uwagę D. Gauthier twierdząc, że moralność pojawia się w sposób naturalny w wyniku ciągłej interakcji osób zorientowanych na własny interes. Wykształcenie się norm moralnych staje się racjonalnym etapem budowy społeczeństwa, korzystnym dla wszystkich jego członków.

Warto dodać, że w ramach kolejnego eksperymentu R. Axelrod przeprowadził „ewolucyjny” konkurs strategii polegający na zwielokrotnianiu typów po-

²² Warto zauważyć, że taka strategia odpowiada niskiemu poziomowi skali rozwoju moralnego Kolberga, w stadium przedkonwencyjnym.

szczególnej strategii w kolejnej rundzie, w zależności od wyników danej strategii w rundzie poprzedniej. Z każdą kolejną rundą (generacją) pewne strategie stawały się coraz liczniejsze, inne coraz rzadsze. Ponownie najlepsze okazały się strategie typu *wet za wet*, charakteryzujące się przyjaznością, dążeniem do współpracy, ale jednocześnie reagujące na zdradę drugiej strony²³.

Autorami interesującego eksperymentu symulacyjnego wykorzystującego założenia iterowanego dylematu więźnia, ale wzbogaconego o różne strategie interakcji i postaw etycznych są R. Hill i A. Watkins²⁴.

Stu wirtualnych uczestników zostało przydzielonych do dwóch rywalizujących grup przedsiębiorstw (każda grupa składała się z pięciu przedsiębiorstw, każde reprezentowane było przez 10 agentów). W pięćdziesięciu iteracjach, losowo wybrani agenci z obu grup przedsiębiorstw rozgrywali między sobą grę (wchodzili w interakcję podejmując decyzję o współpracy lub zdradzie) zgodnie z macierzą wypłat w eksperymencie R. Axelroda (por. tab. 4). W eksperymencie uwzględniono trzy etyczne orientacje przedsiębiorstw oraz pięć typów uproszczonej moralności osobistej agentów reprezentujących te przedsiębiorstwa. Typy moralności były następujące:

- *altruista* – zawsze współpracujący niezależnie od zachowania przeciwnika w poprzedniej interakcji,
- *realistyczny altruista* – współpracujący, jeśli w ostatnich pięciu interakcjach przeciwnicy współpracowali co najmniej dwa razy, zdradzający w przeciwnym razie (w początkowej fazie symulacji, przy braku dostatecznej historii interakcji, realistyczny altruista zawsze współpracował),
- *wet za wet* – agent zawsze odwzajemniający zachowanie przeciwnika z poprzedniej interakcji,
- *realistyczny egoista* – działający symetrycznie w stosunku do realistycznego altruisty tj. zdradzający, jeśli w ostatnich pięciu interakcjach przeciwnicy zdradzali co najmniej dwa razy, w innym przypadku współpracujący (w początkowej fazie symulacji, przy braku dostatecznej historii interakcji realistyczny egoista zawsze zdradzał),
- *egoista* – zawsze zdradzający niezależnie od zachowania przeciwnika w poprzedniej interakcji.

²³ W dwudziestolecie publikacji wyników eksperymentów R. Axelroda zorganizowano konkurs, w którym wyniki okazały się odmienne – wygrały strategie faworyzujące zdradę. Wynikało to jednak z faktu, że dopuszczono możliwość zgłaszania wielu programów (strategii) przez jednego uczestnika. Zespół z Uniwersytetu w Southampton zgłosił kilkadziesiąt programów, które działały różnie w zależności od tego, czy przeciwnikiem był „swój”, czy „obcy” program, maksymalizując wygrane w rozgrywkach między sobą i starając się minimalizować wyniki przeciwników w rozgrywce z „obcymi” programami.

²⁴ R. Hill, A. Watkins, *The Impact of Personal and Organizational Moral Philosophies on Marketing Exchange Relationships: A Simulation Using the Prisoner's Dilemma Game*, „Journal of Business Ethics” 2005, no. 62, s. 253–265.

W części symulacji dopuszczalna była zmiana moralności agenta w zależności od orientacji etycznej reprezentowanej przez przedsiębiorstwo. Wyróżniono trzy orientacje:

- *neutralną* – agent zmieniał swoje postępowanie w interakcjach w kierunku moralności reprezentowanej przez najlepiej zarabiających agentów w jego przedsiębiorstwie,

- *etyczną* – agent zmieniał swoje postępowanie w interakcjach w kierunku moralności najlepiej zarabiających agentów w jego przedsiębiorstwie pod warunkiem, że byli to altruści,

- *nieetyczną* – agent zmieniał swoje postępowanie w interakcjach w kierunku moralności najlepiej zarabiających agentów w jego przedsiębiorstwie pod warunkiem, że byli to egoiści.

Autorzy przeprowadzili szereg symulacji zmieniając proporcje agentów o różnych typach moralności w przedsiębiorstwach oraz orientacje etyczne przedsiębiorstw. Każda symulacja składała się ze stu rund (każda runda z pięćdziesięciu iteracji – interakcji wszystkich agentów obu grup przedsiębiorstw).

Wyniki eksperymentu okazały się generalnie zgodne z wynikami konkursu R. Axelroda. Relatywnie (w zależności od scenariusza symulacyjnego) najlepsze wyniki uzyskiwali agenci o umiarkowanych perspektywach moralnych z tendencją w kierunku zachowań altruistycznych. To spostrzeżenie dotyczyło również symulacji z uwzględnieniem różnych orientacji etycznych przedsiębiorstw. W scenariuszu z orientacją etyczną oraz zróżnicowanymi zespołami agentów w przedsiębiorstwach, bezwzględna przewaga przedsiębiorstw z początkową, zwiększoną frakcją altruistów w stosunku do przedsiębiorstw ze zwiększoną frakcją egoistów, była (zarówno pod względem łącznego dochodu, jak i końcowej liczby altruistów) wyraźnie większa niż przewaga przedsiębiorstw z początkową, zwiększoną frakcją egoistów w stosunku do przedsiębiorstw ze zwiększoną frakcją altruistów w scenariuszu z orientacją nieetyczną.

Dylemat więźnia pozwala na wypracowanie etycznych decyzji, zwłaszcza w sytuacji, kiedy zachodzi konflikt między interesem indywidualnym i grupowym. Wykonywanie interesujących ćwiczeń i analizę procesu podejmowania takich decyzji proponuje H. S. James²⁵. Eksperymenty te są pewnym wariantem *dylematu więźnia* z rozszerzoną liczbą uczestników (grupą). Jeden z nich polega na podejmowaniu przez każdego członka grupy jednej z dwóch decyzji: głosowania „0” lub głosowania „1”. Wielkości wypłat dla każdego uczestnika w tym ćwiczeniu przedstawia tab. 5.

²⁵ H. S. James, *Using the Prisoner's Dilemma to Teach Business Ethics When Personal and Group Interests Conflict*, „Teaching Business Ethics” 1998, vol. 2(2), s. 211–222.

Tabela 5. Wielkości indywidualnych wypłat dla głosujących

Liczba głosujących „0”	Wypłata dla głosującego „0”	Wypłata dla głosującego „1”
0	0	4
1	2	6
2	4	8
3	6	10
4	8	12
5	10	14
6	12	16
itd.	itd.	itd.

Źródło: opracowanie własne na podstawie: H. S. James, *Using the Prisoner's Dilemma to Teach Business Ethics When Personal and Group Interests Conflict*, „Teaching Business Ethics” 1998, vol. 2(2)²⁶.

Każdy z uczestników podejmuje decyzję o głosowaniu w sposób poufny i w tym samym czasie. Wielkość wypłaty zależy od dwóch czynników: indywidualnego wyboru oraz od liczby głosujących „0”. Jak można zauważyć każdy z uczestników ma pokusę, żeby głosować „1” i nadzieję, że inni będą głosować „0”. To stanowi o konflikcie między interesem grupowym i indywidualnym. Jeśli np. grupa liczyła 5 uczestników, z których 4 głosowało „0”, a jeden „1”, wówczas głosujący „0” otrzymali po 8, zaś głosujący „1” otrzymał 12. Łączny wynik grupy to 44. Gdyby wszyscy głosowali „0” każdy otrzymałby po 10, w sumie 50. Warto też zauważyć, że gdyby dodatkowo jeszcze dwie osoby głosowały „1” (w sumie trzy z pięciu), wówczas otrzymałyby one tylko po 8, a łączny wynik wyniósłby tylko 32. Traktując głosowanie „1”, jako postępowanie nieetyczne (brak współpracy) widać wyraźnie, że wpływa ono zarówno na pogorszenie wyniku indywidualnego, a także, i to w znacznie większym stopniu, na ogólny wynik grupy.

Powyższy przykład wskazuje na znaczenie i możliwość kształtowania współpracy i zaufania w funkcjonowaniu przedsiębiorstw, czy ogólnie organizacji. Wypracowanie zasad etycznych w rzeczywistym funkcjonowaniu organizacji zwiększa efektywność jej działania. Jak zauważają M. Kaptein i J. Wempe²⁷ kodeksy etyczne pozostają tylko słabo oddziałującym zbiorem zapisanych zdań, zakazów i nakazów, jeśli nie stanowią odzwierciedlenia obserwowanych zachowań i wartości. Na problem rozbieżności między obserwacją wielu zachowań nieetycznych w instytucjach finansowych a rosnącą liczbą instytucji posiadają-

²⁶ Wielkości wypłat zostały nieco zmienione, bez wpływu na ideę ćwiczenia.

²⁷ M. Kaptein, J. Wempe, *Twelve Gordian Knots When Developing an Organizational Code of Ethics*, „Journal of Business Ethics” 1998, vol. 17, s. 853–869.

cych kodeksy etyczne zwraca uwagę J. Dobson²⁸. Jego zdaniem wynika to z akulturacji, przejmowania systemu wartości moralnych przez pracowników danej instytucji poprzez obserwację akceptowanych zachowań. Często mogą to być wartości sprzeczne z wcześniejszą edukacją i doświadczeniem pracowników. Stąd niezbędne są programy szkoleniowe zmierzające do internalizacji u pracowników norm postępowania zapisanych w kodeksach.

Niezwykle ważnym i złożonym aspektem analizy i wyjaśnienia *dylematu więźnia* od strony etycznej jest kwestia motywów postępowania, rozpatrywana zarówno na gruncie teorii teleologicznych, jak i deontologicznych. Z dotychczasowej analizy wynika, że głównym motywem postępowania w *dylemacie więźnia* jest racjonalne dążenie do maksymalizacji korzyści osobistych bądź grupowych. Działanie z perspektywy konsekwencjalizmu może prowadzić do osiągnięcia optimum Pareta (użyteczność), bądź nie (egoizm). Z perspektywy deontologicznej, przyjmując, że etycznym postępowaniem jest współpraca, trudno jest znaleźć argumenty moralnie uzasadniające przeciwne postępowanie. Przekonany był o tym A. Sen podkreślając, że z pewnością żadna z osób nie chciałaby, żeby zdrada stała się uniwersalną praktyką, zaś aby taką była współpraca²⁹. Wydaje się to być zgodne z imperatywem kategorycznym Kanta i jest akceptowane przez większość badaczy. Jednak M. White³⁰ zwraca uwagę, że warunkiem koniecznym byłoby wcześniejsze zobowiązanie do współpracy. W przypadku braku takiego zobowiązania sprawa jest dyskusyjna, gdyż skłonność do myślenia wskazanego przez A. Sena jest bliższa słabszej, złotej regule etycznej („Traktuj innych tak, jak Ty byś chciał być traktowany”) niż imperatywowi kategorycznemu³¹. Poza tym zachowanie nieetyczne (brak współpracy) nie narusza zasady promowania człowieczeństwa, gdyż (w przypadku braku wcześniejszego zobowiązania) nie oznacza to wykorzystania drugiej osoby, potraktowania jej jako środka do osiągnięcia własnych korzyści.

Podsumowując, warto zauważyć, że postępowanie nieetyczne (brak współpracy) jest korzystne dla takiej osoby tylko wtedy, kiedy inni postępują etycznie. Uogólniając, oszustwa, kradzieże opłacają się tylko wtedy, gdy inni tego nie robią. Gdyby wszyscy oszukiwali i okradali się nawzajem szybko doszliby do wniosku, że korzystniej dla wszystkich będzie tego nie robić³².

²⁸ J. Dobson, *Why Ethics Codes Don't Work*, „Financial Analysts Journal” 2003, vol. 59(6), s. 29–34.

²⁹ A. Sen, *Choice, Orderings and Morality*, [w:] S. Körner (red.), *Practical Reason*, Basil Blackwell, Oxford 1974.

³⁰ M. White, *Kantian Ethics and the Prisoners' Dilemma*, „Eastern Economic Journal” 2009, vol. 35, s. 137–143.

³¹ Kant krytykował złotą regułę etyczną, jako wynikającą z subiektywnej skłonności, a nie obiektywnego i powszechnego obowiązku.

³² Stanowi to proste nawiązanie do teorii umowy społecznej Th. Hobbesa, por. s. 73.

PODSUMOWANIE

Dylemat więźnia jest interesującym problemem pozwalającym na analizę zachowań ekonomicznych rozpatrywanych z uwzględnieniem zasad etycznych. Wprowadzenie takich zasad umożliwia osiągnięcie wyników optymalnych, zarówno z punktu widzenia interesów indywidualnych, jak i grupowych. Motywy postępowania mogą być uzasadnione zarówno na gruncie teorii deontologicznych, jak i teleologicznych. W bardziej złożonych eksperymentach symulacyjnych (R. Axelroda oraz R. Hilla i A. Watkins) opartych na *dylemacie więźnia* widoczne są w najlepszych rozwiązaniach również tendencje do szukania współpracy, czyli zachowań etycznych opartych jednak na praktycznej zasadzie odwzajemniania. *Dylemat więźnia* jest dobrym problemem do analizy i kształtowania zasad etycznych. Ćwiczenie i analiza szeregu eksperymentów wykorzystujących *dylemat więźnia* mogłoby przyczynić się do lepszego poznania problemów etycznych wśród studentów kierunków ekonomicznych i pracowników korporacji. Prowadziłoby to do lepszego zrozumienia sensu racjonalnego postępowania i towarzyszących mu zasad moralnych, zwłaszcza w kontekście konfliktu interesów indywidualnych i grupowych.

BIBLIOGRAFIA

- Axelrod R., *The Evolution of Cooperation*, Basic Books, New York 1984.
- Chang O., Nichols D., Schulz J. *Taxpayer Attitude Toward Tax Audit Risk*, „Journal of Economic Psychology” 1987, no. 8.
- Chong S., Humble J., Kendall G., Li J., You X., *The Iterated Prisoner's Dilemma: 20 Years on*, „Advances in Natural Computation” 2007, vol. 4.
- Copeland T. E., Weston J. F., *Financial Theory and Corporate Policy*, Addison Wesley, New York 1988.
- Dobson, J. *Why Ethics Codes Don't Work*, „Financial Analysts Journal” 2003, vol. 59(6).
- Gauthier D., *Morals by Agreement*, Clarendon Press, Oxford 1986.
- Gigerenzer, G., Todd, P. M., & the ABC Research Group, *Simple Heuristics That Make Us Smart*, Oxford University Press, New York 1999.
- Hill R., Watkins A., *The Impact of Personal and Organizational Moral Philosophies on Marketing Exchange Relationships: A Simulation Using the Prisoner's Dilemma Game*, „Journal of Business Ethics” 2005, no. 62.
- Hobbes Th., *Lewiatan, czyli materia, forma i władza państwa kościelnego i świeckiego*, Fundacja Aletheia, Warszawa 2005.
- James H.S., *Using the Prisoner's Dilemma to Teach Business Ethics When Personal and Group Interests Conflict*, „Teaching Business Ethics” 1998, vol. 2(2).
- Kahneman D., Tversky A., *On the Psychology of Prediction*, „Psychological Review” 1973, no. 80.
- Kahneman D., Tversky A., *Prospect Theory: An Analysis of Decision under Risk*, „Econometrica” 1979, no. 47.
- Kaptein M., Wempe J., *Twelve Gordian Knots When Developing an Organizational Code of Ethics*, „Journal of Business Ethics” 1998, vol. 17.

- Kozielecki J., *Psychologiczna teoria decyzji*, PWN, Warszawa 1977.
- Malawski M., Sosnowska H., Wieczorek A., *Konkurencja i kooperacja. Teoria gier w ekonomii i naukach społecznych*, Wydawnictwo Naukowe PWN 2011.
- Sen A., *Choice, Orderings and Morality*, [w:] S. Körner (red.), *Practical Reason*, Basil Blackwell, Oxford 1974.
- Sen A., *Rational Behaviour*, [w:] J. Eatwell, M. Milgate (red.), *The New Palgrave: Utility and Probability*, W. W. Norton, New York 1990.
- Sen A., *Does Business Ethics Make Economic Sense?*, „Business Ethics Quarterly” 1993, vol. 3, issue 1.
- Simon H. A., *Models of Man: Social and Rational*, Wiley, New York 1957.
- Skyrms B., *The Stag Hunt and the Evolution of Social Structure*, Cambridge University Press, Cambridge 2004.
- Smith A., *Badania nad naturą i przyczynami bogactwa narodów*, Wydawnictwo Naukowe PWN, Warszawa 2007.
- Smith A., *Teorii uczuć moralnych*, Wydawnictwo Naukowe PWN, Warszawa 1989.
- Sunstein C. R., *Moral heuristics*, „Behavioral and Brain Sciences” 2005, no. 28.
- Szulczewski G., *Krytyka „etyki sympatii” i „problem A. Smitha” w świetle lektury Teorii uczuć moralnych*, „Prakseologia” 2015, nr 157, t. 1.
- Tversky A., Kahneman D., *The Framing of Decisions and the Psychology of Choice*, „Science” 1981, no. 211(4481).
- Von Neumann J., Morgenstern O., *Theory of Games and Economic Behavior*, Princeton University Press, Princeton (NJ) 2007.
- White M., *Kantian Ethics and the Prisoners' Dilemma*, „Eastern Economic Journal” 2009, vol. 35.
- Zaleskiewicz T., *Psychologia ekonomiczna*, Wydawnictwo Naukowe PWN, Warszawa 2011.

PRISONER'S DILEMMA AND ETHICAL DILEMMAS

Abstract: *The basis of many economic theories is the fundamental postulate of rationality – to maximize own's self-interest. Many of the economic behaviors concern human interactions eg. concluding agreements and contracts. For the sustainability of such economic system it is necessary to strengthen it with a system of moral principles. The aim of the paper is to indicate the possibility of using the prisoner's dilemma in analyzing and shaping ethical behavior through discussion of selected experiments described in the literature. They show how the rational actions of each of the participants in the game lead to sub-optimal results, while following collective rationality, consistent with ethics theories, allows to achieve optimum result, both from the point of view of the individual and collective interests. Motives can be justified both on the grounds of deontological and teleological theories. Exercise and analysis of a series of experiments using the prisoner's dilemma could contribute to a better understanding of the sense of rational behavior and the accompanying moral principles, especially in the context of conflict between individual and collective interests.*

Keywords: *prisoner's dilemma, game theory, ethics, morality, rationality.*