

# Paweł Więckowski

---

## Dylemat więźnia a kryzys moralności rozumianej jako kontrakt racjonalnych egoistów

---

Annales. Etyka w życiu gospodarczym 10/1, 267-275

---

2007

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej [bazhum.muzhp.pl](http://bazhum.muzhp.pl), gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

Paweł Więckowski  
Szkoła Główna Handlowa, Warszawa  
e-mail: platon2004@gazeta.pl

## **Dylemat więźnia a kryzys moralności rozumianej jako kontrakt racjonalnych egoistów**

### **1. Ideał racjonalnej moralności**

Dlaczego mielibyśmy nie kraść, oddawać długi, nie kłamać, dotrzymywać obietnic? Bo są to rzeczy złe i jako takie powinny być ponieczone – odpowiadała tradycyjna moralność Sokratesa, Platona, Arystotelesa czy św. Tomasza z Akwinu. Bo to się nie opłaca – odpowiedzieli filozofowie oświecenia – a nikt nie będzie działał wbrew własnej korzyści. Smith sądził, że uczucie sympatii oraz niewidzialna ręka rynku zastąpią moralność. David Hume nie był tak naiwny i precyzyjnie argumentował, że reguły poszanowania własności (które nazywał regułami sprawiedliwości), są przestrzegane w społeczeństwie, gdyż są dla wszystkich obywateli korzystne<sup>1</sup>. W następujących sytuacjach reguły poszanowania własności nie byłyby respektowane: gdyby dóbr starczało dla wszystkich bez ograniczeń; gdyby wszyscy chętnie się dzielili z innymi; gdyby zapanował powszechny chaos (wojna, demoralizacja) i nasze zachowania nie spotykałyby się z wzajemnością; gdybyśmy byli doskonale samowystarczalni, a także, gdy mamy do czynienia z istotami zależnymi od nas (np. ze zwierzętami). Te same poglądy, choć mniej precyzyjnie, wyraził współcześnie David Gauthier<sup>2</sup>.

### **2. Dylemat więźnia**

Czy jednak racjonalni egoiści – skoncentrowani na własnej korzyści oraz zdolni do kalkulacji start i zysków – potrafią, bez żadnej zewnętrznej sankcji, wynegocjować i wdrożyć społecznie korzystną moralność? W połowie XX w. Melvin Dresher i Merrill Flood wymyślili grę, która po dopisaniu do niej historyjki przez Alberta W. Tuckera stała się jednym z najczęściej dyskutowanych problemów współdziałania racjonalnych egoistów – był to dylemat więźnia<sup>3</sup>. Oto jedna z licznych wersji tej sytuacji (różniaca się od oryginalnej wersji Tuckera). Dwu więźniom podejrzanym o popełnienie przestępstwa i przebywającym w areszcie w oddzielnych celach przedstawiono propozycję nielojalności wobec współpodejrzanego i złożenia obciążających go zeznań. Jeśli obaj więźniowie pozostaną

---

<sup>1</sup> D. Hume, *Badania dotyczące zasad moralności*, PWN, Warszawa 1975, s. 19–45.

<sup>2</sup> D. Gauthier, *Morals by Agreement*, Oxford University Press, New York 1986.

<sup>3</sup> P.D. Straffin, *Teoria gier*, Scholar, Warszawa 2001, s. 94.

lojalni – każdy otrzyma po dwa lata więzienia. Jeśli tylko jeden zachowa się niełojalnie, wyjdzie po kilku miesiącach, podczas gdy jego lojalny kompan dostanie 20 lat więzienia. Jeśli obaj będą się wzajemnie oskarżać – każdy zostanie skazany na 10 lat. Odchodząc od tradycyjnych w teorii gier kostek, możliwe posunięcia tej gry można przedstawić następująco:

Konfiguracja	Więzień A otrzymuje	Więzień A wybiera	Więzień B wybiera	Więzień B otrzymuje
1	2 lata (R)	lojalność (C)	lojalność (C)	2 lata (R)
2	kilka miesięcy (T)	niełojalność (D)	lojalność (C)	20 lat (S)
3	20 lat (S)	lojalność (C)	niełojalność (D)	kilka miesięcy (T)
4	10 lat (U)	niełojalność (D)	niełojalność (D)	10 lat (U)

Ujmując rzecz nieco bardziej formalnie, gracze mają do wyboru współpracę C (*cooperation*) i uchylenie się od niej, czyli zdradę D (*defection*). Zależnie od zachowania drugiej strony mogą otrzymać wypłatę pokusy T (*temptation*), nagrodę R (*reward*), wypłatę niekooperacyjną U (*uncooperative*) lub wypłatę frajera S (*succer*), gdzie  $T > R > U > S$ .

Patrząc na sprawę zdroworozsądkowo i intuicyjnie, najkorzystniejsza dla więźniów byłaby obopólna współpraca. Jednak więzień decydujący się na współpracę, naraża się na zostanie frajerem, rezygnując zaś z niej, ma szansę na najwyższą wypłatę pokusy – jeśli jego partner zostanie frajerem. Zazwyczaj więc wynikiem gry jest konfiguracja 4, która i dla obu więźniów razem, i dla każdego z osobna jest gorsza niż konfiguracja 1. Jednak uzyskanie konfiguracji 1 wymagałoby gwarancji lojalności, której nikt nie jest w stanie uzyskać w tej grze. Konfiguracja 4 jest stanem równowagi Nasha (skądinąd bohatera filmu *Piękny umysł*), gdyż na gruncie indywidualnej racjonalności opiera się na optymalnym wyborze każdego z graczy – przy danym wyborze drugiego gracza niełojalność jest zawsze korzystniejsza niż lojalność (10 lat zamiast 20 i kilka miesięcy zamiast 2 lat). Innymi ilustracjami dylematu więźnia są dwa sklepy, które konkurują obniżając ceny (jeśli oba obniżą w porównywalnym stopniu, każdy z nich straci marżę, a żaden nie zyska klientów) czy dwa mocarstwa zwiększające wydatki na zbrojenia (jeśli oba zwiększą porównywalnie, stracą pieniądze nie zyskując przewagi).

### 3. Próby rozwiązania problemu dylematu więźnia na gruncie teorii gier

U podstaw liberalizmu leży postulat, by życie społeczne oprzeć na regułach będących wynikiem kontraktu racjonalnych egoistów maksymalizujących własną korzyść. Dylemat więźnia ukazuje, że przynajmniej niekiedy jest to niemożliwe – racjonalni egoiści, jeśli nawet uzgodnią kontrakt, nie będą go przestrzegać. Taka postawa nie musi doprowadzić społeczeństwa do upadku, lecz z pewnością znacznie osłabi możliwości jego rozwoju. A przecież nie to jest celem racjonalnych egoistów. Podjęto liczne próby opracowania strategii gry, która prowadziłaby do korzystniejszego rozwiązania, czyli do wyboru konfiguracji 1.

Jedną z prób było posłużenie się perspektywą ewolucyjną (wzorowaną na rywalizacji jastrzębi i gołębi) w celu wykazania, że gracze stosujący strategię niekooperacyjną nie mu-

szą zdominować społeczeństwa<sup>4</sup>. Wymagało to jednak przyjęcia wielu arbitralnych lub nierealistycznych założeń.

Inną próbą było zdefiniowanie metagier, w których wybór postawy uzależniony jest od wyboru strategii uzależnionej od przekonania o wyborze strategii przez przeciwnika. Owocowało to regułami w rodzaju: „*Kooperuje wtedy i tylko wtedy, gdy jesteś przekonany, że Twój przeciwnik będzie kooperował wtedy i tylko wtedy, gdy jest przekonany, że ty będziesz kooperował*”<sup>5</sup>. Ich użyteczność w realnym świecie jest wątpliwa.

Najpoważniejszą próbą było opracowanie strategii karania przeciwnika za niekooperacyjne zachowanie w dylemacie rozgrywanym wielokrotnie między tymi samymi graczami (iterowanym). Gracze karząc się wzajemnie za postawę niekooperatywną w kolejnych pojedynkach mieli wymuszać na sobie współpracę. Na początku lat 80. XX w. została ona opisana przez Roberta Axelroda, który wcześniej zorganizował kilka pojedynków programów komputerowych<sup>6</sup>. Zwycięzcą okazała się strategia A. Rapaporta „wet-za-wet” (*tit-for-tat*). Stosujący ją gracz rozpoczyna od kooperacji (strategia jest przyjazna), gdy spotka się z jej brakiem, również zaczyna być nielojalny (strategia jest też odwetowa), gdy zaś przeciwnik decyduje się jednak współpracować – w kolejnej grze wet-za-wet powraca do współpracy (strategia jest przebacząca). Ponadto wymagana jest przejrzystość – przeciwnik musi wiedzieć, czego się spodziewać (problemem może być to, dlaczego przeciwnik ma zaufać posiadanym informacjom – tu też powstaje możliwość celowego wprowadzenia w błąd, czyli nadużycia zaufania). Choć najskuteczniejsza, strategia ta ma jednak wiele wad – (1) uniemożliwia zwycięstwo (daje co najwyżej remis); (2) wymaga wielu rozdań, co w realnym świecie nie zawsze się zdarza (gdy oszukana firma bankrutuje, nie może już ukarać partnera w następnym rozdaniu; niekoniecznie też istnieje możliwość wielokrotnego rozgrywania gry z tym samym przeciwnikiem – chyba że są to dwa oligopole, jak producenci *Pesi-coli* i *Coca-coli*); (3) jeśli wiadomo, ile razy będzie rozegrany pojedynek, istnieje ryzyko, że współpraca w ogóle nie nastąpi (w ostatnim pojedynku nie warto współdziałać, bo można zostać frajerem i nie będzie można ukarać partnera za nielojalność; w takim razie nie warto też kooperować w przedostatnim pojedynku, bo w ostatnim i tak zostanie wybrana strategia karząca; w takim razie nie warto kooperować w przed-przedostatnim – i na zasadzie domina czy też wstecznej indukcji matematycznej – nie warto kooperować w żadnej); (4) strategia ta mimo wszystko czasem przegrywa.

Po pewnym czasie opracowano bezpieczniejszą wersję tej strategii – „wet-za-dwa-wety”. Gracz pragnący wymusić współpracę zrywa kooperację dopiero po drugim braku współpracy ze strony partnera, by w ten sposób uchronić się przed wejściem w spiralę odwetową zapoczątkowaną przypadkowym, jednorazowym brakiem współpracy<sup>7</sup>. Strategia ta może się jednak okazać bardzo niekorzystna, gdy przeciwnik będzie często stosował „przypadkowy” brak współpracy (jeśli zrobi to w co drugiej grze, gracz stosujący „wet-za-dwa-wety” będzie w połowie gier frajerem, a jednocześnie sam nigdy nie zerwie współpracy!).

---

<sup>4</sup> Hasło *Evolutionary Game Theory*, [w:] *Stanford Philosophical Encyclopedia*, <http://plato.stanford.edu/entries/game-evolutionary/>

<sup>5</sup> P.D. Straffin, *Teoria gier*, Scholar, Warszawa 2001, s. 100.

<sup>6</sup> R. Axelrod, *The Emergence of Cooperation Among Egoists*, „The American Political Science Review” 1981, No. 75, s. 306–318; R. Axelrod, *The Evolution of Cooperation*, Basic Books, New York 1984.

<sup>7</sup> Hasło *Prisoner's Dilemma*, [w:] *Stanford Philosophical Encyclopedia*, <http://plato.stanford.edu/entries/prisoner-dilemma/>

Obecnie opracowuje się strategie ewolucyjne (uczące się), które podejmują decyzję o współpracy analizując wzorzec zachowań przeciwnika w wielu poprzedzających grach<sup>8</sup>. Wymaga to dużej inwestycji czasu.

#### 4. Rozszerzony dylemat więźnia

Analizy dokonywane na gruncie teorii gier są w realnym świecie niemal bezużyteczne – nie ma strategii gwarantujących współpracę między racjonalnymi egoistami. Może przynajmniej dylematy więźnia nie są zbyt powszechne? Wówczas, choć kłopotliwe, nie wpływałyby znacząco na funkcjonowanie społeczeństwa. Tak jednak nie jest – każdy przypadek oszustwa i wykorzystania gotowości innych do współpracy stwarza podobny problem. Klasyczny dylemat więźnia jest tylko szczególnym przypadkiem dylematów społecznego współdziałania. W 1968 r. Garrett Hardin opisał inny rodzaj dylematów oparty jednak na podobnej zasadzie – kierując się doraźnym interesem własnym, gracze narażają dobro wspólnoty, co po pewnym czasie zemści się także na nich<sup>9</sup>. Nazwał to *Tragedy of the Commons*, co różnie tłumaczone jest na polski – m.in. jako tragedia wspólnego pastwiska, łowiska, wspólnego dobra. Oto kilka przykładów:

- Grupa ludzi wybiera się do restauracji i decyduje, że rachunek podzielą równo między siebie. Każdy zamawia jak najwięcej wiedząc, że jeśli zamówi mało, zapłaci za innych, którzy zamówili dużo. W rezultacie każdy płaci bardzo dużo.
- Każdy z rybaków łowi za dużo, bo boi się zbankrutować (jeśli złowi mało, inni – łowiąc więcej – uzyskają tańsze koszty połowu). W rezultacie wspólnie doprowadzają do zagłady wszystkie ryby.
- Wsie wylewają ścieki do jeziora, by zaoszczędzić na budowie oczyszczalni. Po jakimś czasie wszyscy muszą złożyć się na oczyszczenie jeziora, co kosztuje znacznie więcej niż budowa oczyszczalni.
- Pasażerowie nie płacą za bilet, bo liczą, że inni zapłacą i utrzymają linię. W rezultacie linia staje się deficytowa i zostaje zlikwidowana.
- Granat wpada do pokoju pełnego ludzi. Każdy może uratować wszystkich pozostałych rzucając się na niego i ginąc, lecz nikt się nie decyduje i w rezultacie giną wszyscy.
- Wyborca nie idzie na wybory kalkulując, że jeden głos nie ma wpływu na wynik, można więc zaoszczędzić czas. W rezultacie wybrany zostaje kandydat popierany przez skrajną mniejszość i większość jest niezadowolona.
- Pacjenci domagają się drogich i niepotrzebnych badań w ramach ubezpieczenia, co w rezultacie prowadzi do ciągłego wzrostu opłaty za ubezpieczenie.
- W czasie suszy można by utrzymać dostawy wody, gdyby wszyscy zdecydowali się ograniczyć jej zużycie. Każdy jednak liczy, że to inni je ograniczą. W rezultacie nikt go nie ogranicza i po krótkim czasie woda zostaje odcięta wszystkim.
- W zakładzie rządzonej przez spółkę pracowniczą, gdzie każdy czuje się właścicielem i kontroluje się sam, wszyscy starają się uchylać od pracy. W rezultacie zakład bankrutuje i wszyscy zostają bezrobotnymi.

---

<sup>8</sup> Tamże.

<sup>9</sup> G. Hardin, *Tragedy of the Commons*, „*Science*” 1968, December, vol. 162.

- Polityk może podjąć poważne reformy uzdrawiające stan państwa, co jednak narazi jego osobistą pomyślność materialną. Dlatego często woli wspierać układy korupcyjne pozostawiając reformy następcom.

Żadnej z tych katastrof nie może zapobiec, jeśli decyzje będą podejmować racjonalni egoiści dążący do maksymalizacji własnej korzyści. Mimo to dość często zdarza się, że ludzie współpracują i osiągają rezultaty korzystne dla wszystkich. Od kilkudziesięciu lat badacze omawiający dylemat więźnia starają się odkryć, jak to jest możliwe. Oto kilka rozwiązań wraz z ich krytyką.

## 5. Rozwiązanie Hobbesa – suweren

Rozwiązaniem zaproponowanym przez Hobbesa było powołanie arbitra, który będzie gwarantował dotrzymanie porozumień – czyli rządu. Rządy także mogą być egoistycznie nielojalne. Nie wystarczy powołać rząd, należy go jeszcze kontrolować, a także kontrolować tych, którzy go kontrolują – i tak w nieskończoność. Elinor Ostrom sugeruje, że powołanie rządu zwiększy jedynie liczbę możliwych dylematów – powstaną dylematy drugiego stopnia, zanurzone w dylematach pierwszego stopnia<sup>10</sup>. Co gorsza – obecność państwa ingerującego w relacje międzyludzkie może powodować zanik odpowiedzialności obywatelskiej<sup>11</sup>. Im więcej państwa, tym staje się ono bardziej niezbędne.

Poglądy te wydają się jednak zbyt skrajne. Ich zwolennikami od dawna były jednostki zbuntowane i anarchistyczne. Ostatnio zaś dołączyli do nich zwolennicy wielkich międzynarodowych korporacji, które ze zrozumiałych względów robią wiele, by pozbawić swoich pracowników oparcia, jakie daje im silne państwo. Oczywiście jest, że zdarzają się dobrze funkcjonujące państwa, choć wiele jest też państw złych, despotycznych i niewydolnych. Krytycy państwa jako gwaranta lojalności pytają często retorycznie – Jak wspólnota, która sobie nie ufa, może wyłonić władzę, której zaufa? Łatwo wskazać przynajmniej jedną wiarygodną odpowiedź. To nie wspólnota wyłania władzę, władza wyłania się sama – przejmują ją najsprytniejsi i najenergiczniejsi obywatele. Dlaczego mieliby stać się gwarantami lojalności? Bo leży to w ich interesie. Nawet władca, który okrada i oszukuje swoich poddanych, jeśli tylko nie jest półgłówkiem lub psychopatą, nie jest zainteresowany tym, by poddani okradali się i oszukiwali wzajemnie. Niekiedy też na czele państwa staje charyzmatyczny wódz, który skutecznie narzuca obywatelom reguły współpracy (takie były na przykład początki imperium arabskiego Mahometa). Jest to jednak z pewnością rozwiązanie bardzo niedoskonałe (nawet względnie nieopresyjny despota z trudem radzi sobie z wyzwalaniem twórczej energii obywateli koniecznej do rozwoju państwa) i nietrwałe (jeśli regulacje państwowe nie zyskują aprobaty obywateli upadają wraz ze śmiercią charyzmatycznego wodza – jego następcy na ogół nie potrafią utrzymać jego standardów). Czy są inne podstawy sprawnej organizacji społeczeństwa?

<sup>10</sup> E. Ostrom, *Governing the Commons*, Cambridge University Press, New York 1990.

<sup>11</sup> M. Taylor, *Anarchy and Cooperation*, John Wiley and Sons, New York 1976.

## 6. Wspólnota równych jako optymalne rozwiązanie

Niezwykle pouczające są w tej materii badania przeprowadzone przez Roberta Putnama nad rządami lokalnymi we Włoszech<sup>12</sup>. W roku 1970 odgórną decyzją powołano do istnienia lokalne samorządy we Włoszech, państwie uprzednio silnie scentralizowanym. Putnam śledził przez ćwierć wieku rozwój tych samorządów (których strukturę zaprojektowano identycznie dla każdego regionu) i odkrył różne wzorce ich funkcjonowania. Na Południu nowe władze okazały się aroganckie i nieudolne, podczas gdy na Północy były sprawne i lojalne wobec wyborców. Najwyraźniej pierwsze wybrały strategię nieufności i oszustwa, a drugie strategię rzetelnej współpracy. Putnam zbadał wiele czynników, które mogły mieć na to wpływ, by ostatecznie postawić śmiałą hipotezę – gotowość do współpracy jest cechą zakorzenioną bardzo głęboko w mentalności ukształtowanej przez tradycję. Południowe Włochy były w swej historii zorganizowane przez państwa o hierarchicznej strukturze, ich mieszkańcy byli wobec siebie nieufni, a korzyści czerpali z posłuszeństwa wobec zwierzchników. W północnych Włoszech natomiast już w dojrzałym średniowieczu wykształciły się tradycje współdziałania między równymi sobie obywatelami. W obliczu osłabnięcia władzy królewskiej, mieszkańcy miast tworzyli demokratyczne komuny, w których w drodze dyskusji i rozważania argumentów wypracowywali wspólne sposoby działania i odpowiedzialnie je egzekwowali. Uogólniając rozważania Putnama można uznać, że tradycja współpracy polega na wytworzeniu postaw czy też nawyków, z których najważniejsze to: poczucie braterstwa z pozostałymi współobywatelami; poczucie odpowiedzialności za wspólnotę; gotowość do tropienia i karania tych, którzy postępują nielojalnie. Tymczasem na Południu utrwały się postawy nieufności i bezradności, w obliczu których klientelizm okazywał się indywidualnie optymalnym rozwiązaniem.

Putnam zauważa, że choć wybór postawy współpracy lub nieufności może być przypadkowy, gdy już zostanie dokonany, ma silne tendencje do umacniania się. Nieufni i zależni z czasem stają się coraz bardziej nieufni i zależni, a współpracujący i odpowiedzialni pozostają takimi nawet przy niekorzystnych czynnikach zewnętrznych. W społeczeństwie obywatelskim na Północy na straży współpracy stały wciąż doskonałe mechanizmy obiegu informacji (o nielojalnych graczach). Nie ma wobec tego prostej recepty, jak uleczyć kultury zainfekowane nieufnością. Co gorsza, nawyki społeczeństwa obywatelskiego mogą zostać rozregulowane. W innej swojej książce<sup>13</sup> Putnam analizuje niepokojące przejawy upadku społeczeństwa obywatelskiego w USA. Kraj niegdyś przodujący w pomocy sąsiedzkiej i wyposażony w liczną i odpowiedzialną klasę średnią, staje się obecnie zbiorowością samotnych zindywidualizowanych egoistów.

Na marginesie swych głównych wywodów Putnam przeciwstawia się trzem rozpowszechnionym hipotezom społeczno-politycznym. Po pierwsze, uznaje za wątpliwe, by o wyborze strategii decydowały geny. Miliony ludzi przeniosły się z Południa na Północ i rozwinęły gotowość do współpracy. Ten wniosek może się jednak okazać pochopny, nie jest bowiem wykluczone, że genetycznie ludność Południa była zróżnicowana, a wyemigrowali tylko ci, którzy byli predysponowani do otwartości i współpracy.

---

<sup>12</sup> R.D. Putnam, *Demokracja w działaniu. Tradycje obywatelskie we współczesnych Włoszech*, Znak, Kraków 1995.

<sup>13</sup> R.D. Putnam, *Bowling Alone: The Collapse and Revival of American Community*, Simon & Schuster, New York 2000.

Po drugie, nie jest prawdą, że czynniki gospodarcze wpływają determinująco na przemiany społeczne. Różnice mentalności między Południem a Północą zarysowały się już tysiąc lat temu, a mimo to regiony południowe przez wiele stuleci rozwijały się gospodarczo nie gorzej niż regiony północne. Dopiero od połowy XIX w. gotowość do współpracy okazała się zdecydowanie korzystna ekonomicznie. Jest to kolejny po słynnych analizach protestanckich korzeni kapitalizmu Maxa Webera ważki argument za poglądem, że to nie byt określa świadomość lecz odwrotnie.

Po trzecie, nie wydaje się uzasadniona teza, że aktywność obywateli osłabia państwo lub wymaga jego słabości. Niewątpliwie państwo despotyczne, paraliżujące obywateli, upośledza ich aktywność. Jednak we wspólnotach demokratycznych – dowodzi Putnam – instytucje państwowe i administracyjne funkcjonują sprawnie. (Jest to zapewne model państwa umocowanego demokratycznie i zgodnego z ideałami Locke’a, a nie Hobbesa).

Badania Putnama rzucają nowe światło na problemy dylematu więźnia. Nie jesteśmy racjonalnymi egoistami, kształtuje nas tradycja, która utrwała w nas nawyki umożliwiające, a nawet wymuszające współpracę. Działanie oparte na nawykach nie jest w pełni racjonalne, choć zapewne pod wpływem racjonalnego namysłu można je zmodyfikować.

Inne badania wspierają ten pogląd. Martin A. Novak i Karl Sigmund sugerują, że kooperacja poprawia reputację i dlatego jest faworyzowana nawet wówczas, gdy oddziaływania są jednokrotne. U pewnych gatunków ptaków gdzie częste jest pomaganie sobie, zaobserwowano współzawodnictwo o bycie ‘pomagającym’. Koszt pomagania zwraca (z nadatkiem) wzrost statusu osiągniętego w akcie ‘altruistycznym’<sup>14</sup>.

James Rilling z zespołem zbadali spontaniczną gotowość ludzi do współpracy. Uczestniczki eksperyment (zgłosiły się tylko kobiety) wybierały częściej współpracę niż nieojoalność, a badanie ich mózgow rezonansem magnetycznym wykazywało, że współpraca nawet bez nagrody przynosi im przyjemność<sup>15</sup>.

Rozbudowany eksperyment zaproponowali Ernst Fehr z Uniwersytetu w Zurychu i Simon Gächter z Uniwersytetu w St Gallen. Wymyślili oni finansową grę, w której uczestnicy – obcy sobie – muszą zdecydować, czy zainwestować własne środki we wspólne przedsięwzięcie, czy odmówić, a następnie wykorzystać społeczną postawę pozostałych. Gracze zakładali wspólny fundusz, przypominający opisane przez Putnama rotacyjne spółki kredytowe. Zgromadzone pieniądze zostawały podwojone a następnie rozdzielone równo między wszystkich graczy, niezależnie od tego, ile każdy zainwestował. Grupa jako całość zyskuje najwięcej, gdy wszyscy gracze inwestują maksymalną przyjętą sumę, ale pierwszy zdrajca, który odmówi udziału w składce, osobiście zyskuje najwięcej.

W szwajcarskim eksperymencie brało udział 240 graczy, których dzielono na zmieniające się ciągle grupy, tak by dowolni dwaj gracze spotykali się tylko raz. Nie istniała więc sytuacja dylematu iterowanego ani obieg informacji i budowanie własnej reputacji. Dopóki nie wprowadzono kar finansowych dla tych, którzy nie inwestowali własnych środków dla wspólnego dobra, a tylko korzystali ze szczodrości innych, współpraca załamywała się całkowicie po sześciu rundach. Aby ukarać egoistów inni gracze sami przestawali współpracować.

Kiedy jednak wprowadzono kary dla egoistów, wspólne dobro zatryumfowało: w tym przypadku ponad 90% graczy zwiększyło własne wkłady. Najbardziej skuteczna okazała

---

<sup>14</sup> M.A. Novak, K. Sigmund, *Evolution of Indirect Reciprocity by Image Scoring*, „Nature” 1998, nr 393, s. 573–576.

<sup>15</sup> J.K. Rilling et al., *A Neural Basis for Social Cooperation*, „Neuron” 2002, vol. 35, 18 July, s. 395–405.



się strategia „altruistycznego karania”, w której osoby wymierzające karę same ponoszą pewne niewielkie koszty. Badacze stwierdzili, że mimo ponoszonej straty strategia ta dawała graczom satysfakcję psychiczną, nie zaś materialną, gdyż stwarzała okazję rozładowania złości z powodu pasożytnictwa innych. Zjawisko bezinteresownego karania egoisty zdaje się odgrywać istotną rolę w tworzeniu klimatu społecznej współpracy<sup>16</sup>.

Wyniki międzykulturowych badań nad egoizmem zebrano ostatnio w książce *Foundations of Human Sociality*<sup>17</sup>. Badacze jednoznacznie odrzucili tezę o istnieniu *Homo oeconomicus*, który kieruje się wyłącznie dążeniem do maksymalizacji egoistycznych korzyści materialnych. Zdecydowana większość ludzi równie silnie dąży do umacniania sprawiedliwości, równości i wzajemności. Nie znaleziono żadnego społeczeństwa ani wspólnoty, w których pogoń za interesem własnym byłaby fundamentem życia społecznego.

## 7. Uwagi końcowe

Podsumujmy rozważania nad dylematem więźnia. Społeczeństwa, aby rozkwiąć, potrzebują współpracy, a ta zagrożona jest przez krótkoterminowy egoizm, skłaniający bardziej do oszustwa niż lojalności. W niektórych społeczeństwach udaje się jednak osiągnąć wysoki poziom współpracy. Czasem przyczynia się do tego istnienie silnego scentralizowanego państwa, którego przywódca bądź to we własnym interesie, bądź z motywacji ideowej egzekwują normy lojalnej współpracy. Bardziej stabilną podstawą jest jednak powstanie wspólnoty równych, w której współpraca opiera się na irracjonalnych poczuciu sympatii, wzajemności i sprawiedliwości (m.in. powinności karania jednostek pasożytniczych), które przyczyniają się do ukształtowania odpowiednich nawyków współdziałania, przekazywanych następnie przez tradycję.

Żadne z dotychczas ukształtowanych wspólnot nie funkcjonowały na podstawie kontraktu racjonalnych egoistów kalkulujących maksymalną korzyść własną. Pomysł, że ludzie są racjonalnymi egoistami wraz z naiwną wiarą, iż jest to wystarczający warunek powstania norm moralnych i funkcjonowania społeczeństwa, był sztandarową ideą moralną filozofów europejskiego oświecenia. Wiązano z nią wówczas wielkie nadzieje. Mniemano, że zło obecne w społeczeństwach tradycyjnych, opartych na „zabobonie”, będzie można usunąć, gdy ludzie uświadomią sobie swój egoizm i zaczną postępować racjonalnie. Dziś wiemy, że doprowadzić to może tylko do wyboru niekorzystnej dla wszystkich konfiguracji czwartej w dylemacie więźnia.

Mimo to idea racjonalnego egoizmu cieszy się dużą popularnością wśród zwolenników liberalizmu. Typowy student ekonomii rozumie Smitha w prosty sposób: „Będę dbał tylko o siebie, a niewidzialna ręka rynku zadba, by moje działania były korzystne także dla społeczeństwa.” I paradoksalnie właśnie dziś, gdy okazało się, że racjonalny egoizm prowadzi na manowce, szanse na jego tryumf są większe niż kiedykolwiek. Jeśli wielkie korporacje będą kontynuować swoją ekspansję, w naszym stuleciu może wypełnić się wizja

---

<sup>16</sup> E. Fehr, S.-V. Renninger, *Paradoks samarytanina*, „Świat Nauki – Umysł” 2005, nr 1 specjalny (listopad-grudzień), s. 31–37. E. Fehr, U. Fischbacher, *The Nature of Human Altruism*, „Nature” 2003, vol. 425 (23 Sept.), s. 785–791.

<sup>17</sup> *Foundations of Human Sociality. Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*, J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr and H. Gintis (eds.), Oxford University Press, New York 2004.

kapitalizmu nakreślona w połowie XIX w. przez Karola Marksa – koncentracja kapitału, dominacja niewielkiej liczby ponadnarodowych korporacji, które nie poddane żadnej kontroli będą decydować o losie miliardów ludzi. W zglobalizowanym świecie nie będzie miejsca na demokratyczną wspólnotę równych, nawet o zasięgu lokalnym (ludzie mieszkający na jednym obszarze będą mieć coraz mniej wspólnego ze sobą, a relacje lokalne mogą nie będą istotne dla globalnego procesu produkcji); tym bardziej nie będzie istnieć państwo – ani demokratyczne, ani zcentralizowane, które mogłoby normować relacje między obywatelami.

Ustrój globalnego świata będzie podobny do ustroju Polski szlacheckiej z XVIII w., w którym wielkie korporacje zajmą miejsce rodów magnackich. Relacje między nimi będą zapewne egoistyczne i racjonalne, bo kto jeśli nie decydenci wielkich korporacji miałby pasować do stereotypu racjonalnego egoisty? Pozostałe relacje – na przykład między przełożonymi i podwładnymi w ramach korporacji – będą nacechowane krótkoterminowym egoizmem i klientelizmem. W tych warunkach mogą zawieść wszelkie znane historii mechanizmy rządzące prospołecznymi zachowaniami. Ostatnim ratunkiem, choćby cząstkowym, chroniącym, jeśli nie przed bylejakością życia, to przynajmniej przed popadnięciem w całkowity chaos, pozostaną strategie gry racjonalnych egoistów – jeśli ktoś zdoła je do tego czasu sformułować.

## **The Prisoner's Dilemma and the Crisis of Morality as a Contract of Rational Egoists**

### **Summary**

David Hume assumed that morality and social order in general can be based on rational calculations of selfish agents who try to maximize their self-interest. The widely discussed prisoner's dilemma and the tragedy-of-the-commons dilemma show that under certain conditions rational egoists are unable to avoid a social catastrophe that follows their choices. I examine several solutions to those in the game theory, which, however, seem unsatisfactory. Then I present different empirical data (e.g. collected by R. Putnam and E. Fehr) which demonstrate that, although overcoming the obstacles of social dilemmas is possible, it never results from rational calculations of selfish agents.

**Key words:** *prisoner's dilemma, morality, business ethics, liberalism, civil society*