



## Od-Kantować Rawlsa: ewolucyjne uzasadnienie teorii sprawiedliwości w ujęciu Kena Binmore'a

Sebastian SZYMAŃSKI\*

### ABSTRACT

**De-Kanting Rawls: evolutionary justification of the theory of justice by Ken Binmore:** The article discusses the criticism of John Rawls's theory of justice formulated by Ken Binmore. Binmore criticizes Rawls in his work for inadequate justification of fairness norms. In his opinion, their proper justification requires the rejection of Kant's "metaphysics" in favor of Hume's empiricism. To explain how fairness norms actually work, one must first provide a convincing account of evolutionary processes that could lead to such norms being encoded in our genome. Secondly, it should be explained how this biological mechanism interacts with our cultural heritage and generates a selection of balance points in games that describe our social interactions.

### KEYWORDS

Ken Binmore; John Rawls; theory of justice; justification; evolutionary theory

---

\* Dr, adiunkt, Wydział „Artes Liberales”, Uniwersytet Warszawski. E-mail: s.j.szymanski@uw.edu.pl. Artykuł powstał w ramach grantu nr 2011/03/N/HS1/00469 pt. „Uzasadnienia teorii sprawiedliwości: dziedzictwo Rawlsa” finansowanego przez Narodowe Centrum Nauki.

Teoria sprawiedliwości i model jej uzasadnienia przedstawiony przez Johna Rawlsa od niemal półwiecza stanowi punkt odniesienia dla analiz filozoficznych dotyczących tych zagadnień. Do najbardziej interesujących reakcji, które wywołała Rawlowska koncepcja *justice as fairness*, należy stanowisko zaproponowane przez Kena Binmore'a. W artykule zamierzam przeanalizować sformułowaną przez niego krytykę koncepcji Rawlsa oraz ocenić, na ile jest ona trafna.

Binmore za najpoważniejszą słabość stanowiska Rawlsa uważa wszelkie nawiązania do stanowiska transcendentnego. *Justice as fairness* potrzebuje, jego zdaniem, czysto naturalistycznego uzasadnienia. Za naturalnego sprzymierzeńca w swoim przedsięwzięciu uznaje zatem Davida Hume'a. Jak ujmuje to zwięźle, celem, który stawia sobie w swoich pracach poświęconych sprawiedliwości, było „od-Kantowanie Rawlsa” (Binmore, 1994: 7–12) czy też „zde-Kantowanie [jego] idei do Humowskiej butelki” (Binmore, 2011: 17)<sup>1</sup>.

Binmore zajmuje jednoznaczne stanowisko wobec tradycyjnie rozumianej filozofii moralności: „Sądzę, że ortodoksyjna filozofia moralności do niczego nie doszła, ponieważ stawia złe pytania”<sup>2</sup> (Binmore, 2011: 1). Jego zdaniem pytanie dotyczące kwestii normatywnych, takich jak: „Do czego powinniśmy dążyć?”, „Jakie są nasze obowiązki?” albo „Jakie życie powinniśmy prowadzić?”, są źle postawione. Przemawiają za tym dwa ważne argumenty. Po pierwsze, jak stwierdza Binmore, sami filozofowie zgadzają się, że w filozofii moralności dokonał się niewielki postęp. Po drugie — i to wydaje się znacznie ważniejsze — w filozofii moralności myli się kilka istotnych kwestii. Przede wszystkim nie odróżnia się reguł deklarowanych od reguł faktycznie przestrzeganych. Filozofowie koncentrują się na tych pierwszych i na ich najczęściej nietrafnych uzasadnieniach (by nie powiedzieć: wtórnych racjonalizacjach). Poza tym reguły moralne błędnie ujmuje się jako jasno określone zasady działania, świadomie przyswajane w oparciu o racjonalne argumenty (Binmore, 1998: 145–147). Tymczasem:

Zasady rzeczywiście rządzące naszym zachowaniem składają się z mieszaniny instynktów, zwyczajów i konwencji, które są zarazem bardziej przyziemne i bardziej złożone, niż tradycyjni uczeni chcą przyznać. Są one ukształtowane w dużej mierze przez siły ewolucyjne — zarówno społeczne, jak i biologiczne. Jeśli chce się badać takie reguły, nie jest pomocnym stawianie pytań, jak wspierają one Dobro albo jak chronią Słuszność. Trzeba zamiast tego pytać, jak ewoluowały i dlaczego przetrwały. Oznacza to, że musimy traktować moralność jak naukę (Binmore, 2011: 1).

Binmore, jak sam oświadcza, nie wierzy ani w „metafizyczną Słuszność”, ani w „metafizyczne Dobro” (Binmore, 1994: 16). Jego ambicją jest natomiast

<sup>1</sup> Dekantowanie polega na zlewaniu wina do innego naczynia w celu pozbycia się niepotrzebnego osadu.

<sup>2</sup> Wszystkie cytaty z prac obcojęzycznych w przekładzie autora.

sformułowanie koncepcji moralnej, którą można by uznać za naukową w przywołanym przez niego sensie.

Traktuje on wartości moralne jako konwencje regulujące zachowania członków społeczeństwa. Innym określeniem, jakiego używa w odniesieniu do tych konwencji, są „umowy społeczne”. „Umowa społeczna to zbiór wspólnych porozumień (*understandings*), które pozwalają obywatelom społeczeństwa skoordynować swoje wysiłki” (Binmore, 2011: 3). Te porozumienia albo konwencje mają zróżnicowany charakter. Od zaleceń etykiety po tabu seksualne, od reguł ruchu drogowego po reguły języka i od wymogów mody po reguły ustalające własność — wszystko to są przykłady konwencji regulujących zachowania członków społeczności, w której się je uznaje. Pomimo znacznego na pierwszy rzut oka zróżnicowania wymienionych przykładowo reguł Binmore utrzymuje, że różnice między nimi są jedynie różnicami stopnia. Różnica stopnia dotyczy tu ścisłości wymogów opisywanych przez poszczególne reguły oraz zakresu porządkowanych przez nie zachowań. Wymogi ruchu drogowego są ściślejsze i precyzyjniejsze niż wymogi mody, a reguły etykiety regulują węższy zakres zachowań niż reguły określające, co do kogo należy.

Binmore podkreśla, że tak rozumiane umowy społeczne nikogo nie wiążą w takim sensie, w jakim związane są strony umowy w potocznym rozumieniu tego pojęcia. Za umowami społecznymi nie stoi również żaden zewnętrzny mechanizm, który w razie potrzeby wspiera egzekwowanie ich warunków. Przeciwnie, tego rodzaju mechanizmy w przypadku porozumień między członkami społeczności same opierają się na umowie społecznej regulującej funkcjonowanie społeczeństwa (Binmore, 1994: 5–6). Jednak mogą to robić skutecznie, tylko jeśli taka jest umowa społeczna obowiązująca w danym społeczeństwie. Z punktu widzenia stanowiska Binmore’a istotne jest to, jak faktycznie wygląda praktyka społeczna. System prawny i instytucje egzekwujące jego przestrzeganie mają znaczenie, pod warunkiem że prawo jest faktycznie respektowane, a instytucje te rzeczywiście są w stanie doprowadzić do jego przestrzegania. Jeśli powszechną praktyką jest łamanie obowiązującego prawa, a stojące na jego straży organy są wobec tego bezsilne — jest to po prostu częścią umowy społecznej w rozumieniu Binmore’a (Binmore, 2011: 4).

Co w takim razie zapewnia funkcjonowanie tak rozumianej umowy społecznej? To pytanie jest zdaniem autora *Natural justice* po prostu źle postawione. Stabilna umowa społeczna nie potrzebuje żadnej gwarancji funkcjonowania. Sankcje za jej nieprzestrzeganie — jeśli w ogóle występują — są rozproszone. Działa ona przede wszystkim dzięki temu, że członkowie społeczeństwa dostrzegają swój interes w zgodnym działaniu. Jest to oczywiście idea dobrze znana w filozofii, którą można odnaleźć na przykład w poniższym fragmencie z *Traktatu o naturze ludzkiej* Hume’a:

Jest to [umowa społeczna — przyp. S.S.] tylko ogólne poczucie wspólnego interesu; poczucie, które wzajemnie członkowie społeczności ujawniają i które skłania ich do tego, by regulowali swoje postępowanie pewnymi normami. Stwierdzam, że będzie zgodne z moim interesem, by pozostawić innego człowieka w posiadaniu jego dóbr pod warunkiem, że on zachowa się tak samo względem mnie. On znowu również świadom jest podobnego swojego interesu, gdy reguluje swoje postępowanie. [...] Toż działania każdego z nas liczą się z działaniami drugiego z nas i są dokonywane przy założeniu, że druga strona też musi coś dokonać. Dwaj ludzie, którzy poruszają wiosłami na jednym czółnie, czynią to mocą ugody czy umowy, choć nie dali sobie żadnych przyrzeczeń (Hume, 1963: 286).

Binmore otwarcie przyznaje, że znaczna część jego koncepcji poświęcona jest analizie konsekwencji sformułowanej przez Hume'a intuicji, że umowę społeczną można ujmować jako niewypowiedziane porozumienie prowadzące do skoordynowania gry życia, w którą wszyscy gramy, w oparciu o jakiś konkretny punkt równowagi (Binmore, 2011: 4).

## RÓWNOWAGA I EFEKTYWNOŚĆ

W ujęciu proponowanym przez Binmore'a każdą sytuację związaną z interakcją można opisać jako grę. Gracze dążą do uzyskania jak najkorzystniejszego dla siebie wyniku przy zadanych regułach gry. W tym dążeniu przyjmują pewne plany postępowania nazywane strategiami. Gra znajduje się w równowadze, gdy układ strategii graczy jest taki, że strategia przyjęta przez każdego z nich jest najlepszą odpowiedzią na strategię pozostałych.

Pojęcie równowagi pozwala opisać istotny aspekt gry. Jeśli ta znajduje się w równowadze, posiada rozwiązanie, które jest znane wszystkim graczom. W takim przypadku nie będą mieli oni racjonalnego powodu, żeby zmienić przyjętą strategię gry. Strategia optymalna, czyli taka, która prowadzi do powstania równowagi, jest dla graczy racjonalnym wyborem. Równowaga jest istotna również z innego, jeszcze ważniejszego powodu. Jeśli uznamy, że wypłaty dla graczy odpowiadają temu, jak bardzo są oni dostosowani, to proces ewolucji, który wynagradza lepiej dostosowanych kosztem dostosowanych gorzej, ulegnie zatrzymaniu, gdy gra znajdzie się w stanie równowagi. W równowadze bowiem wszyscy gracze są tak dobrze dostosowani, jak mogą być w danych warunkach (Binmore, 2007: 18–19).

Kolejnym pojęciem charakteryzującym umowę społeczną jest efektywność. Tę rozumie się tu w sensie wprowadzonym przez Vilfreda Pareto. Rozwiązanie jest efektywne wówczas, kiedy nie można poprawić położenia którejkolwiek jednostki bez jednoczesnego pogorszenia położenia innej. Jeśli umowa społeczna w rozumieniu Binmore'a dopuszcza możliwość, że sytuacja którejś ze stron może się poprawić, a nikogo innego się nie pogorszy, zostanie zastąpiona przez inną umowę społeczną, która jest efektywna.

Przykładem efektywnego rozwiązania jest doskonale konkurencyjny rynek osiągnący stan równowagi, w którym popyt pokrywa się z podażą. Na doskonale konkurencyjnym rynku sytuacja żadnego z kupujących ani sprzedających nie może ulec zmianie bez zmiany na niekorzyść kogoś innego, ponieważ na rynku w stanie równowagi sprzedający oferują dobra po maksymalnej cenie, jaką są w stanie zaakceptować kupujący. Doskonale efektywny rynek jest przykładem gry, w której równowaga jest efektywna.

Umowa społeczna, na której opiera się nasze społeczeństwo, musi być grą w równowadze, ale też równowaga musi być efektywna. Binmore zwraca uwagę, że pojęcie odwzajemnionego altruizmu wprowadzone przez Roberta Triversa (1985) pozwala wyjaśnić, jak wyewoluowały nasze umowy społeczne, które są zarazem w równowadze i efektywne. Jednak podobne rozwiązanie zaproponował już Hume w *Traktacie o naturze ludzkiej*:

Uczę się więc oddawać usługi innemu człowiekowi, nie żyjąc do niego żadnej przychylności; a to dlatego, że przewiduję, iż on odwzajemni się za moją usługę, oczekując innej usługi tego samego rodzaju oraz w tym samym celu, iżby zachować tę samą wymianę świadczeń ze mną czy też z innymi. Toteż, gdy ja mu oddałem przysługę i gdy on osiągnął korzyść z mojego działania, skłonny jest on wykonać swoje świadczenie wzajemnie, ponieważ przewiduje konsekwencje swojej odmowy (Hume, 1963: 327).

Wyjaśnienie zaproponowane przez Hume'a ma tę zaletę, że tłumaczy kooperację nie tylko w ramach genetycznie spokrewnionych grup, ale także w zbiorowościach skupiających jednostki niespokrewnione ze sobą.

W teorii gier zazwyczaj jako pewnik przyjmuje się obiegowe twierdzenie (*folk theorem*)<sup>3</sup>, które głosi, że relacje oparte na racjonalnej wzajemności nie będą działać, jeśli interakcje między jednostkami nie będą się powtarzać w nieskończoność. Przez powtarzanie się w nieskończoność rozumie się tutaj, że żaden z uczestników gry nie może być pewien, iż kolejna rozgrywka będzie ostatnią. *Folk theorem* głosi, że w takiej sytuacji osiągnięte porozumienie będzie samoregulującą umową społeczną. Oznacza to, że do jej egzekwowania nie będzie potrzebna żadna instytucja zewnętrzna. W sytuacji równowagi żadna ze stron nie będzie próbowała osiągnąć zysków kosztem innych z obawy, że reakcją pozostałych stron będzie anulowanie tych profitów.

Gry zazwyczaj mają wiele punktów równowagi. Wielość tychże prowadzi do powstania problemu wyboru punktu równowagi, który musi zostać rozwiązany w procesie ewolucyjnym, jeśli gracze mają stworzyć społeczeństwo zdolne do

<sup>3</sup> Choć można spotkać się ze stwierdzeniami, że *folk theorem* jest przyjmowane bez dowodu, dowód został sformułowany przez Roberta Aumanna (Aumann, 1959: 307–317). Badacz ten wykazał, że każdy układ wypłat, w którym każdy z graczy dostaje więcej niż w równowadze gry jednorazowej, wystąpi w równowadze gry iterowanej (powtarzanej nieskończenie wiele razy). Konsekwencją tego jest między innymi to, że optymalna strategia dla tej samej gry iterowanej może być inna niż dla gry jednorazowej.

funkcjonowania i trwania w czasie. Jak twierdzi Binmore, *folk theorem* pokazuje, że problem wyboru spośród wielu punktów równowagi był znacznie trudniejszy do rozwiązania w przypadku naszych ewolucyjnych przodków (Binmore, 1994: 319–328). Skoro punktami równowagi mogły być zarówno efektywne punkty równowagi, na które mogli się zgodzić racjonalni gracze, jak i nieefektywne punkty równowagi, dlaczego ewolucja miałaby wyselekcjonować raczej efektywny niż nieefektywny punkt równowagi?

Odpowiedź na to pytanie wymaga założenia, że grupy naszych ewolucyjnych przodków konkurowały ze sobą. Wyobraźmy sobie dwa społeczeństwa działające w oparciu o odmienne umowy społeczne: jeśli jedna z nich sprawia, że członkowie tego społeczeństwa są lepiej dostosowani, to jest prawdopodobne, że ta umowa społeczna w końcu będzie dominować nad inną, prowadzącą do gorszego dostosowania członków społeczeństwa. W tym ujęciu lepsze dostosowanie oznacza większą liczbę potomstwa, co powoduje, że społeczeństwo opierające się na umowie społecznej dającej przewagę ewolucyjną rozrasta się szybciej. Jeśli przyjąć za Binmore’em, że społeczeństwa radzą sobie ze wzrostem populacji poprzez tworzenie kolonii, czyli odrębnych grup opartych na tej samej umowie społecznej, to oznacza to, że umowa społeczna prowadząca do szybszego wzrostu populacji będzie szybciej się „reprodukować”.

Binmore zastrzega, że jego ujęcie jest dość nieortodoksyjne. Tak rozumiana reprodukcja umowy społecznej ma charakter kulturowy, a nie genetyczny. Jednak, jak zauważa Binmore, przyjęcie socjobiologicznego punktu widzenia, za którym się opowiada, nie oznacza uznawania wszystkich mechanizmów ewolucyjnych za zdeterminowane przez przekaz genów. W przypadku naszych społeczeństw i przynajmniej niektórych społeczeństw naszych ewolucyjnych przodków mamy do czynienia z koewolucją genetyczno-kulturową (Binmore, 2011: 12–14).

Choć ewolucja naszego gatunku następowała na tych samych zasadach, co ewolucja innych gatunków, kultura wprowadza istotną różnicę. Zdaniem Binmore’a różnica ta musiała pojawić się na dość wczesnym etapie ewolucji naszego gatunku. Jednak sam fakt posiadania kultury nie jest tu kluczowy, ponieważ istnieje wiele innych gatunków, niekiedy dość odległych od nas ewolucyjnie, które wytworzyły mechanizmy transmisji kulturowej. Zdaniem Binmore’a rolą kultury w podziale przysłowiowego ciastka jest przechowywanie oraz przekazywanie wiedzy i umiejętności pozwalających na eliminowanie konfliktów skutkujących marnotrawieniem cennych zasobów. Jednak wyjaśnienie, jak do tego dochodzi, nie może ograniczać się — jak to często jest u socjologów i antropologów — do analizy zakazów i nakazów moralnych stanowiących trzon naszych kultur. Binmore uważa, że poszukiwanie uniwersaliów moralnych pozwalających wyjaśnić efektywną kooperację wymaga wyjścia poza zróżnicowanie kulturowe, na którym koncentrują się nauki społeczne, i zbadania głębszej struktury umów społecznych wpisanych w nasze dziedzictwo genetyczne (Binmore, 1994: 319–320).

Jakie uniwersalia moralne ma na myśli Binmore? Pojęcia charakteryzujące umowę społeczną w jego ujęciu, które dotychczas omówiłem, pozbawione są wydzwiku moralnego. Stabilność systemów kooperacji społecznej przemawia za tym, że umowy społeczne muszą być punktami równowagi w grze życia. Ponieważ gry mają więcej niż jeden punkt stabilności, pojawia się problem wyboru jednego z nich. Pojęcie efektywności pozwala częściowo wyjaśnić, jak ten problem jest rozwiązywany, ale *folk theorem* pokazuje, że efektywnych punktów równowagi także może być wiele. Społeczeństwo, aby móc trwać i funkcjonować, musi wybrać jeden z nich.

Binmore stawia tezę, że rozstrzygnięcia, które uznajemy za fair, są wybranymi przez ewolucję rozwiązaniami problemu wyboru punktu równowagi w grze życia, w którą grali nasi ewolucyjni przodkowie. Na poparcie swojej tezy przytacza wiele przykładów zaczerpniętych głównie ze społeczeństw badanych przez antropologów kulturowych. Twierdzi, że większość istniejących jeszcze społeczeństw zbieracko-łowieckich opiera się na podobnej umowie społecznej o takiej samej strukturze głębokiej. Jej dwie najważniejsze cechy to brak jednostek dominujących nad innymi i bardzo egalitarny charakter relacji społecznych (Binmore, 2011: 14).

Binmore zastrzega, że te egalitarystyczne społeczeństwa zaczęto badać w schyłkowym okresie ich istnienia. Z tego względu nie jesteśmy w stanie ustalić, jak ważną rolę odegrała kultura w ustalaniu, co uznawane jest w tych społeczeństwach za fair. Jednak pewne zróżnicowanie przekonań dotyczących *fairness* w różnych kulturach wskazuje, że również element kulturowy ma wpływ na wybór jednego z możliwych efektywnych punktów równowagi. Zatem w ustalaniu, co uznawane jest za fair, bierze udział element zarówno genetyczny, jak i kulturowy. Skoro jednak przekonania dotyczące *fairness* zmieniają się w zależności od kultury, uniwersalia moralne — zwłaszcza te dotyczące sprawiedliwości — muszą być w jakiś sposób zakodowane w genach wspólnych wszystkim przedstawicielom naszego gatunku. Binmore wysuwa więc przypuszczenie, że głęboka struktura umowy społecznej przekazywana jest genetycznie. Prowadzi go to do pytania, jaką formę ta postulowana przez niego struktura przybiera (Binmore, 1994: 183–186).

Binmore sformułował swoją koncepcję jako rozstrzygnięcie problemu wyboru zasad sprawiedliwości w sytuacji początkowej, postawionego przez Rawlsa, który jednak — zdaniem autora *Natural justice* — błędnie go rozwiązał. Binmore twierdzi, że powszechna głęboka struktura norm *fairness* została trafnie scharakteryzowana w Rawlowskiej sytuacji początkowej, ale autor *Teorii sprawiedliwości* mylił się co do tego, dlaczego zostałyby w niej wybrane dwie zasady sprawiedliwości.

Rawls traktuje swoje stanowisko jako próbę zoperacjonalizowania Kantowskiego imperatywu kategorycznego. Jego zdaniem wybór zasad sprawiedliwości w sytuacji początkowej pokazuje, jak w praktyce wyglądałoby sprawdzanie, czy

proponowane zasady moralne spełniają kryterium imperatywu kategorycznego. Jednak Binmore jest przekonany, że propozycja Rawlsa jest atrakcyjna z zupełnie innego powodu. Jego zdaniem dwie zasady sprawiedliwości wybierane w sytuacji początkowej za zasłoną niewiedzy odpowiadają po prostu normom *fairness*, które są wykorzystywane w społeczeństwach do rozwiązywania problemu selekcji efektywnego punktu równowagi (Binmore, 1998: 7–12).

Binmore w przeciwieństwie do Rawlsa traktuje normy *fairness* jako narzędzie koordynacji zachowań jednostek w codziennych interakcjach zachodzących w mikroskali. Tymczasem Rawls swoje dwie zasady sprawiedliwości sformułował z myślą o regulowaniu funkcjonowania najważniejszych instytucji działających w makroskali. Dwie zasady miały normować przede wszystkim „sposób, w jaki główne instytucje społeczne dystrybuują zasadnicze prawa i obowiązki oraz określają podział korzyści płynących ze społecznej kooperacji” (Rawls, 2009: 34). Binmore zauważa jednak, że nasze normy *fairness* nie wyewoluowały jako narzędzie wykorzystywane w takiej skali. Normy *fairness* są stosowane przede wszystkim do koordynowania codziennych zachowań jednostek. Zazwyczaj następuje to w sposób niezauważalny dla tych, które tymi normami się kierują. Internalizacja norm *fairness* sprawia, że koordynacja codziennych zachowań zachodzi bez wysiłku i najczęściej bezrefleksyjnie. Jednak z punktu widzenia trwania i funkcjonowania społeczeństwa mają one pierwszorzędne znaczenie. Gdyby koordynacja nie następowała właśnie w taki sposób, każda sytuacja dopuszczająca możliwość więcej niż jednego zachowania jednostek biorących w niej udział byłaby potencjalnym zarzewiem konfliktu, a istnienie społeczeństwa byłoby po prostu niemożliwe (Binmore, 2011: 16–17).

Umowa społeczna, żeby trwać, musi odpowiadać punktowi równowagi w grze życia, w której wszyscy bierzemy udział. Traktowanie naszej moralności jako wytworu ewolucji biologicznej i społecznej prowadzi do uznania jej za narzędzie służące wyważeniu władzy w ludzkich społecznościach. Teoria gier wydaje się szczególnie dobrym narzędziem do opisu takiej sytuacji. Poniżej skoncentruję się na tych elementach koncepcji Binmore’a, które są istotne z filozoficznego punktu widzenia, starając się w jak najmniejszym stopniu wchodzić w szczegóły czysto techniczne.

Zdaniem Binmore’a teoria gier wydaje się szczególnie dogodnym narzędziem, ponieważ pozwala uchwycić najważniejsze elementy kooperacji i konfliktu. Przy minimalnych założeniach dotyczących racjonalności graczy, takich jak dążenie do maksymalizacji wypłat, niechęć do straty i unikanie ryzyka, pozwala ustalić, jaką strategię przy danej strukturze wypłat powinni przyjąć, żeby uzyskać maksymalnie korzystny wynik gry. Strategia oznacza tu po prostu to, jakich wyborów spośród dostępnych możliwości dokonuje gracz, uwzględniając możliwe wybory dokonywane przez pozostałych graczy.

Z punktu widzenia ewolucjonistycznego ujęcia moralności szczególnie istotne są strategie mieszane (Binmore, 1994: 211–216). Strategia mieszana polega



na randomizacji wyboru strategii, co oznacza, że przy podejmowaniu decyzji przypisuje się dostępnym strategiom pewne prawdopodobieństwa. Dlaczego właśnie takie rozwiązanie okazuje się najlepszym sposobem opisu ewolucji? W najprostszym jej opisie za pomocą teorii gier przypisuje się prawdopodobieństwo wyboru gracza z jednej lub więcej populacji. Nie ma potrzeby przyjmowania założenia, że gracze są racjonalni — wręcz przeciwnie, zakłada się, że są genetycznie zaprogramowani do stosowania jednej strategii za każdym razem, gdy biorą udział w grze.

Ewolucja następuje w ten sposób, że w kolejnych rozgrywkach coraz częściej wykorzystywane są strategie dające wyższe wypłaty kosztem strategii prowadzących do niższych wypłat. Na przykład ewolucja biologiczna dokonuje się dzięki temu, że lepiej przystosowani gracze przekazują swoje strategie liczniejszemu potomstwu. Natomiast ewolucja kulturowa polega na tym, że gracze uzyskujący większe wypłaty są naśladowani częściej niż ci, którzy uzyskują niższe wypłaty, a tym samym stosowane przez nich wzorce zachowania rozpowszechniają się skuteczniej niż zachowania mniej efektywnych graczy.

W długiej perspektywie czasowej cały układ dąży w kierunku czystej strategii, która w danej sytuacji stanowi najlepszą odpowiedź na aktualne warunki. Strategia, która zostanie w ten sposób osiągnięta, będzie stanowić punkt równowagi gry. Jednak nie każdy punkt równowagi prowadzi do osiągnięcia strategii stabilnej ewolucyjnie. Dlatego samo pojęcie równowagi nie wystarcza do rozwiązania gry opisującej ewolucję moralności.

### FOLK THEOREM

Wspomniany wyżej *folk theorem* pozwala zidentyfikować strategie prowadzące do efektywnej równowagi dzięki scharakteryzowaniu wszystkich możliwych punktów równowagi dla wszystkich gier iterowanych (Binmore, 1994: 317–319). W nieformalnym ujęciu na przykładzie dylematu więźnia wyglądałoby to następująco. Alicja i Bob rozgrywający dylemat więźnia mogliby się targować, jaki powinien być wynik gry. Mogliby się zgodzić na każdy z możliwych wyników. Mogliby się jednak porozumieć, że wynik gry zostanie ustalony inaczej, na przykład w drodze rzutu monetą, ciągnięcia losów czy w jakikolwiek inny sposób, na który zgodzą się strony. W ten sposób możliwe do przyjęcia porozumienia będą tworzyć najmniejszy wypukły zbiór zawierający wypłaty zawarte w macierzy.

Jednak targowanie się ma tę wadę, że do realizacji osiągniętego w ten sposób porozumienia potrzebna jest zewnętrzna instytucja egzekwująca jego przestrzeganie. W razie jej nieobecności porozumienie zawarte między Alicją i Bobem musi egzekwować się samoczynnie. Jak już była mowa, tylko punkty równowagi w grze wyznaczają porozumienia, które będą mogły się samoczynnie utrzymać.

W przypadku jednorazowego dylematu więźnia punktem równowagi jest sytuacja, w której obie strony wybierają strategię zdrady. Jednak w iterowanym dylemacie więźnia jest inaczej: każdy wynik, na który racjonalny gracz mógłby się zgodzić przy założeniu istnienia zewnętrznej instytucji egzekwującej porozumienie, może stanowić efektywną równowagę, gdy gra jest powtarzana. Jest tak dlatego, że dowolny wynik gry można osiągnąć dzięki temu, że podmiot egzekwujący porozumienie będzie po prostu karał każdego gracza, który odchodzi od strategii niezbędnej do osiągnięcia tego punktu.

Co się jednak stanie w przypadku, gdy nie ma takiego podmiotu? Karanie nadal jest możliwe, przy czym polegać będzie na tym, że druga strona w odwecie zastosuje strategię, która zmniejszy korzyści gracza odmawiającego kooperacji. W przypadku iterowanego dylematu więźnia karanie polegać będzie na tym, że druga strona również będzie stosować strategię zdrady, co doprowadzi do najgorszego możliwego wyniku, jakim druga strona może ukarać gracza nieskłonного do kooperacji. Gdyby bowiem pozostała przy strategii kooperacji, „nagrodziłaby” partnera za zdradę. Zbiorem wyników, które mogą stanowić samoczynnie podtrzymywaną równowagę w iterowanej grze bez zewnętrznej instytucji egzekwującej porozumienie, będzie zatem zbiór wyników nie gorszych niż najgorszy możliwy wynik, jakim druga strona może ukarać gracza nieskłonного do kooperacji. Każdy z graczy będzie preferował którykolwiek z tych wyników, ponieważ alternatywą jest najgorszy możliwy wynik<sup>4</sup>.

Binmore uważa, że *folk theorem* pozwala zredukować problem umowy społecznej do kwestii targu prowadzącego do selekcji jednego z możliwych samoczynnie podtrzymujących się punktów równowagi w grze. Jeśli aktualna umowa społeczna nie należy do zbioru wyników nie gorszych niż najgorszy możliwy wynik, jakim druga strona może ukarać gracza nieskłonного do kooperacji, obie strony będą preferować najgorszy możliwy wynik, co oznacza, że umowa społeczna będzie się opierać co najmniej na najgorszym możliwym samoczynnie podtrzymującym się punkcie równowagi w grze. Jednak zawsze istnieje możliwość przejścia do bardziej efektywnej równowagi.

*Folk theorem* pozwala również udzielić odpowiedzi na pytanie stawiane wielokrotnie w historii myśli politycznej: dlaczego ludzie podporządkowują się władzy, której siła opiera się tylko na konwencji i nawyku? W tym celu należy pokazać, że *folk theorem* obowiązuje również dla doskonałej równowagi w podgrach. Doskonała równowaga to profil strategii wyznaczający równowagę każdej podgry gry pierwotnej. Jeśli założymy, że istnieje szansa, iż gracze będą popełniać z pewnym niewielkim prawdopodobieństwem dowolne możliwe do popełnienia błędy w podgrach gry pierwotnej, doskonałe równowagi stają

---

<sup>4</sup> Przy założeniu, że gracze są wolni od zawiści, to znaczy nie są skłonni ponosić straty tylko dlatego, żeby druga strona uzyskała jeszcze gorszy wynik. Jednak standardowo to założenie uwzględnione jest w założeniu o racjonalności stron.

się dogodnym sposobem opisu naszych umów społecznych z uwagi na ludzką skłonność do popełniania błędów (Binmore, 1998: 306–307).

W przypadku idealnej równowagi optymalnym zachowaniem jest trzymanie się pierwotnej strategii prowadzącej do równowagi, nawet jeśli inny uczestnik gry odchodzi od tej strategii, a w rezultacie kolejna podgra nie jest przypadkiem równowagi. Dzieje się tak dlatego, że optymalnym rozwiązaniem jest dla nas ukaranie osoby odchodzącej od strategii prowadzącej do równowagi, w szczególności jeśli strategia ta nakazuje karać graczy, którzy jej nie stosują. Jeśli bowiem tego nie zrobimy, w następnej podgrze optymalnym rozwiązaniem dla innego gracza będzie ukaranie nas za niespełnienie ciążącego na nas obowiązku ukarania gracza niestosującego strategii prowadzącej do równowagi. Jeśli w kolejnej podgrze inny gracz nas nie ukarze, on sam zostanie ukarany w następnej podgrze i tak dalej. W ten sposób, jeśli liczba graczy jest skończona, otrzymujemy wyjaśnienie, skąd bierze się posłuszeństwo wobec władzy opartej na konwencji.

Na pierwszy rzut oka taka spirala samopotwierdzających się przekonań wydaje się zbyt krucha, żeby wesprzeć cokolwiek solidnego. To prawda, że te przekonania tworzą zamknięty krąg, ale *folk theorem* pokazuje, że ich kruchość jest iluzją, ponieważ zachowanie generowane przez te przekonania tworzy spójną całość jako doskonała równowaga (Binmore, 2011: 86).

Binmore uważa, że za tę iluzję odpowiada przede wszystkim skłonność do koncentrowania się na zewnętrznych przejawach współczesnych złożonych instytucji politycznych takich jak państwo, co sprawia, że nie dostrzegamy, iż posłuszeństwo wobec nich opiera się właśnie na takim zamkniętym kręgu samopotwierdzających się przekonań. Formalne reguły obowiązujące w tych instytucjach są jedynie narzędziami pomagającymi w osiągnięciu doskonałej równowagi, która następnie podtrzymuje się samoczynnie. W celu skutecznego koordynowania naszych działań współczesne umowy społeczne oddają pewien zakres władzy w ręce jednostek i instytucji, jednak władza ta jest im tylko „użyczona”, a tak naprawdę jest sprawowana kolektywnie przez wszystkich tych, którzy żywią te samopotwierdzające się przekonania (Binmore, 1998: 335–336).

Binmore sprzeciwia się popularnemu pogładowi, że struktury władzy i przywódcy są po prostu niezbędni. Gdyby tak było, społeczeństwa skrajnie egalitarne nie mogłyby istnieć. Tymczasem mamy dobre argumenty teoretyczne i empiryczne przemawiające za tym, że mogą one istnieć i sprawnie funkcjonować. Jednak Binmore znajduje w swojej teorii miejsce dla obu tych fenomenów. Jego zdaniem funkcją przywódców jest wyznaczenie doskonałej równowagi w grze, która rozgrywana jest w społeczeństwie podlegającym ich władzy. Jeśli umowa społeczna opiera się na respektowaniu decyzji podejmowanych przez przywódcę,

optymalnym zachowaniem będzie postępowanie właśnie w taki sposób. Istotne jest również to, że z tego punktu widzenia tylko działania jednostek wyłamujących się z powszechnie przyjętej konwencji są suboptymalne. Jeśli decyzjom przywódcy sprzeciwia się większość jednostek uznających dotychczas jego władzę, umowa społeczna ulega zmianie. Zastąpienie przywódcy innym przywódcą oznacza zatem zmianę punktu równowagi, jednak funkcja przywódcy pozostaje niezmienna (Binmore, 1998: 417).

## OBOWIĄZKI I DOSKONAŁA RÓWNOWAGA

Binmore dzieli teorie moralne na trzy grupy: teorie dobra, słuszności i „tego, co się przyjmuje” (*seemly*). Jak już wcześniej wspominałem, krytykuje on klasyczne teorie za to, że opierają się na pojęciach o niejasnym z punktu widzenia nauki statusie. Utrzymuje również, że zwolennicy krytykowanych przez niego teorii mają wyraźne predylekcje polityczne. Zwolennicy teorii opierających się na pojęciu dobra skłonni są przedkładać dobro wspólne nad partykularne interesy jednostek. Zwolennicy teorii opartych na pojęciu słuszności mają skłaniać się raczej ku konserwatyzmowi, który Binmore utożsamia ze stanowiskami deontologicznymi. Co więcej, deontolodzy mają być zwolennikami koncepcji praw naturalnych (Binmore, 2011: 93).

Przedstawiona systematyka jest w najlepszym razie bardzo uproszczona i niewolna od błędów, jednak dla mojego wywodu istotne jest zwrócenie uwagi na jej cel, którym jest przede wszystkim skontrastowanie stanowiska naturalistycznego formułowanego przez Binmore’a z najważniejszymi nurtami filozofii moralności. W przeciwieństwie do dostępnych alternatyw stanowisko naturalistyczne jest „przyziemne”:

W teorii „tego, co się przyjmuje” (*seemly*), rzeczy nie są dobre ani słuszne same w sobie: są dobre albo słuszne, ponieważ powszechnie utrzymuje się, że są dobre albo słuszne w konkretnym społeczeństwie. Na przykład, jestem przekonany, że społeczeństwa odnoszące sukces muszą traktować pewne przywileje, jak gdyby były one niezbywalne, a pewne przywileje, jak gdyby były obligatoryjne. Jednak nie zachowujemy się w ten sposób dlatego, że jesteśmy obdarzeni jakimś wrodzonym pojęciem Słuszności, nie bardziej niż pawiany kierowane są jakąś Wyższą Moralnością, gdy decydują, że nadeszła ich kolej, by iskać albo być iskanymi (Binmore, 2011: 94).

Binmore twierdzi, że źródłem naszych koncepcji dobra jest obserwacja, jak realne społeczeństwa w praktyce wybierają punkty równowagi. Zatem aby odpowiedzieć na pytanie, co jest dobrem w danym społeczeństwie, musimy wiedzieć, jak działają w nim normy *fairness*. O tym będzie mowa w dalszej części rozdziału. Natomiast normy *fairness* nie są niezbędne do wytłumaczenia, co jest uznawane za słuszne (*right*), czyli jak funkcjonują uprawnienia (*rights*)

i obowiązki (*duties*)<sup>5</sup>, ponieważ ich źródłem jest — jego zdaniem — obserwacja tego, jak równowaga jest podtrzymywana (Binmore, 2011: 95).

Binmore zastrzega, że proponowane przez niego rozróżnienie można uznać za kontrowersyjne, ponieważ traktuje przeciwstawienie konsekwencjalizmu i deontologii jako pozbawione sensu, gdyż w jego ujęciu oznaczałoby ono przeciwstawienie optymalności i możliwości realizacji (*feasibility*) umów społecznych. Tymczasem te dwie kwestie są ze sobą związane. Do stwierdzenia, które umowy społeczne są możliwe do zrealizowania, niezbędna jest wiedza, jak podtrzymywana jest równowaga w grze życia. Natomiast by ustalić, które umowy społeczne są optymalne, musimy stwierdzić, jak wybierane są punkty równowagi. W teorii „tego, co się przyjmuje”, dobro i słuszność opisują różne aspekty umowy społecznej.

Zdaniem Binmore’a stwierdzenie, że coś jest słuszne (*right*), ma takie samo znaczenie, jak stwierdzenie, że ktoś ma obowiązek to zrobić. Natomiast ktoś ma prawo (*right*) coś zrobić, tylko jeśli nie ma obowiązku powstrzymać się od zrobienia tego. Zatem w jego ujęciu obowiązek polegał będzie na robieniu tego, co słuszne, a to, co słuszne, wyznaczone jest przez obowiązującą umowę społeczną. „W kategoriach teorii gier obowiązek Adama polega na nieodchodzeniu od reguł albo strategii wyznaczonych dla niego przez umowę społeczną” (Binmore, 2011: 95). Jednak nie oznacza to, że mamy obowiązek podporządkowania się regułom, ponieważ nie możemy złamać tych obowiązujących w naszym społeczeństwie. Gdybyśmy byli w stanie je złamać, oznaczałoby to jedynie, że nie są to właściwe reguły. Oczywiście nic nie uniemożliwia nam działania odbiegającego od strategii, jaką wyznacza aktualna umowa społeczna. Jednak ta strategia nakłada wówczas na innych członków społeczeństwa wymóg ukarania nas za odejście od niej, o czym była mowa wcześniej.

Przyjęte przez Binmore’a rozumienie uprawnień znacznie ogranicza sensowne użycie tego pojęcia w porównaniu z potocznym jego rozumieniem. Żeby posiadać uprawnienia, trzeba bowiem być uczestnikiem gry. Tylko bycie nim umożliwi przyjmowanie na siebie obowiązków, a dopiero określenie, co jest, a co nie jest obowiązkiem, pozwala ustalić, jakie są nasze uprawnienia.

Ujęcie obowiązku jako stosowania strategii prowadzącej do doskonałej równowagi pozwala Binmore’owi również scharakteryzować odpowiedzialność moralną w kategoriach proponowanej przez niego koncepcji. W tym celu specyfikuje w ramach swojego stanowiska pojęcia winy i zasługi. Przypisanie winy jest w nim równoważne zidentyfikowaniu jednostki ponoszącej winę jako właściwego obiektu karania w ramach umowy społecznej obowiązującej w danym społeczeństwie. Analogicznie, przypisanie zasługi jest równoważne z uznaniem jednostki za właściwy obiekt nagradzania. Takie ujęcie ma jednak

---

<sup>5</sup> Wywód Binmore’a opiera się na wieloznaczności angielskiego *right*, oznaczającego zarówno słuszność, jak i uprawnienie.

tę konsekwencję, że przypisanie winy przestaje być wyjaśnieniem, dlaczego ktoś jest karany, natomiast staje się po prostu opisem osoby, której zachowanie nie jest zgodne z umową społeczną obowiązującą aktualnie w jej społeczeństwie. Wina i zasługa są w ujęciu Binmore'a zjawiskami emergentnymi, co oznacza, że nie są potrzebne do określenia równowagi tworzącej umowę społeczną, natomiast okazują się przydatne, gdy opisujemy, jak funkcjonują strategie umowy tę podtrzymujące (Binmore, 2011: 97–98).

## GENEZA ZŁOTEJ REGULY

Binmore twierdzi, że sytuacja początkowa jest jedynie bardziej rozbudowaną postacią Złotej Reguły, głoszącej: „postępuj tak, jak chciałbyś, by postępowano wobec ciebie”; sytuacja ta prowadzi do przeformułowania reguły do postaci: „nie postępuj wobec innych tak, jak chciałbyś, żeby oni postępowali wobec ciebie — mogą nie mieć takich samych upodobań jak ty”. Podstawą takiego sformułowania jest to, że wszyscy chcielibyśmy, by inni postępowali wobec nas w taki sposób, w jaki chcieliby, by postępowano wobec nich pod warunkiem, że mieliby nasze preferencje, a nie swoje własne. Inaczej rzecz ujmując, jesteśmy zwolennikami Złotej Reguły pod warunkiem, że stosujący ją wykorzystują preferencje empatyczne. W efekcie otrzymujemy zasadę: postępuj tak, jak chciałbyś, by postępowano wobec ciebie, jeśli byłbyś osobą, wobec której się tak postępuje. Zdaniem autora *Natural justice* sytuacja początkowa jest zoperacjonalizowaniem tej ostatniej wersji Złotej Reguły, a nie — jak twierdził Rawls — Kantowskiego imperatywu kategorycznego (Binmore, 2011: 129–130).

Binmore stwierdza, że jest to przypuszczalnie najbardziej rozpowszechniona reguła moralna, którą można napotkać w bardzo oddalonych od siebie systemach etycznych. Powołuje się na świadectwa antropologów, którzy mieli zarejestrować jej występowanie w najprostszych społeczeństwach zbieracko-łowieckich. Szczególnie uderzający jest, jego zdaniem, fakt, że podobną regułę odkrywano w społeczeństwach bardzo oddalonych geograficznie, co do których nie istniały wiarygodne przesłanki, że mogły nawiązywać kontakty kulturowe w przeszłości. Co istotne, społeczeństwa te znacznie się różnią, jeśli chodzi o inne elementy kultury. Jednak zasada, na której opierają się ich umowy społeczne, pozostaje niezmienna.

Ważne jest to, że stosunki w tych społecznościach nie przypominają utopijnej idylli. Ich członkowie, jak wszyscy inni ludzie, są równie skłonni do naruszania umowy społecznej obowiązującej w ich społeczeństwach. Egalitarne reguły tworzące te umowy społeczne są w większości przypadków respektowane, ponieważ inni członkowie społeczeństwa ukaraliby osoby niestosujące się do nich. Dzielenie się podstawowymi zasobami jest przestrzegane, ponieważ każdy

wie, że jeśli nie będzie się dzielił, inni nie będą chcieli podzielić się z nim, gdy on będzie w potrzebie.

Te spostrzeżenia dotyczą społeczeństw egalitarnych niezbyt oddalonych od nas w czasie — współczesnych lub w najgorszym razie sprzed 100–150 lat. Z punktu widzenia ewolucji (również kulturowej) jest to oczywiście bardzo krótki okres. Co gorsza, nie jest pewne, czy społeczeństwa zbieracko-łowicze opisane przez antropologów można zasadnie traktować jako „zamrożony w czasie” obraz społeczeństw prehistorycznych, w których ewoluowały nasze współczesne normy *fairness*. W końcu mają one za sobą równie długą historię jak nasze społeczeństwa.

Istnieją jednak pewne przesłanki pozwalające sądzić, że nasi przodkowie tworzyli grupy zdobywające środki utrzymania dzięki polowaniu i zbieractwu. Okres, w którym stanowiło to ich główne źródło utrzymania, był wystarczająco długi, żeby mogły zachodzić procesy ewolucyjne. Binmore, przywołując wprost Marksowską ideę, że sposoby produkcji kształtują formy życia społecznego, zauważa, że pewne umowy społeczne mogą działać, jeśli społeczeństwo jest niewystarczająco produktywne. Zarazem systemy produkcyjne mogą działać, tylko jeśli aktualna umowa społeczna dostarcza odpowiednich bodźców (Binmore, 2011: 131).

Gospodarka zbieracko-łowicka z pewnością nie jest szczególnie produktywna. Możemy więc bez obawy popełnienia istotnego błędu przyjąć, że umowa społeczna, która funkcjonowała w tych społeczeństwach, wymagała stosunkowo prostych, niehierarchicznych relacji społecznych. Ponieważ nie produkowano istotnych nadwyżek, nie było możliwe ich gromadzenie, a co za tym idzie, budowanie relacji władzy opartych na nierównościach ekonomicznych. Prehistoryczne społeczeństwa zbieracko-łowicze musiały być zatem dość egalitarne, jeśli chodzi o stosunki zarówno ekonomiczne, jak i „polityczne”.

Dla koncepcji Binmore’a istotne jest to, że okres „egalitarny” trwał wystarczająco długo, żeby zachodziły w nim procesy ewolucyjne. Poważna zmiana w sposobach produkcji zaszła zaledwie około 10 tysięcy lat temu, co jest zdecydowanie zbyt krótkim okresem, żeby ewolucja w stosunkach społecznych znalazła odbicie w naszym wyposażeniu genetycznym, zasadne jest zatem przypuszczenie, że nasze wyposażenie genetyczne pochodzi z czasu, w którym relacje społeczne były zdecydowanie egalitarne. Okres, w którym pojawiły się hierarchiczne relacje społeczne, musiał być domeną ewolucji kulturowej, a nie biologicznej i to właśnie ewolucja kulturowa odpowiada za sukces, jaki nasz gatunek odniósł po porzuceniu (w większości) gospodarki zbieracko-łowiczej. Jednak nasze wyposażenie genetyczne pozostało niezmienione. Dlatego wykazujemy predyspozycje do takiego rodzaju zachowań społecznych, które są odpowiednie w społeczeństwach egalitarnych, choć nad tymi predyspozycjami zostały „nadbudowane” wzorce kulturowe, dzięki którym jesteśmy w stanie tworzyć

znacznie bardziej zróżnicowane i zhierarchizowane organizacje odpowiadające naszym bardziej efektywnym sposobom produkcji (Binmore, 2011: 133–135).

W społeczeństwach zbieracko-łowieckich, które przetrwały, powszechnie występują dwa elementy. W tego rodzaju społecznościach nie ma wodzów. Niektóre jednostki zyskują większy szacunek albo poważanie, ale nigdy nie są w stanie sprawować osobistej władzy. Co więcej, starannie unikają sprawiania wrażenia, że roszczą sobie do niej pretensje, ponieważ istnieje mechanizm zapobiegający dominacji pojedynczych jednostek. W razie gdy ktoś wykazuje skłonność do wynoszenia się nad innych i narzucania swojego autorytetu, następuje eskalacja nieformalnych kar od wyśmiewania do jawnego bojkotu i otwartej przemocy, co skutecznie powstrzymuje przed tendencjami do tworzenia hierarchicznych struktur władzy.

Pomimo braku ośrodków władzy, które w naszych społeczeństwach rozwiązują problem redystrybucji, w społeczeństwach zbieracko-łowieckich istnieją wyraźne reguły podziału nadwyżek. Zdaniem Binmore'a generuje je poczucie *fairness*, które wyewoluowało wraz ze wspomnianymi wyżej mechanizmami zapobiegającymi dominacji, a które w tych społeczeństwach jest narzędziem selekcji punktów równowagi.

Drugim powszechnie występującym elementem jest dzielenie się żywnością, jednak tylko w przypadku mięsa reguły podziału wydają się fair z naszego punktu widzenia. Co istotne, zasoby te są dostarczane wyłącznie przez mężczyzn. Skuteczność myśliwego w znacznym stopniu zwiększa jego szanse reprodukcyjne, a to stanowi istotny bodziec do podnoszenia efektywności. Jednak efektywność poszczególnych myśliwych, a tym samym ich wkład w pozyskanie cennych zasobów, jest zróżnicowana. Zasoby są dzielone w miarę równo, co prawdopodobnie ma zwiększyć szanse grupy na przetrwanie w mniej sprzyjających okresach. Może to wywoływać wrażenie paradoksu — skoro podział cennych zasobów odbywa się wedle zasady „od każdego wedle jego możliwości, każdemu wedle jego potrzeb”, to czy nie mamy do czynienia z pewną wersją dylematu więźnia? Skoro uczestnictwo we wspólnym polowaniu przynosi niewielki udział, bo cenne zasoby dzielone są w miarę równo, to po co w ogóle w nim uczestniczyć, skoro można na własną rękę pozyskać mniej cenne zasoby, które nie podlegają takiemu podziałowi?

Binmore zwraca uwagę, że jest to problem pozorny, ponieważ grupa operuje w równowadze gry, którą rozgrywają jej członkowie. W związku z tym kooperacja przy polowaniu prowadzi do uzyskania ogólnej wyższej sumy cennych zasobów. Niezbędne jest jedynie podtrzymywanie równowagi, w której polowanie oparte na kooperacji jest najbardziej efektywnym rozwiązaniem, co jest możliwe dzięki temu, że efektywni myśliwi odnoszą sukces reprodukcyjny. Mamy zatem dość dobre podstawy, by przyjąć, że Złota Reguła stanowiła podstawę umowy społecznej naszych prehistorycznych przodków przez wystarczająco długi okres, żeby ukształtować nasze dziedzictwo genetyczne.



W naturalistycznym ujęciu Binmore'a moralność stanowi efekt podejmowanych przez naturę prób rozwiązania problemów selekcji punktów równowagi w drodze ewolucji. Naturalistyczne ujęcie sytuacji początkowej wymaga zatem wskazania procesu ewolucyjnego, który mógł doprowadzić do wyboru takiej, a nie innej umowy społecznej. Autor *Natural justice* odwołuje się do zaproponowanej przez Petera Singera koncepcji rozszerzającego się kręgu (*expanding circle*) (Singer, 2011). Zdaniem tego badacza ewolucja moralności polega na rozszerzaniu dziedziny, w której zasady moralne znajdują zastosowanie. Binmore sądzi, że dziedzina moralności ulega rozszerzeniu, kiedy gracze błędnie odczytują sygnały płynące ze środowiska, wskutek czego mylnie stosują zachowanie albo sposób myślenia, które wyewoluowały w grupie ludzi lub pewnym typie gry, do szerszej grupy ludzi lub innego typu gry. W takiej sytuacji gracze usiłują zachowywać się tak, jak powinni się zachowywać, żeby utrzymać równowagę w pierwotnej grze, i nie dostrzegają, że w szerszym kręgu mogą obowiązywać inne reguły. Przykładowo, uczestnicy gry mogą zachowywać się tak, jak gdyby mieli do czynienia z grą iterowaną, mimo że jest ona jednorazowa. Zazwyczaj strategia, która prowadzi w węższym kręgu do równowagi, nie będzie tworzyć równowagi przy zastosowaniu jej w kręgu szerszym. Mogą jednak zdarzać się wyjątki. Jeśli w szerszym kręgu powstanie równowaga w grze przy wykorzystaniu strategii z węższego kręgu, stanie się tak dzięki temu, że gracze będą się zachowywać tak, jak gdyby grali w grę w kręgu węższym, choć tak naprawdę będą uczestniczyć w grze w kręgu szerszym.

Taki mechanizm, zdaniem Binmore'a, doprowadził do ukształtowania się sytuacji początkowej jako narzędzia służącego do selekcji punktów równowagi. Takie narzędzie jest niezbędne w przypadku gier, w których efektywne punkty równowagi odpowiadają długookresowym porozumieniom dotyczącym podziału żywności, zawartym w celu wzajemnego ubezpieczenia się na wypadek głodu. Rozszerzenie kręgu nastąpiło przez włączenie do grupy, w której obowiązują zasady moralne, ludzi obcych, co było możliwe dzięki traktowaniu ich, jak gdyby byli krewnymi. Rozszerzenie grupy przez sztuczne rozciągnięcie więzi pokrewieństwa było jedyną dostępną możliwością, ponieważ na nich opierały się początkowo więzi grupowe, czego świadectw dostarczyły liczne badania antropologiczne. Tym samym preferencje empatyczne, które początkowo stosowane były w wąskim kręgu, zaczęto wykorzystywać również w kręgu szerszym. Przejście do sytuacji początkowej za zasłoną niewiedzy wymagało połączenia tych dwóch kroków.

Aby rozstrzygnąć tę kwestię, rozważamy, co wchodzi w grę, gdy racjonalni gracze rozważają różne przypadki, które mogą się pojawić, gdy planują przyszłość. Aby je ocenić, gracze kalkuluje swoją oczekiwaną użyteczność jako średnią ważoną wypłat wszystkich przyszłych ludzi — którym się powiodło albo nie powiodło — którymi mogą się okazać, kiedy rzucona kostka przestanie się toczyć. Kiedy wybierają strategię

w grze rozgrywanej w rodzinie, gracze w podobny sposób przyjmują, że ich wypłaty będą średnią ważoną przystosowania każdej osoby w rodzinie. W celu zmiany naszej zdolności do negocjowania umów ubezpieczenia w zdolność wykorzystywania *fairness* jako bardziej ogólnego narzędzia koordynacji w grze życia potrzeba wówczas jedynie hybrydyzacji tych dwóch procesów przez umożliwienie zastąpienia jednej z przyszłych osób, którą wskutek rzutu kością mogą się stać, przez osobę w innym ciele. Preferencje empatyczne, które są potrzebne do oceny tej możliwości, wymagają nie więcej, niż żeby traktować tę osobę w innym ciele w prawie taki sam sposób, w jaki traktuje się siostry, kuzynów albo ciotki (Binmore, 2011: 142).

Wagi przystosowania potrzebne do określenia wypłat są uzyskiwane „w jakiś sposób” na podstawie oceny stopnia pokrewieństwa. Natomiast wagi potrzebne do określenia preferencji empatycznych są określane przez ewolucję społeczną. Poniżej omówię tę kwestię szerzej, wyjaśniając, dlaczego — zdaniem Binmore’a — sytuacja początkowa prowadzi do wyboru raczej dwóch zasad sprawiedliwości Rawlsa niż zasady użyteczności.

Binmore jest przekonany, że szczególnie rolę odgrywa tu obecność lub nieobecność podmiotu egzekwującego porozumienie. Kwestię tę poruszałem już skrótowo wcześniej. W społeczeństwach zbieracko-łowieckich tego rodzaju podmiot nie istnieje. Z uwagi na sposób pozyskiwania żywności przeciętna wielkość tego rodzaju społeczności nie może być duża. Sprawia to, że kontrola społeczna jest w nich znacznie ułatwiona i skuteczna. Członkowie takich społeczeństw w miarę łatwo mogą wykryć nadużycia związane z nieprzestrzeganiem norm *fairness* i odpowiednio ukarać naruszających reguły.

Binmore jest przekonany, że istnieją przynajmniej dwa istotne powody, dla których ewolucja kulturowa doprowadza za każdym razem, gdy tylko jest to możliwe, do powstania umów społecznych zapewniających najbardziej skuteczne mechanizmy utrzymywania porządku. Przede wszystkim społeczeństwa tego rodzaju nie rozrastają się. Liczebność populacji ma pierwszorzędne znaczenie dla ich przetrwania, gdy dostępne są tylko ograniczone źródła pożywienia, których w żaden sposób nie można kontrolować. Warunki, w jakich one funkcjonują, przeciwdziałają zwiększeniu liczebności społeczeństw zbieracko-łowieckich. Zatem z uwagi na warunki życiowe są niewielkie szanse na pojawienie się hierarchicznych struktur władzy.

Drugą ważną przyczyną, dla której tego rodzaju struktury są niebezpieczne dla społeczeństw balansujących najczęściej na granicy przeżycia, jest to, że w okresach szczególnego niedoboru pożywienia, kiedy istnienie grupy jest zagrożone, nawet jeśli część członków społeczności zginie z głodu, reszta może znaleźć schronienie w innych grupach działających w oparciu o podobną umowę społeczną. W okresach bardziej sprzyjających, które oddzielają dość rzadkie okresy szczególnego niedoboru, dzięki nadwyżkom pożywienia jednostki o wyróżnionej pozycji społecznej mogłyby doprowadzić do zmiany podstaw organizacji społeczeństwa, reformując umowę społeczną w kierunku innego sposobu

redystrybucji. Prowadziłoby to do zaburzenia mechanizmu zapewniającego przeżycie okresów niesprzyjających, a tym samym groziło całkowitym unicestwieniem grupy w tym czasie. Innymi słowy, mechanizmy społeczne zapobiegające ukształtowaniu się pozycji władzy powstały w wyniku presji ewolucyjnej, ponieważ społeczeństwa zorganizowane w ten sposób są w stanie przetrwać okresy szczególnego niedoboru (Binmore, 1998: 403–405).

Binmore zauważa, że choć nie ma podmiotu egzekwującego umowę społeczną w skali całego społeczeństwa i dlatego na tym poziomie potrzebny jest mechanizm zapewniający samoczynne jej egzekwowanie, taki podmiot istnieje w odniesieniu do różnych podgrup funkcjonujących w jego ramach. Jest nim koalicja wszystkich tych członków społeczeństwa, którzy nie należą do danej podgrupy. W związku z tym może zachodzić sytuacja, kiedy w skali całego społeczeństwa narzędziem rozwiązywania problemów dotyczących koordynacji jest sytuacja początkowa, natomiast „lokalnie” zasadą pełniącą analogiczną funkcję może być zasada użyteczności.

Przypuszczalna ewolucyjna historia formowania się zasad sprawiedliwości prezentuje się w ujęciu Binmore’a w największym uproszczeniu następująco. Za zasłoną niepewności zawierane są umowy ubezpieczenia ustalające zasady dzielenia się pożywieniem. Niezbędne jest, by strony umowy były zdolne do wzajemności. Zawarcie umowy za zasłoną niepewności jest możliwe dzięki temu, że strony sympatyzują ze swoim przyszłym Ja — troszczą się o swoją przyszłość. Za zasłoną niewiedzy w sytuacji początkowej zawierana jest umowa społeczna zawierająca zasady *fairness*. Jest to możliwe dzięki temu, że strony sympatyzują z innymi — są w stanie postawić się w ich sytuacji. W zależności od tego, czy istnieje zewnętrzna instytucja egzekwująca umowę społeczną, wybrane zostaną różne zasady *fairness*: zasada użyteczności, jeśli taki podmiot istnieje, albo dwie zasady sprawiedliwości Rawlsa, jeśli taki podmiot nie istnieje. Ewolucja kulturowa prowadzi w obu przypadkach do ustalenia zasad porównań międzyosobowych (Binmore, 2011: 142).

## ZASADA ZRÓŻNICOWANIA

Binmore, choć aprobeuje sformułowane przez Rawlsa zasady, ma jednoznaczną opinię na temat ich uzasadnienia zaproponowanego przez autora *Teorii sprawiedliwości*:

John Rawls twierdził, że racjonalny targ w sytuacji początkowej doprowadzi do umowy społecznej, w której pewne podstawowe wolności i swobody zostają zabezpieczone w pierwszej kolejności, a następnie pozostałe kwestie dystrybucyjne zostają rozwiązane przy wykorzystaniu jego zasady różnicowania. Sądzę jednak, że Rawls nie jest uprawniony do żadnego z dwóch wyciągniętych przez siebie wniosków (Binmore, 2011: 166).

Zdaniem Binmore'a pierwszy wniosek — o czym już wspominałem — nie wynika z samej sytuacji początkowej. Przede wszystkim wymaga założenia, że gra życia, którą prowadzą uczestnicy sytuacji początkowej, jest wystarczająco zbliżona do tej, którą znamy, żeby uzasadnić wybór dwóch zasad sprawiedliwości. Jeśli gra życia miałaby być wojną wszystkich ze wszystkimi, wybór Hobbesowskiej umowy społecznej byłby bardziej uzasadniony. Nie wydaje się jednak, żeby ten zarzut Binmore'a był trafny. Rawls zaznacza, że elementem definiującym sytuację początkową jest Humowskie założenie umiarkowanego niedoboru. Oznacza to, ujmując rzecz w terminologii Binmore'a, że gra życia prowadzona przez uczestników sytuacji początkowej toczy się w okolicznościach, które nie wymagają, żeby priorytetem uczestników było fizyczne przetrwanie. W takich okolicznościach, jak zauważa Rawls, kwestie sprawiedliwości schodzą na dalszy plan. Z punktu widzenia Binmore'a oznacza to po prostu przemycanie normatywnych założeń ograniczających z góry zbiór możliwych do zrealizowania umów społecznych. Jednak nawet jeśli przyjąć, że gra życia jest wystarczająco zbliżona do naszej, żeby uznać wnioski Rawlsa za zasadne, to — jak była mowa wyżej — kwestie dystrybucji byłyby rozstrzygane przy użyciu zasady użyteczności. Według Binmore'a, jeśli wnioski wyciągnięte przez Rawlsa mają zostać utrzymane, potrzebna jest dość daleko idąca modyfikacja sytuacji początkowej.

Binmore uważa, że najistotniejszą modyfikacją powinno być doprowadzenie Rawlsofskiej koncepcji wysiłku zaangażowania (*strains of commitment*) do jej „ostatecznych logicznych konsekwencji”. Rawls zauważył, że aby strony umowy zawieranej w sytuacji początkowej mogły traktować się nawzajem serio jako partnerów porozumienia, muszą przyjąć pewne założenie dotyczące swojej (i swoich partnerów) zdolności do respektowania warunków tego porozumienia. Tylko wówczas strony

mogą być pewne, że ich zobowiązanie nie jest podejmowane na próżno. Zakładając, że biorą wszystko pod uwagę, włącznie z ogólnymi faktami psychologii moralnej, mogą na sobie polegać, że będą się do przyjętych zasad stosowały. Biorą więc pod uwagę wysiłek zaangażowania. Nie będą przystępować do umów mogących rodzić takie konsekwencje, jakich nie mogłyby zaakceptować. Będą też unikać takich, których mogłyby dotrzymać jedynie z wielką trudnością (Rawls, 2009: 262–263).

W ujęciu Rawlsa jest to szczególnie ważne, ponieważ rozstrzygnięcie dokonywane w sytuacji początkowej jest ostateczne i nie ma możliwości jego późniejszego skorygowania, a wybierane zasady mają mieć istotny wpływ na perspektywy życiowe jednostek. Binmore również przypisuje duże znaczenie możliwości respektowania przyjętych reguł, choć z innych względów niż sam Rawls. Zdaniem Binmore'a koncepcja *fairness* musi brać pod uwagę, czy strony są w stanie respektować zawarte porozumienie, ponieważ przy braku zewnętrznej instytucji egzekwującej umowę społeczną *fairness* musi stanowić

samoczynnie podtrzymującą się równowagę gry, niezależnie od tego, jaka jest to gra (Binmore, 1998: 436–439).

Druga kwestia, którą Binmore chce zmodyfikować, dotyczy roli praw i swobód. W ujęciu Rawlsa pierwsza zasada równych wolności ma stanowić wynik targu w sytuacji początkowej za zasłoną niewiedzy. Autor *Natural justice* traktuje równe wolności nie jako wynik targu, lecz jako warunek wstępny niezbędny do tego, by reguły *fairness* same pełniły funkcję skutecznego narzędzia selekcji punktów równowagi.

Ostatnia modyfikacja, którą Binmore wprowadza do koncepcji Rawlsa, dotyczy takiej zmiany problemu targu, żeby zasada zróżnicowania mogła być jego egalitarnym rozwiązaniem. Co istotne, Binmore redukuje drugą zasadę sprawiedliwości do reguły maksyminowej głoszącej, że najlepszą umową społeczną jest taka umowa, która najbardziej niekorzystnie usytuowanego gracza stawia w możliwie najlepszym położeniu. Takie ujęcie zasady zróżnicowania wydaje się zbyt daleko idącym jej uproszczeniem, ponieważ nie uwzględnia zawartej w niej zasady autentycznej równości szans i zasady sprawiedliwego oszczędzania (z odpowiednią regułą pierwszeństwa) (Rawls, 2009: 438). Jednak dzięki takiemu zredukowaniu drugiej zasady i przy założeniu, że wypłaty uczestników sytuacji początkowej są mierzone za pomocą skal empatycznych, których zerem jest *status quo*, a zbiór możliwych do zrealizowania umów społecznych jest zbiorem wypukłym, reguła maksyminowa wskazuje bezpośrednio egalitarne rozstrzygnięcie problemu targu w sytuacji początkowej.

Binmore twierdzi, że wynik uzyskany przez Rawlsa można osiągnąć, nie poświęcając ortodoksyjnej teorii sprawiedliwości. Trzeba jednak poświęcić niektóre jej założenia, żeby uniknąć wniosku, że w sytuacji początkowej za zasłoną niewiedzy strony rozumowałyby jak ortodoksyjni utilitaryści. W tym celu należy odrzucić założenie, że prawdopodobieństwo stania się którąkolwiek ze stron umowy po podniesieniu zasłony niewiedzy jest takie samo. W sytuacji, w której zostałyby wybrana zasada użyteczności, możemy przyjąć, że rozstrzyga o tym rzut rzetelną monetą w chwili podniesienia zasłony niewiedzy<sup>6</sup>. Jednak by wybór padł na egalitarną zasadę zróżnicowania, musimy rozważyć pewne kwestie dotyczące tego hipotetycznego rzutu monetą. A przede wszystkim to, dlaczego strony umowy czują się zobowiązane do respektowania wyniku rzutu niezależnie od tego, jaki on będzie (Binmore, 1998: 436–439).

Binmore zwraca uwagę, że gdy istnieje zewnętrzna instytucja egzekwująca rozstrzygnięcia, odwołania do zasad *fairness* w chwili podejmowania decyzji mogą zostać zastąpione wcześniejszym rozstrzygnięciem decydującym, jaki wynik będzie uznany za fair. Jeśli ustalimy wcześniej, że wybór strony boiska

---

<sup>6</sup> Binmore dla uproszczenia przyjmuje, że społeczeństwo składa się z dwóch osób, więc rozstrzygnięcie za pomocą rzutu monetą adekwatnie obrazuje sytuację wychodzenia zza zasłony niewiedzy.

przed meczem piłkarskim zostanie uzgodniony za pomocą rzutu monetą, to w przypadku niekorzystnego wyniku nie można skarżyć się na tej podstawie, że gdyby wybrało się drugą stronę monety, wynik byłby dla nas korzystny. W takiej sytuacji uznano by, że strona żądająca powtórzenia losowania w celu uzyskania korzystnego dla niej rozstrzygnięcia nie respektuje wcześniejszego rozstrzygnięcia polegającego na wyborze strony monety przed losowaniem. Zewnętrzna instytucja (sędzia, druga strona, kibice) nie pozwoliłaby na podważenie na takiej podstawie wyniku losowania. Co jednak stanie się w sytuacji, gdy nie ma zewnętrznej instytucji, a losowanie ma przynieść rozstrzygnięcie znacznie poważniejszej kwestii — na przykład jakie będą nasze perspektywy życiowe mające wpływ na wszystkie aspekty naszej egzystencji.

Binmore przypomina, że w jego naturalistycznym ujęciu zasady *fairness* nie stanowią alternatywy dla władzy (jak, jego zdaniem, jest w przypadku tradycyjnych koncepcji sprawiedliwości), lecz służą jako instrument wyważania władzy w społeczeństwie. W takim ujęciu strona niezadowolona z rozstrzygnięcia nadal posiada całą władzę, która nie została scedowana na zewnętrzną instytucję egzekwującą porozumienia, i może ją wykorzystać, by uzyskać korzystne dla siebie rozstrzygnięcie. Oczywiście takiego postępowania nie można uznać za działanie nie fair, ponieważ dopiero ma zostać ustalone, co będzie, a co nie będzie za takie działanie uznawane. Również stwierdzenie, że w takim razie nic nie zostanie rozstrzygnięte, nic nie wnosi, ponieważ brak rozstrzygnięcia jest jedną z możliwych opcji, które strony mogą wybrać. Innymi słowy, jeśli stronom będzie zależało w różnym stopniu na wyniku losowania, przy braku zewnętrznej instytucji egzekwującej egalitarna zasada zróżnicowania nie zostanie wybrana.

Rozwiązanie wydaje się więc trywialnie proste — stronom w sytuacji początkowej musi w równym stopniu zależeć na wyniku losowania, żeby przy braku zewnętrznej instytucji egzekwującej wybór padł na drugą zasadę sprawiedliwości Rawlsa. Z rozważań poświęconych wyborowi zasady użyteczności w sytuacji początkowej wynika, że stanie się tak, gdy strony będą miały równe użyteczności empatyczne. Dla uproszczenia można przyjąć, że gdy stronom zależy na wyniku losowania w równym stopniu, są one indyferentne wobec wyniku losowania.

## RAWLS POPRAWIONY CZY ODRZUCONY?

Stanowiska Rawlsa i Binmore'a są sobie przeciwstawne w tym sensie, że Rawls prezentuje koncepcję normatywną, która pokazuje, jakie powinny być zasady i mechanizmy selekcji, natomiast brytyjski ekonomista tworzy teorię czysto opisową, która wyjaśnia, skąd biorą się te zasady i jak działają mechanizmy selekcyjne. Binmore, jak podkreślałem, traktuje swoje stanowisko

jako koncepcję czysto opisową. Kiedy przechodzi do formułowania pewnych wniosków normatywnych, prezentuje je wyłącznie jako subiektywne zalecenia, które mają się opierać na ustaleniach wypływających z badań naukowych. Pod tym względem stanowiska Rawlsa i Binmore'a mają wiele wspólnego. Autor *Teorii sprawiedliwości* wielokrotnie podkreślał, że na wszystkich etapach argumentacji na rzecz zasad sprawiedliwości powinniśmy w maksymalnym stopniu odwoływać się do ustaleń formułowanych przez różne dyscypliny naukowe. Warunek dostępu do pełnej wiedzy o charakterze ogólnym w sytuacji początkowej ma to właśnie zapewnić. Jednostki dokonujące wyboru zasad sprawiedliwości, mimo że nie wiedzą, kim są i jakie są ich indywidualne plany życiowe, posługują się w namyśle wiedzą dostarczaną przez nauki szczegółowe. Podobnie jest na innych etapach argumentacji. W argumencie o możliwości realizacji zasad sprawiedliwości (*feasibility*) Rawls odwołuje się do ustaleń nauk ekonomicznych i politycznych, żeby pokazać, jakie efekty przyniosłoby zastosowanie zasad wybranych w sytuacji początkowej do konstruowania w praktyce systemu społecznego.

Autor *Natural justice* traktuje sytuację początkową zupełnie inaczej niż sam jej twórca. W ujęciu Rawlsa sytuacja początkowa rozumiana jest jako „narzędzie reprezentacji” pozwalające przedstawić wszystkie istotne aspekty, które muszą zostać uwzględnione w publicznej argumentacji dotyczącej przyjęcia zasad sprawiedliwości. Można ją traktować jako model takiej argumentacji tylko w tym sensie, że wyznacza ona normatywne wymogi narzucone temu rozumowaniu. W ujęciu Binmore'a sytuacja początkowa jest modelem w tym sensie, że przedstawia proces prowadzący do powstania równowagi w grze. Funkcja, jaką ta idea spełnia w obu koncepcjach, jest więc zupełnie odmienna.

Choć obaj autorzy traktują problem wyboru zasad *fairness* jako problem targu w odpowiednio zdefiniowanej sytuacji wyboru, proponowane przez nich rozwiązania diametralnie się od siebie różnią. Propozycja Rawlsa redukuje problem targu do namysłu prowadzonego przez pojedynczą jednostkę. Warunki sytuacji początkowej są tak zdefiniowane, że problem wyboru prezentuje się identycznie z perspektywy każdej ze stron. W efekcie problem targu zostaje zastąpiony jednostkową decyzją. Tymczasem Binmore ustala warunki sytuacji początkowej w taki sposób, że problem targu zostaje zachowany. Jednak sytuacja targu zostaje zmodyfikowana, ponieważ strony nie znają swoich preferencji indywidualnych. Dlatego rozumowanie zmierzające do rozstrzygnięcia prowadzone jest tak, jak gdyby znajdowały się one za zasłoną niepewności, która uniemożliwia im dowiedzenie się, którą ze stron okażą się po dokonaniu rozstrzygnięcia. Jednak targ pozostaje targiem, ponieważ strony mają preferencje empatyczne. Jeśli preferencje te się różnią, rozstrzygnięcie targu jest nietrywialne.

Koncepcje Rawlsa i Binmore'a zostały stworzone w innych celach i spełniają odmienne funkcje. Choć Binmore głosi, że jego celem było zaproponowanie alternatywnego uzasadnienia zasady zróżnicowania, w gruncie rzeczy przedstawił

alternatywną wizję uprawiania refleksji nad problematyką sprawiedliwości. Rawlsowi chodziło o sformułowanie przekonujących argumentów za określoną zasadą sprawiedliwości, a Binmore dążył do stworzenia modelu, który pokazuje, że zasada zróżnicowania jest tą zasadą, którą rzeczywiście respektujemy w codziennych interakcjach generujących problem koordynacji. W związku z tym model uzasadnienia zaproponowany przez Binmore'a można potraktować jako alternatywę dla modelu Rawlowskiego tylko o tyle, o ile jesteśmy gotowi uznać, że rola teorii moralnej jest zasadniczo odmienna od tej, jaką przypisywał jej autor *Teorii sprawiedliwości*. Jeśli jednak chcemy pozostać wierni tradycyjnemu rozumieniu teorii moralnej, pozostaje nam uznać koncepcję Binmore'a za interesującą propozycję z zupełnie odmiennego obszaru badań nad kwestią sprawiedliwości.

## BIBLIOGRAFIA

- Aumann, R. (1959). *Acceptable points in general cooperative n-person games* (s. 307–317). W: H. Kuhn, A.W. Tucker (Red.). *Contributions to the Theory of Games IV. Annals of Mathematics Study 40*. Princeton: Princeton University Press.
- Binmore, K. (1994). *Game theory and the social contract* (t. 1: *Playing fair*). Cambridge & London: MIT Press.
- Binmore, K. (1998). *Game theory and the social contract* (t. 2: *Just playing*). Cambridge & London: MIT Press.
- Binmore, K. (2007). *Playing for real: A text for game theory*. Oxford: Oxford University Press.
- Binmore, K. (2011). *Natural justice*. Oxford: Oxford University Press.
- Hume, D. (1963). *Traktat o naturze ludzkiej* (t. 2: *O uczuciach*). (Przeł. Cz. Znamierowski). Warszawa: Państwowe Wydawnictwo Naukowe.
- Rawls, J. (2009). *Teoria sprawiedliwości. Wydanie nowe*. (Przeł. M. Panufnik, J. Pasek, & A. Romaniuk, przekł. popr. S. Szymański). Warszawa: Wydawnictwo Naukowe PWN.
- Singer, P. (2011). *The expanding circle: Ethics, evolution, and moral progress*. Princeton: Princeton University Press.