

JAMES GRIFFIN

Intersubiektywne porównania użyteczności*

Przedmiot rozważań

Jak powinno się mierzyć nierówność? Jedną z możliwych odpowiedzi brzmi: powinno się mierzyć użyteczności i ustalić ich rozkład. Odpowiedź ta była i nadal jest poddawana w wątpliwość. Jednym z wysuwanych zarzutów jest zastrzeżenie, że porównywanie użyteczności przypadających różnym osobom — co jest nieuniknionym następstwem przyjęcia powyższej odpowiedzi wiedzie do nierozwiązywalnych trudności. W niniejszym artykule chcę odeprzeć ten zarzut.

Przed wszystkim podejmowanie podobnych porównań jest nieuniknione. Nie osiągnęlibyśmy żadnych rezultatów w rozważaniach na temat sprawiedliwości i wolności, gdybyśmy nie byli w stanie stwierdzić, że ta oto osoba jest w gorszej sytuacji niż tamta, lub że straty poniesione przez tę właśnie osobę są większe niż straty poniesione przez inną. Większość ekonomistów zajmuje jednak bardzo sceptyczne stanowisko na temat możliwości intersubiektywnych porównań użyteczności i poszukuje takich sposobów dokonywania wyborów społecznych, które nie uwzględniają w ogóle użyteczności. Z tego też powodu, wielu ekonomistów przyjmuje kryterium potencjalnej optymalności w sensie Pareto: badamy sytuację społeczną, stawiając pytanie: czy istnieje taka zmiana, w której osoby odnoszące korzyść mogłyby tytułem odszkodowania dla osób ponoszących stratę ustalić ich sytuację na aktualnym poziomie powodując jednocześnie, że przynajmniej jedna osoba odniesie korzyść w wyniku zmiany. Stosując tę metodę unika się intersubiektywnych porównań. Wystarczy byśmy zapytali każdego, kto zyskuje, i każdego, kto traci, czy uwzględniając kompensację znajdzie się w lepszej czy w gorszej sytuacji niż dotąd. Test potencjalnej optymalności Pareto nie zapewnia

* Jest to tekst referatu wygłoszonego w Warszawie na polsko-brytyjskiej konferencji etycznej: 3 X — 10 X 1979.

nam jednak nic więcej [A. K. Sen 1973, s. 12—15]. Może być stosowany do oszacowania zmian dotyczących powszechnego dobrobytu, lecz nie mówi nic na temat sprawiedliwości i równości. Autorzy opierający się na analizie nakładów i efektów stwierdzają, że potencjalna optymalność w sensie Pareto jest uzasadnieniem dla całości ich doktryny [por. E. J. Mishan 1972, s. 14; 1975, s. XVIII]. Sądzą oni, iż ma ona maksymalizować efekty w stosunku do nakładów, wśród których głównym czynnikiem — przy pewnych dodatkowych założeniach — jest równość. Innymi słowy, analiza nakładów i efektów eliminuje porównania międzyjednostkowe jedynie w tym sensie, że usuwa je z obrębu teoretycznych rozważań pozostawiając to zadanie do wykonania jednostkom, które same mają ferować pewne sądy na temat sprawiedliwości i równości.

Istota problemu

Jakie istotne zagadnienia wiążą się z porównywaniem użyteczności przypadającej różnym jednostkom? Ekonomiści często określają owe porównania jako „bezsensowne”, ponieważ nie podlegają one weryfikacji¹. Trudno jednak takie zastrzeżenie brać serio. Gdyby intersubiektywne porównania użyteczności były rzeczywiście pozbawione sensu, to nic z tego, co przytrafia się danej osobie, nie byłoby nigdy lepsze lub gorsze od tego, co przytrafia się innym. Podobnie uproszczone wersje weryfikacjonistycznej teorii znaczenia ze zrozumiałych powodów mają obecnie niewielu zwolenników.

Ekonomiści wysuwają jednak mniej radykalny argument twierdząc, że intersubiektywne porównania użyteczności nie są stwierdzeniem faktu, a więc nie są sądami należącymi do nauki, lecz są sądami wartościującymi. Tak właśnie argumentuje Lionel Robbins [1935, rozdz. 6] w swych klasycznych rozważaniach na ten temat. Podobnego argumentu, choć w zmodyfikowanej formie, użył też ostatnio John Rawls [1972, rozdz. 49]. Trudno jednak zastrzeżenie to traktować poważnie. Gdyby było ono słuszne, trzeba by przyjąć, że za każdym razem, gdy ktoś mówi, że spośród dwóch osób jedną ząb boli bardziej niż drugą, wypowiada jedynie swe oceny i nie opisuje niczego, co rzeczywiście istnieje. Bez wątplenia jednak ani Robbins, ani Rawls nie nadają swym twierdzeniom takiego sensu. Mają raczej na myśli coś innego, a mianowicie, że przechodząc od systemu pomiarów użyteczności przypadającej jednej osobie do systemu pozwalającego dokonywać intersubiektywnych porównań musimy się odwo-

¹ Por. K. J. Arrow 1963, s. 9: Przyjmie się, że intersubiektywne porównania użyteczności nie mają żadnego sensu...”

łać do sądów wartościujących, a tym samym intersubiektywne porównania użyteczności wcale nie są stwierdzeniami faktów².

Bez wątplenia w jakimś sensie sądy wartościujące muszą być wprowadzone do porównań. Weźmy na przykład całkiem sensowną propozycję von Neumanna i Morgensterna, by dla określenia indywidualnych użyteczności używać skali liczbowej. W jaki sposób można taką skalę wykorzystać dla porównywania użyteczności? Zazwyczaj przyjmuje się, że wystarczy taką skalę wyzerować [por. A. K. Sen 1970, s. 98—99; J. Rawls 1972, s. 323—325]. To znaczy, że dla każdej z interesujących nas osób wystarczy przypisać wartość z e r o tej sytuacji, którą określa ona jako sytuację najgorszą oraz przypisać wartość jeden sytuacji określonej przez nią jako najlepsza. Następnie, opierając się na liczbowych pomiarach użyteczności pozostałych sytuacji, rozmieszczamy je odpowiednio między dwoma ekstremami. Na pierwszy rzut oka procedura ta wydaje się godna przyjęcia z uwagi na swe zalety moralne: wydaje się, że u jej podstaw leży demokratyczne założenie, wedle którego każdej osobie przysługuje jeden i tylko jeden głos w decydowaniu na temat wyborów społecznych. Choć z pozoru wydaje się, że jest to założenie dobre, chwila zastanowienia wystarczy, abyśmy spostrzegli, że w istocie jest to założenie rażąco arbitralne. Jest ono bowiem równoważne stwierdzeniu, że każdy człowiek ma tę samą zdolność zaspokajania pragnień — a to nie jest prawdą. Przy bliższym badaniu znika nawet pozór zalet moralnych. Jak podkreśla Rawls, zasada zerowania zakłada, że większa użyteczność społeczna płynie z takiego wytresowania ludzi, by poprzestawali na prostych potrzebach, dających się łatwo zaspokoić: gdy nikt nie mierzy wysoko, sytuacje dające się łatwo osiągnąć są cennie wyżej. Jednak niezależnie od tego, jak arbitralna czy jak moralnie trafna jest to zasada, opiera się ona na założeniu ustalającym, jaką wagę przypiszemy użytecznościom przypadającym poszczególnym jednostkom, i to jest właśnie sąd wartościujący, o którym mówili Robbins i Rawls.

Inny dobrze znany sposób wprowadzenia porównywania użyteczności wygląda następująco: najpierw wprowadzamy skalę użyteczności dla jednej osoby, przyjmując całkiem prawdopodobne założenie, że nikt nie potrafi z nieograniczoną dokładnością odróżniać poziomów użyteczności. Następnie ustalamy, że dana osoba nie ma żadnych preferencji na temat zajścia poszczególnych zdarzeń z tego samego poziomu, oraz że każdy poziom jest preferowany w stosunku do poprzedniego zawsze z taką samą mocą. Założenia te wystarczają do wprowadzenia skali licz-

² Por. O. Morgenstern 1972, s. 1169: „Podstawową zaletą pojęcia optymalności w sensie Pareto jest rzekomo to, że można uniknąć intersubiektywnych porównań użyteczności, które, jak dobrze wiadomo, zawsze muszą się opierać na sądach etycznych, nie mieszczących się w teorii deskryptywnej”.

bowej przypisanej dla każdej jednostki. By zapewnić sobie nadto porównywalność poszczególnych skal, wystarczy założenie, że odstęp między poszczególnymi poziomami są równe dla wszystkich jednostek. Znow, na pierwszy rzut oka wydaje się, że jest to doskonałe, demokratyczne założenie: każdy człowiek jest „w środku” podobny do innego. Konsekwencje tego założenia są jednak fatalne. Jak wykazał Rothenberg [1961, s. 199; por. też A. K. Sen 1970, s. 93—94], pewni ludzie reagują tylko w sytuacjach krańcowych, uznając, że wszystko na świecie jest albo „cudowne” albo „nie do zniesienia”. Inni skłaniają się do wprowadzenia bardziej subtelnych odróżnień, uznając pewne rzeczy za „doskonałe”, inne tylko za „dobre”, jeszcze inne za „średnie”, i wreszcie pozostałe bądź za „dostateczne”, bądź za „niedostateczne”. Jest więc niesłuszne przypisywać tę samą wagę przeniesieniu osoby pierwszego rodzaju z jej jedyne go niższego poziomu na wyższy, co przeniesieniu osoby drugiego rodzaju z jednego z jej niższych poziomów na poziom bezpośrednio wyższy. Założenie, że ludzie są równi pod względem zdolności i gotowości wprowadzenia pewnych rozróżnień nie jest ani poprawne ani potrzebne, choć to ostatnie stwierdzenie samo też jest przykładem arbitralnego sądu wartościującego na temat szacowania użyteczności przypadających różnym ludziom.

Teraz jednak sposób użycia pojęcia „sąd wartościujący” w argumentacji tym musi się wydać dziwny. Robbins mówi o „wielkiej różnorodności założeń faktycznie poczynionych w różnym czasie i miejscu” biorąc ją za dowód „konwencjonalnego charakteru” tych założeń. W zachodnich demokracjach przyjmuje się, mówi Robbins, że różni ludzie zdolni są w tym samym stopniu do zaspokajania swych pragnień; w innych społeczeństwach przyjmuje się jednak, że pewne rasy, klasy lub kasty odczuwają lub reagują słabiej [L. Robbins 1936, s. 140]. W tym przypadku „sąd wartościujący” trzeba rozumieć jako „arbitralne założenie”, co jest bardzo szczególnym i zdegenerowanym przypadkiem użycia tego wyrażenia, wywodzącym się, jak sądzę, z pozytywistycznego podziału twierdzeń na empirycznie sensowne i pozostałe, gdzie „pozostałe”, ponieważ nie odnoszą się do faktów, mogą być łatwo uznane za wyraz subiektywnych odczuć, a tym samym, za w pewnym sensie „wartości”. Wyrażenie „sąd wartościujący” nie powinno być używane, moim zdaniem, w taki sposób, by mogło równie dobrze oznaczać „domysł”, „upraszczające założenie” czy „arbitralną decyzję”. Weźmy przykład Robbinsa, że ludzie mają tę samą zdolność zaspokajania pragnień. Jest prawdą, że takie założenie jest uproszczeniem. Może to być jednak uproszczenie, które w pewnych okolicznościach, gdy nie można przyjąć lepszych założeń, warto przyjąć. Oczywiście, można też przyjąć inne założenie na podobnych warunkach; i tym samym, każde założenie okazuje się kon-

wencjonalne. Jednakże status założenia, nawet w przypadku, gdy założenie nie wydaje się nieuniknione, nie nadaje założeniu automatycznie statusu sądu wartościującego. Twierdzenie, że każdy ma tę samą zdolność zaspokajania pragnień pozostaje twierdzeniem faktycznym; i co więcej jest fałszywe³.

Ten sam argument w sformułowaniu Rawlsa [1972, rozdz. 15 i 49] nie wydaje mi się bardziej przekonujący. Wygląda on tak: istnieją rozmaite metody zdefiniowania intersubiektywnych porównań użyteczności (jest to bez wątpienia uwaga trafna, jeśli ją interpretować jako obserwację historyczną); metody te opierały się na uderzająco różnych założeniach (to znów jest, historycznie rzecz biorąc, zgodne z prawdą); jest sprawą moralności wybrać którąś z tych metod, jeśli w ogóle którąkolwiek z nich należy wybrać. Wywód ten kończy uwaga: „choć jest oczywiste, że uznanie zasady użyteczności jest sprawą teorii moralności, jest mniej widoczne, że sama procedura mierzenia dobrobytu rodzi podobne problemy. Ponieważ procedur tych jest więcej niż jedna, wybór którejkolwiek z nich zależy od przewidywanego użycia procedury, a tym samym zależy ostatecznie od ustaleń o charakterze etycznym” [J. Rawls 1972, s. 324].

Z tego właśnie powodu niepewność związana z zasadą użyteczności nie da się przezwyciężyć przez zastosowanie bardziej precyzyjnych mierników użyteczności przypadającej różnym osobom, gdyż wybór mierników jest determinowany względami etycznymi. Rawls nie twierdzi więc, że intersubiektywne porównania użyteczności są sędami wartościującymi, lecz mówi, że porównania te opierają się na sądach wartościujących. Jeśli przyznamy Rawlsowi rację, będziemy zmuszeni uznać nie tylko to, że przyjęte metody intersubiektywnych porównań opierają się na arbitralnych założeniach — co jest faktem historycznym — ale również będziemy musieli uznać znacznie silniejsze twierdzenie, mianowicie, że metody te z konieczności muszą być arbitralne. Sprawa ta wymaga jednak dalszych badań. Być może istnieją metody dokonywania porównań intersubiektywnych, które nie wymagają przyjęcia arbitralnych założeń, a tym bardziej — nierealistycznych lub nieprawdziwych. Być może istnieją założenia jednocześnie realistyczne i dające się uzasadnić. Podobnie nawet jeśli wszystkie metody tego rodzaju opierają się na arbitralnych założeniach, choćby dlatego, że założenia te stanowią swoiste uproszczenie (i że uproszczenia mogą być dokonane w różny sposób), zmuszeni jesteśmy przyjąć, że upraszczające założenia same są lub opierają się na sądach wartościujących. Ale twierdzenie to także wy-

³ L. Robbins [1935, s. 139] dokonuje przejścia od „opierają się one [intersubiektywne porównania] na ocenach zawierających umowne elementy” do „są wobec tego zasadniczo normatywne”.

maga dowodu, ponieważ istnieją inne możliwości. Wybór założeń upraszczających może być dokonany nie pod wpływem etycznych racji, takich jak chęć zaprowadzenia równości, ale przez uwzględnienie upraszczających technik, które na dłuższą metę dawałyby w możliwie największym stopniu takie same efekty, jakie otrzymalibyśmy w warunkach idealnego porównywania użyteczności między jednostkami bez przyjmowania założeń upraszczających. Wybierając tę możliwość opieramy się na ideale epistemologicznym, a nie moralnym.

Tak więc ani Robbins, ani Rawls, nie podali ostatecznego powodu do uznania, że wbrew pierwotnemu wrażeniu intersubiektywne porównania użyteczności same są lub muszą się opierać na sądach wartościujących. Jest jeszcze jedna możliwość, którą musimy pobieżnie rozważyć. Możemy uznać, że intersubiektywne porównania są sensowne i że polegają na stwierdzeniu faktów, będąc jednocześnie zdania, że nie wiemy, kiedy te porównania są prawdziwe, a kiedy fałszywe. Z tą sprawą wiążą się obszernie zagadnienia, którymi zajmę się tylko pobieżnie. Przyjmijmy, że porównują siebie i Jonesa. Mówię, że odczuwa on ból, a ja nie. Pytanie o to, czy w ogóle mam prawo stwierdzić to właśnie, co stwierdziłem, sprowadza się do pytania, czy mam prawo w ogóle stwierdzić, że Jones odczuwa ból. Każdy, kto chciałby bronić radykalnego sceptycyzmu w tej kwestii, musi wyjaśnić, dlaczego czasem możliwe jest wyłącznie poznanie stanu własnego umysłu — nigdy zaś innych osób. Różni filozofowie bezskutecznie próbowali znaleźć odpowiedź na to pytanie. Są oczywiście osoby, które świadomie opierają się poznaniu. Pewien autor piszący za życia Balfoura skomentował kiedyś nieprzeniknione oblicze premiera: „Nikt nie odgadnie jak wielki pożar ukrywa twarz tego człowieka, nikt nie odgadnie, co ten pożar wznieciło, i nikt nie wie, czy rzeczywiście twarz ta ukrywa pożar”⁴. Były to pytania w pełni uzasadnione, jednak przypuszczenie, że wszystkie umysły są równie nieprzeniknione, nie jest niczym poparte.

W moim przekonaniu istotne pytania dotyczą zarówno teorii: jak dokonujemy intersubiektywnych porównań i na czym je opieramy? jak i praktyki: jak często możemy dokonywać takich porównań, w jakich sprawach, kiedy możemy na nich polegać, a przede wszystkim, czy rola, jaką im przypisujemy w teorii społeczeństwa i teorii moralności, istotnie zakłada coś niewykonalnego? Warto tu podkreślić, aby uchylić dość rozpowszechnione nieporozumienie, że problemy te stają nie tylko przed teorią utylitaryzmu. Jak widzieliśmy, wszelkie rozważania na temat sprawiedliwości wcale nie posuną się naprzód, dopóki nie zdobędziemy umiejętności porównywania sytuacji między jednostkami.

⁴ Cyt. na podstawie R. G. Martin 1972, s. 93.

Porównania użyteczności — propozycje i zastrzeżenia

Rozważyliśmy dwie propozycje: zasadę zerowania skali i zasadę równej odległości między poszczególnymi poziomami użyteczności dla wszystkich jednostek. Wady obu tych propozycji są takie same: rezygnują one z próby zastosowania realistycznej intersubiektywnej miary użyteczności i przyjmują założenia nieprawdziwe i moralnie nietrafne.

John Harsanyi [1976, rozdz. 2; 1977, rozdz. 4], Amartya Sen [1973, s. 14—15, 45] i Kenneth Arrow [1977; 1978] przedstawili propozycje wolne od tych wad: nie są one ani fałszywe, ani moralnie błędne. Propozycje te nie są zbieżne, natomiast wychodzą z tego samego założenia: bez wątpienia wypowiadamy się czasem na temat tego, czy wolelibyśmy być Smithem w warunkach, w jakich Smith się znajduje, z jego upodobaniami i zainteresowaniami, bądź czy wolelibyśmy być Jonesem, wraz z jego upodobaniami i zainteresowaniami. Pod tym względem propozycje wspomnianych autorów są identyczne: dowolna osoba stawia Jonesa i Smitha na tym samym poziomie użyteczności, wtedy i tylko wtedy, gdy jest jej obojętne, czy miałyby stać się Smithem wraz z jego poglądami, czy miałyby stać się Jonesem i przyjąć jego poglądy. Postulat ten jest ograniczony pod jednym ważnym względem: odwołuje się on do preferencji tylko jednej określonej osoby. Nie wyklucza się, że postulat ten może doprowadzić do uzyskania obiektywnego wzorca, który zostanie zaakceptowany przez wszystkich; choć nie jest to rezultat gwarantowany. Zaletą tej propozycji jest to natomiast, że nie wymaga ona dokonywania intersubiektywnych porównań w całej ich złożoności. Nikt nie musi stwierdzać, czy Smith będąc w sytuacji, w której się znajduje, jest w lepszej sytuacji niż Jones. Wystarczy w zupełności, bym tylko zdecydował we własnym imieniu, czy wolę przyjąć zapatrywania Smitha i znaleźć się w jego sytuacji, czy wolę przyjąć zapatrywania Jonesa i znaleźć się w jego sytuacji. Porównania intersubiektywne są tym samym sprowadzone do porównań intrasubiektywnych [por.: J. C. Harsanyi 1976, s. 50, 79—80; 1977, s. 59; K. J. Arrow 1977, s. 220]. Jednocześnie propozycja ta łagodzi w jakiś sposób niepokój Robbinsa dotyczący sensowności porównań intersubiektywnych, bo w myśl omawianej propozycji porównania dotyczą zachowań (co prawda hipotetycznych, a nie realnych) objawianych w sytuacji wyboru [por. K. J. Arrow 1978, s. 227]. Wydaje mi się jednak, że ten pozór epistemologicznej skromności i semantycznej zasadności jest fałszywy. Omawiana propozycja wymaga bowiem, bym wprowadził pewną, intersubiektywną miarę na mój własny użytek, polegającą na uporządkowaniu stanów, będących w istocie stanami mnie samego. Czy rzeczywiście jednak są to moje własne stany? Decyduję na przykład, że

wolałbym być Smithem alpinistą, pomimo tego, że ryzykuje on swym życiem, niż Jonesem naukowcem, który wie, że życie jest bezpieczne. By dokonać wyboru, muszę przyjąć dwa poglądy na świat: Smitha i Jonesa i wtedy pojawiają się dwa problemy. Pierwszy, to oczywisty problem tożsamości osobowej. Czy mogę sensownie powiedzieć, że całkowicie przyswoiłem sobie upodobania Smitha, jego postawy, charakterystyczne sposoby myślenia, ambicje i cechy fizyczne, a mimo wszystko nadal pozostałem sobą? Być może jest to wykonalne. Być może, jak sądził Henry James, jest to zadanie stojące przed biografem: „Nie da się przeżyć życia innego człowieka, jeśli się nie dzieli jego odczuć, nie rośnie wraz z nim, nie rozwija i nie doświadcza zmiennej intensywności powtarzających się zdarzeń, bo z tych właśnie zdarzeń składało się jego życie”⁵.

Być może jest prawdą, że pozostaję sobą, gdy patrzę na świat wysiłkiem swej wyobraźni, tak jak to potrafili zrobić tylko najwięksi biografowie, oczami innego człowieka. Jest jednak i druga trudność, i na niej chciałbym się zatrzymać. Czy przyjmując pogląd Smitha i Jonesa na świat — tak jak muszę to zrobić — nie dokonuję czegoś zupełnie innego niż ustalenie preferencji?

Gdy zwykle decyduję, co wolałbym: znaleźć się w skórze Smitha, czy Jonesa, odwołuję się do moich własnych wartości i upodobań. Mogę na przykład być Jonesem, ponieważ zdecydowanie potępiam ryzykowanie życiem w górach przez ludzi, którzy jak Smith, mają na utrzymaniu dzieci. Nie jest to jednak oszacowanie użyteczności życia Smitha i Jonesa. Chcąc takiego oszacowania dokonać, muszę się bowiem wyrzec swych poglądów moralnych. Dopiero wtedy, pozbywszy się poglądów moralnych, a w istocie pozbywszy się wszelkich swoich poglądów, poza tym, co jaką ma użyteczność, mogę, moim zdaniem, zdecydować, że wolałbym raczej być w skórze Smitha niż Jonesa, bo lubię ryzyko. To jednak też nie byłoby intersubiektywnym porównaniem użyteczności, lecz raczej twierdzeniem mówiącym, co wybieram będąc sobą. Muszę się więc wyzyść wszystkich swych upodobań i skłonności. Jeśli się wyrzeknę wszystkich preferencji wiążących się z moją własną użytecznością, mogę ustalić, że wolałbym raczej być Smithem i znaleźć się w jego sytuacji niż być Jonesem i znaleźć się w sytuacji Jonesa. Ten wysiłek wyobraźni nie jest zwykle przesadnie wielki. Mogę sobie wyobrazić, że moje życie przebiegało inaczej (powiedzmy, że byłem adoptowany w dzieciństwie przez ludzi mających zupełnie inny charakter niż moi rodzice) i że wyrosłem na człowieka kochającego ryzyko, jak Smith; lub że wyrosłem na człowieka unikającego ryzyka, jak Jones. Chcę tu ra-

⁵ Cyt. na podstawie L. Edel 1978 — motto do tomu I biografii Henry Jamesa.

czej podkreślić inną trudność. W jaki sposób, będąc pozbawiony własnych preferencji, mogę dokonywać wyboru między stanami, które mają być rzekomo stanami mnie samego: mnie jako Smitha i mnie jako Jonesa? Do czego mam się odwołać dokonując wyboru? Bez własnych upodobań mogę się oprzeć tylko na tym, co każdy ma do dyspozycji: na ogólnej wiedzy o naturze ludzkiej, na szczegółowych informacjach na temat Smitha i Jonesa. Mogę się odwołać do nauk społecznych i mogę wziąć pod uwagę informacje określającą stopień zaspokojenia interesów Smitha i Jonesa. Wtedy jednak dodatkowe odwoływanie się do moich preferencji lub do preferencji osoby, którą jestem w swej wyobraźni, staje się czymś zbytecznym. Przynajmniej w pierwszym przypadku wybór taki nie jest sprawą moich preferencji, lecz raczej sądem opartym na danych dostępnych dla każdego i dotyczących życia Smitha i Jonesa. Ja jestem ważny jedynie jako osoba wypowiadająca sąd, że Smith jest w lepszej sytuacji niż Jones. I jest to dodatkowy powód do pozostawienia bez odpowiedzi pytania o to, czy można poprawnie zapytać, jak bym się czuł, gdybym był Smithem, lub jak bym się czuł, gdybym nim nie był — tzn. pytania o to, czy poprawnie można pytać o swą własną tożsamość. Zresztą w interesujących nas rozwiązaniach pytanie to w ogóle się nie pojawia. Nie wybieram między stanem Smitha i stanem Jonesa. Wypowiadam jedynie sąd na podstawie przesłanek dostępnych i dotyczących względnej użyteczności stanów Smitha i Jonesa.

Mam więc wrażenie, że stanowisko, które zaproponowali Harsanyi, Sen i Arrow, wymaga korekty. Jeśli intersubiektywne porównanie jest preferencją sformułowaną przy omówionych zastrzeżeniach, to w istocie nie jest ono w ogóle preferencją, lecz jest raczej sądem o faktach, opartym na jakiejś ogólnej wiedzy na temat natury ludzkiej i dodatkowych informacjach na temat poszczególnych osób. Mój sąd może być błędny w pewnych przypadkach; może się różnić od sądu innych osób, ale to samo można powiedzieć na temat wszystkich sądów. Jeśli więc różnię się z kimś w poglądzie na ten temat, czy lepiej jest być Smithem, czy Jonesem, różnica zdań między nami nie wygląda tak, jak zwykle wygląda różnica zdań na temat preferencji, pojawiająca się wśród ludzi mających różne upodobania i postawy. Na temat charakteru tego sądu o faktach będziemy musieli powiedzieć jeszcze coś więcej, i zaraz tym się zajmę. Nie wydaje się jednak, bym porównując Smitha i Jonesa mógł oprzeć się na jakichś pewniejszych i mniej problematycznych sądach niż sądy o naturze ludzkiej i szczegółowe informacje o jednostkach.

Odpowiedź na zastrzeżenie: funkcja ekstrapolowanej użyteczności

W tym miejscu mógłby ktoś powiedzieć, że John Harsanyi znalazł jednak odpowiedź na to, co jest istotne w moim zastrzeżeniu, znajdując tym samym sposób na podtrzymanie przekonania, że porównania intersubiektywne wyrażają preferencje. Wydaje się bowiem, że udało mi się pogodzić oba cele: porównania są dla niego z a r ó w n o preferencjami, j a k i sądami o faktach. Stanowisko Harsanyiego przedstawia się następująco [por. J. C. Harsanyi 1977, s. 58—59]: Gdyby świat był prosty i preferencje każdej jednostki takie same, to funkcja użyteczności byłaby dla każdej jednostki identyczna, a intersubiektywne porównania dawałyby się realizować bez przeszkód przez odwołanie się do tej funkcji. W ten sposób porównania intersubiektywne sprowadziłyby się do porównań intrasubiektywnych. W rzeczywistości jednak ludzie mają różne preferencje. Na szczęście preferencje każdej osoby powstają pod wpływem tych samych zmiennych przyczynowych. Różnice preferencji dadzą się więc w zasadzie przewidzieć na podstawie różnic owych zmiennych. Jeśli dwie osoby mają tę samą biologiczną konstytucję i tę samą jednostkową historię, to, ponieważ podlegają tym samym prawom psychologii, które określają powstawanie preferencji, ich preferencje będą identyczne, gdyż obie osoby w ten sam sposób przypisywać będą użyteczność wszystkim rzeczom. Formuła opisująca intersubiektywne porównania użyteczności stwierdza, że osoba oceniająca wolałaby być raczej w obiektywnej sytuacji osoby M mając poglądy M , niż być w obiektywnej sytuacji osoby N mając poglądy N . Ponieważ jednak użyteczność, jaką oceniający przypisuje stanowi M opiera się na ogólnej wiedzy o tym, co ktoś z wyposażeniem biologicznym M i z jednostkową historią M swobodnie wybiera, użyteczność, jaką różne osoby oceniające w imieniu M przypisywać będą rzeczom, jest taka sama, jeśli tylko wiedza każdego z oceniających jest pełna. Jeśli użyteczność, jaką osoba oceniająca przypisuje znalezieniu się w sytuacji M lub w sytuacji N nazwiemy użytecznością ekstrapolowaną, to funkcja użyteczności ekstrapolowanej jest dla każdego taka sama. Choć więc zwykła funkcja użyteczności jest dla każdej jednostki inna, ekstrapolowana funkcja użyteczności jest dla wszystkich jednaka. To nam pozwala powrócić do tego samego stanowiska, w którym porównania intersubiektywne zostały sprowadzone do porównań intrasubiektywnych. Harsanyi twierdzi, że ponieważ porównania intersubiektywne są sądami o faktach, możemy je sprowadzić do sądów wyrażających jednostkowe preferencje.

Trudno zaprzeczyć, że w większości spraw Harsanyi ma rację. Wydaje się całkiem prawdopodobne, że preferencje każdej jednostki determino-

wane są przez te same zmienne przyczynowe. Mówiąc swobodnie znaczy to, że gdybym miał tę samą konstytucję biologiczną i tę samą biografię co Smith, to miałbym także te same preferencje co Smith. Nie widzę też powodu, by odmówić racji Harsanyiemu, gdy twierdzi on, że mając dostatecznie dużo informacji na temat Smitha i Jonesa, a także mając ogólną wiedzę na temat przyczyn, mogę zrekonstruować preferencje każdego z nich. Mam jednak wrażenie, że rekonstrukcja taka byłaby zbyt doskonała. W jaki sposób miałbym bezpośrednio konfrontować moją wiedzę na temat stanu Smitha z moją wiedzą na temat stanu Jonesa w intersubiektywnym porównaniu? Niezależnie od tego, że Harsanyi ma dużo racji w tym, co mówi na temat ogólnych prawidłowości przyczynowych rządzących powstawaniem preferencji, nie wiadomo jak przejść od preferencji Smitha do preferencji Jonesa — jest tu wyraźna luka. Harsanyi twierdzi, że lukę tę wypełniają moje preferencje. I w tym punkcie, jak mi się wydaje, poglądy nasze się rozchodzą. Usunąwszy z pola widzenia moje własne preferencje przez przyjęcie stanowiska, że intersubiektywne porównania są sądami o faktach, Harsanyi próbuje nadać im ponownie znaczenie, argumentując, że ponieważ sądy o faktach są takie same dla wszystkich, stanowią one podstawę dla wprowadzenia ekstrapolowanej funkcji użyteczności, która też jest identyczna dla wszystkich. Tak w istocie nie jest. Twierdzenie, że istnieje funkcja ekstrapolowanej użyteczności, głosi jedynie, że preferencje są kształtowane przez ogólne prawidłowości przyczynowe i że, wobec tego, każdy ostatecznie uzna te same sądy przyczynowe. Funkcja ekstrapolowanej użyteczności nie jest więc funkcją mojej użyteczności. Porządkując użyteczności Smitha i użyteczności Jonesa nie porządkuje moich użyteczności. Problem nie sprowadza się do tego co wolę: być sobą w ten czy w tamten sposób? Jeśli biologiczne cechy Smitha i Jonesa są różne od moich — jak to bez wątpienia zwykle bywa — to trudności związane z określeniem tożsamości jednostki pojawiają się na nowo, bo nie można wtedy odwoływać się do stanu jednego czy drugiego z nich, nazywając je moimi stanami. Chcę jednak podkreślić, że, niezależnie od tej ostatniej trudności, już wcześniej zostałem pozbawiony jakichkolwiek podstaw do formułowania preferencji. Moja wiedza przyczynowa pozwala mi opisać stan Smitha i stan Jonesa, i opisy te stanowią jedyną podstawę dla intersubiektywnych porównań. Nie są one w żadnym stopniu podstawą do formułowania preferencji. Pozwalają mi one stwierdzić, że Smith jest na wyższym poziomie użyteczności niż Jones, i że sąd ten może być podstawą do tego, bym uznał, że wolę być w sytuacji Smitha niż Jonesa. Ale nawet ta preferencja bierze pod uwagę moje wartości, które nie miały się w ogóle pojawić.

Harsanyi przechodzi od przyczynowej interpretacji preferencji do stwierdzenia, że istnieje wspólna funkcja ekstrapolowanej użyteczności i tym samym redukuje porównania intersubiektywne do porównań intrasubiektywnych. Poprawna interpretacja wspólnej funkcji ekstrapolowanej użyteczności nie uzasadnia jednak tej redukcji, a tym samym nie pozwala na nowo oprzeć się na własnych preferencjach.

Kolejna odpowiedź na zastrzeżenie: funkcja głębokiej użyteczności

By uporać się z tymi trudnościami, możemy przekształcić funkcję ekstrapolowanej użyteczności w funkcję głębokiej użyteczności, zaproponowaną niedawno przez Rawlsa. Postępując w ten sposób zachowamy to, co najważniejsze z koncepcji intersubiektywnych porównań Harsanyiego, a w szczególności sugestię, byśmy polegali na wyborze między znalezieniem się w sytuacji Smitha a znalezieniem się w sytuacji Jonesa. Poszerzymy teraz podstawę porównania tych preferencji. Dopóki ludzie autentycznie pragną tych samych rzeczy, różnice jakie najprawdopodobniej odkryjemy między ich funkcjami użyteczności nie będą miały większego znaczenia. Jeśli ty i ja, i wszyscy inni dążą do tego samego celu, który można by określić jako „zadowolenie z życia”, i jeśli cel ten stanowi wartość, za pomocą której oceniamy stan Smitha i Jonesa, to na jakimś głębszym poziomie porównań wszyscy mamy tę samą funkcję użyteczności. Funkcja ekstrapolowanej użyteczności Harsanyiego odwoływała się nie do wartości, lecz wyłącznie do wiedzy intuicyjnej. Funkcja głębokiej użyteczności zakłada wspólną podstawę wartości i na niej opiera porównania intersubiektywne. Wartości wspólne są również moimi wartościami, i tego właśnie potrzebowaliśmy, by móc traktować porównania intersubiektywne jako wyraz własnych preferencji.

W gruncie rzeczy propozycja ta jest nawrotem do hedonizmu. Najbardziej rozpowszechniony obecnie sposób interpretowania użyteczności nie utożsamia jej ze stanem umysłu (przyjemnością), lecz sprowadza użyteczność do pewnych stanów w świecie (do tego, co jest w nim pożądanym). „Użyteczność” to „zaspokojenie pragnień”, gdzie przez „zaspokojenie” rozumie się spełnienie określonych warunków, tak jak np. w „spełnieniu warunków umowy”. Zatem Smith czerpie więcej użyteczności z pewnej rzeczy niż Jones z innej wtedy i tylko wtedy, gdy każdy z nich znając do pewnego stopnia obie rzeczy dokonuje wyboru. Gdy jednak interpretując funkcję głębokiej użyteczności mówimy, że każdy dąży do „zadowolenia z życia”, to „zadowolenie” nie może być interpretowane w czysto formalnym sensie „spełnienia warunków”.

Gdybyśmy bowiem przyjęli tę interpretację, twierdzilibyśmy jedynie, że każdy przede wszystkim pragnie, aby były spełnione wszelkie jego potrzeby. Takie czysto formalne pragnienie w żaden sposób nie wzbogaca naszych podstawowych pragnień i z pewnością nie jest dostatecznie wyraźne, by stanowić podstawową wartość życia. „Użyteczność” zdefiniowana w taki sposób nie może być traktowana jako podstawa wartości, ani nawet jako jedna z wartości. Powinniśmy raczej powiedzieć, że definicja taka mówi nam, na czym polega owa zależność między rzeczą a człowiekiem, którą się opisuje, mówiąc, że człowiek ów ceni sobie tę rzecz. Tak zdefiniowana użyteczność daje się pogodzić z nieredukowalną wielkością wartości, która wyklucza możliwość uznania twierdzenia, że w głębi serca wszyscy wyznajemy te same wartości. Musimy więc oprzeć się na twierdzeniu, że wszyscy dążymy do zaspokojenia pragnień w materialnym sensie tego stwierdzenia, tzn. dążymy do określonego stanu, który nazywamy zaspokojeniem pragnień. I to właśnie jest hedonizm. Mówiąc ściśle jest to hedonizm psychologiczny. Odczuwanie zadowolenia nie jest na gruncie omawianej teorii jedyną wartością, którą powinniśmy uznać, lecz jest jedyną wartością faktycznie przez nas uznawaną.

Jeśli intersubiektywne porównywanie użyteczności wymaga przyjęcia funkcji głębokiej użyteczności, to nieuchronnie prowadzi to do hedonizmu. Jest to konkluzja niepokojąca. Po pierwsze, oznacza ona, że wiele osób uważających, że można oprzeć się na pojęciu użyteczności zakładając porównywalność sytuacji międzyludzkich przy jednoczesnym odrzuceniu hedonizmu, ludzi samych siebie. Osoby te są w istocie nieświadomymi hedonistami. Po drugie, konkluzja ta oznacza, że choć porównywanie jednostkowych użyteczności jest często stosowane w pluralistycznych teoriach moralności, pluralizm ten niefortunnie wiąże się z monistycznym hedonizmem. I po trzecie, konkluzja ta oznacza, że każda teoria moralności, jak i każda teoria społeczeństwa odwołująca się do pojęcia porównywania jednostkowych użyteczności musi się opierać na prymitywnej i godnej pożałowania koncepcji jednostki ludzkiej, pragnącej w głębi duszy wyłącznie własnego zadowolenia, dla którego zawsze gotowa jest porzucić wszelkie swe zobowiązania i wartości. Mam wrażenie, że żadnej z tych konsekwencji nie możemy przyjąć.

Nie wiadomo właściwie, czy Harsanyi, Sen lub Arrow zechcieliby się oprzeć na funkcji głębokiej użyteczności, żeby obronić koncepcję porządkowania użyteczności na podstawie intersubiektywnych porównań. Gdyby jednak któryś z nich zechciał przyjąć to rozwiązanie, zająłby stanowisko błędne pod dwoma względami. Po pierwsze, błędem jest przyjmować hedonizm psychologiczny. Po drugie, odwołanie się do funkcji głębokiej użyteczności jest w istocie rzeczą zbyteczną. Intersubiek-

tywne porównania nie muszą się opierać na preferencjach. Pierwszy błąd omówiliśmy wystarczająco jasno, drugi wymaga natomiast dłuższej dyskusji.

Proponowane rozwiązanie

Jeśli funkcja ekstrapolowanej użyteczności i funkcja głębokiej użyteczności nie stanowią poszukiwanej podstawy porównań, to co mogłoby stanowić taką podstawę? Jak już powiedziałem, intersubiektywne porównania użyteczności są czystymi sądami o faktach, bez elementów preferencji. Czy sądy tego typu są wystarczającą podstawą porównań? Lepiej być nieszczęśliwym Sokratesem niż szczęśliwym głupcem. Zapewne, ale w jaki sposób, przez odwołanie się do jakich prawidłowości ludzkiej psychiki możemy dojść do podobnego stwierdzenia?

Moim zdaniem „użyteczność” najlepiej możemy zinterpretować uznając ją za zaspokojenie pragnień w formalnym sensie tego wyrażenia. Wyobraźmy sobie, że Jones cierpi ból, a Smith nie, choć żaden z nich nie chce cierpieć. Pod tym względem Smith jest więc w lepszej sytuacji niż Jones. Trudno uznać to porównanie indywidualnych użyteczności za sąd wartościujący. Oczywiście istnieje więcej niż jedna definicja sądu wartościującego. Nie ma też jednomyślności na temat tego, czym różni się sąd wartościujący od sądu o faktach; nie jest na przykład jasne, czy te dwa rodzaje sądów wykluczają się wzajemnie. Przypuszczam, że większość dzisiejszych filozofów nie zgodziłaby się z tym, co Robbins powiedział w latach trzydziestych, dopóki jednak trzymamy się przykładów, co do których wszyscy jesteśmy jednomyślni, możemy prowadzić nasz wywód dalej. I, jak sądzę, zdaniem każdego z nas, sąd mówiący, że w opisanym przypadku pragnienia Smitha są lepiej zaspokojone niż pragnienia Jonesa, uznamy za sąd niewartościujący. Co więcej, wydaje się wątpliwe, by taki sąd mógł w ogóle opierać się na sądach wartościujących. Przypomnijmy sobie, że porównywalność wprowadziliśmy jako wyjaśnienie sposobu rozumienia słowa „użyteczność”, kiedy próbowaliśmy określić, w jakim stopniu pragnienia Smitha i Jonesa są zaspokojone. Trudno odgadnąć, co mogłoby nasunąć myśl, że porównanie takie opiera się na sądach wartościujących. Wydaje się, że można mieć dwojakiemu rodzaju wątpliwości na temat intersubiektywnych porównań: ogólne i szczegółowe. Ogólne wątpliwości, na przykład te, o których mówił Robbins, wypływają z powszechnego sceptycyzmu dotyczącego możliwości poznania innych umysłów, a co za tym idzie również możliwości dokonania jakichkolwiek intersubiektywnych porównań stanów umysłowych lub uczuć. Szczegółowe wątpliwości, takie na przykład, o jakich mówił Rawls, nie wiążą się ze sceptycyzmem co do możliwości dokonywania porównań w ogóle, lecz kwestionują możliwość porów-

nywania użyteczności. Problem dotyczy więc nie tego, że cudzy umysł jest inny, lecz właściwego sposobu użycia pojęcia „użyteczność”. Jest jednak wtedy jasne, że nie wszystkie porównania intersubiektywne są lub opierają się na sądach wartościujących. W przypadku zdań typu: „Smith cierpi bardziej niż Jones”, „Smith bardziej się zezłościł niż Jones”, „Smith bardziej jest tym zainteresowany niż Jones”, zasadniczy sceptycyzm nie daje się utrzymać. Jednakże jeśli porównania intersubiektywne dotyczą użyteczności i przyjmują formę: „M jest w lepszej sytuacji niż N”, wypowiedź taka o wiele bardziej przypomina sąd wartościujący. Mimo wszystko, nasze porównanie mówiące, że Smith, który nie doświadcza żadnego bólu, jest pod tym względem w lepszej sytuacji niż Jones, nie jest sądem wartościującym. A więc nie wszystkie intersubiektywne porównania użyteczności są sądami wartościującymi lub opierają się na nich. Nasz przykład przypomina sąd wartościujący, ponieważ zawiera pojęcie „lepszy”, co sugeruje, że sąd ten opiera się na ogólnym przekonaniu, że lepiej jest, gdy się nie doznaje bólu. Bez wątplenia sąd mówiący, że lepiej jest, gdy się nie doznaje bólu, dokładnie odpowiada temu, co wielu autorów uznaje za sąd wartościujący. Nasz przykład nie opiera się jednak na tym sądzie. Jest również prawdą, że jeśli sądy o użyteczności są przesłankami decyzji na temat powinności, to konkluzja opiera się na sądach wartościujących. Wyprowadzenie wniosku na temat powinności z sądów o użyteczności opiera się wtedy na sądzie wartościującym, takim jak: „powinno się, jeśli inne względy nie wchodzą w grę, powiększać, a nie pomniejszać użyteczność”. Intersubiektywne porównania użyteczności poprzedzają jednak taki wniosek.

Rozpatrzmy trudniejszy przykład. Wyobraźmy sobie, że Smith wspiął się po górach, tak jak to lubi robić, natomiast Jones nie ma dość czasu na studia. Wydaje się, że tutaj znowu porównanie nie nastęrcza trudności. Jeśli chodzi o ulubione zajęcie w życiu, Smith jest w lepszej sytuacji niż Jones. W istocie możemy w tym wypadku mieć jednak znowu dwie wątpliwości.

Po pierwsze, gdy Mill przestrzegał, że lepiej być nieszczęśliwym Sokratesem niż szczęśliwym głupcem, używał słowa „lepiej” w jego istotnym sensie, mając na myśli „na wyższym poziomie użyteczności”. Słusznie jednak podkreślał, że sąd taki opiera się na preferencjach popartych informacją. Uważał bowiem, mówiąc bardziej szczegółowo, że sąd taki opiera się na szczególnych preferencjach popartych informacją, na preferencjach osoby określonego rodzaju, osoby, która zna obie możliwości. Z konieczności więc czyjeś preferencje musiały być brane pod uwagę. Upodobanie chodzenia po górach objawione przez Smitha możemy jednak uznać za szaleństwo. Jeśli porównamy zadowo-

lenie naukowca z zadowoleniem alpinisty, takie porównanie wyda nam się czymś zdecydowanie odmiennym od porównywania preferencji jednostek. Większość ludzi jest bowiem zdania, że to, czego pragną głupcy, jest głupotą. Gdyby głupiec wiedział więcej na temat tego, czego pragnie, lub gdyby znał inne możliwości, zapewne zmieniłby zdanie. Właśnie ten warunek, opisujący sytuacje odbiegające od faktów, charakteryzuje po części to, czym jest użyteczność. Jeśli ktoś pragnie czegoś mając mylne wyobrażenie o tym, czego pragnie i czego by nie pragnął, gdyby lepiej to znał, to użyteczność przypadająca na ową osobę jest mniejsza niżby była wtedy, gdyby pragnęła ona czegoś innego niż to, czego pragnie. Użyteczność jednostki nie sprowadza się do zaspokojenia aktualnie odczuwanych pragnień. Pełna charakterystyka użyteczności wyznacza pewną rolę, którą trudno scharakteryzować zwięźle, obejmującą to, czego dana osoba pragnęłaby, gdyby była lepiej poinformowana. Głupiec w istocie rzeczy chciałby żyć jak Sokrates, podobnie jak dziecko cierpiące ból zgodziłoby się na podskórny zastrzyk znieczulający, gdyby rozumiało, co doktor robi. Twierdzenie, że zarówno dziecko, jak i głupiec chcieliby czegoś innego, gdyby byli lepiej poinformowani jest twierdzeniem o faktach. Ale nawet teraz, gdy Sokrates i głupiec są w pełni poinformowani, pozostają nam dwa szeregi preferencji i choćby były one identyczne, musimy znaleźć sposób na sprowadzenie ich do jednego szeregu. Przyjmijmy, że w wybranym przypadku (co nie musi być prawdą w innych okolicznościach) pragnienia Sokratesa i głupca mają rozkład o identycznych intensywnościach; że to, czemu obaj przypisują najwyższą preferencję, jest przez każdego z nich upragnione w równym stopniu. Takie przypuszczenie samo jest przypadkiem intersubiektywnego porównania, ale opiera się na ogólnych poglądach na temat rozwoju ludzkiej osobowości, a w szczególności na przekonaniach dotyczących wzrastania lub zmniejszania się intensywności pragnień. Reguła zerowania bywa w tym wypadku często stosowana, ale jej użycie powinno być zawsze empirycznie uzasadnione, a nie przyjęte jako założenie motywowane życzliwymi uczuciami demokratycznymi. Tak więc w przypadku szczęśliwego Smitha i nieszczęśliwego Jonesa, jeśli poprawnie możemy zastosować zasadę zerowania, tzn. jeśli przyjmiemy, że każdy z nich ceni sobie równie wysoko to, co najchętniej robi, i jeśli każdy z nich dobrze zdaje sobie sprawę z tego, co wybrał, a co odrzucił, to lepiej być szczęśliwym Smithem niż nieszczęśliwym Jonesem. Wniosek ten nie świadczy jednak o tym, że naprawdę zawsze będziemy mogli dojść do podobnej konkluzji, jeśli pragniemy ustalić, czy lepiej być szczęśliwym Smithem, czy lepiej nieszczęśliwym Jonesem? Czasem nie będziemy mogli zdecydować, kim być lepiej.

Powiedziałem jednak, że musimy się uporać z dwiema trudnościami.

Oto jak wygląda druga. Istnieje wiele powodów, które sprawiają, że lepiej czasem być nieszczęśliwym niż szczęśliwym. Jeśli ktoś na przykład pragnie robić rzeczy odrażające, lepiej może, by mu się to nie udało. Popelnilibyśmy jednak błąd, gdybyśmy starali się utrzymać tę konkluzję opierając się na przedstawionej metodzie dokonywania intersubiektywnych porównań. Koncepcja ta nie może być podstawą dla oceny różnych modeli życia. Bez wątplenia lepiej być nieudolnym dusicielem niż wytrawnym mordercą. Interesują nas intersubiektywne porównania, które prowadzą do oceny pewnego modelu życia i uznają go za lepszy nie w tym sensie, że „biorąc wszystko pod uwagę jest on moralnie lepszy”, lecz w tym sensie, że „znajduje się na wyższym poziomie użyteczności”. Kwestie pojawiające się przy podejmowaniu decyzji, czy wytrawny morderca lub nieudolny dusiciel są na wyższym poziomie użyteczności, różnią się w każdym systemie moralnym (w tym również na gruncie utylitaryzmu) od kwestii, jakie się pojawiają przy próbie oceny, który model życia jest moralnie lepszy.

Propozycja moja sprowadza się więc do tego, że porównania są twierdzeniami o faktach, mówiącymi przede wszystkim o tym, w jakim stopniu różni ludzie chcą (lub w jakim stopniu chcieliby, gdyby dysponowali odpowiednią wiedzą) tego, czego chcą. Ale twierdzenie to winno być przedstawione bardziej szczegółowo i zasługuje na osobne rozważenie, którego nie możemy w tym miejscu podjąć⁶.

*Czy intersubiektywne porównania
są możliwe w skali społecznej?*

Wszystkie propozycje, które rozważaliśmy dotąd, mają pewną cechę wspólną: opierają się na podstawie subiektywnej, na pragnieniach o określonej intensywności, zależnej od jednostek przeżywających owe pragnienia. John Rawls uważa jednak, że zamiast zajmować się subiektywną intensywnością pragnień, możemy przyjąć za wskaźnik szczęścia coś, co nazywa podstawowymi dobrami społecznymi. Bez względu na to, jakie są poszczególne cele jednostki, mówi Rawls, zawsze będzie sobie ona cenić pewne rzeczy wyżej niż inne, ponieważ w zależności od tego, ile będzie miała tych cenionych wyżej rzeczy, łatwiej jej będzie osiągnąć założone cele [J. Rawls 1972, s. 92—93, 396]. Twierdzenie to sprowa-

⁶ Interesujące omówienie tego, jak twierdzenia na temat pragnień dadzą się objąć w szerszym systemie wyjaśnień ludzkiego postępowania znajduje się u J. Waldnera 1972, w szczególności na stronach 95—96. Ogólne uwagi Waldnera na temat statusu tych twierdzeń wydają mi się bardziej przekonujące niż wskazane przez niego powiązania między pragnieniami a innymi mniej zróżnicowanymi formami zachowania. W każdym razie jednak ani jego, ani moje stanowisko nie jest behawiorystyczne.

dza się, jak mówi Rawls [1972, s. 95], do wyrażenia zgody na porównania sytuacji jednostek, biorąc pod uwagę to, ile każda z nich posiada owych rzeczy, o których można sądzić, że każdy chciałby mieć ich jak najwięcej. Jeśli założenie to jest usprawiedliwione i powszechnie zostanie uznane za takie, to może być ono źródłem uzasadnionej nadziei znalezienia powszechnie akceptowanej, obiektywnej miary dokonywania porównań między jednostkami. Teoria utylitarystyczna ma mniejsze szanse zdobycia podobnego uznania. W teorii utylitarystycznej nie tylko musimy brać pod uwagę zmienne intensywności całego szeregu pragnień każdej osoby, ale ponadto musimy założyć powszechnie stosowalną intersubiektywną korespondencję między nimi. Rawls mówi na ten temat: „Problemy, które się pojawiają, gdy się chce to zrobić, są tak wielkie, a przybliżenia tak nieścisłe, że jawnie przeciwstawne opinie mogą się wydać różnym osobom równie prawdopodobne. Jedni mogą twierdzić, że zyski pewnej grupy znacznie przewyższają zyski innej, inni mogą twierdzić coś wprost przeciwnego” [J. Rawls 1972, s. 320]. Sądzę, że Rawls ma dwa zastrzeżenia przeciwko tej propozycji. Po pierwsze, sądzę, że błędem jest maksymalizowanie użyteczności. Powinniśmy się starać maksymalizować zaspokojenie podstawowych potrzeb, a nie zaspokojenie pragnień, których intensywność może słabo się wiązać z tym, co cenimy sobie w życiu, i które mogą się kierować w stronę przedmiotów niegodnych realizacji [por. J. Rawls 1975a; 1975b; T.M. Scanlon 1975]. Po drugie, projekt sumowania użyteczności za bardzo obciąża naszą zdolność dokonywania intersubiektywnych porównań. Jest to bez wątpienia poważne zastrzeżenie. Procedura, która wymaga czegoś, co nie daje się zrealizować, nie może być serio traktowana jako procedura nadająca się do praktycznego zastosowania. Chciałbym teraz rozważyć tylko to drugie zastrzeżenie, ponieważ wiąże się ono bezpośrednio z porównywalnością sytuacji poszczególnych jednostek.

Zastrzeżenie to trafia przede wszystkim w schemat intersubiektywnych porównań zaproponowany przez Harsanyi'ego, Sena i Arrowa, a także w poprawioną przeze mnie wersję tego schematu. Jest to doprawdy niezwykle osiągnięcie, gdy ktoś potrafi się znaleźć w skórze drugiego człowieka i jest to możliwe tylko w przypadku bliskiej zażyłości albo w wyniku długotrwałych starań utalentowanego biografy. Ogólnie biorąc, im bliżej się żyje z osobą, którą się ocenia, tym więcej można dokonać intersubiektywnych porównań na jej temat i tym bardziej są one trafne. Rodzice mogą mieć pewność, że wiedzą wiele o swych dzieciach. Wrażliwi nauczyciele znają dobrze swych uczniów. Jeśli jednak przejdziemy na drugi koniec skali intersubiektywnego dystansu, szanse trafnej oceny maleją. Prawodawca może zapoznać się bezpośrednio tylko z jedną spośród miliona osób, których interesu powinien bronić. Nie jest w stanie

poznać wszystkich. I tu pojawia się dylemat. Zasada równej różnicy między poziomami i zasada zerowania dadzą się wprawdzie pogodzić z problemem dystansu między jednostkami, ale nie są poprawne. Nasz schemat, choć jest poprawny, nie daje się pogodzić z problemami dystansu między jednostkami.

Założenia upraszczające

Nie należy jednak wyolbrzymiać różnicy w sposobie rozumowania z dwóch stron skali. Jest prawdą, że rodzice mogą się opierać na bezpośredniej obserwacji dzieci, jednak ich sądy na temat dzieci dokonywane są zawsze w ramach teoretycznego zrozumienia natury ludzkiej. Weźmy pod uwagę, jak bardzo od czasów Freuda zmieniły się poglądy rodziców na stresy przeżywane przez ich dzieci. Podobnie też prawodawcy opierają się na teoretycznym zrozumieniu natury ludzkiej. Brakuje im jedynie pełni spostrzeżeń na temat indywidualnych przypadków. Ich sposób traktowania tematu musi być pod tym względem prostszy niż podejście rodziców. W każdym razie mogą oni postąpić na dwa sposoby. Mogą przyjąć pewne założenia, by nadrobić brak bezpośredniej obserwacji albo mogą zrezygnować z tych założeń i poszukiwać takiego miernika ludzkiego szczęścia, który nie wymaga tych informacji. Z wielu powodów Rawls wybrał drugie wyjście. Rozpatrzmy jednak przez chwilę wyjście pierwsze. Założenia, jakie prawodawca musi przyjąć, by zrekompensować brak szczegółowych informacji, możemy nazwać „założeniami upraszczającymi”. Rawls analizuje problemy sprawiedliwości podchodząc do nich z jednej strony skali dystansu społecznego. Jego opis nie może więc, rzecz jasna, odwoływać się do poszczególnych przypadków, lecz musi się opierać na prostszych założeniach, takich jak wskaźnik podstawowych dóbr społecznych. Każdy intersubiektywny miernik przeznaczony do zastosowania w skali społecznej musi się opierać na takich uproszczeniach. Można jednak przyjąć inne uproszczenia niż te, które przyjął Rawls.

Jakie upraszczające założenia powinna przyjąć teoria utylitaryzmu? Jest to obszerny problem i mogę na ten temat powiedzieć tylko kilka niewiążących uwag.

Istnieją różnego rodzaju założenia upraszczające. Jednym z nich jest system głosowania, na przykład założenie, że każdy człowiek ma jeden głos. Innym ograniczeniem są pewne instytucje polityczne, np. rząd reprezentacyjny. Nas interesują jednak nie dowolne ograniczenia upraszczające, lecz tylko takie, które ułatwiają sumowanie użyteczności w praktycznych sytuacjach. Pewne upraszczające warunki pojawiające się w praktyce życia politycznego uzasadnione są innymi racjami niż maksy-

malizacja szczęścia. Na przykład zasada, że każdy człowiek ma jeden głos, charakteryzuje społeczeństwa, które zobowiązały się bronić równych praw. I analogicznie, nie poszukujemy wówczas dowolnej procedury dla podejmowania decyzji społecznych, lecz szukamy takiej, która maksymalizowałaby użyteczność, gdy użyteczność jest tylko jednym z celów realizowanych przez społeczeństwa.

Wiemy, jak ma wyglądać doskonała wiedza potrzebna do sumowania użyteczności. Przyjmujemy (i istnieją powody, by tak przyjąć), że istnieją liczbowe skale indywidualnej użyteczności. Przyjmijmy (choć nie ma powodów, by przyjąć takie założenie w odniesieniu do kogokolwiek poza małymi grupami), że dysponujemy dostateczną ilością czasu, by wyobrazić sobie, że potrafimy się wczuć w sytuację rozmaitych ludzi i porównać sytuację każdego człowieka z sytuacją przynajmniej jednego człowieka, a także, że potrafimy znaleźć dostatecznie dużo zbieżnych punktów między porównywanymi osobami, by móc ustalić w miarę wiarygodny współczynnik przekształcania wielkości użyteczności każdego z nich w wielkość odpowiadającą użyteczności drugiego. Nie ma przy tym znaczenia, czy jest to wykonalne, czy nie — tak wygląda ideał. Otóż wybór uproszczających założeń dla sumowania użyteczności powinien brać pod uwagę następujący wzgląd: jaka procedura może nas maksymalnie przybliżyć do procedury idealnej? Na skali użyteczności społecznej nie znajdziemy idealnego wyniku, mogącego służyć jako kryterium trafności przybliżeń. Wiemy jednak, jak wygląda procedura idealna, i możemy ocenić, z jakim prawdopodobieństwem inne procedury mogą prowadzić do tych samych rezultatów. W szczególnych przypadkach, gdy mamy do czynienia z małymi grupami, możemy bezpośrednio porównywać wyniki procedury idealnej z wynikami procedury aproksymatywnej. Możemy stwierdzić na przykład, czy dany system głosowania dał grupie to, czego ona pragnęła.

Jakie są więc alternatywne procedury sumowania użyteczności? Mam wrażenie, że w różnych okolicznościach muszą być używane różne procedury. Wszystkie te procedury opierają się, podobnie jak i sądy rodziców na temat dzieci, na ogólnym teoretycznym zrozumieniu natury ludzkiej. W prostej sytuacji prawodawca może się oprzeć na czymś takim, jak lista podstawowych dóbr Rawlsa. Zasada konstrukcji tej listy musiałaby być jednak inna, i sama lista musiałaby być dłuższa. Nie poszukujemy bowiem tego, czego ludzie chcą mieć jak najwięcej, by posłużyć się tym jako środkiem do realizacji dowolnego racjonalnego celu, ale szukamy tego, co potencjalnie może być pożądane przez każdego człowieka, i co przyczynia się do jego szczęścia. Lista ta musiałaby więc obejmować: dochód, zdrowie, bezpieczeństwo, niezależność, czas wolny i możliwość wypełniania go wedle uznania, wykształcenie, in-

teresującą pracę, szacunek dla samego siebie itp. W prostych przypadkach prawodawca może przyjąć, że zwiększenie wielkości którejkolwiek z tych zmiennych przyczynia się bezpośrednio do powiększenia sumy użyteczności. Rząd może na przykład dążyć do powiększenia dochodu narodowego po to, by powiększyć którąkolwiek wielkość odnotowaną na liście. W bardziej złożonych przypadkach prawodawca musi jednak zdecydować, którą zmienną powinien wybrać, jeśli każdą można realizować tylko kosztem innych. W tym celu musi on się posłużyć tabelą wymiany, której w istocie każdy człowiek używa dla własnych celów. Czasami różnice między ludźmi będą tak wielkie, że podobna tabela będzie musiała pominąć pewne zmienne; jest jednak pewien zakres rzeczy, co do których prawie wszyscy ludzie mają to samo zdanie, i gotowi są do kompromisów. Na przykład, większość ludzi ceni sobie wysoko zdrowie i uwolnienie od niepokoju. Czy rząd powinien więc popierać szczepienia przeciw chorobie Heinego-Mediny, choć tym samym będzie zmuszony do ograniczenia wydatków na naukę i sztukę? Skutki tej choroby są bardzo poważne, a niepokój jaki powoduje myśl o epidemii jest dotkliwy i powszechny. Koszt powszechnego szczepienia nie jest wysoki i nie przekracza w normalnych warunkach profilu kompromisu, który chcielibyśmy zaproponować po to, by wykazać, że strata ponoszona przez sztukę jest mniejsza niż zysk płynący z uwolnienia od choroby i niepokoju. W pewnych przypadkach jednakże nie możemy się odwołać do ogólnej teorii natury ludzkiej, by skonstruować normalny profil kompromisu. Trudno na przykład powiedzieć, przy jakim poziomie dochodu ludzie wybrałyby raczej inne rzeczy niż wzrost dochodu. Prawodawcy muszą jednak czasem dokonywać wyboru pomiędzy inwestycjami w przedsięwzięcia ekonomiczne dające przyrost dochodu w przyszłości a wzrostem nakładów na takie usługi, jak opera czy galeria sztuki. Ponieważ nie wiemy, co należy wtedy wybrać, rząd zrobiłby zapewne dobrze, podejmując odpowiednie badania lub przeprowadzając ankietę.

Czasem jednak badania lub ankietę byłyby zbyt drogim rozwiązaniem w porównaniu z wartością przedmiotu kontrowersji albo nie dawałyby nadziei na osiągnięcie wyniku dostatecznie jednoznacznego. Nie mając lepszego wyjścia, prawodawca powinien wtedy wyobrazić sobie, że znalazł się w sytuacji tych, dla których jego prawo będzie wiążące i odgadnąć wówczas, co byłoby użytecznością dla osób zależnych od jego decyzji. Pod tym względem Rawls ma rację: wnioski do jakich dojdą prawodawcy w tym myślowym eksperymencie mogą być różne. Jeśli więc procedura ta nie prowadzi do jednomyślnego wyniku, musimy się odwołać do innej procedury. W tych warunkach nie można jednak powiedzieć, że użycie nowej procedury jest najlepszym przybliżeniem do maksymalizacji użyteczności.

Na tym zakończymy charakterystykę założeń upraszczających. Powrócę na koniec do dwóch twierdzeń Rawlsa dotyczących intersubiektywnego miernika użyteczności. Po pierwsze, czy porównania intersubiektywne dopuszczone przez tę procedurę ostatecznie opierają się na sądzie wartościującym? Sądzę, że nie. Założenia upraszczające opierają się na pewnym ideale, ale jest to ideał epistemologiczny, a nie moralny. Po drugie, czy propozycja ma charakter praktyczny? Na to pytanie trudniej odpowiedzieć. Prawodawcy mogą być zmuszeni do domysłów na temat, czy zysk jednej grupy jest większy niż strata innej, i domysły te mogą być rozbieżne. Jak mówi Rawls, mogą to być nawet „głęboko konfliktowe opinie”. Ponadto Rawls nie szukał subiektywnych domysłów, lecz obiektywnego miernika dla wszystkich sytuacji. Czy bez takiego miernika jest możliwe, zastanawia się Rawls, by ludzie o rozbieżnych pragnieniach jednomyślnie wybrali ten sam miernik intersubiektywnych porównań? A dopóki nie ma w tej sprawie jednomyślności, czy jakikolwiek miernik jest propozycją praktyczną? Są to nadal kwestie otwarte i jak sądzę, nie będziemy potrafili znaleźć na nie odpowiedzi, dopóki nasza wiedza nie zostanie poszerzona. Ale jestem zdecydowanie przekonany, po pierwsze, że zawsze, gdy pojawia się rozbieżność zdań, prawodawca może się odwołać do innej procedury, która rozbieżność tę usuwa — np. do reguły większości głosów, i po drugie, że wprowadzenie pewni sumienni prawodawcy stosują procedurę, którą tu przedstawiłem w zarysie, większość prawodawców nadal oddaje się raczej mniej bezinteresownym kalkulacjom. Skoro więc nie przyprawia nas o niestrawność to, co się dzieje teraz, dlaczego mielibyśmy się poczuć gorzej, gdybyśmy poprawili istniejący stan rzeczy?

Keble College, Oxford

Tłum. J. Hołówka

Literatura

- Arrow, K. J. 1963, *Social Choice and Individual Values*, 2nd ed. New Haven and London: Yale University Press.
- Arrow, K. J. 1977, *Extended Sympathy and the Possibility of Social Choice*, „The American Economic Review, Papers and Proceedings” vol. 67.
- Arrow K. J. 1978, *Extended Sympathy and the Possibility of Social Choice*, „Philosophia” vol. 7.
- Edel, L. 1978, *Henry James*, vol. I. *The Untried Years: 1843—1870*. New York: Avon Books.
- Harsanyi, J. C. 1976, *Essays on Ethics, Social Behaviour, and Scientific Explanation*, Dordrecht—Boston: D. Reidel.
- Harsanyi, J. C. 1977, *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*, Cambridge: Cambridge University Press.
- Martin, R. G. 1972, *Lady Randolph Churchill, A Biography*, vol. II, London: Cassell.

- Mishan, E. J. 1972, *Elements of Cost-Benefit Analysis*, London: George Allen and Unwin.
- Mishan, E. J. 1975, *Cost-Benefit Analysis*, 2nd ed. London: George Allen and Unwin.
- Morgenstern, O. 1972, *Thirteen Critical Points in Contemporary Economic Theory: An Interpretation*, „Journal of Economic Literature” vol. 10.
- Rawls, J. 1972, *A Theory of Justice*, Oxford: Clarendon Press.
- Rawls, J. 1975a, *A Kantian Conception of Equality*, „Cambridge Review” vol. 96.
- Rawls, J. 1975b, *Fairness to Goodness*, „Philosophical Review” vol. 84.
- Robbins, L. 1935, *An Essay on the Nature and Significance of Economic Science*, 2nd ed. London: Macmillan.
- Rothenberg, J. 1961, *The Measurement of Welfare*, Englewood Cliffs: Prentice-Hall.
- Scanlon, T. M. 1975, *Preference and Urgency*, „The Journal of Philosophy” vol. 72.
- Sen, A. K. 1970, *Collective Choice and Social Welfare*, San Francisco: Holden-Day; Edinburgh: Oliver and Boyd.
- Sen, A. K. 1973, *On Economic Inequality*, Oxford: Clarendon Press.
- Waldner, I. 1972, *The Empirical Meaningfulness of Interpersonal Utility Comparisons*, „Journal of Philosophy” vol. 69.

Джемс Грифин

ПРОБЛЕМА СРАВНИМОСТИ ПОЛЕЗНОГО ДЛЯ РАЗНЫХ ЛЮДЕЙ

Когда мы говорим о равенстве, то имеем в виду также и равенство благосостояний. Однако так говорить можно только тогда, когда признается, что можно сравнивать эти благосостояния между собой, проводить сравнение между полезным для меня и полезным для другого. Такой подход не раз критиковался. Одним из главных аргументов рассматриваемых в данной статье и выдвигаемых против него, является утверждение, что сравнение полезного для различных людей: 1) либо невозможно, 2) если же возможно, то настолько малодостоверно, что попросту бесполезно для анализа реальных общественных ситуаций, 3) либо представляет собой лишь способ выражения чьих-либо ценностных ориентаций, вследствие чего это не имеет большого значения даже для тех, кто имеет такие же моральные убеждения.

Исходным пунктом статьи является анализ важнейших проблем, касающихся межличностного сравнения. Одна из них состоит в том, что экономисты обычно отрицают такое сравнение как «бессмысленное», что тут нет действительной проблемы. Другой вопрос касается другого, более слабого тезиса экономистов, а именно утверждения, что сравнение межличностных полезностей всегда или почти всегда приводит к оценочным суждениям данных личностей. В связи с этим рассматривается та версия данного аргумента, которая отстаивается Лайонелом Роббинсом и Джоном Ролсом. Делается вывод, что это — неубедительный аргумент, что важнейшая проблема сравнения полезностей для разных людей — это 1) теоретическая проблема (в чем критерий такого сравнения?) и 2) в то же время практическая проблема (настолько такое сравнение достоверно?).

Рассматривая ее ближе как теоретическую проблему, автор анализирует решения, предложенные Джоном Харсани, Арматья Сенем и Кэннетом Эроу, а именно то утверждение, что сравнение полезного для разных людей — это лишь выражение ожиданий, уподоблений, склонностей той личности, которая

выносит оценку действий других людей и значимости полезного для них. В статье отвергается это решение. Межличностное сравнение не может опираться, как считают Харсани, Сен и Эроу, на рефлексивных суждениях оценивающей личности, ее ожиданиях и склонностях. Автор предлагает собственное решение: суждение, в котором сравнивается полезное для различных людей, это суждение о фактах, которое опирается на представления о том, насколько интенсивно разные люди хотят или хотели бы обладать определенными вещами; это суждение не выражает ожиданий личности как субъекта сравнения, а потому нельзя межличностное сравнение редуцировать к *внутриличностному* суждениям.

Что же касается практической проблемы, то ее трудно решить, опираясь на широкую шкалу суждений. Однако и здесь, надо полагать, существуют способы, которые позволяют преодолеть трудности, не обращаясь к оценочным суждениям.

James Griffin

INTERPERSONAL COMPARISONS OF UTILITY

When we think about equality, one thing (not necessarily the only thing) we have in mind is equality of welfare. How should we carry on that sort of thinking? One answer is that we should compare people's utilities. That answer is attacked for many reasons, but the attack that I consider in this paper is the one that says: interpersonal comparisons of utilities are either 1) impossible, or 2) if possible, so unreliable as to be unusable in actual social situations, or 3) merely expressions of one's own personal values, so have no standing except with people who share one's moral outlook.

I begin by asking what are the live issues about interpersonal comparisons. One issue is this: economists have often condemned these comparisons as „meaningless”, but I argue that this is not a serious issue. Another is this: economists have often argued less drastically that interpersonal comparisons either are, or depend upon, value judgments. I consider the versions of this argument advanced by Lionel Robbins and by John Rawls, and I argue that neither succeeds. I conclude that the live issues about interpersonal comparisons are, first, a theoretical issue (what are they like?) and, second, a practical issue (how reliable they are?)

On the theoretical question, I begin by considering an answer proposed by John Harsanyi, Armatya Sen, and Kenneth Arrow, viz. that interpersonal comparisons of utility express the preference of the person making the judgement as between entering another's state with his probably different tastes, nature, etc. I argue against this proposal. Interpersonal comparisons cannot, as Harsanyi, Sen and Arrow agree, rest upon the judger's own tastes and values, therefore they cannot, I argue, be expressions of the judger's preference. So I make a proposal of my own: interpersonal comparisons are factual judgements, based upon judgements about how much different people want, or would want, certain things; they are not expressions of preference, and therefore *interpersonal* comparisons cannot be reduced to less problematic *intrapersonal* comparisons.

On the practical question, I admit that my account of the nature of interpersonal comparisons makes their use on the social scale in certain ways especially difficult. But I suggest ways of meeting the difficulty, without resort to value judgements.