

Artur Szutta

Metoda refleksyjnej równowagi. Część I: prezentacja metody

Diametros nr 35, 129-149

2013

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

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

METODA REFLEKSYJNEJ RÓWNOWAGI. CZĘŚĆ I: PREZENTACJA METODY

– Artur Szutta –

Abstract. The article concerns the method of reflective equilibrium which, according to its advocates, allows for correcting and justifying our ethical views. Generally speaking, it consists in the identification of our considered moral judgements, the formulation of moral principles that explicate these judgements, and the elimination of conflicts between them and any relevant background theories. The result of this process of adjustment is an equilibrium in which these elements support and justify each other. Despite a number of advocates (J. Rawls, N. Daniels, R. Dworkin), the method has been heavily criticized (for example by R.M. Hare, P. Singer, or R. B. Brandt), the objections being that it is an example of a vicious circle or that it justifies ethical relativism. This article opens a series of three papers which aim to defend the method. The first step to realize this goal is to present and explain the method itself.

Keywords: reflective equilibrium, considered moral judgements, John Rawls.

1. Wstęp

W polskiej literaturze filozoficznej zagadnieniu metody refleksyjnej równowagi nie poświęcono, według mojej wiedzy, ani jednej poważniejszej publikacji. Jak mogła przejść niezauważona, tocząca się już od kilkadziesiąt lat dyskusja w filozofii anglosaskiej, w której głos zabierali tacy autorzy jak John Rawls, Ronald Dworkin, Richard M. Hare, Richard B. Brandt, Peter Singer? Są dwie możliwe odpowiedzi na to pytanie. Pierwsza jest taka, że we współczesnej filozofii istnieje duża liczba wartych podejmowania problemów, natomiast liczba filozofów w Polsce, szczególnie autorów zajmujących się metaetyką, jest dość ograniczona. Oczywiście nie oznacza to, że polscy filozofowie nie są zaznajomieni z metodą refleksyjnej równowagi. W polskim środowisku byłem świadkiem dwóch prezentacji poświęconych temu zagadnieniu. Tadeusz Szubka w trakcie 50. Tygodnia Filozoficznego Katolickiego Uniwersytetu Lubelskiego z przekonaniem bronił metody refleksyjnej równowagi jako rozsądnego sposobu uprawiania filozofii, zaś Krzysztof Saja na VIII Polskim Zjeździe Filozoficznym w Warszawie prezentował tę metodę jako kryterium wyboru teorii etycznej. Niemniej dla obu zagadnienie to było

tematem pobocznym napotkanym w kontekście zajmowania się inną problematyką¹. Pomimo świadomości tego zagadnienia nikt nie poświęcił refleksyjnej równowadze przynajmniej artykułu. Druga możliwość jest taka, że problem ten uznano za mało interesujący, że metoda refleksyjnej równowagi nie jest obiecującą metodą uzasadniania przekonań etycznych i w związku z tym szkoda czasu na bezowocne badania. Obie możliwości motywują do tego, aby zwrócić uwagę na tę problematykę, co czynię poniżej.

Zacznijmy od określenia ogólnego problemu, w kontekście którego pojawia się zagadnienie refleksyjnej równowagi. John Rawls powszechnie uważany za tego, który jako pierwszy nazwał i zaproponował metodę refleksyjnej równowagi dla filozofii moralnej (i politycznej)², w jednej z pierwszych swoich prac postawił następujące pytanie:

[...] czy istnieje rozsądna [reasonable]³ procedura decyzyjna, wystarczająco silna, która, przynajmniej w niektórych przypadkach, określałaby sposób rozstrzygania pomiędzy konkurującymi interesami [...]; czy można ustalić istnienie takiej procedury jak i jej rozsądnosc [reasonableness] za pomocą racjonalnej metody dociekań?⁴

Problem ten, niezależnie od intencji Rawlsa, można określić nieco szerzej, jako pytanie o możliwość sformułowania takiej metody dociekań moralnych, która jeśli nie dawałaby nam gwarancji dochodzenia do prawdziwej teorii moralnej⁵, lub nie

¹ Np. T. Szubka [2009] s. 222-227.

² Bliższe prawdy jest stwierdzenie, że Rawls zapożyczył termin „refleksyjna równowaga” od Nelsona Goodmana, który wprowadził ten termin w kontekście logiki indukcyjnej (zob. Goodman [1955]), niemniej sama idea, którą Rawls określił za pomocą tego terminu została przezeń zaproponowana na kilka lat przed publikacją Goodmana (zob. Rawls [2001a]). Chociaż poniżej będę często odwoływał się do tego, co na temat metody refleksyjnej równowagi pisze Rawls (i inni autorzy), celem moim nie jest charakterystyka metody zawężona jedynie do Rawlowskiej koncepcji. Chcę zarysować, niezależnie od autorów piszących na jej temat, pewną dającą się obronić ideę refleksyjnej równowagi.

³ Angielskie terminy *reasonableness* jak i *reasonable* tłumaczę tutaj jako „rozsądnosc” i „rozsądny”, chociaż w polskich tłumaczeniach tekstów Rawlsa używa się raczej terminu „rozumność”, „rozumny.” Rozumność kojarzy się bardziej z racjonalnością, którą Rawls przeciwstawia rozsądności. Niuanse związane z rozumieniem pojęcia rozsądnosci dookreślam gdzie indziej (zob. Szutta [2010]).

⁴ Rawls [2001a] s. 177.

⁵ Termin „teoria moralna” (czasami używany zamiennie z terminem „teoria etyczna”) rozumiem tutaj w luźnym sensie, jako zbiór przekonań moralnych (wyrażanych zarówno w postaci zasad, jak i partykularnych sądów moralnych), które pretendują do miana przekonań krytycznie wypracowanych za pomocą racjonalnych dociekań. W sformułowanym powyżej pytaniu można właściwie zrezygnować ze słowa „teoria” i pytać po prostu o metodę, która pozwalałaby dochodzić do uzasadnionych przekonań moralnych, nie rozstrzygając, czy stanowią one teorię czy nie. Chociaż

zwiększałyby naszych szans na sformułowanie teorii zbliżonej do ideału prawdziwości, to przynajmniej umożliwiałyby konstruowanie teorii racjonalnej bądź w jakimś sensie rozsądnej, tj. takiej teorii, która mogłaby być podzielana przez osoby, których dane zasady moralne mają dotyczyć, o ile osoby te dążą do autentycznej kooperacji, kierują się rozumem, oraz życzliwie uwzględniają interesy i racje pozostałych osób, których stanowiące dane rozwiązanie zasady mają obowiązywać⁶.

W świetle współczesnej literatury filozoficznej kandydatem na tego rodzaju satysfakcjonującą metodę dociekań etycznych jest właśnie metoda refleksyjnej równowagi⁷. Najogólniej mówiąc jest to opierająca się na refleksji (rozważnym namyśle) metoda identyfikowania godnych wstępnego zaufania sądów moralnych (zwanych rozważnymi [considered] sądami moralnymi), formułowania zbioru zasad (zwanych teorią moralną), które owe sądy wyjaśniają (eksplikują), oraz usuwania wszelkich istotnych sprzeczności pomiędzy tymi sądami, zasadami, ale także wszelkimi teoriami filozoficznymi i niefilozoficznymi (zwanymi teoriami tła), które mogą mieć wpływ na rewizję rozważnych sądów i zasad moralnych. Wynikiem tej metody jest osiągnięcie punktu ekwilibrium, w którym rozważne sądy moralne, teoria moralna i teorie tła wzajemnie się wspierają (uwiarogodniają). Można więc powiedzieć, przynajmniej zdaniem zwolenników tej metody, że sądy i teorie podzielane przez daną osobę w stanie refleksyjnej równowagi są uzasadnione, bądź że ich utrzymywanie jest racjonalne.

Niemniej metoda ta została poddana także ostrej krytyce. Dwa najpoważniejsze i najczęstsze zarzuty brzmią: a) metoda refleksyjnej równowagi stanowi klasyczny przykład błędnego koła – jej wynik końcowy jest w sposób istotny uwarunkowany jej punktem wyjścia – intuicyjnie przyjmowanymi sądami moralnymi, które same jako odzwierciedlenie naszych stronniczych pragnień, kultury, wychowania czy indoktrynacji, nie zasługują na wiarygodność⁸; b) zgodnie z drugim zarzutem metoda ta nie jest obiecująca, ponieważ nie spełnia istotnego warunku – nie daje szans na konwergencję poglądów moralnych wśród osób ją stosujących,

w kontekście uprawiania filozofii moralnej, ambicją naszą jest oczywiście sformułowanie takiego spójnego zbioru przekonań, który zasługiwałby na miano teorii etycznej.

⁶ Na temat Rawlowskiego rozumienia osoby rozsądnej zob. Szutta [2010].

⁷ Jest to zdanie wielu współczesnych filozofów, spośród których warto wymienić przede wszystkim J. Rawlsa i N. Danielsa. Za różnymi wersjami tej metody opowiadali się również tacy autorzy jak Michael DePaul, Ronald Dworkin. Niektórzy także w akceptowanych przez siebie metodach dochodzenia do uzasadnionych przekonań moralnych a metodą refleksyjnej równowagi dopatrują się wielu istotnych podobieństw (Zob. np. Audi [2007]; Huemer [2005] s. 106).

⁸ Zarzut ten w różnych wersjach formułują np. R.M. Hare (Hare [1989]); P. Singer (Singer [1974], [2005]); czy R.B. Brandt (Brandt [1979], [1990]).

nawet jeśli ich punktem wyjścia były te same rozważne sądy moralne⁹; ponadto uznając prawomocność metody refleksyjnej równowagi, musielibyśmy się zgodzić, że nawet wykluczające się wzajemnie poglądy różnych osób, o ile każda z nich indywidualnie znajduje się w stanie refleksyjnej równowagi, należy traktować jako równie uzasadnione.

Satysfakcjonująca prezentacja metody refleksyjnej równowagi, obok wyczerpującego jej opisu i wyjaśnienia powinna zatem zawierać także dogłębną analizę zarzutów oraz ewentualną próbę ich odparcia. W jednym artykule satysfakcjonująco zrealizować tak szeroki projekt nie sposób. Wymaga on albo obszernej monografii, albo serii artykułów, z których każdy stanowiłby analizę poszczególnych elementów zagadnienia. Artykuł ten zatem będzie stanowił część serii trzech publikacji, które kolejno będą poświęcone: 1) prezentacji metody refleksyjnej równowagi; 2) analizie i próbie odparcia zarzutu błędnego koła i braku wiarygodności intuicji moralnych; oraz 3) analizie i próbie odparcia zarzutu z niezgody moralnej i moralnego relatywizmu.

Opisując proces dochodzenia do refleksyjnej równowagi można wymienić jego trzy etapy, którym będą odpowiadać trzy kolejne części tekstu: a) selekcję rozważnych sądów moralnych; b) eksplikację tych sądów poprzez zasady moralne i uzyskiwanie wąskiej refleksyjnej równowagi; c) dojście do szerokiej refleksyjnej równowagi. Pewne dopełnienie obrazu metody refleksyjnej równowagi będzie stanowić krótka analiza możliwych jej interpretacji. W zakończeniu postaram się przytoczyć argumentację na rzecz metody refleksyjnej równowagi, jako jedynej racjonalnej metody formułowania teorii etycznej, oraz zarysować pobieżnie dalsze kroki w dwóch odrębnych artykułach stanowiących z tym tekstem zamierzoną całość obrony metody refleksyjnej równowagi.

2. Selekcja rozważnych sądów moralnych

Pierwszym krokiem w procesie dochodzenia do refleksyjnej równowagi jest selekcja tzw. rozważnych sądów. Tej selekcji dokonujemy spośród własnych sądów wartościujących, zarówno o charakterze partykularnym jak i ogólnym¹⁰,

⁹ Taki zarzut wysuwają np. P. Singer (Singer [1974] s. 493n); Michael Philips (zob. Philips [1994] s. 37-42). Klarownie zarzut ten prezentuje także T.M. Scanlon (zob. Scanlon [2003] s. 151-153).

¹⁰ Pogląd samego Rawlsa ewoluował w tym aspekcie. Pierwotnie uważał on, że sądy rozważne winny dotyczyć tylko konkretnych sytuacji (zob. Rawls [2001a] s. 182); później jednak odszedł od tak wąskiej interpretacji sądów, przyznając, że „ludzie posiadają rozważne sądy na każdym poziomie ogólności”. (Zob. Rawls [2001b] s.289; por. także Scanlon [2003]). Z pewnością czynią tak inni autorzy, dopuszczając także przynajmniej niektóre sądy odnoszące się do hipotetycznych sytuacji, np. DePaul, (Depaul [1993] s. 16). Uzasadnienie takiego poszerzenia spektrum sądów jest takie, że również sądy dotyczące sytuacji hipotetycznych, także o pewnym stopniu ogólności, o ile

dotyczących (realnych lub hipotetycznych) działań, postaw, sytuacji lub osób, w których oceniamy dane działanie jako słuszne lub niesłuszne, moralnie dobre lub złe, oceniamy postawy osób jako odpowiedzialne lub nieodpowiedzialne¹¹. Sądy, spośród których dokonujemy selekcji sądów rozważnych, wydajemy w odniesieniu do sytuacji z codziennego życia, czytanych powieści, oglądanych sztuk teatralnych czy filmów. Twierdzimy na przykład, że nasz sąsiad oddając nam zgubione przez nas pieniądze zachował się przyzwoicie, Frodo (bohater powieści Tolkiena), litując się nad Golumem, postąpił szlachetnie, że Hitler był nikczemnym człowiekiem, że wojna jest czymś złym. Twierdzimy też, że w takich samych pod każdym istotnym pozamoralnym względem sytuacjach, ocena moralna będzie taka sama, albo że umów należy dotrzymywać itp.¹². Wydaje się, że wyliczanie takich przekonań może ciągnąć się praktycznie w nieskończoność. Owszem, możemy skupiać się na przekonaniach dotyczących tylko pewnego wąskiego zagadnienia, np. sprawiedliwości, kary śmierci czy eutanazji – niemniej może okazać się, że dojście do stanu refleksyjnej równowagi w kwestiach o mniejszym zakresie będzie wymagało odwołań do przekonań, które pozornie nie mają z nimi nic wspólnego.

Nie wszystkie wstępne sądy moralne mają charakter sądów rozważnych i zasługują na równe zaufanie. Istnieje długa lista przyczyn, które mogą na nas oddziaływać deformując nasze przekonania moralne. Formułując sądy moralne możemy zachowywać się w sposób stronniczy, działać pod wpływem pragnień, które każą nam pewne działania, pewne osoby oceniać łagodniej lub surowiej, niż na to zasługują, możemy kierować się nadzieją na spodziewane korzyści lub obawą przed spodziewaną stratą; możemy też źle zinterpretować okoliczności (przeoczyć istotne elementy sytuacji, które mogłyby diametralnie zmienić naszą ocenę). To przeoczenie może być wynikiem wysokiego skomplikowania sytuacji, naszego pośpiechu w trakcie wydawania sądu lub bycia pod wpływem silnych emocji itp. Możemy również opierać swój sąd na fałszywych informacjach, ukrytych przesłankach, których sobie nie uświadamiamy. Także cechy naszego charakteru takie jak upór, lenistwo, roztargnienie mogą mieć istotny wpływ na naszą błędną ocenę. Istotnym czynnikiem może być również wpływ innych osób, błędnie wziętych za autorytety. Pewne przekonania mogą w nas tkwić jako wielokrotnie powtarzane

nie dotyczą one sytuacji skomplikowanych, mogą stanowić wiarygodną (w słabym przynajmniej sensie wiarygodności) podstawę do formułowania teorii.

¹¹ Zob. DePaul [1993] s. 16.

¹² Por. *ibidem*.

w naszym środowisku, kulturze itp. Nie można także wykluczyć możliwości iluzji będącej wynikiem oddziaływania naszego języka czy samego rozumu¹³.

Rozważne sądy moralne stanowią tę kategorię sądów, w których, ujmując to słowami Rawlsa, „nasze zdolności moralne mają największą szansę ujawnić się bez zniekształcenia”¹⁴. Tak więc, biorąc pod uwagę, że niektóre z naszych wstępnych sądów moralnych (może nawet większość) nie były formowane w sprzyjających epistemicznie warunkach, konieczne jest wyselekcjonowanie takich sądów, które powstały w warunkach wolnych od zakłóceń. Rawls wymienia kilka cech, jakie rozważne sądy powinny posiadać: a) winny być bezstronne, wydane w warunkach wolności od dających się przewidzieć ich konsekwencji, rozumianych w terminach spodziewanych korzyści (np. nagrody) lub straty (np. kary); b) winny być także formułowane w warunkach wolności od stanów emocjonalnego wzburzenia bądź przerażenia; c) nie powinny być zbyt skomplikowane; d) powinny być poprzedzone namysłem [inquiry] nad wszystkimi istotnymi faktami; e) winny być odczuwane jako pewne; f) winny odznaczać się stabilnością, tzn. opierać się na świadectwie, które w innym czasie, w podobnych przypadkach, także w przypadku innej osoby dokonującej sądu, prowadziłyby do wydania takiego samego sądu¹⁵; g) sądy te winny być intuicyjne względem zasad etycznych, tzn. przyjmowane niezależnie od ich wynikania logicznego z tych zasad, które rozważne sądy mają wspierać – jest to konieczne dla uniknięcia popadania w błędne koło¹⁶.

¹³ Powyższa lista przyczyn błędów w ocenach moralnych w dużej mierze jest inspirowana tekstem Michaela Huemera (Huemer [2005] s. 137-139).

¹⁴ Rawls [1993] s. 70; por. Rawls [1999] s. 42. Uzasadnione ze strony czytelnika może być pytanie, o jakich zdolnościach poznawczych tutaj mowa. Zamiast o rozważnych sądach należałoby być może od razu pisać o leżących u podstaw tych sądów intuicjach moralnych, ponieważ jednak zarówno Rawls, jak i inni autorzy na etapie opisu omawianej tu metody ograniczają się do mówienia o rozważnych sądach, o intuicjach pisząc dopiero w kontekście dyskusji z krytykami metody refleksyjnej równowagi, tak też i w tym tekście postanowiłem (z małymi wyjątkami) na tym etapie nie odwoływać się do intuicji moralnej jako źródła formułowanych przez nas rozważnych sądów moralnych.

¹⁵ Warunek ten wydaje się sprzeczny z założeniem, że proces refleksyjnej równowagi (przynajmniej na tym etapie opisu) dotyczy jednej osoby. Dlaczego jedna osoba miałaby kierować się tym, czy inne (kompetentne) osoby uznałyby świadectwo, na którym opiera ona swój sąd? Sądzę, że można podać dobre racje za taką interpretacją stabilności sądu – szczególnie, jeśli chcemy, aby indywidualna refleksyjna równowaga była sprzężona z refleksyjną równowagą innych w taki sposób, aby możliwe było porozumienie (międzyosobowa refleksyjna równowaga). Kwestię tę zostawiam jednak otwartą.

¹⁶ Por. Rawls [2001a] s. 181-183. We fragmencie, do którego odnosi się powyższe wyliczenie warunków, jakie winny spełniać rozważne sądy, Rawls wymienia także warunek konkretności i niehipotetyczności tych sądów. Warunek ten z powodów już wymienionych pomijamy. Warunek (b) wolności od stanów wzburzenia i przerażenia jest dodany przez Rawlsa w: Rawls [1993] s. 70.

W *Outline of a Decision Procedure for Ethics* Rawls dostrzegał także konieczność pewnego stopnia kompetencji osoby dokonującej rozważnych sądów moralnych. Jak pisał, osoba zdolna do formułowania tego rodzaju sądów winna posiadać przynajmniej przeciętną inteligencję, normalną wiedzę na temat faktów dotyczących świata, dających się przewidzieć konsekwencji podejmowanych decyzji. Winna także być osobą rozsądną [reasonable], czyli gotową stosować kryteria logiki indukcyjnej, posiadać otwarty umysł, tj. przejawiać gotowość przyjrzenia się wszystkim dostępnym „za i przeciw” danej decyzji, oraz gotową do zmiany decyzji w świetle ewentualnych nowych rozważań; świadomą swoich własnych predylekcji, emocjonalnych czy intelektualnych mających na nią wpływ podczas ważenia racji. Co równie istotne, w swoich rozważaniach osoba rozsądna kieruje się w sposób życzliwy dla innych osób wiedzą na temat tego, co leży w interesie tych osób, jak i na temat sposobu, w jaki jej decyzje mogą mieć wpływ na ich los¹⁷. Niemniej w późniejszych tekstach (po *Teorii sprawiedliwości*) Rawls zdaje się osłabiać wymóg kompetentnego sędziego: każdy normalny dorosły człowiek, o przeciętnej inteligencji i podstawowej wiedzy o świecie i zdolności przewidywania skutków swojego działania winien być uznany za osobę zdolną do formułowania rozważnych sądów moralnych¹⁸. To złagodzenie wymogu kompetencji można wyjaśnić tym, że Rawlsa, autora *Teorii sprawiedliwości* i *Liberalizmu politycznego* mniej interesowały dociekania filozoficzne zawodowych teoretyków próbujących konstruować teorię moralną, a bardziej metoda konstruowania zasad sprawiedliwości dostępna hipotetycznym obywatelom znajdującym się w sytuacji pierwotnej. Należy jednak zgodzić się z M. DePaulem, że idea kompetentnego sędziego ma pewne znaczenie dla metody refleksyjnej równowagi: wydaje się oczywiste, że wstępna wiarygodność rozważnych sądów jest tym większa, im wyższy poziom kompetencji prezentuje dzielająca je osoba¹⁹.

Na koniec tej części jeszcze dwie uwagi. Po pierwsze, chociaż rozważne sądy moralne cieszą się wysokim stopniem zaufania, to ich wiarygodność ma charakter *prima facie*. Nie są traktowane jako absolutnie nieomyłne i mogą ulec rewizji. Po drugie jak sądzę, selekcji rozważnych sądów nigdy nie należy traktować jako zakończonej – i nie chodzi tu tylko o odrzucanie sądów, które uprzednio uznane za rozważne, nie ostały się w dalszych etapach procesu dochodzenia do refleksyjnej równowagi. W każdym momencie możemy odwołać się do kolejnych, nieuwzględnionych wcześniej przekonań, które możemy ocenić jako wstępnie

¹⁷ Zob. Rawls, [2001a] s. 178-181.

¹⁸ Por. DePaul [2006] s. 605.

¹⁹ Por. ibidem.

wiarygodne i dołączyć do zbioru rozważnych sądów ustalonych wcześniej. Rozsądne wydaje się także stwierdzenie, że zbiór ów jest potencjalnie nieskończony, tak jak nieskończona może być ilość różnych przypadków działań, sytuacji, postaw, w odniesieniu do których możemy formułować nasze sądy moralne.

3. Eksplikacja rozważnych sądów moralnych i dochodzenie do wąskiej refleksyjnej równowagi

Po dokonaniu selekcji rozważnych sądów moralnych, kolejnym krokiem jest sformułowanie zasad moralnych, które eksplikowałyby rozważne sądy zachowując jednocześnie spójność zarówno z rozważnymi sądami, jak i ze sobą nawzajem. Skupmy się na procesie formułowania zasad i pojęciu eksplikacji. Po pierwsze, niektóre rozważne przekonania same mogą mieć charakter sądów ogólnych, czyli zasad (i to na różnym poziomie ogólności). Po drugie, analiza poszczególnych przekonań może prowadzić nas do pewnych wglądów, aktów zrozumienia, uchwycenia pewnych zasad ogólnych. Np. analizując swoje przekonania typu „Sasha Raskolnikow nie powinien był zabijać staruszki”; „Golum źle uczynił zabijając swojego przyjaciela” itd. mogę (o ile już wcześniej nie podzielałem tego przekonania) dojść do uznania ogólnej zasady, że zabijanie zasadniczo jest moralnie złe. To przekonanie nie musi mieć charakteru zwykłej indukcyjnej generalizacji. Może być i tak, że refleksja nad jednym przypadkiem pozwala nagle dojrzeć, uchwycić czy zrozumieć, że każdy inny przypadek, podobny pod istotnym względem, powinien być w ten sam sposób oceniony²⁰.

Na czym polega eksplikacja rozważnych sądów? Zdaniem Rawlsa dana zasada eksplikuje określony rozważny sąd moralny, gdy każdy inteligentny człowiek, świadomie ją stosując do konkretnych przypadków, jest w stanie sformułować taką samą ocenę, jak ta wyrażona w rozważnym sądzie²¹. Sens eksplikacji, trafnie a zarazem bardzo plastycznie oddaje przykład DePaula odwołujący się do pojęcia krzywej funkcji na wykresie. Zgodnie z tym przykładem nasze konkretne rozważne sądy byłyby reprezentowane przez odpowiednie punkty na wykresie współrzędnych, zaś zasada byłaby reprezentowana przez odpowiednią krzywą. Z reprezentacją eksplikacji mamy do czynienia, gdy krzywa pokrywa się z punktami, pozwalając nie tylko (przy pomocy ogólnej formuły) identyfikować te punkty, ale także identyfikować nowe punkty nie zaznaczone uprzednio na wykresie²². Relacja eksplikacji wydaje się wielopoziomowa. Ogólne zasady eksplikują

²⁰ Por. Audi [2007].

²¹ Zob. Rawls [2001a] s. 184.

²² Zob. DePaul [1987] s. 466.

nie tylko partykularne sądy, ale też wszelkie inne mniej ogólne zasady, same też mogą być eksplikowane przez zasady jeszcze bardziej ogólne.

Charakterystykę eksplikacji Rawls dookreśla poprzez zastrzeżenia, czego przez eksplikację nie rozumie, a także przez wskazanie pozytywnych cech satysfakcjonującej eksplikacji. Tak więc nie jest ona analizą znaczenia terminów etycznych, nie zajmuje się też intencjami tych, którzy wyrażają sądy etyczne, ani nie jest teorią faktycznych przyczyn rozważnych sądów²³. Satysfakcjonująca eksplikacja pozwala dochodzić na zasadzie wnioskovania do pełnego zakresu sądów, które są tożsame z rozważnymi sądami lub przynajmniej do nich przybliżone²⁴. Chociaż Rawls tego wprost nie mówi, wydaje się, że zasady satysfakcjonująco eksplikujące rozważne sądy, jednocześnie pozwalają formułować nowe sądy, które albo nam najzwyczajniej nie przyszły do głowy (po dłuższej refleksji może byśmy je sformułowali jako własne przekonania), albo byłyby dla nas niedostępne poznawczo. Jeśli mielibyśmy do czynienia z doskonałą²⁵ eksplikacją, to oznaczałoby, że udało nam się z jednej strony sformułować jedną (lub najwyżej kilka spójnych) najogólniejszych zasad, które eksplikują wszystkie nasze rozważne sądy, nie pozostając z żadnymi w sprzeczności; z drugiej zaś strony nasze rozważne sądy efektywnie wspierają naszą refleksję nakierowaną na formułowanie takich zasad. Taki stan optymalnej eksplikacji przez pryzmat późniejszych niż *The Outline...* prac Rawlsa można określić mianem wąskiej refleksyjnej równowagi²⁶.

Stanu tego nie osiągamy jednak od razu, na zasadzie jednego kroku czy nawet kilku kroków polegających na sformułowaniu i zebraniu zasad eksplikujących rozważne sądy. Dochodzenie do wąskiej refleksyjnej równowagi jest procesem. W trakcie formułowania zasad eksplikujących rozważne sądy natrafiamy bowiem na trudności: a) niespójność (konflikty, rozbieżności) między zasadami a partykularnymi sądami, a także między samymi zasadami, jak również pomiędzy samymi partykularnymi sądami, które te zasady eksplikują; b) zasady mają ograniczony zasięg eksplikacji, tzn. pozostawiają jakąś grupę rozważnych sądów poza eksplikacją; c) same zasady mogą też budzić nasze zaufanie w różnym stopniu. Stąd proces dochodzenia do wąskiej refleksyjnej równowagi wymaga eliminacji konfliktów lub rozbieżności pomiędzy zasadami a rozważnymi sądami,

²³ Rawls [2001a] s. 184-185.

²⁴ Zob. ibidem, s. 184n.

²⁵ Sam Rawls nie używa określenia „doskonała eksplikacja”, mówi jedynie o pozytywnych cechach satysfakcjonującej eksplikacji.

²⁶ Chociaż pojęcie wąskiej refleksyjnej równowagi przeciwstawionej szerokiej refleksyjnej równowadze jest już implicite obecne w *Teorii sprawiedliwości*, wprost tego terminu Rawls używa dopiero w późniejszym artykule *The Independence of Moral Theory* (Rawls [2001b]).

wymaga także ciągłego poszukiwania zasad, które lepiej spełniałyby swoją funkcję eksplikowania rozważnych sądów.

Usuwanie konfliktów opiera się na wzajemnym dopasowywaniu zasad i rozważnych sądów, rezygnacji z jednych zasad na rzecz innych, formułowanych na nowo, ale także rezygnacji z niektórych rozważnych sądów. W tego typu wyborach dana osoba musi kierować się wycuciem tego, z jaką siłą dana zasada lub sąd ją przekonują, czy też wycuciem stopnia, w którym osoba ta czuje się przywiązana do danej zasady lub sądu. Także ważne jest, jak centralną rolę dany sąd czy zasada odgrywa dla pozostałych sądów i zasad danej osoby. Ów proces wzajemnego dopasowywania zasad i rozważnych sądów ma charakter chodzenia „w tę i z powrotem” aż do momentu, w którym oba zbiory (zasad i rozważnych sądów) osiągają stan odpowiedniego zestrojenia, zwanego właśnie wąską refleksyjną równowagą²⁷.

4. Szeroka równowaga refleksyjna

Wąska refleksyjna równowaga nie jest czymś ostatecznym ani trwałym. Jak pisze Rawls, „może zostać zachwiana” w wyniku dalszego badania rozważnych sądów, refleksji nad konkretnymi przypadkami, które mogą skłonić daną osobę do dalszej rewizji przekonań²⁸. Celem jest dążenie do określenia danej teorii moralnej w wyniku pełnej refleksji. Teorie moralne sformułowane na tym etapie opisywanej tu metody stanowią jedynie odbicie naszych sądów moralnych, które co prawda, zdaniem zwolenników tej metody, cechuje pewien poziom wstępnej wiarygodności, niemniej nie można ich traktować jako nieomyłne. Zatrzymując się na wąskiej refleksyjnej równowadze, możemy popaść w dogmatyzm, nawet pomimo pewnego stopnia rewidowalności rozważnych sądów moralnych. Istnieje bowiem niebezpieczeństwo, że chociaż uzyskanie spójności pomiędzy rozważnymi sądami moralnymi a moralnymi zasadami pozwala uporządkować i rzucić nieco światła na nasze poglądy moralne, to nasze początkowe (w tym rozważne) sądy moralne mogą stanowić jedynie odzwierciedlenie naszej kultury, interesów klasowych czy historycznych przypadków²⁹. Innym powodem wyjścia poza wąską refleksyjną równowagę jest fakt, że różne osoby osiągające stan wąskiego ekwilibrium mogą formułować teorie moralne wzajemnie się wykluczające bez możliwości odwołania się do jakiejś teoretycznej podstawy, dzięki której mogłyby w trakcie dyskusji

²⁷ Por. Rawls [1993] s. 34-37.

²⁸ Ibidem, s. 36.

²⁹ Por. Daniels [1996] s. 21.

uzgodnić swoje stanowiska. Stąd należy pytać o dalsze racje, dla których mielibyśmy akceptować zasady moralne³⁰.

Kluczowym elementem czy ideą metody refleksyjnej równowagi jest, jak już zauważyliśmy, wzajemne wspieranie się naszych przekonań polegające na ich wzajemnej koherencji i wyjaśnianiu. Zgodnie z tą ideą wzajemne wspieranie się przekonań jest tym mocniejsze, im szerszy jest zbiór przekonań osiągających ekwilibrium, dlatego też, aby odpowiedzieć na pytanie o dalsze racje uznania przekonań moralnych należy, parafrazując słowa Danielsa, poszerzyć koło uzasadniania o zbiór wszelkich możliwych moralnych i pozamoralnych przekonań w poszukiwaniu ich wzajemnego wsparcia³¹. Takie poszerzenie metody, zdaniem Danielsa, z jednej strony pozwala na głębszą rewizję naszych przekonań moralnych i tym samym przekroczenie dogmatyzmu, z drugiej zaś strony daje nadzieję na łatwiejsze rozwiązywanie problemu moralnej niezgody, pozwala bowiem odwoływać się do pozamoralnych przekonań, które dając się łatwiej rozstrzygać w oparciu o obiektywne empiryczne kryteria, mogą pośrednio wpływać także na zmianę przekonań moralnych³².

Ideę wyjścia poza wąską refleksyjną równowagę Rawls wyraża już w *Teorii sprawiedliwości*. Twierdzi tam, że zasady sprawiedliwości (do przyjęcia których ma prowadzić metoda refleksyjnej równowagi) winny być wybrane nie tylko w oparciu o proces prostej eksplikacji rozważnych sądów moralnych, ale także w oparciu o analizę porównawczą tych zasad w odniesieniu do wszelkich możliwych alternatywnych koncepcji sprawiedliwości. Z powodów praktycznych Rawls ogranicza liczbę alternatyw do listy kilku tradycyjnych koncepcji sprawiedliwości. Chcąc upewnić się, że proponowane przez niego zasady sprawiedliwości zasługują na miano najlepszej propozycji, konstruuje pojęcie hipotetycznej sytuacji pierwotnej, w której, mówiąc w wielkim uproszczeniu, analiza porównawcza i ostatecznie wybór zasad sprawiedliwości byłyby dokonywane w oparciu o wiele pozamoralnych, filozoficznych, psychologicznych, socjologicznych i ekonomicznych argumentów³³.

W pełni rozwiniętą, Rawlowską w duchu koncepcję szerokiej refleksyjnej równowagi proponuje uczeń Rawlsa, Norman Daniels w serii artykułów z lat 1979-1980, wydanych później wraz z innymi jego tekstami poświęconymi metodzie refleksyjnej równowagi w jednym tomie pt. *Justice and Justification. Reflective*

³⁰ Por. Daniels [1996] s. 1, wstęp.

³¹ Ibidem.

³² Por. ibidem, s. 7, 21-25.

³³ Por. Rawls [1993] rozdz. III, s. 167-264.

*Equilibrium in Theory and Practice*³⁴. W jego ujęciu szeroka refleksyjna równowaga oznacza wzajemne dopasowanie rozważnych sądów, zbioru zasad moralnych (stanowiących daną teorię moralną) z tzw. relewantnymi teoriami tła. Teorie tła mogą odnosić się do całej gamy zagadnień, np. koncepcji osoby, teorii gier, koncepcji racjonalności, roli moralności w naszym życiu czy społecznej kooperacji itp.³⁵. Ich zadaniem jest pokazać, w sposób stosunkowo niezależny od rozważnych sądów moralnych, że dany zbiór zasad (teoria moralna) jest bardziej akceptowalna, niż jej alternatywy. Podobnie jak w przypadku wąskiej refleksyjnej równowagi, w razie konfliktu (niekoherencji), żaden z elementów tych trzech zbiorów nie jest poza możliwością rewizji. W trakcie dopasowywania ich do siebie dana osoba, podobnie jak w przypadku dochodzenia do wąskiej refleksyjnej równowagi, kieruje się siłą swojego przywiązania do danych przekonań i logicznymi relacjami pomiędzy nimi³⁶.

Daniels zdaje sobie sprawę z trudności zestawiania ze sobą sądów moralnych z przekonaniem pozamoralnym, uznaje bowiem, że sądy moralne są nieredukowalne do sądów pozamoralnych. Niemniej często (szczególnie na niskim poziomie abstrakcji) nasze sądy moralne dotyczą fizycznych stanów rzeczy i zdarzeń. Twierdzimy np., że arszeniku nie należy podawać ludziom do jedzenia, że osoby będące nosicielami pewnego typu bakterii nie powinny przygotowywać posiłków w restauracjach, itp. Sądy te są w pewnym sensie połączeniem elementów czysto normatywnych (moralnych) i opisowych (pozamoralnych). Uznanie lub nieuznanie ich zasadności wymaga wiedzy na temat faktów pozamoralnych (choć ta wiedza nie jest wystarczająca), np. czy arszenik jest trucizną, co to znaczy, że jest trucizną, jakie są skutki zakażenia żywności pewnego typu bakteriami i kiedy do takiego zakażenia może dojść, itp. Spory moralne, spory o zasady moralne czy o konkretne sądy moralne często okazują się dotyczyć nie tyle komponentu czysto normatywnego, ale właśnie kwestii opisowych. Na przykład w wielu przypadkach sporów o aborcję nie chodzi o to, czy można zabijać ludzi, ale o to, czy zabijany płód jest człowiekiem. Niemniej rola teorii tła, według Daniela, nie

³⁴ Daniels [1996]. Chodzi tu przede wszystkim o *Wide Reflective Equilibrium and Theory Acceptance in Ethics*, s. 21- 46 (pierwotnie wydany w „*Journal of Philosophy*” (76) 1979, s. 256-282); oraz *Reflective Equilibrium and Archimedean Points*, s. 47-65 (pierwotnie wydany w „*Canadian Journal of Philosophy*” (10) 1980, s. 83-103).

³⁵ Por. Daniels [1996] s. 6.

³⁶ O możliwości odwołania się do poczucia przywiązania do poglądów (dotyczących teorii tła) na etapie dążenia do szerokiej refleksyjnej równowagi Daniels nie pisze. Przyjęcie takiego kryterium rozstrzygnięcia konfliktów na tym etapie wydaje się jednak zasadne. Skoro mówimy tu o metodzie stosowanej przez daną osobę indywidualnie, wszelkie rozstrzygnięcia winny być w zgodzie ze standardami epistemicznymi tej właśnie osoby (por. DePaul, *Balance and Refinement* [1993] s. 17-20).

ogranicza się do tego rodzaju funkcji. Gdyby tak było, teorie tła miałyby znaczenie drugorzędne względem zasad i sądów etycznych, tymczasem mają one umożliwić bardziej fundamentalną rewizję tego rodzaju sądów i zasad. Ponieważ jednak nie można z poziomu czysto deskryptywnego wyciągać wniosków normatywnych, teorie tła, aby wchodzić w głębsze relacje koherencji i wyjaśniania z zasadami moralnymi i rozważnymi sądami moralnymi, same muszą zawierać w sobie element normatywny³⁷. To zaś oznacza, że teorie tła są w istotny sposób determinowane przez moralne przekonania danej osoby. Taka interpretacja rozważnych sądów otwarta jest jednak na zarzut, że teorie tła nie są niczym więcej jak reformulacjami rozważnych sądów, że być może poszerzają sieć koherujących ze sobą przekonań, ale nie jest to istotna zmiana jakości, w związku z czym można argumentować, że szeroka refleksyjna równowaga posiada zasadniczo te same słabości, co wąska refleksyjna równowaga.

Dla uniknięcia tego rodzaju trudności Daniels wprowadza tzw. wymóg niezależności [independence constraint]³⁸. Zgodnie z tym wymogiem teorie tła winny dawać wsparcie teorii moralnej w sposób niezależny od rozważnych sądów, dzięki którym teoria ta została sformułowana na poziomie wąskiego ekwilibrium. Ponieważ, jak już zauważyliśmy powyżej, teorie tła same nie mogą się obejść bez pewnej grupy moralnych przekonań, zawierają komponent normatywny, konieczne jest wyodrębnienie przekonań moralnych determinujących teorie tła jako podzbiór sądów moralnych rozłączny (bądź niezależny) względem nietrywialnej części rozważnych sądów wspierających daną teorię moralną w stanie wąskiej refleksyjnej równowagi. Warunek ten zdaniem Daniela może być spełniony, gdy na przykład w teoriach tła odwołujemy się do innych pojęć moralnych (np. pojęcie bezstronności, czy pojęcie osoby) niż w zasadach moralnych i rozważnych sądach z etapu dochodzenia do wąskiej refleksyjnej równowagi. To podzielenie zbioru rozważnych sądów na dwa względnie, ale w sposób nietrywialny, rozłączne podzbiory, pozwala uniknąć pułapki biernego (nie wnoszącego istotnego novum) systematyzowania rozważnych sądów moralnych. Uzyskanie koherencji pomiędzy teorią moralną wraz z rozważnymi sądami, które ona eksplikuje, a wprowadzonymi do refleksji niezależnymi teoriami tła (zdeterminowanymi pewnymi moralnymi przekonaniem, ale stanowiącymi odrębny względem rozważnych sądów moralnych użytych do osiągnięcia wąskiej równowagi zbiór) stanowi zdaniem Daniela dodatkowy silny argument za daną teorią moralną.

³⁷ Świadomość tego problemu w tekstach Daniela sugerują niektóre fragmenty artykułu *Reflective Equilibrium and Archimedean Points* (por. Daniels [1996] s. 49-50, 52.)

³⁸ Zob. Daniels [1996] s. 22-23, 49-51.

Powyższe rozwiązanie nie wydaje się jednak przekonujące. Wymóg niezależności zmusza nas już od samego początku stosowania metody refleksyjnej równowagi do sztucznego trzymania na boku pewnej grupy rozważnych sądów - tak, aby nie wpływały one na formułowanie zasad moralnych. To odseparowanie można postrzegać jako arbitralne (skąd bowiem mamy wiedzieć, które sądy mają być odłożone na później, skoro jeszcze nie poddaliśmy refleksji teorii tła), poza tym sformułowana w ramach wąskiej równowagi teoria moralna z góry zostaje skazana na istotną niekompletność³⁹. Czy nie lepiej zaangażować cały (dostępny nam w danym momencie) zbiór rozważnych sądów od samego początku stosowania metody? Zdaniem niektórych autorów nie musi to oznaczać, że szeroka refleksyjna równowaga będzie zbędna: wyjście poza koherencję moralnych sądów i zasad w kierunku szerszego kręgu wspierających się przekonań, obejmującego także przekonania pozamoralne, zgodnie z rozumowaniem koherencjonistycznym, wzmocni stopień uzasadnienia sądów i zasad moralnych⁴⁰.

Przyjęcie powyższej konkluzji ma swoją cenę. Po pierwsze, należy zgodzić się, że stopień uzasadnienia i głębokość rewizji, jakie umożliwia szeroka równowaga refleksyjna, nie jest taki wysoki, jak tego się spodziewał Daniels. Owszem teorie dotyczące świata fizycznego, psychologii itd. pozwalają nam rozstrzygać niektóre spory moralne, w których niezgoda właściwie dotyczy pozamoralnych komponentów, pozwalają nam też lepiej zrozumieć czynniki warunkujące nasze zdolności do wydawania rozważnych sądów moralnych, przez co pośrednio wpływają na ich doskonalszą selekcję, nie dają jednak możliwości głębszej rewizji czysto normatywnych elementów naszych moralnych przekonań, jak zdaje się sugerować Daniels. Po drugie, skoro teorie tła w aspekcie, w którym stają się istotne dla korygowania i uzasadniania moralnych przekonań, nie są niezależne od rozważnych sądów moralnych, to ostateczny sukces metody refleksyjnej równowagi zależy w znacznie większym stopniu od epistemicznej wartości rozważnych sądów moralnych, niż skłonny byłby przyznać Daniels⁴¹.

³⁹ Niniejsze krytyczne uwagi na temat warunku niezależności w dużej mierze inspirowane są lekturą DePaula (DePaul [1993] s. 21-23). Por. także Halset [1987], szczególnie s. 307-308.

⁴⁰ Por. DePaul [1993] s. 22.

⁴¹ Kładąc nacisk na konieczność szerokiej refleksyjnej równowagi oraz nadając tak duże znaczenie pozamoralnym przekonaniom, Daniels miał nadzieję na znaczne pomniejszenie roli rozważnych sądów moralnych w uzasadnianiu zasad moralnych. Był przekonany, że początkowe rozważne sądy moralne można traktować jako coś w rodzaju „Wittgensteinowskiej drabiny”, potrzebnych jako coś, od czego trzeba zacząć, ale później można całkowicie odrzucić. Tymczasem w świetle krytyki wymogu niezależności widać, że całkowita rewizja początkowych sądów rozważnych jest niemożliwa. Nawet jeśli odrzucilibyśmy (zrewidujemy) wszystkie rozważne sądy moralne użyte w dochodzeniu do wąskiej refleksyjnej równowagi, to zostałyby się rozważne sądy moralne (lub ich część, raczej większa niż mniejsza) determinujące teorie tła. Ta druga kategoria rozważnych są-

5. Spory interpretacyjne

Prezentacja metody refleksyjnej równowagi wymaga także, chociaż siłą rzeczy skrótowego, przedstawienia kilku problemów związanych z interpretacją tej metody⁴², zarysowania przynajmniej różnych możliwości jej rozumienia.

Po pierwsze, istnieje spór o to, czy metoda refleksyjnej równowagi ma charakter deskryptywny, czy deliberatywny (preskryptywny). W *Teorii sprawiedliwości* Rawls określił filozofię moralną jako próbę opisaną naszej zdolności moralnej [*moral faculty*]; stwierdził, że istnieje pewna analogia pomiędzy tego rodzaju zdolnością a naszym wycuciem gramatyczności, czy kompetencją językową⁴³. Tak jak w językoznawstwie chodzi o opis poczucia poprawności gramatycznej jednostek, przez który można wyabstrahować jakąś ogólną strukturę języka, tak też zgodnie ze słowami Rawlsa, w filozofii moralnej na podstawie opisu poczucia sprawiedliwości konkretnych (wykształconych) osób można odnaleźć „gotową formę koncepcji moralnej”⁴⁴. Można zatem interpretować teorię sprawiedliwości jako teorię „naszego”⁴⁵ poczucia sprawiedliwości, a skoro metoda refleksyjnej równowagi ma na celu odkrycie i wierny opis przekonań, jakie podzielamy kierując się tym poczuciem⁴⁶, to jak się wydaje, właściwą interpretacją metody refleksyjnej równowagi, jest interpretacja deskryptywna. Konsekwencją przyjęcia takiej interpretacji, przynajmniej zdaniem niektórych jej krytyków, jest stwierdzenie, że rezultaty metody refleksyjnej równowagi nie mają charakteru normatywnego; teoria taka nie mówi nam, co powinniśmy czynić, albo co jest sprawiedliwe, ale jedynie opisuje, co większość z nas sądzi na temat danej powinności czy sprawiedliwości, lub sądziłaby gdyby miała możliwość dokonania pełnej refleksji nad swoim systemem pojęć moralnych⁴⁷. Interpretację deskryptywną zdaje się wzmacniać również późniejsza ewolucja stanowiska Rawlsa, jaka się dokonała po napisaniu *Teorii sprawie-*

dów, chociaż w jakimś sensie „przetrzymanych na później” tak naprawdę również stanowi element zbioru naszych początkowych rozważnych przekonań moralnych.

⁴² Dla uniknięcia ewentualnych nieporozumień co do terminu „interpretacje,” chcę zaznaczyć, że nie chodzi tutaj tylko o interpretacje tekstów Rawlsa, tego, co mógł mieć na myśli pisząc na temat refleksyjnej równowagi, ale również o to, w jaki sposób należy interpretować samą metodę, jak ją rozumieć niezależnie od intencji Rawlsa.

⁴³ Zob. Rawls [1993] s. 74.

⁴⁴ Ibidem.

⁴⁵ Można się jeszcze spierać, czy to „nasze” poczucie ma charakter uniwersalny, czy jest wspólne jedynie w ramach pewnej wspólnoty kulturowej.

⁴⁶ W podobny sposób o celach teorii moralnej Rawls pisze także w: *The Independence of Moral Theory* (zob. Rawls [2001b] s. 288).

⁴⁷ Tak metodę refleksyjnej równowagi postrzega na przykład R.M. Hare (por. Hare [1989], szczególnie s. 86).

dliwości, której ostateczny wyraz dał w *Liberalizmie politycznym*, gdzie stwierdza, że jego celem jest wydobycie naszych najgłębszych przekonań na temat wolności i równości⁴⁸.

Niemniej zgodnie z tradycyjnym rozumieniem teorii moralnej, metodę refleksyjnej równowagi należałoby interpretować jako normatywną, bowiem tylko tak rozumiana ma znaczenie dla formułowania teorii moralnej sensu stricto, tylko tak rozumiana może prowadzić do uzasadnienia i rzeczywistej rewizji poglądów moralnych. Jeśli metoda, którą stosuję, jedynie opisuje moją wrażliwość moralną, to jak na jej podstawie miałbym twierdzić lub zmieniać zdanie w sprawie tego, co powinniśmy czynić? Co ostatecznie Rawls miał na myśli, pozostawiam tutaj jako kwestię otwartą⁴⁹. Nawet jeśli ostatecznie sam skłaniałby się do interpretacji deskryptywnej, zakładam, że metodę refleksyjnej równowagi można także interpretować normatywnie, jako metodę dochodzenia do obiektywnych zasad moralnych. Jako argument za takim wyborem niech wystarczy twierdzenie, że tylko w takiej interpretacji metoda ta pozwala formułować teorie moralne sensu stricto, w innym razie nie mogłaby ona zasługiwać na miano metody teorii etycznej, co najwyżej, jak słusznie zauważył Richard M. Hare, teoria taka mogłaby zasługiwać na miano antropologii moralnej⁵⁰.

Po drugie, można zastanawiać się, czy ambicją stosującego tę metodę jest dojście do prawdziwych przekonań na temat realnych własności świata moralnego, sformułowanie obiektywnych w jakimś szerokim sensie tego słowa zasad racjonalności, czy też jest nią jedynie skonstruowanie rozsądnych z praktycznego punktu widzenia zasad współdziałania grupy. Ronald Dworkin na przykład argumentuje, że bardziej właściwą interpretacją metody refleksyjnej równowagi jest właśnie „model konstrukcyjny,” który Dworkin przeciwstawia obarczonemu jego zdaniem trudnościami „modelowi naturalnemu”⁵¹. Z perspektywy modelu naturalnego, metoda refleksyjnej równowagi jest analogiczna do metod nauk empirycznych (czy też naturalnych). Rozważne sądy moralne odgrywają w niej rolę zdań obserwacyjnych, danych, na bazie których budowana jest teoria, zasady moralne są w tym modelu „odkrywane,” tak jak prawa przyrody są odkrywane

⁴⁸ Por. Rawls [1998] s. 38-41.

⁴⁹ Tim Scanlon w swoim artykule na temat stanowiska Rawlsa w kwestii uzasadniania dokonuje analizy newralgicznych fragmentów pism Rawla i dochodzi do wniosku, że w ich świetle obie interpretacje mogą być uzasadnione, jeśli jednak metodę refleksyjnej równowagi mielibyśmy traktować jako metodę uzasadniania naszych przekonań moralnych, wówczas zdaniem Scanlona, właściwa jest interpretacja deliberatywna (zob. Scanlon [2003] s. 141-144).

⁵⁰ Zob. Hare [1989] s. 86.

⁵¹ Zob. Dworkin [1989], szczególnie s. 27-37. W ciekawy sposób dyskutuje z Dworkinem, broniąc modelu naturalnego, N. Daniels (zob. Daniels [1996] s. 74-77).

w fizyce. Rozważne sądy są wynikiem działania intuicji, która traktowana jest jako pewna władza moralna [moral faculty], zaś sądy moralne mogą być oceniane w kategoriach prawdziwości.

Tymczasem w modelu konstruktywnym, którego wzorcem dla Dworkina jest model praktyki prawniczej, intuicje traktowane są bardziej jako projekcje na temat tego, jak można konstruować zbiór zasad działania. Zasady te są postrzegane nie jako realnie istniejące fakty moralne, które można odkryć, ale jako pewne praktyczne wskazówki działania, szczególnie w kontekście społecznym, które można skonstruować. Tak właśnie Dworkin interpretuje proponowany przez Rawlsa mechanizm sytuacji pierwotnej. Ma on ułatwić ustalanie przez hipotetyczne strony kontraktu akceptowalne dla nich zasady sprawiedliwości. Zdaniem Dworkina model konstruktywny ma przewagę nad modelem naturalnym szczególnie w kontekście praktyki społecznej, po prostu jest bardziej przystosowany do konstruowania zasad sprawiedliwości, które mają powszechnie w ramach danego społeczeństwa obowiązywać, aby mogło ono skutecznie działać i cechować się stabilnością.

Nie rozstrzygając, czy Dworkin ma rację (a wydaje się, że jeśli chodzi o interpretację stanowiska Rawlsa, szczególnie w świetle jego późniejszych publikacji, Dworkin ma mocne argumenty)⁵², w niniejszym tekście przyjmuję model naturalny. Sam Dworkin nie odrzuca możliwości takiego modelu, ale jedynie uważa, że nie nadaje się on do określania zasad politycznej kooperacji. Ponieważ jednak, przynajmniej w tym artykule, przedmiotem naszych rozważań nie jest metoda wieloosobowej refleksyjnej równowagi, tym bardziej w kontekście politycznym, interesuje nas jedynie możliwość stosowania metody refleksyjnej równowagi przez jedną osobę dążącą do uzasadnienia swoich przekonań moralnych (szczególnie w kontekście uprawiania etyki jako dyscypliny filozoficznej), model naturalny wydaje się jak najbardziej właściwy.

Po trzecie, można zastanawiać się, czy metoda ta określa paradygmat uprawiania filozofii w ogóle, czy też stosuje się jedynie do etyki i filozofii politycznej. Tadeusz Szubka analizując charakterystykę metody refleksyjnej równo-

⁵² Uznanie, że Rawls przyjmuje konstruktywistyczny model refleksyjnej równowagi nie oznacza, że odrzuca on całkowicie możliwość mówienia o obiektywnym uzasadnieniu sądów moralnych. Chociaż można twierdzić, że uznając tzw. ciężary sądu, Rawls stał na stanowisku sceptycznym (tj. uznawał niemożność ustalania prawd moralnych), to tacy znawcy i uczniowie Rawlsa jak Samuel Freeman, wyraźnie stwierdzają, że „Rawls wierzył w obiektywność sądów moralnych, oraz [...] w to, że sądy moralne mogą być mniej lub bardziej poprawne, włączając w to ich prawdziwość lub fałszywość” (Freeman [2007] s. 29-30). Na temat relacji pomiędzy konstruktywizmem a obiektywizmem, oraz możliwością orzekania prawdziwości lub fałszywości naszych przekonań moralnych w ramach konstruktywistycznych teorii etycznych zobacz Brink [1989], szczególnie s. 15-16.

wagi w ujęciu Daniela, zauważa, że równie dobrze może być ona potraktowana jako charakterystyka uprawiania filozofii w ogóle, tak właściwie postępujemy, rozważając różne kwestie filozoficzne, nie tylko etyczne⁵³. Wydaje się, że tak właśnie jest, niemniej aby uniknąć konieczności szerokich rozważań metafizycznych, ograniczam się do charakterystyki omawianej tu metody jako metody uprawiania etyki (także filozofii społecznej), w których centralną część stanowią rozważania dotyczące kwestii normatywnych.

7. Dlaczego refleksyjna równowaga?

Powyższy opis metody refleksyjnej równowagi nie wystarcza do uzasadnienia tezy, że gwarantuje ona dochodzenie do racjonalnych, uzasadnionych, czy jeszcze mocniej, prawdziwych przekonań moralnych, że warto stosować i zajmować się tą metodą. Cel ten, mam nadzieję, zostanie bardziej satysfakcjonująco zrealizowany w trakcie dyskusji z zarzutami wobec tej metody, niemniej już na tym etapie można, przynajmniej skrótowo zaprezentować pewną argumentację na rzecz metody refleksyjnej równowagi. Po pierwsze, wydaje się, że zaprezentowany powyżej obraz stanowi wyidealizowaną wersję rzeczywistych praktyk etyków prezentujących swoje teorie etyczne. W naszych praktykach (teza ta być może odnosi się do uprawiania filozofii w ogóle) w dużej mierze bazujemy na naszych wątpliwościach (tym, jak pewne rzeczy nam się jawią) oraz na refleksji polegającej na próbie eliminowania nieścisłości bądź sprzeczności. Po drugie wydaje się, że nie mamy dla metody refleksyjnej równowagi racjonalnej alternatywy⁵⁴. Wszelka racjonalna refleksja obejmuje badanie relacji logicznych i oczywistościowych pomiędzy intuicyjnymi sądami oraz innymi przekonaniami i teoriami, zaś w sytuacji konfliktu lub niekoherencji rewizję lub eliminację pewnych sądów w świetle tego, co jawi się bardziej prawdopodobne⁵⁵. Wszelkie dające się pomyśleć alternatywy to albo całkowite, albo częściowe zarzucenie refleksji skazujące na irracjonalną arbitralność. Tak więc skoro nie mamy żadnej alternatywy dla metody refleksyjnej równowagi, warto jej poświęcać uwagę również po to, aby umożliwić jej doskonalenie, tym samym prowadząc pośrednio do doskonalenia naszych teorii etycznych.

⁵³ Zob. Szubka [2009] s. 227.

⁵⁴ Por. DePaul [1998].

⁵⁵ Por. *ibidem*, s. 301.

8. Zakończenie

Jak zaznaczyłem we wstępie, metoda refleksyjnej równowagi została poddana ostrej krytyce. Najważniejsze zarzuty zasadniczo sprowadzają się do dwóch: a) argumentacji, że metoda refleksyjnej równowagi stanowi typowy przykład błędnego koła, ponieważ albo opiera się jedynie na błędnym koherencjonistycznym założeniu, że sam fakt wzajemnej koherencji przekonań daje asumpt do uznania tych przekonań za prawdziwe lub uzasadnione, albo opiera się na niewiarygodnych intuicjach moralnych; b) dowodzeniu, że mamy do czynienia z faktem tzw. uporczywej niezgody co do kwestii moralnych, oraz że zastosowanie metody refleksyjnej równowagi nie gwarantuje konwergencji poglądów moralnych, nawet wówczas, gdy istnieje zasadnicza zgoda co do określenia początkowych rozważnych sądów, a nawet (przy założeniu, że odmienne poglądy moralne różnych osób będące w stanie szerokiej refleksyjnej równowagi mają charakter uzasadniony) może doprowadzić do przyjęcia relatywizmu moralnego.

Są to zarzuty poważne, których nie należy ignorować. Jeśli są one słuszne, metoda refleksyjnej równowagi nie zasługuje na to, aby stanowić wzór uprawiania etyki, nawet jeśli nie mamy dla niej racjonalnej alternatywy⁵⁶. Sądzę jednak, że można w stosunku do obu tych zarzutów zaproponować przekonującą odpowiedź, co będzie celem kolejnych dwóch artykułów poświęconych zagadnieniu refleksyjnej równowagi. W pierwszym z nich będę próbował dowodzić wstępnej wiarygodności intuicji moralnych leżących u podstaw formułowanych przez nas rozważnych sądów moralnych, dzięki czemu bezpodstawne jest mówienie o błędzie zamkniętego koła. W drugim artykule skupię się na pokazaniu, że fakt niezgody moralnej da się wytłumaczyć w sposób niepodważający wiarygodności intuicji moralnych, jak i wartości metody refleksyjnej równowagi.

Chociaż do czasu przedstawienia polskiemu czytelnikowi satysfakcjonujących odpowiedzi na wspomniane powyżej zarzuty, nawoływanie do zajmowania się metodą refleksyjnej równowagi można by uznać za przedwcześnie, to mam jednak nadzieję, że nawet ten pierwszy z planowanych trzech artykułów na temat metody refleksyjnej równowagi stanowi wystarczającą zachętę, aby omawianemu tutaj zagadnieniu poświęcić w naszych polskich dyskusjach etycznych i metaetycznych więcej uwagi.

⁵⁶ Zawsze można twierdzić, że nie istnieją żadne obiektywne fakty moralne, wobec czego nie istnieje też konieczność konstruowania metody budowania teorii dotyczącej takich faktów.

Bibliografia

- Audi [2007] – Robert Audi, *Intuitionism, Pluralism, and the Foundations of Ethics*, [w:] *Foundations of Ethics. An Anthology*, R. Shafer-Landau (ed.), Blackwell 2007, s. 402-419.
- Brandt [1979] – R.B. Brandt, *A Theory of the Good and the Right*, Oxford University Press, Oxford 1979.
- Brandt [1990] – R.B. Brandt, *The Science of Man and Wide Reflective Equilibrium*, “Ethics” (100) 1990, s. 259-278.
- Brink [1989] – D.O. Brink, *Moral Realism and the Foundations of Ethics*, Cambridge University Press, Cambridge 1989.
- Daniels [1996] – N. Daniels, *Justice and Justification, Reflective Equilibrium in Theory and Practice*, Cambridge University Press, Cambridge 1996.
- DePaul [1987] – M. DePaul, *Two Conceptions of Coherence Methods in Ethics*, “Mind” (96) 1987, s. 463-481.
- DePaul [1993] – M. DePaul, *Balance and Refinement: Beyond Coherence Methods of Moral Inquiry*, Routledge, Florence KY 1993.
- DePaul [1998] – M. DePaul, *Why Bother with Reflective Equilibrium?*, [w:] *Rethinking Intuition*, M. DePaul, W. Ramsey (eds), Rowman & Littlefield Publishers Inc., 1998.
- DePaul [2006] – M.R. DePaul, *Intuitions in Moral Inquiry*, [w:] *Ethical Theory*, David Copp (ed.), Oxford University Press, Oxford 2006, s. 596-623.
- Dworkin [1989] – R. Dworkin, *The Original Position*, [w:] *Reading Rawls. Critical Studies on Rawls' 'A Theory of Justice'*, N. Daniels (ed.), Basic Books, New York 1989, s. 124-140.
- Freeman [2007] – S. Freeman, *Rawls*, Routledge, London 2007.
- Goodman [1955] – N. Goodman, *Fact, Fiction, and Forecast*, Harvard University Press, Cambridge 1955.
- Halsett [1987] – W. Halsett, *What is Wrong with Reflective Equilibria?*, “The Philosophical Quarterly” (37) 1987, s. 305-311.
- Hare [1989] – R.M. Hare, *Rawls's Theory of Justice*, [w:] *Reading Rawls, Critical Studies on Rawls' 'A Theory of Justice'*, N. Daniels (ed.), Basic Books, New York 1989, s. 81-107.
- Huemer [2005] – M. Huemer, *Ethical Intuitionism*, Palgrave Macmillan, 2005.
- Philips [1994] – M. Philips, *Between Universalism and Skepticism: Ethics As Social Artifact*, Oxford University Press, Cary, NC 1994.
- Rawls [1993] – J. Rawls, *Teoria sprawiedliwości*, tłum. M. Panufnik, Wydawnictwo Naukowe PWN, Warszawa 1993.
- Rawls [1998] – J. Rawls, *Liberalizm polityczny*, tłum. A. Romaniuk, PWN, Warszawa 1998.
- Rawls [1999] – J. Rawls, *A Theory of Justice*. Revised Edition, The Belknap Press of Harvard University Press, Cambridge, Mass. 1999.
- Rawls [2001a] – J. Rawls, *The Outline of a Decision Procedure for Ethics*, “The Philosophical Review” (60) 1951, s. 177-197; [przedrukowany w:] J. Rawls, *Collected Papers*, S. Freeman (ed.), Harvard University Press, Cambridge Mass. 2001, s. 1-19.

- Rawls [2001b] – J. Rawls, *The Independence of Moral Theory*, “The Philosophical Review” (60) 1951, s. 177-197; [przedrukowany w:] J. Rawls, *Collected Papers*, S. Freeman (ed.), Harvard University Press, Cambridge Mass. 2001, s. 286-302.
- Scanlon [2003] – T.M. Scanlon, *Rawls on Justification*, [w:] *The Cambridge Companion to Rawls*, S. Freeman (ed.), Cambridge University Press, New York 2003, s. 139-167.
- Singer [1974] – P. Singer, *Sidgwick and Reflective Equilibrium*, “Monist” (58) 1974.
- Singer [2005] – P. Singer, *Ethics and Intuition*, “The Journal of Ethics” (9) 2005, s. 331-352.
- Szutta [2010] – A. Szutta, *Cnota rozsądnosci*, [w:] *Współczesna etyka cnót: możliwości i ograniczenia*, red. Natasza Szutta, Wydawnictwo Naukowe Semper, Warszawa 2010, s. 195-204.
- Szubka [2009] – T. Szubka, *Filozofia analityczna. Koncepcje, metody, ograniczenia*, Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław 2009.