

S t a n i s ł a w J ę d r c z a k

Idea transcendentalizmu według Johna Rawlsa. Problem uzasadnienia w teorii sprawiedliwości

Słowa kluczowe: *argumentacja transcendentalna, filozofia polityczna, sprawiedliwość jako bezstronność, teoria moralna, zasady sprawiedliwości*

1. Zwrot transcendentalny

Aż do połowy minionego stulecia system filozoficzny Immanuela Kanta nie stanowił wyróżnionego punktu odniesienia dla myślicieli zaliczających się do tradycji analitycznej. Wyjaśnienie tego stanu rzeczy można dostrzec we wpływowym poglądzie Bertranda Russella na nowożytną historię idei. Przełomowym wydarzeniem w jego biografii intelektualnej stał się sprzeciw wobec idealizmu dominującego w filozofii brytyjskiej jeszcze na przełomie wieków XIX i XX. Dojrzałe poglądy Russella kształtowały się w opozycji do neohegłowskiego programu Francisa H. Bradleya, Johna M.E. McTaggarta, Thomasa H. Greena.

Krytykując klasyczny idealizm niemiecki, Russell zakładał, że to Kant był właściwym założycielem tej szkoły filozoficznej. Na kantyzm patrzył więc, podobnie jak Hegel, przez zniekształcającą soczewkę pism Johanna Gottlieba Fichtego. Interpretację tę Russell przejął być może od McTaggarta, jednego ze swoich nauczycieli w Cambridge. „Kant, a jeszcze bardziej Fichte” – czytamy na kartach *Dziejów filozofii Zachodu* – „doprowadził zapoczątkowaną przez Kartezjusza tendencję subiektywistyczną do nowych skrajności” (Russell

2000, s. 801). Dając wyraz ambiwalencji, w następujących liniijkach Russell powściągliwie akcentował swoje zdanie: „Immanuel Kant powszechnie uchodzi za największego z filozofów nowożytnych. Nie mogę zgodzić się z tą opinią [...]” (Russell 2000, s. 802).

W myśli angielskojęzycznej zwrot transcendentálny nastąpił już w okresie powojennym. Za jego prekursorów, odpowiednio w kulturze amerykańskiej i brytyjskiej, zasłużenie uchodzą Wilfrid Sellars oraz Peter F. Strawson. To prace tych autorów (Sellars 1956, 1967, 1968; Strawson 1959, 1966) zapoczątkowały w metafizyce analitycznej nurt transcendentálny, który dopiero w ostatnich dwóch dekadach traci powoli na znaczeniu. Tym, co ich różniło, był stosunek do Kantowskiej filozofii umysłu.

Strawson oceniał ją negatywnie, sugerując, by w systemie Kanta teorię szczerze odseparować od metody – argumentację transcendentálną od transcendentálnego idealizmu. Do tego stopnia podzielał więc uprzedzenie Russella. Czym jednak miałyby być owa argumentacja? Mówiąc najkrócej, wzmiankowane rozumowanie prowadzi od opisu struktury świadomości, doświadczenia i języka do sądów, które dotyczą – przykładowo – istnienia trwałych przedmiotów fizycznych lub innych osób w czasie i przestrzeni (zob. np. Gryganiec 2019, s. 137–140). Argument transcendentálny czyni tym samym krok od wnętrza umysłu do świata zewnętrznego, odpowiadając na pytanie, jaka musi okazać się rzeczywistość, która daje się poznać w tak skonceptualizowanej postaci (choćby jako jednorodny system czasoprzestrzenny). Według Strawsona, wśród ogółu kategorii wyróżniają się te, które stanowią fragment każdego schematu pojęciowego, ahistoryczny rdzeń wszelkiego myślenia. O użyteczności argumentacji transcendentálnych przesądza więc okoliczność, że pozwalają one określić uniwersalne granice poznania, a zwłaszcza – sensownych wątpliwości mnożonych przez sceptyków. Teoria Kanta przeciwstawia natomiast to, co się nam jawi (fenomeny), temu, co istnieje naprawdę, niezależnie od postrzegającego podmiotu, czyli rzeczom samym w sobie (noumenom). Opozycję tę – jako prowadzącą na metafizyczne manowce – w mniemaniu Strawsona trzeba z kolei odrzucić.

Nie wdając się w szczegóły, zaznamy, że Sellars – który pozytywnie ustosunkował się do psychologii transcendentálnej – nie dokonywał chirurgicznego rozdzielenia systemu Kanta na teorię i metodę. Dlatego też więcej uwagi poświęcił studiom nad Kantowską filozofią umysłu, piętnując chybione, jak sądził, interpretacje kantyizmu – wśród nich zaproponowaną przez Strawsona. Ta historyczna różnica zdań wciąż wytycza linię podziału między dwoma nurtami metafizyki analityczno-transcendentálnej, których prominentnymi przedstawicielami są między innymi Quassim Cassam i Christopher Peacock (z jednej strony) oraz Robert Brandom i John McDowell (z drugiej).

2. Kantowska interpretacja sprawiedliwości jako bezstronności

Opublikowana po raz pierwszy w roku 1971 *Teoria sprawiedliwości* stanowi monumentalny wykład liberalizmu deontologicznego – koncepcji, która w myśli autora dojrzewała w ciągu poprzednich kilkunastu lat. Od tamtej pory książka Rawlsa, licząca około sześciuset stron, zdążyła stać się drugim spośród najchętniej cytowanych dzieł filozoficznych (jeżeli ograniczyć się do powojennych wydawnictw anglojęzycznych), ustępując jedynie *Strukturze rewolucji naukowych* Thomasa S. Kuhna¹. Na przestrzeni mijającego półwiecza nadawała ton filozofii politycznej, a polemikę z nią podjęli Ronald Dworkin, Herbert L.A. Hart czy Robert Nozick.

Poczynając od 1962 roku, przez następnych 29 lat Rawls wygłaszał w Uniwersytecie Harvarda wykłady z historii filozofii moralnej. W połowie lat 70. w ich centrum znalazł się system Kanta (Rawls 2000, s. XV–XVIII). Wspominając ten okres na krótko przed śmiercią, Rawls oszczędnie scharakteryzował dawną strategię hermeneutyczną. Czytając, starał się mianowicie dociec nie tylko tego, co dany autor naprawdę głosił, ani nawet – czy się w tym nie mylił, ale przede wszystkim, co twierdziłby, gdyby miał rację. „Kantowska interpretacja sprawiedliwości jako bezstronności” nie byłaby przeto, ujmując rzecz precyzyjnie, interpretacją doktryny Kanta, lecz reinterpretacją sprawiedliwości jako bezstronności Rawlsa w świetle kantowskiej metody (Rawls 2013, s. 376–377).

Pod względem oryginalności i doniosłości reinterpretacja ta nie ustępowała pokrewnym przedsięwzięciom Sellarsa i Strawsona. To za pośrednictwem Rawlsa i jego uczniów (wśród których możemy wymienić Paula Guyera, Christine Korsgaard, Onorę O'Neill czy Susan Neiman) wpływ kantyizmu na analityczną filozofię moralną stał się przemożny (Tampio 2007, s. 81; Ameriks 2001, s. 12). Akcentując admirację, jaką darzył dzieło Kanta, Rawls wyznał, że nigdy nie był usatysfakcjonowany przedłożoną przez siebie wykładnią. Lektura Kanta, utrzymywał, wydaje się wręcz niewyczerpalna (Rawls 2000, s. XVII). Całość jego systemu, mówiąc metaforycznie, stanowi niejako granicę możliwych interpretacji w nieskończoności.

Rawls zdawał się akceptować stanowisko Strawsona w odniesieniu do systemu Kanta. Choć wprost nie odwoływał się do tamtej wykładni, przyjmował, że kantyizm, rozumiany wyłącznie jako metoda, należy pogodzić z realistyczną metafizyką empiryzmu, odrzuciwszy uprzednio zwodniczy idealizm (Rawls 1977, s. 165). Z tego powodu zaproponował czysto proceduralne

¹ Zob. raport Leitiera, <https://leiterreports.typepad.com/blog/2009/11/the-most-cited-books-in-postwwii-anglophone-philosophy.html> [09.09.2021].

odczytanie filozofii praktycznej Kanta – a więc związane jedynie z trybem postępowania przy wyborze subiektywnych maksym woli, w tym instytucjonalnych zasad sprawiedliwości. Dotyczy to zarówno kontraktualizmu, jak i konstruktywizmu Rawlsa: alternatywnych strategii uzasadnienia reguł priorytetu, wypracowywanych przez niego w latach 70. i 80. Podczas gdy Strawson starał się wyzyskać metodę transcendentalną na gruncie ontologii, a Sellars – filozofii umysłu, Rawls zastosował ją, by dostarczyć racji na rzecz zasad leżących u podstaw demokracji liberalnej.

Jego postawa wobec idei transcendentalizmu zmieniała się w czasie. Pierwsze wydanie *Teorii sprawiedliwości* od ukazania się jej rekapitulacji w roku 2001 oddzielają trzy dziesięciolecia. W twórczości ich autora okres transcendentalny trwał do roku 1980, kiedy wciąż jeszcze koncepcję sprawiedliwości jako bezstronności Rawls uważał za dyscyplinę teorii moralnej. W miarę postępu rozważań nad moralnym konstruktywizmem, inspirowanym zwłaszcza *Krytyką praktycznego rozumu*, a nie *Uzasadnieniem metafizyki moralności*, jak miało to miejsce przedtem, stopniowo zaczął dostrzegać słabe strony kantowskiej interpretacji. Wszelako dopiero w *Liberalizmie politycznym*, ogłoszonym drukiem w 1993 roku, dysponował pojęciami, które pozwoliły mu zrozumieć, że zawiera ona wewnętrzną niespójność. W książce tej Rawls wyraźnie przeciwstawił kantowski konstruktywizm – jako rozległą doktrynę filozoficzną – własnemu konstruktywizmowi politycznemu. Na stronach *Justice as Fairness: A Restatement*, pracy redagowanej przez Erin Kelly, Rawls swoją koncepcję określił mianem „hobbesowskiej raczej niż kantowskiej” (Rawls 2001, s. 82, przyp. 2), a o filozofie z Królewca wspominał przede wszystkim na marginesie – w kontekście „liberalizmu Kanta i Milla”.

Co spowodowało tę ewolucję poglądów? W kolejnych akapitach tekstu zastanowię się nad przyczynami, dla których Rawls w końcu postanowił zerwać z kantyzmem. Zanim jednak przejdę do najistotniejszej części tej sprawy, omówię pokrótce specyficzność metody Kanta. Wywód mieszczący się w sekcji trzeciej pełni funkcję pomocniczą wobec moich dalszych prób interpretacyjnych.

3. Pojęcie transcendentalizmu

Co właściwie kryje się pod pojęciem tego, co transcendentalne? Ssensu omawianego terminu Kant nie wyłuszczył jednoznacznie. Komentatorzy, starając się o jego rozjaśnienie, sięgają zazwyczaj do „Wstępu” poprzedzającego wydanie B *Krytyki czystego rozumu* (1787). W jednym z akapitów Kant mianem tym „nazywa wszelkie poznanie, które zajmuje się w ogóle nie tyle przedmiotami, ile naszym sposobem poznania przedmiotów, o ile sposób ten ma być *a priori* możliwy” (B 25; Kant 1957, s. 86).

Na użytek dalszych rozważań pożyteczniejsze wydaje mi się jednak odwołanie do rzadziej cytowanego ustępu, który znajduje się w pierwszym poddziale „Transcendentalnej nauki o metodzie”. Nawiązując do wniosku ustalonego w drodze argumentacji transcendentalnej, Kant powiada tam, co następuje: „[c]hoć musi być dowiedziony, nazywa się go raczej zasadą, a nie twierdzeniem [*Grundsatz und nicht Lehrsatz*], ponieważ ma tę szczególną cechę, że sam stanowi podstawę swojego dowodu [...] i musi być w nim uprzednio zawsze założony” (A 737/B 765; za: Kant 1999, s. 642).

Przytoczony passus może czytelnika wprawić w konsternację. Niezależnie od tego powinno się rozumieć go dosłownie. Aby wychwycić tkwiący w nim sens, posłużmy się ilustracją zapożyczoną z „Estetyki transcendentalnej”.

Przedmiotem rozważań Kanta była tam między innymi możliwość geometrii jako nauki, która „określa własności przestrzeni w sposób syntetyczny, a mimo to *a priori*” (A 25/B 40; Kant 1957, s. 100). Czym musi okazać się przestrzeń – brzmiało pytanie transcendentalne – aby tego rodzaju poznanie było możliwe? W odpowiedzi Kant wskazywał, że przestrzeń musi czynić zadość warunkom składającym się na jej metafizyczne pojęcie. Musi być zatem czystą naocznością zewnętrzną, która wobec obiektów napotykanym w doświadczeniu pełni funkcję konstytutywną. Wiadomo, że naukę Kant definiował jako poznanie zarazem aprioryczne i syntetyczne: gdyby przestrzeń stanowiła jedynie pojęcie, a nie naoczność, jej poznanie miałoby charakter analityczny. Gdyby zaś była naocznością empiryczną, a nie czystą – owo poznanie okazałoby się niemożliwe *a priori*. Innymi słowy, tylko treść metafizycznego omówienia pojęcia przestrzeni czyni zrozumiałą (inteligibilną) naukowość geometrii. Tak, zdaniem Kanta, przedłożone wnioskowanie zyskuje apodyktycznie pewność, a dane pojęcie przestrzeni – transcendentalne uzasadnienie.

W ten sposób ustalono więc transcendentalny warunek naukowości geometrii. Warunek ten – na przekór enigmatycznie brzmiącemu określeniu – stanowi po prostu *conditio sine qua non* niesprzeczności opisu twierdzeń geometrycznych jako sądów syntetycznych *a priori*. Przyjmijmy, że geometria euklidesowa oraz przynajmniej niektóre geometrie nieeuklidesowe, służące za model teorii fizycznych, mają status nauki. Filozofia krytyczna – czym różni się od dogmatycznej – nie podaje tego faktu w wątpliwość. Wobec tego możemy przeprowadzić rozumowanie, które w XX-wiecznej filozofii analitycznej pod wpływem Strawsona przyjęło się nazywać argumentacją transcendentalną.

- (α) Koniecznym warunkiem możliwości posiadania przez geometrię statusu nauki jest istnienie przestrzeni będącej czystą formą ludzkiej naoczności.
- (β) Geometria ma status nauki.
- (γ) A zatem przestrzeń stanowi czystą formę ludzkiej naoczności.

Odnotujmy, że kierunek wynikania odpowiada tu schematowi: skoro α i β , to γ . Kierunek rozumowania jest przy tym odmienny: aby skonstatować to, że β , musimy przedtem dowiedzieć się, że zachodzi γ . Wnioskowanie to ma więc charakter redukcyjny (kierunek rozumowania jest tu przeciwny kierunkowi wynikania). Będąc wnioskowaniem entymematycznym, zawiera ono ponadto błędne koło w sensie epistemologicznym, ponieważ znajomość konkluzji (γ) poprzedza w nim znajomość jednej z przesłanek (β). Co więcej, wyłącznym sposobem zapewnienia mu niezawodności jest uprzednie założenie samej konkluzji (γ). Błędne koło okazuje się więc nieusuwalne. Przesłanką entymematyczną jest bowiem istnienie przestrzeni według metafizycznego omówienia jej pojęcia, które w ogóle warunkuje problematyzację naukowości geometrii. Wyrażając to samo inaczej, powiedzielibyśmy, że możliwość geometrii jako nauki (niezależnie od tego, czy jest ona faktem) implikuje pewien sąd egzystencjalny dotyczący przestrzeni fizycznej (por. Palmer 1985, s. 61–67).

Osobliwość tej formy argumentacji – zgodnie z trafną uwagą Strawsona – wyraża się w tym, że już możliwość postawienia problemu empirycznego pociąga za sobą faktyczne istnienie jego rozwiązania (Białek 2017, s. 28). Gdyby więc zdanie oznaczone powyżej literą γ było fałszywe, to zdanie β straciłoby swój sens. Krótko mówiąc, następnik przesłanki transcendentalnej – okres warunkowy markowany jako α – wyraża semantyczną presupozycję jej poprzednika. Wydaje się, że to właśnie miał na myśli Kant, gdy zwracał uwagę, iż zdanie transcendentalne może zostać wyprowadzone jako zasada, ale nie jako twierdzenie.

4. Argument transcendentalny w *Teorii sprawiedliwości*

Dotychczas analizowałem zagadnienie należące do krytyki rozumu teoretycznego. W filozofii praktycznej Kant zainteresowany był nade wszystko możliwością sądów syntetycznych *a priori* orzekających o tym, co być powinno – nie zaś o tym, co jest. Zastanawiał się więc nad możliwością trafnego rozpoznawania przez rozum praktyczny odpowiednich sądów moralnych, czyli maksym, których każdy z nas może zawsze pragnąć bezwarunkowo. Odpowiedzią jest w tym wypadku przypisanie woli autonomiczności w specyficznym sensie, jaki temu terminowi nadał Kant, oraz wyjaśnienie związku owej autonomiczności z prawem moralnym. W *Teorii sprawiedliwości* rekonstrukcji toku Kantowskiego rozumowania, wciąż odsyłającej zwłaszcza do *Uzasadnienia metafizyki moralności*, poświęcono § 40, uchodzący za nader trudny i sporny fragment książki Rawlsa. Rekonstrukcja ta przeciwstawia się wykładni Henry'ego Sidgwicka, która do czasów Rawlsa silnie oddziaływała na pokolenia brytyjskich myślicieli.

Według Kanta, wolność woli polega na braku zdeterminowania jej przez przyczyny zewnętrzne. Przyczynowość – a sąd ten ma charakter analityczny i jest prawdziwy na mocy znaczenia samych pojęć – posiada strukturę nomologiczną. W następstwie tego ogólne i powszechne prawa, którymi kieruje się wola, muszą pochodzić z jej wnętrza. Musi ona dyktować je sama sobie, będąc dla siebie swoistą *causa sui*. Rawls uważał, że autonomia oznacza tym samym postępowanie według takich zasad, które manifestują tożsamość praktyczną każdego z nas jako *osoby*, czyli wolnego i równego innym racjonalnego podmiotu. Być wolnym to działać zgodnie ze swoją naturą. Tożsamość podmiotu transcendentalnego – bez względu na sytuację osobistą, społeczną, ekonomiczną jednostki – znajduje zaś wyraz, gdy postępuje on jedynie według maksym, które nie opierają się na prywatnych przekonaniach lub pragnieniach, lecz na obiektywnym poczuciu obowiązku. Maksymy te odpowiadają imperatywowi kategorycznemu, podstawowej zasadzie skłaniającej czystą, a nie zaledwie empiryczną wolę do moralnego działania. Imperatyw kategoryczny opisuje więc to, co musi czynić człowiek, by w uczynkach uzewnętrznić własne człowieczeństwo.

W tym miejscu można postawić zasadne pytanie, dlaczego to zasady moralne najlepiej wyrażają naszą osobowość. Czy nie mogłaby ona przejawiać się na równi w wyborze spójnego zbioru zasad niemoralnych lub amoralnych? Według Rawlsa, Kant nie rozstrzygnął tej kwestii, ponieważ nie dość detalicznie opracował procedurę selekcji zasad. Jej kompletny wykład miałby się znaleźć dopiero w *Teorii sprawiedliwości*. Opinia ta zdaje się pomijać fakt, że Kant – zdradzając tym światopogląd oświeceniowy – hołdował intelektualizmowi etycznemu oraz melioryzmowi. Czysty rozum praktyczny ze swojej natury kieruje się w jego mniemaniu zawsze tym, co być powinno, a nigdy – tym, co być nie powinno. Stąd także synonimiczność pary pojęć: prawa moralnego oraz zasad, których podmiot transcendentalny mógłby niesprzecnie pragnąć (por. Korsgaard 1995, s. 1169–1172).

Dlaczego jednak Rawls zdawał się nie dostrzegać tej okoliczności? Ponieważ sam – dzieląc z Kantem oświeceniowy światopogląd – wyznawał intelektualizm, nie czyniąc istotnego rozróżnienia między tym, co rozsądne, a tym, czego każdy z nas pragnąłby w okolicznościach sprzyjających podjęciu racjonalnej decyzji. Te jego poglądy w znacznej mierze ukształtowały paradygmat racjonalistyczny we współczesnych naukach społecznych i politycznych.

Rawls zasugerował kongruencję problemów rozważanych przez siebie i – dwieście lat wcześniej – przez Kanta. Przystawanie to opiera się na strukturalnej odpowiedniości pojęcia imperatywu kategorycznego i zasad sprawiedliwości oraz położenia podmiotu transcendentalnego i stron hipotetycznej umowy społecznej (Rawls 2013, s. 369–370). Sparafrazujemy wyartykułowaną już wątpliwość. Skąd bierze się wobec tego przekonanie, że tożsamość osoby – dokonującej tym razem wyboru kolektywnego – znajduje pełen wyraz w selek-

cji zasad sprawiedliwych? Kantowska interpretacja sprawiedliwości jako bezstronności podsuwa w odpowiedzi opis sytuacji początkowej. „Skoro wszyscy są podobnie wolni i racjonalni” – powiada Rawls – „każdy musi mieć równy głos przy przyjmowaniu publicznych zasad etycznej wspólnoty. Znaczy to, że wszyscy, jako ja noumenalne, mają zgodzić się na te zasady. Dopóki nie zgodzono się na zasady łotra, nie mogą być one wyrazem wolnego wyboru, jakkolwiek skłonne byłoby je przyjmując jakieś pojedyncze ja” (Rawls 2013, s. 376). Rawls daje tym do zrozumienia, że Kant nie dostarczył przekonujących powodów, dla których człowiek może przejawiać prawdziwą tożsamość tylko, gdy działa według prawa moralnego. Łotr może niesprzecznie pragnąć, by ogólne maksymy, którymi się kieruje, stały się powszechnym prawem. Jednomyślna zgoda, do której odwołuje się Rawls, ma usuwać ten mankament. Wydaje mu się bowiem niemożliwe, by ogół podmiotów racjonalnych faworyzował zasady godzące w poczucie moralności. O ile tylko strony zawiązujące umowę społeczną nie postępują wbrew własnej naturze, o tyle będą kierowały się tym, co szlachetne.

Członkowie wspólnoty mają swobodę w wyborze reguł. Jeżeli jednak – będąc osobami – pragną dać wyraz swojej naturze, to zdecydują się na maksymy, które zwiększają społeczną wolność, sprawiedliwość i równość szans. To właśnie te maksymy wyrażają zasady sprawiedliwości dla instytucji i jednostek, czyli prawo moralne. „Tak więc pragnienie działania sprawiedliwego, gdy je właściwie rozumieć” – podsumowuje Rawls – „wywodzi się po części z pragnienia jak najpełniejszego wyrażenia tego, kim jesteśmy lub kim możemy być, mianowicie naszej natury wolnych i równych istot racjonalnych mających swobodę wyboru” (Rawls 2013, s. 374).

W ten sposób Rawls ustalił transcendentálny warunek istnienia społeczeństwa dobrze urządzonego *per se*, niezależnie od jakichkolwiek aspektów empirycznych. Jest nim „przedmiotowa ważność” pojęcia sprawiedliwości jako wybieralności w sytuacji początkowej (analogicznie możliwość twierdzeń geometrycznych jako sądów syntetycznych *a priori* implikowała u Kanta istnienie przestrzeni według jej metafizycznego omówienia). Pojęcie to stanowi po prostu *conditio sine qua non* niesprzecznosci opisu zasad instytucjonalnych jako tych maksym, w których dochodzi do głosu tożsamość praktyczna wszystkich obywateli. Przyjmijmy, że demokracja konstytucyjna, przedstawicielska, liberalna stanowią przykłady społeczeństw dobrze urządzonych. Teoria moralna nie podaje tego faktu w wątpliwość. Jest wprost przeciwnie. Zjawisko to stanowi część *explanandum* koncepcji Rawlsa, w tym sensie, że wyjaśniać ma ona, w jaki sposób zmysł sprawiedliwości potrafi rozpoznawać tego rodzaju sądy – zbiór liberalnych zasad – jako adekwatne. Toteż możemy przeprowadzić rozumowanie, które przypomina transcendentálne uzasadnienie geometrii, przytoczone w sekcji trzeciej niniejszego artykułu.

(α') Koniecznym warunkiem możliwości dobrego urządzenia wspólnego życia jest wybieralność rządzących nim zasad w sytuacji początkowej. (β') W rzeczywistości istnieją społeczeństwa dobrze urządzone (tak, jak istnieje zapoczątkowana przez Euklidesa nauka geometrii). Przykładem takiego reżimu jest chociażby nowoczesne państwo demokratyczne. (γ') A zatem pojęcie sprawiedliwości jako bezstronności zyskuje legitymizację (również pojęcie przestrzeni według jej omówienia metafizycznego jest ważne ze względu na faktyczne zastosowanie geometrii euklidesowej bądź nieeuklidesowej w opisie zjawisk fizycznych). W myśl przewidywań Rawlsa rezultatem przeprowadzenia wyboru zgodnie z tą procedurą byłoby wskazanie trzech reguł priorytetu: wolności, sprawiedliwości i równych szans. W ten sposób zyskiwałyby one, pod warunkiem, że owo przewidywanie nie jest błędne, transcendentalne uzasadnienie. To samo dotyczy zresztą dowolnego zbioru zasad wybranych w sytuacji początkowej.

Na czym polega transcendentalność tego argumentu? Zauważmy, że sformułowane najpierw kryterium sprawiedliwości zasad zakłada wyłącznie opis położenia stron pierwotnego kontraktu (podobnie w „Estetyce transcendentalnej” Kant założył jedynie strukturę ludzkiego poznania zmysłowego). W dalszej kolejności przyjęto, że reguły priorytetu właściwe liberalnej demokracji konstytucyjnej są tego rodzaju zasadami. Skoro zaś pryncypia istniejącego naprawdę ustroju spełniają warunek akceptacji za zasłoną niewiedzy, to pojęcie sprawiedliwości jako bezstronności okazuje się przedmiotowo uzasadnione. Mówiąc inaczej, możliwość społeczeństwa dobrze urządzonego pociąga za sobą realność sprawiedliwości według jej proceduralnego omówienia. Zastanówmy się jednak, czy powyższy argument transcendentalny właściwie oddaje intencje Rawlsa. Rozstrzygnięcie tej kwestii ściśle zależy od tego, jakimi wiadomościami dysponować mają strony hipotetycznej umowy społecznej, znalazłszy się za zasłoną niewiedzy.

5. Zwrot naturalistyczny?

W *Liberalizmie politycznym* Rawls zdystansował się od nakreślonego powyżej argumentu transcendentalnego. Okoliczność tę może tłumaczyć fakt, że w treść interpretacji kantowskiej – występującej zarówno pod postacią teorii umowy społecznej, jak i późniejszego konstruktywizmu – wpisany jest nieodłącznie pewien pogląd na naturę ludzką. Przekonanie, że w idealnych warunkach wyboru strony pierwotnego kontraktu pragnęłyby ustanowić akurat racjonalne reguły kooperacji, pociąga za sobą mianowicie tezę intelektualizmu etycznego. W świetle tego poglądu, to rozum dostarcza woli, czyli władzy chcenia, wystarczających pobudek do działania. Toteż pozostaje tylko przedłożyć odpo-

wiednią procedurę, która pozwoliłaby bez skrzepowania manifestować się naturze ludzkiej, by społeczność niejako automatycznie przesądziła o priorytecie wolności, sprawiedliwości i równości szans. Zależność ta pokazuje, że Rawlsowi nie udało się wydestylować z systemu Kanta metody transcendentalnej i konsekwentnie odseparować jej od teorii. W głośnej książce, która zainicjowała spór między liberalizmem a komunitaryzmem, Michael Sandel przedstawił podobną obiekcję, dostrzegając wewnętrzną niespójność deontologicznego i empirystycznego komponentu teorii moralnej Rawlsa (Sandel 1985, s. 13).

Jednak to Rawls wyłożył najistotniejszy powód przesądzający o fiasku interpretacji kantowskiej. Rację tę uświadomił sobie poniewczasie. Wyśłowienie jej okazało się bowiem możliwe dopiero z chwilą, gdy autor *Liberalizmu politycznego* czytelnie odróżnił od siebie pojęcia rozległej doktryny oraz rozumu publicznego, nieobecne *explicite* w jego pismach przed rokiem 1993. Wraz z ich wprowadzeniem, stało się już całkiem jasne, że interpretacja kantowska – niezależnie nawet od wartości logicznej intelektualizmu – nie może dostarczyć adekwatnych podstaw dla zasad sprawiedliwości. A oto krótkie wyjaśnienie tego stanowiska.

Otóż w warunkach społeczeństwa liberalnego panuje pluralizm rozumnych rozległych doktryn. Żadna z nich – jako usytuowana w apolitycznej sferze prywatnej – nie jest wyróżniona przez rozum publiczny. Jeżeli mamy zatem rację, że interpretacja kantowska milcząco zakłada intelektualizm, to rozum publiczny nie może odwołać się do niej w celu uzasadnienia jakichkolwiek pryncypiów. Intelektualizmowi przeciwstawia się przecież woluntaryzm – głoszący, że ostatecznie to wola kieruje myśleniem, a nie rozum działaniem – a te partykularne koncepcje, choć wzajemnie sprzeczne, z punktu widzenia uniwersalnego rozumu pozostają równoważne. Ugruntowanie zasad sprawiedliwości w jednym ze światopoglądów byłoby pozbawione sensu. Światopoglądów jest bowiem wiele, a rozum publiczny – dla którego owe zasady powinny być zrozumiałe same przez się – tylko jeden. Takie uzasadnienie budziłoby więcej kontrowersji niż rzekomo wspierany przez nie wniosek. Jednym słowem, ufundowanie reguł priorytetu w interpretacji kantowskiej stanowi swoisty błąd kategorialny.

Zmiana postawy względem kantowskiej interpretacji rzuca światło na odmienność teorii moralnej i filozofii politycznej. W *Teorii sprawiedliwości* Rawls zgłębiał transcendentalne warunki społeczeństwa dobrze urządzonego, utożsamiając je z wyborem maksym w sytuacji początkowej. Antycypacja rezultatu tej procedury – która sama budzi niebezpieczne wątpliwości – pozwalała mu uzasadnić uniwersalność reguł priorytetu wolności, sprawiedliwości i równości szans. W *Liberalizmie politycznym* fenomen będący przedmiotem wyjaśnienia został istotnie zawężony. Nie jest już nim możliwość wydawania przemyślanych sądów przez wszystkie istoty racjonalne (z wyłączeniem bogów i aniołów, Rawls 2013, s. 376), ale te zasady, które naprawdę zostały zaakcep-

towane przez obywateli liberalnej demokracji konstytucyjnej, ustroju osadzonego w konkretnej sytuacji historycznej, społecznej, ekonomicznej i prawnej. Odchodząc tym samym od transcendentalnej teorii podmiotu, Rawls coraz wyraźniej zaznaczał empiryczny charakter swojego przedsięwzięcia.

Konsekwencją takiego postawienia sprawy jest całkowita rezygnacja z próby uzasadnienia sprawiedliwych maksym. Filozofia polityczna – obierając skromniejszy cel – za wyłączne zadanie ma ich artykulację, czyli ujawnienie treści pryncypiów, które warunkują racjonalność działania obywateli nowoczesnego społeczeństwa demokratycznego. W porządku uzasadnienia zasady liberalizmu, demokracji, pluralizmu bądź egalitaryzmu (by wymienić tylko niektóre) mają pierwszeństwo przed należącymi do kultury podłoża systemami filozoficznymi. To spójność z owymi zasadami służy odąd za kryterium rozumności rozległych doktryn filozoficznych, moralnych lub religijnych. Tymczasem intuicyjne podstawowe idee, którymi podszyta jest polityczna kultura demokratycznego społeczeństwa, stanowią niejako uzasadnienie same dla siebie. Próba ich filozoficznego zakwestionowania zostaje zepchnięta na margines przestrzeni publicznej jako nierozumna.

Domniemany zwrot naturalistyczny w biografii intelektualnej Rawlsa można by porównać do tego, który nastąpił w twórczości Strawsona. Po raz ostatni odwołajmy się do przykładu metafizyki deskryptywnej. Zgodnie z dominującym przeświadczeniem, Strawson pod wpływem krytyki uznał, że filozofia nie opisuje transcendentalnych warunków wszelkiego poznania empirycznego (Stern 2011). Wydobywa co najwyżej na wierzch założenia, które *implicite* czynią kompetentni użytkownicy języków o strukturze podmiotowo-orzecznikowej w rzeczywistej praktyce komunikacyjnej. Ponieważ punktem wyjścia refleksji Strawsona jest pewien fakt empiryczny (np. to, że postrzegamy rzeczywistość jako jednorodny system czasoprzestrzenny), nie może mieć ona prawdziwe transcendentalnego charakteru. Nie odpowiada więc na kantowskie pytanie o własności świata mające wytłumaczyć pewną zdolność mentalną przypisywaną racjonalnym podmiotom jako takim. Z tego powodu – przynajmniej od roku 1985 (por. Callanan 2011) – autor *The Bounds of Sense* własne „argumenty transcendentalne” skłonny był reinterpretować w duchu Davida Hume’a. Wbrew nazbyt wygórowanym oczekiwaniom Kanta, filozofia jest niezdolna orzekać z pieczęcią naukowości o uniwersalnych warunkach konceptualizacji świata zewnętrznego, wyprowadzając z nich rzeczowe twierdzenia ontologiczne. W konsekwencji należy zawęzić ją do odkrywania faktycznych związków pojęciowych łączących poszczególne elementy naszego, a więc lokalnego, aparatu pojęciowego.

Przytoczone dotychczas uwagi wspierają tezę o naturalizacji teorii sprawiedliwości, ku której zwracał się Rawls od momentu, gdy interpretację kantowską porzucił na rzecz filozofii politycznej. Obraz ten komplikuje się, jeśli tylko bliżej przyjrzeć się metodologicznym wyjaśnieniom udzielonym przez

Rawlsa jeszcze w 1971 roku. Wówczas historyczny bieg wydarzeń wyda się daleki od dramatycznych zwrotów akcji. Zasadnicze tezy wysuwane przez filozofa nie zmieniły się na przestrzeni lat. Ich Kantowskie odczytanie przestał on natomiast uważać za ważne.

W § 9 *Teorii sprawiedliwości* Rawls charakteryzuje własne przedsięwzięcie nade wszystko jako teorię empiryczną, której celem jest wyjaśnienie, w jaki sposób zmysł sprawiedliwości może wydawać przemyślane sądy o podstawowej strukturze społecznej (Rawls 2013, s. 89–91). Koncepcję tę przyrównuje do gramatyki generatywno-transformacyjnej Noama Chomsky'ego. Jak wiadomo, tłumacząc umiejętność rozpoznawania zdań poprawnych gramatycznie, Chomsky postulował istnienie skończonego zbioru reguł wytwarzających poprawne syntaktycznie ciągi znaków. Rawls starał się zaś wyselekcjonować zbiór reguł, które pozwalają rozpoznać sprawiedliwe zasady. Jego teoria posiada zbiór empirycznych falsyfikatorów, którymi są nasze przemyślane sądy o podstawowej strukturze społecznej nowoczesnej demokracji. Wy tłumaczenia faktu, że mylimy się niekiedy w ocenie tego, co słuszne i niesłuszne – podobnie jak popełniamy błędy gramatyczne – należy poszukiwać w koncepcji ciężarów sądów. Teoria Rawlsa pod tym kątem nie różni się od innych wyjaśnień empirycznych. W takim rozumowaniu moment transcendentálny (prowadzący do ustalenia najogólniejszych przesłanek – takich, jak α i α') jest w ogóle nieobecny.

Powróć teraz do pytania, które wydaje mi się kluczowe. Kim są strony umowy społecznej? Wbrew domniemaniu Sandela, nie odpowiadają one Kantowskiemu podmiotowi transcendentálnemu. Odpowiedzialnością za interpretacyjne nieporozumienie należy zapewne obarczyć niefortunną terminologię wykorzystaną przez Rawlsa. Jednym z dualizmów, które stanowczo odrzucił on już w *Teorii sprawiedliwości*, było przecież przeciwstawienie prawd apriorycznych i aposteriorycznych. Wbrew mylącej etykietce kolektywnej jaźni noumenalnej, zasłona niewiedzy nie bierze w nawias faktów socjologicznych ani historycznych. Jak zauważył Richard Rorty, „ci, którzy są postawieni w sytuacji pierwotnego wyboru, reprezentują pewien współczesny typ człowieka, a nie jakąś ahistorycznie rozumianą naturę ludzką” (Rorty 1999, s. 269, przyp. 11). Instytucjonalne zasady sprawiedliwości nie wiążą więc wszystkich istot rozumnych (inaczej niż Kantowskie prawo moralne), a jedynie obywateli liberalnej demokracji konstytucyjnej. Wynikałoby stąd, że Rawls nigdy nie starał się o jej „transcendentálną dedukcję”.

Sądzę, że projekt Rawlsa nie jest odległy od wizerunku filozofowania zaproponowanego przez Robina G. Collingwooda, jednego z czołowych reprezentantów idealizmu w brytyjskiej filozofii XX wieku. Przypomnijmy, że Collingwood traktował metafizykę jako odsłanianie absolutnych przesłanek poszczególnych obszarów wiedzy – a szerzej, racjonalnej działalności ludzkiej – w ich historycznej postaci. W przypadku filozofii politycznej oznaczałoby to poszukiwanie zasad empirycznych, a nie transcendentálnych, którymi kierują

się obywatele państwa demokratycznego. Czy sprawiedliwość jako bezstronność rzeczywiście warunkuje racjonalność działania jego obywateli? W miejsce badania rzeczywistości Collingwood podsunął analizę partykularnych i zmieniających się w czasie schematów pojęciowych. XX-wieczny transcendentalizm – próbując oderwać metodę Kanta od kłopotliwych założeń idealizmu – zatacza więc pełne koło. Jedynym przedmiotem filozoficznego poznania okazuje się bowiem nałożona na rzeczywistość siatka pojęciowa. Pragnąc ustrzec się przed idealizmem – dualizmem zjawisk i rzeczy samych w sobie – Rawls (a przed nim również Strawson) przyjął więc stanowisko krańcowo idealistyczne. Z tego punktu widzenia metafizyka nie polega już na badaniu świata zewnętrznego, lecz analizowaniu tego, jak jawi się on w zwierciadle świadomości, doświadczenia i języka.

Czy odkrywanie przesłanek absolutnych demokracji konstytucyjnej – partykularnych, nie uniwersalnych – wciąż wpisuje się w ideę transcendentalizmu? Uchylając się poniekąd od rozstrzygnięcia, wskazałbym jedynie, że odpowiedź musi być twierdząca, jeśli przydawką „transcendentalne” pragniemy opatrzyć projekty pokrewne Rawlowskiemu: metafizykę deskryptywną Strawsona czy też, na gruncie filozofii kontynentalnej, nauki rekonstrukcyjne Karla-Ottona Apla, Wolfganga Kuhlmann, Jürgena Habermasa (por. Białek 2017, s. 71–114). Pod względem formalnym operacja myślowa zaprojektowana w *Krytyce czystego rozumu* przez Kanta, a wykorzystana przez Rawlsa, pozostaje niezmienna. Jej istotą wciąż jest poszukiwanie kryteriów niesprzeczności opisu pewnej zdolności ludzkiego umysłu (presupozycji semantycznych wyróżnionej klasy sądów, np. twierdzeń geometrycznych lub sprawiedliwych maksym). Modyfikacji ulega natomiast wizerunek nauki. Efektem poznania naukowego – wbrew sloganowi Newtona, który powtarzał Kant: „hipotez nie wymyślam” (*hypotheses non fingo*) – nie są apodyktycznie pewne i konieczne sądy syntetyczne *a priori*. Różnica między dedukcją empiryczną a transcendentalną tym samym zaciera się, a wraz z nią rozmyta zostaje również granica między niezależną od umysłu prawdą a uzasadnieniem. Badanie transcendentalne – przynajmniej w postaci, jaką nadał mu Kant – wymaga zatem reinterpretacji. Zamierzeniem Rawlsa w *Teorii sprawiedliwości* było uczynienie kilku pierwszych kroków w tym kierunku.

6. Podsumowanie

W swoim artykule podjąłem próbę naszkicowania postaci, jaką idea transcendentalizmu przybrała w pismach Johna Rawlsa. Moim podstawowym celem było czytelne zrekonstruowanie argumentacji transcendentalnej na rzecz zasad sprawiedliwości. To na niej opiera się bowiem kantowska interpretacja przed-

łożona przez Rawlsa w jego *opus magnum*. Rozstanie z kantyzmem – na które Rawls zdecydował się wraz z przejściem od teorii moralnej do filozofii politycznej – wiązało się z porzuceniem owej interpretacji. Z jej perspektywy efektywność transcendentального uzasadnienia pojęcia sprawiedliwości jako bezstronności miała zależeć bowiem od przystania na rozległą rozumną doktrynę metafizyczną, jaką jest intelektualizm etyczny. „Naturalizowanie Kanta” – jak można by rzec, posługując się wyrażeniem Guyera (Guyer 2001, s. 59–84; Tampio 2007, s. 97, przyp. 21) – zakończyło się fiaskiem, ponieważ metoda Kanta okazała się nierozdzielna od jego teorii. Morał płynący z filozofii politycznej Rawlsa jest więc negatywny. W myśl *Liberalizmu politycznego* interpretacja kantowska jako uzasadnienie reguł priorytetu zostaje skazana na niepowodzenie.

Zestawmy w punktach główne ustalenia tekstu:

- (1) Przedłożona przez Rawlsa interpretacja kantowska jest przykładem oryginalnej i doniosłej recepcji dziedzictwa filozofa z Królewca.
- (2) W okresie transcendentальnym została ona pomyślana jako uzasadnienie procedury normującej postępowanie stron umowy społecznej w sytuacji początkowej.
- (3) Z upływem lat Rawls uprzytomnił sobie, że interpretacja kantowska stanowi rozumną rozległą doktrynę metafizyczną, a zakładając intelektualizm etyczny nie może odpowiednio uzasadnić pojęcia sprawiedliwości jako bezstronności.
- (4) Wyrazem zwrotu naturalistycznego w jego twórczości stała się rezygnacja z próby oparcia zasad sprawiedliwości na swoistym argumencie transcendentальnym; w *Liberalizmie politycznym* Rawls dążył jedynie do ich wyartykułowania, a nie uzasadnienia.
- (5) W konsekwencji jego zamysł zdradza raczej pokrewieństwo z metodą Collingwooda, oddalając się od Kantowskiej idei transcendentalizmu.
- (6) Fiasko Rawlowskiej strategii uzasadnienia było skutkiem niemożności systematycznego rozdzielania metody Kanta od jego doktryny: argumentacji transcendentальной od założonego przez nią *implicite* intelektualizmu etycznego.

Bibliografia

- Allison H. (2016), *Transcendental Deduction and Transcendental Idealism*, „European Journal of Philosophy” 4 (24), s. 920–933.
- Ameriks K. (2001), *Text and Context: Hermeneutical Prolegomena to Interpreting a Kant Text*, w: D. Schönecker, T. Zwenger (red.), *Kant verstehen / Understanding*

- Kant. *Über die Interpretation philosophischer Texte*, Darmstadt: Wissenschaftliche Buchgesellschaft, s. 11–31.
- Białek P. (2017), *Pycha i skromność rozumu. Współczesny spór o argumenty transcendentalne a filozofia Kanta i Fichtego*, Kraków: Universitas.
- Callanan J.J. (2011), *Making Sense of Doubt: Strawson's Anti-Scepticism*, „Theoria” 77 (3), s. 261–278.
- Grygianiec M. (2019), *Status argumentacji transcendentalnej*, „Przegląd Filozoficzny – Nowa Seria” 4 (112), s. 131–160.
- Guyer P. (2001), *Naturalizing Kant*, w: D. Schönecker, T. Zwenger (red.), *Kant verstehen / Understanding Kant. Über die Interpretation philosophischer Texte*, Darmstadt: Wissenschaftliche Buchgesellschaft, s. 59–84.
- Kant I. (1957), *Krytyka czystego rozumu*, przeł. R. Ingarden, t. I, Warszawa: Państwowe Wydawnictwo Naukowe.
- Kant I. (1999), *Critique of Pure Reason*, przeł. i opr. P. Guyer, A.W. Wood, Cambridge: Cambridge University Press.
- Kędziora K. (2019), *John Rawls. Uzasadnienie, sprawiedliwość i rozum publiczny*, Bibliotheca Philosophica 5, Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Korsgaard Ch. (1995), *Rawls and Kant: On the Primacy of the Practical*, „Proceedings of the Eighth International Kant Congress”, t. I, Memphis 1999, s. 1165–1173.
- Palmer H. (1985), *Presupposition and Transcendental Inference*, New York: Routledge.
- Poręba M. (2008), *Możliwość rozumu. Ćwiczenia z metafizyki*, Warszawa: Fundacja Aletheia.
- Rawls J. (1977), *The Basic Structure as Subject*, „American Philosophical Quarterly” 2 (14), s. 159–165.
- Rawls J. (1980), *Kantian Constructivism in Moral Philosophy*, „The Journal of Philosophy” 9 (77), s. 515–572.
- Rawls J. (1985), *Justice as Fairness: Political not Metaphysical*, „Philosophy & Public Affairs” 3 (14), s. 223–251.
- Rawls J. (1989), *Themes in Kant's Moral Philosophy*, w: E. Förster (red.), *Kant's Transcendental Deductions. The Three Critiques and the Opus postumum*, Stanford: Stanford University Press, s. 81–113.
- Rawls J. (1993), *Political Liberalism*, New York: Harvard University Press.
- Rawls J. (2000), *Lectures on the history of moral philosophy*, red. B. Herman, Cambridge, MA – London: Harvard University Press.
- Rawls J. (2001), *Justice as Fairness: A Restatement*, red. E. Kelly, Cambridge, MA: Harvard University Press.
- Rawls J. (2013), *Teoria sprawiedliwości*, przeł. M. Panufnik, J. Pasek, A. Romaniuk, Warszawa: Wydawnictwo Naukowe PWN.
- Rorty R. (1999), *Pierwszeństwo demokracji wobec filozofii*, w: tenże, *Obiektywność, relatywizm i prawda. Pisma filozoficzne*, przeł. J. Margański, t. I, Warszawa: Fundacja Aletheia, s. 261–291.
- Russell B. (2000), *Dzieje filozofii Zachodu i jej związki z rzeczywistością polityczno-społeczną od czasów najdawniejszych do dnia dzisiejszego*, przeł. T. Baszniak, A. Lipszyc, M. Szczubiałka, Warszawa: Fundacja Aletheia.

- Sandel M. (1982), *Liberalism and the Limits of Justice*, Cambridge: Cambridge University Press.
- Sellars W. (1956), *Empiricism and the Philosophy of Mind*, „Minnesota Studies in the Philosophy of Science” I, s. 253–329.
- Sellars W. (1967), *Some Remarks on Kant's Theory of Experience*, „Journal of Philosophy” 64, s. 633–647.
- Sellars W. (1968), *Science and Metaphysics. Variations on Kantian Themes*, London: Routledge and Kegan Paul.
- Strawson P.F. (1959), *Individuals: An Essay in Descriptive Metaphysics*, London: Methuen.
- Strawson P.F. (1966), *The Bounds of Sense*, London: Methuen.
- Tampio N. (2007), *Rawls and the Kantian Ethos*, „Polity” 1 (39), s. 79–102.
- Wolniewicz B. (2017), *Aksjomat Elzenberga*, „Przegląd Filozoficzny – Nowa Seria” 4 (110), s. 277–288.

S t a n i s ł a w J ę d r z a k

The idea of transcendentalism according to John Rawls: The problem of justification in the theory of justice

Keywords: *justice as fairness, moral theory, political theory, principles of justice, transcendental arguments*

This paper aims at presenting a transcendental argument, so termed and constructed by John Rawls, as a justification of his theory of ‘justice as fairness’. The crucial stage in the chain of his reasoning is to establish the necessary condition of the political arrangement of the basic structure of society. This condition turns out to be acceptability of the publicly endorsed principles in the original position. However, the procedure of exercising free choice, as described by Rawls, presupposes a philosophical view of human nature, and consequently undermines the presumably purely theoretical basis for the principles of justice. The author discusses the impact of Kantian moral philosophy on Rawls’s theory of justification. He tries to show that the rejection of moral theory in favour of political philosophy was the result of a profound change in Rawls’s attitude to the idea of transcendentalism, as it is evidenced by his later thought.