

Imperatyw kategoryczny i intuicjonistyczne dziedzictwo Sidgwicka

Artykuł przedstawia interpretację imperatywu kategorycznego podaną przez H. Sidgwicka w *The Methods of Ethics* i wpływ tej interpretacji na kolejne pokolenia filozofów anglosaskich. Interpretacja ta jest Kantowi całkowicie obca, ponieważ w swojej lekturze Kanta Sidgwick korzystał z zespołu idei, które od dawna były obecne w brytyjskiej filozofii moralnej i znacznie różniły się od idei Kantowskich. Autorytet Sidgwicka spowodował, że jego błędne opinie o imperatywie kategorycznym bezkrytycznie powtarzano bez solidnych badań nad tekstami Kanta. Sidgwickowska interpretacja imperatywu kategorycznego żyje własnym życiem w wielu podręcznikach historii etyki i pośrednio kształtuje myślenie moralne części dzisiejszych etyków.

Henry Sidgwick — jeden z najbardziej wpływowych anglosaskich moralistów i etyków — nie poświęcał wiele miejsca analizie filozofii moralnej Kanta — jednego z najbardziej wpływowych „kontynentalnych” etyków i moralistów. Jednym z powodów tego stanu rzeczy jest zapewne fakt, że Sidgwick rozumiał Kanta przez pryzmat doskonale znanej sobie etyki anglosaskiej, i dlatego nie widział w etyce królewieckiego filozofa istotnej nowości teoretycznej. Streszczając etykę Kanta „dla czytelnika angielskiego” Sidgwick posługiwał się poglądami Richarda Price’a i Thomasa Reida¹, zakładając (skądinąd słusznie), że tą drogą lepiej dotrze do swojego odbiorcy. Jednak postępowanie to — uzasadnione względami dydaktycznymi — stało się zarazem podstawą wyrobienia sobie przez „czytelników angielskich” pewnego obrazu etyki Kanta i (jak się wydaje, wbrew intencjom Sidgwicka) przesądziło wiele w ich rozumieniu jej.

O filozofii moralnej Kanta Sidgwick mówi stosunkowo niewiele. Obszerniejsze fragmenty prac Sidgwicka poświęcone filozofii Kanta dotyczą głównie zagadnień ontologicznych². Z pism poświęconych etyce Kanta i zagadnieniom jej pokrewnym

¹ *Outlines of the History of Ethics for English Readers*, (b.m.) 1896, s. 271–272.

² *The So-called Idealism of Kant*, „Mind” 4 (15) 1879; *Kant’s Refutation of Idealism*, „Mind” 5 (17) 1880; *A Criticism of the Critical Philosophy, I*, „Mind” 8 (29) 1883; *A Criticism of the Critical Philosophy, II*, „Mind” 8 (31) 1883; *Kant’s View of Mathematical Premises and Reasonings*, „Mind” 8 (31) 1883; *Kant’s View of Mathematical Premises and Reasonings*, „Mind” 8 (32) 1883; *The Metaphysics of Kant*, [w:] *Lectures on the Philosophy of Kant and Other Philosophical Lectures and Essays*, London, Macmillan 1905.

na uwagę zasługują trzy. Pierwsze to artykuł *The Kantian Conception of Free Will*,³ w którym Sidgwick omawia dwa pojęcia wolności obecne w etyce Kanta. Jest to praca z pogranicza etyki i metafizyki, która, jak się wydaje, w istotny sposób wpłynęła na ogólne wyobrażenie etyków anglosaskich o Kanta filozofii moralności. Drugi fragment to krótkie przedstawienie etyki Kanta w *Outlines of the History of Ethics for English Reader*. Poza tym uwagi Sidgwicka o etyce Kanta to rozproszone w *The Methods of Ethics*, zwykle krytyczne, wzmianki o imperatywie kategorycznym.

Z punktu widzenia recepcji Kanta w etyce w ogóle, a w etyce anglosaskiej w szczególności, artykuł o jego koncepcji wolnej woli ma mniejsze znaczenie. Problem wolności woli szeroko dyskutowano w etyce anglosaskiej w dwudziestym wieku, lecz na ogół dyskusje te nie miały wiele wspólnego z egzegezą i krytyką Kanta. Sprawa ta dość długo żyła własnym życiem. Wpływ Sidgwicka był tu jednak istotny, ponieważ za pośrednictwem G. E. Moore'a i jego następców nadał ton problematyce wolności woli. Podejście to charakteryzuje rezygnacja z badań metafizycznych i nacisk na kwestie pojęciowe oraz relację możliwości alternatywnych do faktycznych wyborów jednostek.⁴

Zdecydowanie większe znaczenie dla recepcji Kanta w etyce anglosaskiej miały dwa pozostałe — objętościowo mniejsze — fragmenty o etyce Kanta. Chociaż *Outlines* poszły w zapomnienie, to wydaje się, że miały duży wpływ na postrzeganie Kanta przez etyków anglosaskich. Książka ta utrzymała — obecny już u J. S. Milla⁵ i co najmniej do piątego wydania *The Methods of Ethics* podzielany przez Sidgwicka (*ME*, xvii)⁶ — obraz Kanta jako filozofa spekulatywnego, którego prace są subtelne, czasem nawet wnikliwie, lecz z powodu swojej spekulatywności nie mają większego znaczenia dla rozwiązania głównych problemów filozofii moralności. Sidgwick dostrzega, że prace Kanta czasem wpływały na myśl etyczną Anglosasów (wspomina Samuela Taylora Coleridge'a i Williama Whewella⁷), lecz — jak sugeruje choćby to, jak niewiele miejsca poświęca etyce Kanta — chyba nie uznaje tego wpływu za szczególnie istotny.

Największe znaczenie dla recepcji Kanta w etyce anglosaskiej wydają się mieć uwagi Sidgwicka o imperatywie kategorycznym zawarte w *The Methods of Ethics*. Chociaż objętościowo najkrótsze, uwagi te — ze względu na uznanie dla *The Methods* w myśli etycznej Anglosasów — wywarły największy wpływ na ich wyobrażenie o etyce Kanta. Fragmenty te dotyczą głównie imperatywu kategorycznego, a obecną tam jego interpretację i krytykę można znaleźć w pracach wielu autorów

³ *The Kantian Conception of Free Will*, „Mind” 13 (51), 1888.

⁴ Zbiór artykułów podejmujących kwestię wolności woli zgodnie z tą tradycją znajduje się np. w książce pt. *Filozofia moralności. Postanowienie i odpowiedzialność moralna*, wybór i wstęp J. Hołówka, przeł. J. Nowotniak, W. J. Popowski, Warszawa 1997, Wydawnictwo Spacja–Fundacja Aletheia.

⁵ J. S. Mill, *Utylitaryzm*, [w:] *Utylitaryzm. O wolności*, przeł. M. Ossowska, A. Kurlandzka, Warszawa 1956, Państwowe Wydawnictwo Naukowe, s. 8.

⁶ Odsyłacze do *The Methods of Ethics*, London 1907, Macmillan, zostały umieszczone w tekście i po skrócie tytułu (*ME*) zawierają numer strony w wydaniu siódmym.

⁷ *Outlines*, dz. cyt., s. 271 i 277–278.

anglojęzycznych czerpiących z teoretycznych zasobów *The Methods*. Dość często krytyka ta jest (czasem bezkrytycznie) powtarzana i uznawana za druzgocącą⁸.

Poniżej zajmę się tymi najkrótszymi, lecz niezwykle ważnymi komentarzami Sidgwicka do etyki Kanta. Nie będę jednak dowodził, że Sidgwicka interpretacja Kanta weszła do większości anglosaskiej literatury z zakresu etyki. Teza taka jest fałszywa na pewno od ukazania się w 1948 r. *The Categorical Imperative* H. J. Patona⁹ i komentarza do *Krytyki praktycznego rozumu* Lewisa White Becka w 1960 r.¹⁰ Prawdziwy przełom dokonał się za sprawą *Teorii sprawiedliwości* Johna Rawlsa po 1971 r.¹¹. Wprowadzając elementy kantowskiego myślenia moralnego, Rawls wykorzystuje prace Patona i Becka i zrywa z interpretacją utrwaloną przez Sidgwicka i z protekcyjnym stosunkiem wielu Anglosasów do filozofa z Królewca. Rawls zainicjował w filozofii anglosaskiej badania nad Kantem na szerszą skalę, które obecnie kontynuują autorzy tacy jak np. Karl Ameriks, Paul Guyer, Barbara Herman, Thomas E. Hill Jr., Christine M. Korsgaard, Onora O'Neill.

W niniejszym artykule zrekonstruuję interpretację imperatywu kategorycznego podaną przez Sidgwicka w *The Methods of Ethics* i pokrótce przedstawię wpływ tej interpretacji na kolejne pokolenia filozofów anglosaskich. Pokażę, że interpretacja Sidgwicka jest całkowicie Kantowi obca. W swojej lekturze Kanta Sidgwick — pomimo znajomości języka i literatury niemieckiej — korzystał z zespołu idei, które były obecne w brytyjskiej filozofii moralnej na długo przed jego wstąpieniem na arenę filozoficzną. Korzystanie z tych idei, a nie z idei Kanta, ukształtowało w specyficzny sposób interpretację i krytykę etyki królewieckiego filozofa podaną przez tego niezwykle rzetelnego badacza idei moralnych. Autorytet Sidgwicka spowodował, że jego opinie bezkrytycznie powtarzano bez solidnych badań nad tekstami Kanta. Wykrycie źródeł Sidgwickowskiej interpretacji imperatywu kategorycznego jest ważne nie tylko ze względów historyczno-filozoficznych, ale i dlatego, że interpretacja ta żyje własnym życiem w wielu podręcznikach historii etyki i pośrednio kształtuje myślenie moralne części dzisiejszych etyków.

I. Intuicja i podstawy moralności w brytyjskiej etyce XVIII i XIX w.

Dla trafnego zrozumienia komentarzy Sidgwicka o imperatywie kategorycznym niezwykle istotne znaczenie ma znajomość inspiracji i źródeł terminologicznych

⁸ Tak o Sidgwicku pisze np. B. Hooker (*Sidgwick and Common-sense Morality*, „Utilitas” Vol. 12, No 3, November 2000): „Sidgwicka ujęcie sądu moralnego (s. 23–38) jest lepsze od właściwie wszystkiego, co pojawiło się od tamtej pory. [...] Sidgwicka ujęcie relacji między szczęściem a powinnością (s. 162 — 175) nadal pozostaje najlepsze w historii. Jego krytyka Kanta (s. 389 n.) wydaje się idealnie trafiona (*exactly right*) i często ją powtarzano”.

⁹ *The Categorical Imperative. A Study in Kant's Moral Philosophy*, Chicago 1948, The University of Chicago Press.

¹⁰ L. White Beck, *A Commentary on Kant's Critique of Practical Reason*, Chicago & London 1960, The University of Chicago Press.

¹¹ *Teoria sprawiedliwości*, przeł. M. Panufnik, J. Pasek, A. Romaniuk, Warszawa 1994, Wydawnictwo Naukowe PWN.

decyzji autora *The Methods of Ethics*. Decyzje te, po pierwsze, ujawniają drogi, jakimi Sidgwick doszedł do swoich ostatecznych poglądów filozoficznych i, po drugie, sens, jaki nadawał wypowiedziom dyskutowanych filozofów oraz kierunki podawanych krytyk. Z tej perspektywy najważniejsze wydaje się stosowanie przez Sidgwicka dwóch pojęć, które pojawiają się w kontekście krytyk imperatywu kategorycznego: intuicji i rozumu w kontekstach etycznych. Pojęciom tym towarzyszą dwa inne. Jedno — w oczywisty sposób wywodzące się od Kanta, lecz, jak pokażę poniżej, nie-Kantowskie — to pojęcie imperatywu kategorycznego, a drugie to — w oczywisty sposób obce Kantowi — pojęcie aksjomatu moralnego.

Dwa pierwsze pojęcia Sidgwick przejął między innymi od Williama Whewella (1794–1866), który przewodniczył Trinity College, Cambridge, w czasie, gdy Sidgwick był tam studentem od października 1855 r. Chociaż Sidgwick nie studiował wówczas filozofii, do obowiązujących go lektur należała jedna z prac Whewella poświęconych etyce¹². W swoich dwóch książkach z zakresu filozofii moralności (*The Elements of Morality, Including Polity*), których pierwsze wydanie ukazało się w 1845 r., i *Lectures on Systematic Morality* z 1846) Whewell zajmuje stanowisko racjonalistyczne, zgodnie z którym podstawowe normy moralne można ustalać jak aksjomaty w geometrii¹³. Chociaż później porzuca tę analogię jako mylącą, to jednak nadal argumentuje tak, aby wykazać, że pierwsze zasady etyki są naocznie oczywiste¹⁴.

Wedle Whewella przedmiotem rozpoznawania rozumowego są prawidła (*rules*) i zasady moralne. Pierwsze z nich dotyczą działania, natomiast zasady nadają prawidłom status prawdziwości. Zasady moralne są zatem prawidłami wyższego rzędu i mają się do prawideł moralnych jak aksjomaty geometrii do dowodów geometrycznych¹⁵. Owe aksjomaty moralności zostają — jak mówi Whewell w czwartym i ostatnim wydaniu *The Elements* z 1864 r. — „ustalone w naszych umysłach przez samą kontemplację naszej natury i kondycji moralnej; czy też, co wyraża to samo, przez wgląd (*intuition*)”¹⁶.

Konsekwentne korzystanie z analogii między poznaniem geometrycznym a moralnym, które kulminuje w pojęciu intuicji rozumianej jako rozumowe rozpoznanie zasad moralnych, jasno pokazuje, że dla Whewella intuicja nie była spontanicznym i nie opartym na racjonalnej refleksji oglądem, jak to miało miejsce np. w teoriach zmysłu moralnego Shaftesbury’ego i Hutchesona, lecz czynnością racjonalną, której rezultatem jest odkrycie koniecznych prawd moralnych¹⁷. Na podstawie analiz zależności pojęciowych Whewell twierdzi, że rozróżniając rozmaite władze (*faculties*) umysłu nazywamy różne czynności czy rodzaje czynności, a nie

¹² J. B. Schneewind, *Sidgwick’s Ethics and Victorian Moral Philosophy*, Oxford 1997, Clarendon Press, s. 41.

¹³ *The Elements of Morality, Including Polity, volume one*, New York 1861, Harper & Brothers, s. vii.

¹⁴ Schneewind, dz. cyt. s. 106–107.

¹⁵ *The Elements*, dz. cyt. s. 191.

¹⁶ *The Elements*, wydanie czwarte z 1864 roku, s. 11

¹⁷ *The Elements*, dz. cyt., s. 76–77.

oddzielne władze. Ponieważ odkrywanie prawd moralnych jest czynnością umysłu, który w poznawaniu zwie się rozumem¹⁸, a prawidła moralne są przedmiotem świadomego rozpoznawania, Whewell utożsamia świadomość (*consciousness*) prawideł z sumieniem (*conscience*)¹⁹, wyraźnie wykorzystując wspólny źródłosłów *consciousness* i *conscience* i twierdząc, że sumienie polega na znajomości samego siebie (*self-knowledge*)²⁰. Dla Whewella intuicja to czynność sumienia, które jest władzą rozpoznawania podstaw moralności.

Whewell nie był pierwszym autorem tych związków pojęciowych. Podobnie jak myślenie w kategoriach analogii geometrycznych w filozofii moralności nie było obce brytyjskiemu empiryzmowi co najmniej od czasów J. Locke'a²¹, który — nie podając, co prawda, stosownych racji i dowodów — sądził, że prawa moralne, nieśmiertelność duszy i istnienie Boga można udowodnić jak twierdzenia geometryczne. Z analogii między „rozumowaniem naturalnym” (tj. poznawaniem) a moralnym korzystał młodszy od Locke'a Samuel Clarke (1675–1729), który twierdził, że podstawową zasadę moralną odkrywamy z taką samą koniecznością jak prawdy geometryczne. Zasadę tę Clarke nazywał zasadą prawości (*righteousness*). W odniesieniu do stosunków między ludźmi przyjmuje ona postać zasady uczciwości (*equity*) i wymaga „abyśmy tak postępowali z każdym człowiekiem, jak w podobnych okolicznościach moglibyśmy oczekiwać, że on będzie postępował wobec nas”²². Konieczność tej zasady ujawnia się dzięki analizie zawartości niezepsutego umysłu, który zawiera sumienie dyktujące człowiekowi jego powinności z nieodpartą oczywistością²³.

To utożsamienie rozpoznawania podstawowych zasad moralnych z sumieniem i rozumieniem tych zasad jako przedmiotu racjonalnego oglądu, czasem zwanego intuicją, pojawia się później u wpływowego poety i filozofa brytyjskiego — Samuela Taylora Coleridge'a (1772–1834). Jest on autorem tym ważniejszym, że w czasie podróży do Niemiec jako jeden z pierwszych Brytyjczyków zapoznał się z niemiecką filozofią moralną, a zwłaszcza z etyką Kanta. W wyniku późniejszych studiów pewne elementy niemieckiego myślenia filozoficznego przeniósł na grunt brytyjski. Idee zaczerpnięte z filozofii niemieckiej Coleridge ujmuje w sposób, który można później rozpoznać w brytyjskiej filozofii moralności i w teorii Sidgwicka.

W *Essay on Faith*²⁴ Coleridge omawia najwyższą zasadę moralności, której, jak twierdzi, jesteśmy bezpośrednio świadomi, tak jak jesteśmy świadomi faktów. Jest to świadomość nie konkretnej jednostki, ale świadomość wspólna wszystkim

¹⁸ Tamże, s. 30–1

¹⁹ Tamże, s. 259–260.

²⁰ Tamże, dz. cyt. s. 260.

²¹ J. Locke, *Rozważania dotyczące rozumu ludzkiego*, przeł. B. Gawecki, Warszawa 1955, PWN, tom 2, s. 231 i 374.

²² Samuel Clarke, *A Discourse of Natural Religion* z 1704/28 r. przedrukowany [w:] *British Moralists 1650–1800*, ed. D. D. Raphael, Oxford 1969, The Clarendon Press, tom I, s. 207, por. też s. 200.

²³ Tamże, s. 202–203.

²⁴ *Essay on Faith*, [w:] *The Complete Works of Samuel Taylor Coleridge*, ed. W. G. T. Shedd, New York 1853, Harper, tom V.

ludziom jako ludziom. Świadomość (*consciousness*) tę Coleridge nazywa sumieniem (*conscience*), co można rozumieć jako rodzaj wiedzy wspólnej wszystkim ludziom. Nietrudno dostrzec, że Coleridge wyraźnie odwołuje się, z jednej strony, do etymologii słów *conscience* i *consciousness*, a także do wpływowego w tamtym czasie pojęcia *common sense* stanowiącego centrum filozofii Thomasa Reida. Świadomość tę Coleridge widzi jako wspólną wszystkim ludziom wiedzę, której oczywistość nie musi być bezpośrednio dostępna dla każdego, lecz którą — jak mówi w innym miejscu²⁵ — zna się intuicyjnie w doświadczeniu powinności. Świadomość powinności Coleridge charakteryzuje jako wiedzę samodzielną i źródłową, a nie pochodną i wtórną. Jest to wiedza rozumowa, a rozum i jego przedmioty, pisze Coleridge w *Essay on Faith*, „nie są rzeczami refleksji, skojarzenia, rozmowy, dyskursu w dawnym sensie tego słowa, w którym przeciwstawia się go intuicji”. Coleridge utożsamia rozum z sumieniem, a sumienie ze znajomością podstawowych prawideł, które mają się do konkretnych rozstrzygnięć moralnych jak aksjomaty do twierdzeń pochodnych.

Jest rzeczą szczególnie interesującą, że treścią opisywanego w *Essay on Faith* sumienia czy intuicji, która jest zdaniem Coleridge’a źródłem znajomości podstawowego aksjomatu moralności, jest zasada głosząca, że należy czynić dla innych to, co chcieliby się od nich otrzymać. Zasadę tę Coleridge nazywa *imperatywem kategorycznym*. Objaśniając ją, Coleridge pisze, że wymaga ona, aby maksymy można było bez sprzeczności chcieć jako praw dla wszystkich moralnych istot rozumnych. Wpływ studiów Coleridge’a nad filozofią niemiecką, a zwłaszcza nad etyką Kanta, jest tu jasny. Widoczne jest jednak i to, że wspomniane przez Coleridge’a dwie zasady to dwie różne rzeczy. Pierwsza to złota reguła, która wymaga od jednostki postępowania wobec innych tak, jakby chciała, aby oni postępowali wobec niej; druga — przypomina imperatyw kategoryczny, który wymaga postępowania na podstawie zasad, których można bez sprzeczności chcieć jako praw powszechnych.

Zastanawia, że mimo oczywistych różnic sformułowań i niewątpliwych talentów filozoficznych Coleridge’a zasady te uznał on za identyczne. Zdziwienie jednak znika z chwilą wzięcia pod uwagę tradycji filozoficznej, w której kształcił się Coleridge w Cambridge. W rozmaitych sformułowaniach złota reguła funkcjonowała w brytyjskiej literaturze etycznej od dłuższego czasu, w odpowiedzi — jak się wydaje — na sceptycyzm Hume’a. Często przedstawiano ją jako oczywistą regułę dojrzałego rozumu i sumienia, porównywano do aksjomatu, którego nie można dowieść demonstratywnie, który jednak uobecnia się w intuicji czy oglądzie. Tak w 1788 r. postępował Thomas Reid (1710 — 1796), który podając przykłady pierwszych zasad moralności, przytacza zasadę zakazującą nam czynić innym to, co osądziłibyśmy jako niesłuszne, gdyby zostało nam uczynione w podobnych okolicznościach. Jest to jedna z oczywistych samych przez się (*self-evident*) zasad,

²⁵ W *Philosophical Lectures*, 1818–1819, ed. K. Coburn, London 1949 — odsyłacz podają za Schneewind, dz. cyt. s. 93.

którą po chłodnym namyśle (*in his cool moments*) akceptuje każdy, kto dysponuje pełnią zdrowej władzy moralnej zwanej sumieniem²⁶.

W etyce brytyjskiej z czasów przed wkroczeniem na scenę Sidgwicka — tj. w etyce, której uczył się w Cambridge w latach pięćdziesiątych XIX w. — występuje zespół wzajemnie powiązanych idei i tez filozoficznych, którego obecność i ewolucję można prześledzić u wpływowych filozofów brytyjskich — co najmniej od Reida, przez Coleridge'a po Whewella. Na zespół ten składały się m.in. następujące tezy: 1) sumienie jest tożsame z rozumem (lub co najmniej z częścią rozumu); 2) sumienie (*resp.* rozum) dostarcza pierwszych zasad moralnych, które ani nie są dane w intelektualnym oglądzie, ani za pośrednictwem dowodów demonstratywnych (nawet jeśli ich ustalenie wymaga refleksji czy zastanowienia), lecz ujawniają się w wyniku zastanowienia jako oczywiste; 3) w namyśle moralnym zasady te odgrywają rolę analogiczną do aksjomatów w matematyce w tym sensie, że są podstawą pozostałych tez moralnych; 4) ogół czynności wiodących do uobecnienia się pierwszych zasad moralnych to intuicja; 5) jedną z zasad uobecnianych w intuicji jest złota reguła lub jej odpowiedniki.

Biorąc pod uwagę dość częste występowanie omówionego zespołu idei oraz to, że należał on do instrumentarium pojęciowego przekazywanego Sidgwickowi w czasie edukacji filozoficznej w Cambridge, można postawić hipotezę, iż zespół ten odegrał istotną rolę w kształtowaniu się myślenia filozoficznego Sidgwicka. Prawdopodobne jest, że stanowiska czytanych autorów postrzegał przez pryzmat właśnie tego zespołu idei. Sam Sidgwick był świadom obecności tych idei i ich luźnego związku z etyką Kanta. Clarke'a metodę badawczą określał jako skrajną formę intuicjonizmu (*ME*, 384), a zasadę uczciwości utożsamiał ze złotą regułą (*ME*, 384 — 385). Swoją początkową niechęć do intuicjonizmu Sidgwick tłumaczył własnym brakiem krytycyzmu dla krytyki Whewella przedstawionej przez J. S. Milla (*ME*, xvi), a docenienie zalet tej metody etyki zawdzięczał wyzwoleniu się spod wpływu Milla i ponownej lekturze Kanta (*ME*, xvii).

Wydaje się, że intuicjonistyczna tradycja brytyjska przesłoniła Sidgwickowi doniosłe różnice zarówno między ich metodą a metodą Kanta jak i między ich tezami teoretycznymi a tezami Kanta. Sidgwick przyznawał, że na grunt brytyjski etykę Kanta przeniósł Coleridge, a potem pewne jej elementy podchwycił Whewell²⁷. Pozostając pod wpływem opisanego zespołu idei, Sidgwick czytał Kanta tak, jakby królewiecki filozof był jednym z moralistów brytyjskich, przez co imperatyw kategoryczny to dla Sidgwicka wersja złotej reguły bądź zbliżona do niej zasada. Argumenty na rzecz tej ostatniej tezy znajdują się w następnej części niniejszego artykułu.

²⁶ T. Reid, *Essays on the Active Powers of Man*, (III.iii.6), [w:] *The Works of Thomas Reid, D. D. Now Fully Collected, with Selections from His Unpublished Letters*, Edinburgh 1872, MacLachlan and Stewart, s. 590–592; por. też *Rozważania o władzach poznawczych człowieka*, przeł. oraz wstępem i przyp. opatrzył M. Hempoliński, s. 691–693.

²⁷ *Outlines*, dz. cyt., s. 271; por. też *Lectures on the Philosophy of Kant*, dz. cyt., gdzie stanowisko to zostaje powtórzone na s. 267.

II. Imperatyw categoryczny według Sidgwicka

Powyższa charakterystyka zespołu idei funkcjonujących w brytyjskiej filozofii moralnej pozwala bez trudu rozpoznać inspiracje i pojęcie metody, którą Sidgwick nazywa intuicyjną. Oprócz metod egoizmu i utylitaryzmu w *The Methods of Ethics* Sidgwick wyróżnia metodę intuicyjną, która w stadium filozoficznym (w rozdziale XIII księgi III) pozwala wyróżnić trzy naczelne abstrakcyjne zasady moralne. Zasady te nie zawsze są jawne dla potocznego myślenia moralnego (*the morality of common sense*), lecz mają walor oczywistości. Metoda filozoficznego intuicjonizmu polega na badaniu moralności potocznej w celu wykrycia zasad, które moralność ta zakłada i za pomocą których można ją uporządkować i usystematyzować (*ME*, 373 — 374).

Sidgwick znajduje trzy zasady metody intuicyjnej: zasadę sprawiedliwości, zasadę roztropności (*prudence*) i zasadę racjonalnej dobroczynności (*benevolence*). Te aksjomaty moralności nie dają się wzajemnie do siebie zredukować, o czym świadczy sprzeciw Sidgwicka wobec prób wyprowadzania zasady dobroczynności z zasady sprawiedliwości, w krótkiej uwadze (*ME*, 386) i w dołączonej do rozdziału XIII księgi III obszerniejszej nocie poświęconej etyce Kanta (*ME*, 389–40). Aksjomaty te charakteryzują się też tym, że „są zbyt abstrakcyjnej natury i zbyt ogólny jest ich zakres, by za pomocą bezpośredniego zastosowania ich umożliwić nam rozpoznanie, co powinniśmy czynić w konkretnym przypadku; konkretne powinności i tak trzeba określić jakąś inną metodą” (*ME*, 379). Zasady te mają charakter zasad logicznych, a nie zawierających treść moralną pouczeń, z których można by wydedukować przepisy działania.

Właśnie w tym kontekście pojawia się jedna z nielicznych i zdawkowych krytyk imperatywu categorycznego, który — wbrew wyraźnemu ostrzeżeniu Kanta, aby zasady tej (w postaci formuły celu samego w sobie) nie mylić ze złotą regułą (*Gr*, 430p; 63)²⁸ — Sidgwick utożsamia z zasadą sprawiedliwości czy prawości (*ME*, 380 — 381), a tę ze złotą regułą. Sidgwick wyznaje, że fundamentalna zasada etyki Kanta „nadała »złotej regule« z ewangelii (»Czyń innym tak, jak byś chciał, aby inni czynili tobie«) postać, która przemawiała do mego rozumu” (*ME*, xvii). Zasada ta głosi, zdaniem Sidgwicka, że przypadki podobne należy traktować podobnie. W najściślejszej wersji zasada ta brzmi jego zdaniem, następująco: „nie może być rzeczą słuszną dla A traktować B w sposób, w który byłoby rzeczą niesłuszną dla B traktować A jedynie na tej podstawie, że są oni dwiema różnymi jednostkami i pod nieobecność jakiegokolwiek różnicy między ich naturą bądź okolicznościami,

²⁸ Odsyłacze do tekstów Kanta znajdują się w tekście. Zawierają skrót tytułu, po którym następuje numer strony w *Kants gesammelte Schriften: herausgegeben von der Deutschen Akademie der Wissenschaften*, Berlin 1902, Walter de Gruyter — (AA), a po średniku numer strony w wydaniu polskim. Skróty: *Gr.* — *Grundlegung zur Metaphysik der Sitten* (AA, t. IV; *Uzasadnienie metafizyki moralności*, przeł. M. Wartenberg, Warszawa 1984, Państwowe Wydawnictwo Naukowe), *KpV* — *Kritik der praktischen Vernunft* (AA, t. V; *Krytyka praktycznego rozumu*, przeł. J. Galecki, Warszawa 1984, Państwowe Wydawnictwo Naukowe).

w których się znajdują, a którą można by przedstawić jako rozsądny powód różnego traktowania” (ME, 380).

Sidgwick, często wychwalany przez pokolenia filozofów moralności za precyzję i troskę o jasność, utożsamia zatem trzy różne zasady: zasadę podobnego traktowania przypadków podobnych ze złotą regułą, a te z imperatywem kategorycznym Kanta. Jest to obserwacja niezwykle doniosła, ponieważ pozwala zrozumieć i ocenić jego krytykę imperatywu kategorycznego oraz wartość argumentów filozofów idących w jego ślady. Bezzasadność krytyk opartych na tej interpretacji stanie się ewidentna z chwilą zauważenia, że zasady, z którymi Sidgwick utożsamia imperatyw kategoryczny, nie są identyczne z tym, co terminem tym określał Kant.

W dwóch najważniejszych miejscach, w których Sidgwick krytykuje imperatyw kategoryczny, pojęcie to uzyskuje inną interpretację. Na stronach 209–210 *The Methods of Ethics* przedstawia imperatyw kategoryczny jako zasadę podobnego traktowania przypadków podobnych, nadaje jej jednak interpretację, w której oznacza ona konsekwencję w ocenianiu, czy gotowość do obejmowania swymi zasadami wszystkich ludzi. Imperatyw kategoryczny nie może, zdaniem Sidgwicka, dostarczyć żadnych godnych uwagi zasad moralności, ponieważ nie zawiera istotnej treści moralnej, a przez to nie wyklucza sytuacji, w której różne osoby szczerze przywiązane do odmiennych zasad będą chciały, aby te zasady stały się powszechne. Poprawny sprawdzian moralny powinien dawać identyczne lub co najmniej zbieżne zasady dla różnych osób. Sidgwick diagnozuje tę sytuację w sposób podobny do Hegłowskiego oskarżenia o pusty formalizm²⁹ (jest to „zarzut abstrakcyjności”). Korzystanie z zasady konsekwencji w moralności porównuje do posługiwania się logiką formalną tak, jakby miała dostarczać, jak mówi Sidgwick, „zupełnego kryterium prawdy.”

Natomiast w nocie na stronach 398–340 *The Methods of Ethics* Sidgwick krytykuje Kantowski imperatyw nie za to, że zasada Kanta jest pusta, lecz, przeciwnie, za to, że zawiera zbyt wiele. Sidgwick twierdzi, że imperatyw kategoryczny opiera się na założeniach o człowieku, które nie są prawdziwe uniwersalnie, wobec czego znowu imperatyw kategoryczny nie jest w stanie dostarczyć podzielanych zasad moralnych. Sidgwick twierdzi, że, po pierwsze, z imperatywu kategorycznego

²⁹ Najobszerniej sformułowane zarzuty o formalizm można znaleźć we wczesnej pracy Hegla (z 1802 roku) pt. *Über die wissenschaftlichen Behandlungsarten des Naturrechts, seine Stelle in der praktischen Philosophie und sein Verhältnis zu den positiven Rechtswissenschaften*, [w:] *Hegel: Werke*, Frankfurt 1970, Surkamp, tom 2, s. 459–466. Hegel twierdzi, że uogólnienie maksymy każe pomijając w namyśle to, co szczegółowe i dla danej sytuacji niepowtarzalne, a wobec tego pozwala jedynie posługiwać się ogólnymi kategoriami moralnymi, które zakładają powinność. Imperatyw kategoryczny nie generuje zatem powinności szczegółowych, lecz zakłada z góry przyjęte powinności ogólne, które nie wystarczają do kierowania działaniem w konkretnych sytuacjach. Celem tych zarzutów Hegla jest wykazanie, że nie można sformułować konkretnych pouczeń moralnych bez odwołania się do świadomości moralnej członków określonego społeczeństwa o danej strukturze. Skróconą i nieco zmienioną wersję tych zarzutów zawierają *Zasady filozofii prawa*, przeł. A. Landman, Warszawa 1969, Państwowe Wydawnictwo Naukowe, § 135, s. 138–9.

nie da się wyprowadzić zasady pomocy bliźniemu, ponieważ osoby szczególnie silne, zaradne i wierzące w swoje siły mogą odrzucić zasadę udzielania pomocy potrzebującym uznając, że w dłuższym terminie lepiej jest dla nich nie nakładać na siebie ciężarów dobroczynności, nawet ryzykując, że zostaną pozbawione pomocy, gdy ich rachuby okażą się błędne. Kant natomiast zakłada, zdaniem Sidgwicka, identyczność pragnień u wszystkich ludzi niezależnie od ich położenia życiowego i oceny własnych szans (odtąd zarzut ten będę nazywał: „zarzutem idealizacji antropologicznej”).

Po drugie, formuła imperatywu kategorycznego w wersji, która wymaga traktowania wszystkich ludzi jako cele same w sobie, jest zagadkowa, gdyż słowa cel Kant używa w osobliwy sposób. Argumentacja Kanta polega według Sidgwicka na obserwacji, że wszyscy ludzie uważają się za cele same w sobie, wobec czego będą chcieli, aby wszyscy uznawali zasadę traktowania ich jako takie cele. Zdaniem Sidgwicka cele to rzeczy, które się realizuje, a nie, jak chce tego Kant, cele samoistne. Ponadto, ludzie miewają rozmaite cele, które zależą od „poza-racjonalnych” impulsów. Trudno zrozumieć, jak w praktyce może dojść do jakiegokolwiek zbieżności takich celów. Podobnie jak w krytyce opartej na zarzucie idealizacji antropologicznej, Kant przypisuje ludziom cechy, których nie można o nich trafnie orzec na podstawie empirii (odtąd zarzut ten będę nazywał „zarzutem idealizacji racjonalnej”)³⁰.

Fakt, że w jednym miejscu Sidgwick krytykuje imperatyw kategoryczny za niezdolność do generowania pouczeń o postępowaniu, a w innym za zawieranie zbyt wiele „materii”, nie musi dowodzić błędu w tej krytyce. Może być tak, że teoria Kanta jest niespójna. Ważniejsze jest to, że Sidgwick przedstawia imperatyw kategoryczny jako dwie odmienne zasady w zależności od tego, jak go krytykuje. Zarzucając Kantowi abstrakcyjność, Sidgwick uznaje imperatyw kategoryczny za formalną zasadę, która dopuszcza wielość możliwych zasad moralnych w zależności od upodobań jednostek (pod tym względem krytyka ta wiąże się z zarzutem idealizacji racjonalnej). Centrum argumentacji Sidgwicka tkwi w twierdzeniu, że ludzie są zdolni chcieć bardzo rozmaitych rzeczy jako powszechnych reguł postępowania (tu tkwi kolejna zbieżność z oskarżeniem o idealizację racjonalną). Ocena tej argumentacji zależy od pojęcia *możliwości chcenia*, które stanowi specyfikę imperatywu kategorycznego.

W interpretacji Sidgwicka możliwość ta to możliwość psychologiczna. Zarówno w krytyce opartej na zarzucie abstrakcyjności, jak i w krytyce wykorzystującej zarzut idealizacji racjonalnej Sidgwick zwraca uwagę na zróżnicowanie „empirycznych pragnień i awersji” (*ME*, 390) ludzi, które sprawia, że nie można dojść do podziela-nych zasad moralnych. Jeżeli jednak uznać, że o trafności zasady miałyby decydować to, czy można jej chcieć jako prawa powszechnego, to, po pierwsze, łatwo zauważyć,

³⁰ Zarzut ten powtarza np. R. P. Wolff w *The Autonomy of Reason. A Commentary on Kant's Groundwork of the Metaphysic of Morals*, New York, Evanston, San Francisco, London 1973, Harper & Row, s. 174–176.

że chcieć można wszystkiego, ponieważ, jak zauważa Sidgwick, impulsy są nieracjonalne, a po drugie — dołączenie do możliwości chcenia kwalifikacji „jako prawa powszechnego” nie wyróżnia żadnych zasad. Jeżeli bowiem możemy być irracjonalni chcąc cokolwiek, to nie widać powodów, dla których ta irracjonalność miałaby nas opuszczać w chceniu czegoś jako prawa powszechnego.

Problem interpretacji imperatywu kategorycznego przez Sidgwicka w krytyce opartej na oskarżeniu o abstrakcyjność tkwi w tym, że zbyt lekko traktuje kluczowe pojęcia etyki Kanta: „możliwość chcenia” i „prawo powszechne”. Pisząc o prawie powszechnym Kant przeprowadza analogię do praw przyrody, które opisują stałe i uniwersalne prawidłowości (Gr, 421; 51). Sprawdzian imperatywu kategorycznego wymaga od jednostki, która ma pewien plan działania (maksymę) czy wyznaje pewną zasadę moralną, aby wyobraziła sobie, że wszyscy ludzie znajdujący się w okolicznościach istotnie podobnych do tych, w których znajduje się ta jednostka, postępowali na zasadzie uniwersalnej prawidłowości tak, jak proponuje to ten plan czy zasada. Użytkownik imperatywu kategorycznego ma sobie wyobrazić postępowanie według proponowanej zasady przez każdego, kto znajduje się w podobnej sytuacji.

Może zdarzyć się tak, że wykonanie tego sprawdzianu kończy się popadnięciem w sprzeczność. Sprzeczność ta nie musi być od razu jawna dla użytkownika imperatywu kategorycznego, lecz dopiero po rozważeniu większej liczby konsekwencji logicznych uogólnienia maksymy, mimo to jednak pozostanie ona sprzecznością. W takim wypadku rozważana maksyma nie da się „bez sprzeczności nawet *pomyśleć* jako ogólne prawo przyrody; cóż dopiero, żebyśmy mogli jeszcze *chcieć*, iżby ono się miało stać takim” (Gr, 424; 54/55).

Kant nie mówi o tym, czego można chcieć jako powszechnego prawa w sensie psychologicznym, jaki przypisuje mu Sidgwick. Być może niektóre osoby mogą — *psychologicznie rzecz biorąc* — w sposób jawny dla siebie samych chcieć rzeczy niemożliwych: chcieć zjeść ciastko i zachować je na potem. Nie to jednak jest przedmiotem Kantowskiego sprawdzianu, który wymaga oceny, czego można chcieć *bez sprzeczności*. W to zaś, że mogą istnieć plany sprzeczne, zdaje się Sidgwick nie wątpić. Na to, że Sidgwick rozumie możliwość chcenia maksymy jako prawa powszechnego w sposób psychologiczny, wyraźnie wskazuje fakt, że nie mówi o imperatywie kategorycznym Kanta, lecz o złotej regule. To ta reguła wymaga odwołania się do tego, czego ktoś chce faktycznie, a następnie sprawdzenia, czy jest się gotowym chcieć tego samego dla siebie ze strony innych. Złota reguła zasadza się właśnie na możliwości psychologicznej: czy dla danej jednostki rozważane działanie jest do przyjęcia z punktu widzenia jej — żeby posłużyć się określeniem Gilberta Harmana — „układu motywacyjnego”? Przyjmując, iż pragnienia składają się na tożsamość jednostki, pewne rzeczy mogą być niemożliwe z perspektywy jej układu motywacyjnego w tym sensie, że zgoda na nie wymagałaby bycia kim innym, niż się jest faktycznie, czy np. zmiany uznawanych zasad działania.

Królewieckiego filozofa taka niemożliwość nie interesuje. Pytanie zawarte w sprawdzianie imperatywu kategorycznego nie dotyczy możliwości psychologicznych, lecz — jak Kant by to ujął — rozumowych: chociaż pewne rzeczy sprzeczne mogą być chciane jako prawa powszechne, to chcenie ich jest po prostu nieracjonalne. To zaś oznacza, że nie musi być tak, że imperatyw kategoryczny dopuszcza wszelkie zasady. Imperatyw kategoryczny dopuszcza tylko zasady, które są niesprzeczne po uogólnieniu. Choć więc być może znajdują się jednostki, których konstytucja psychiczna pozwala im chcieć rzeczy sprzecznych, to samo przez się nie stanowi to dowodu błędu teorii Kanta.

Ponieważ Kanta nie interesują możliwości i niemożliwości psychologiczne, do jego stanowiska nie stosuje się sławna teza Hume'a, że nie ma nic nieracjonalnego w chceniu, aby raczej to cały świat się zawałił, niż bym miał zadrasnąć się w mały palec.³¹ Posiadanie takich preferencji być może jest psychologicznie możliwe, lecz trwanie przy nich (a więc np. zgoda na utratę życia po to, by uniknąć drobnej dolegliwości) oznacza odmowę brania pod uwagę względów racjonalności. Oskarżenie o abstrakcyjność jest więc trafne, o ile sprawdzian imperatywu kategorycznego traktuje się jako sprawdzian tego, co możliwe bez uwzględnienia wymogów racjonalności. Z chwilą dostrzeżenia, że Kant nie zajmował się tym, co możliwe psychologicznie, lecz tym, co możliwe racjonalnie, zarzut abstrakcyjności staje się bezprzedmiotowy. Równocześnie staje się jasne, że utożsamiając imperatyw kategoryczny ze złotą regułą, Sidgwick pomija kontekst, w którym zasada etyki Kantowskiej występuje, tj. kontekst pewnej wizji racjonalności praktycznej. Kant może się mylić w kwestii natury tej racjonalności, lecz oskarżenie o abstrakcyjność w żaden sposób tego nie dowodzi, ponieważ ignoruje Kantowskie rozumienie racjonalności praktycznej.

Pomijając Kantowskie rozumienie rozumu praktycznego, Sidgwick nie dostrzega drugiego jej aspektu: możliwości chcenia. Wedle Kanta, aby zasada była prawomocna, nie wystarczy, aby można było pomyśleć jej uogólnienie. Może być tak, że działanie da się pomyśleć jako podejmowane przez każdego, kto znajdzie się w sytuacji istotnie podobnej do tej, w której znalazł się użytkownik imperatywu kategorycznego, a mimo to może zdarzyć się tak, że „niemożliwą rzeczą jest *chcieć*, żeby maksyma ... była wyniesiona do ogólności prawa przyrody, ponieważ wola taka zaprzeczałaby sama sobie” (*Gr*, 424; 55). W tym wypadku problem z uogólnioną maksymą nie polega na tym, że jest sprzeczna, ale że wymaga chceń, które są sprzeczne.

Tej tezy Kanta Sidgwick nie dostrzega, przedstawiając zarzut idealizacji antropologicznej, w której przyjmuje, że mogą mieć rozmaite pragnienia i awersje, jednostki mogą uznawać najrozmaitsze zasady. Tak istotnie by było, gdyby tezę o tym, czego istoty rozumne chcą, wraz z Sidgwickiem traktować jak empiryczną. Kant tego jednak nie robi, lecz tak jak w pierwszej części sprawdzianu za pomocą imperatywu kategorycznego traktuje ją jako wynik namysłu nad racjonalnymi

³¹ D. Hume, *Traktat o naturze ludzkiej*, przeł. Cz. Znamierowski, Kraków 1952, Nakładem Polskiej Akademii Umiejętności, tom 2, s. 142.

rygorami nakładanymi na chcenie. W przykładzie obowiązku pomocy bliźniemu Kant nie stawia, jak mu to przypisuje Sidgwick, tezy empirycznej o tym, czego ludzie, czy istoty rozumne w ogóle, chcą. Teza Kanta dotyczy konsekwencji logicznych podlegania pewnym rygorom racjonalności w namyśle nad działaniem.

Jednym z tych rygorów jest oczywiście imperatyw kategoryczny. Lecz nie jest on jedyny. Kant wymienia oprócz niego także imperatyw hipotetyczny, który wymaga dobierania środków wiodących do realizacji planowanych celów. Z punktu widzenia właściwych człowiekowi ograniczonych zdolności osiągnięcia swych celów jest kwestią roztropności dyktowanej rygorami imperatywu hipotetycznego zadbać o środki wiodące do osiągnięcia tych celów. Jednym z takich celów może być współpraca innych nawet dla tych, którzy w danej chwili uważają — jak zauważa Sidgwick — że są dość silni by dać sobie radę samemu, i są gotowi podjąć zbędne ryzyko liczenia *zawsze* tylko na siebie. Niezależne i pewne swego losu osoby z przykładu Sidgwicka nie spełniałyby zatem wymogów imperatywu hipotetycznego, a w ten sposób dowodziłyby pewnego rodzaju nieracjonalności. Jeżeli zaś mają być racjonalne, to — zdaniem Kanta — będą musiały dążyć do współpracy innych ludzi i w ten sposób jako skończone i niesamowystarczalne istoty rozumne przyjąć zasadę oczekiwania przynajmniej okazjonalnej, czy w wyjątkowych sytuacjach, pomocy od innych.

Teza o potrzebie oczekiwania pomocy od innych przez każdego człowieka nie jest więc u Kanta tezą empiryczną — jak chce Sidgwick — ale konsekwencją stosowania imperatywu hipotetycznego będącego jedną z zasad namysłu nad działaniem. Chcenie uzyskania pomocy od innych nie jest konieczne psychologicznie, ale jedynie dla tych, którzy nie odrzucają racjonalności, tj. nie akceptują Hume'owskiego *dictum*. Chcenie to stanowi dla Kanta tło, na którym odbywa się sprawdzian imperatywu kategorycznego. Uogólnienie planu niepomagania osobom znajdującym się w potrzebie, a więc przyjęcie zasady pozostawiania bez pomocy wszystkich, którzy jej chcą, oznacza zgodę na pozbawienie tej pomocy również samego siebie. Po uogólnieniu maksymy odmawiania pomocy potrzebującym skończona istota rozumna, która musi chcieć uzyskania pomocy, miałaby sprzeczne chcenia: na mocy imperatywu hipotetycznego chciałaby pomocy, zaś na mocy uogólnienia nie mogłaby jej chcieć. Konkluzja Kanta jest więc następująca: choć możliwe jest wolne od sprzeczności uogólnienie zasady odmawiania pomocy wszystkim potrzebującym, to niemożliwe jest chcenie takiej uogólnionej zasady bez popadania w sprzeczność z wymaganym przez roztropność chceniem uzyskania jej.

Podobnie jak w przypadku oskarżenia o abstrakcyjność, tak i w zarzucie idealizacji antropologicznej Sidgwick nie docenia Kantowskiego pojęcia możliwości chcenia³². Kantowską możliwość tj. niesprzeczność w planie działania i w chceniu

³² Podobną opinię wyraża Onora O'Neill mówiąc, że Sidgwick nie docenia podwójnej modalności imperatywu kategorycznego i że jego interpretacja tej zasady jest niezwykle słaba i nie uwzględnia Kantowskiego pojęcia rozumu praktycznego; *Sidgwick on Practical Reason*, [w:] Henry Sidgwick, ed. R. Harrison, Oxford 2001, Oxford University Press.

planu (maksymy), zastępuje psychologiczną zdolnością żywienia pewnych pragnień. Kant może się naturalnie mylić w kwestii tego, czy skończone istoty rozumne muszą mieć pewne chcenia, teza ta jednak jest inna od prostej konstatacji, którą przypisuje mu Sidgwick: że ludzie zawsze pragną pomocy lub że muszą jej pragnąć na mocy jakichś prawidłowości psychologicznych.

Formułując zarzuty abstrakcyjności i idealizacji antropologicznej, Sidgwick w istocie izoluje pojęcie imperatywu kategorycznego od reszty teorii Kanta, podobnie jak to się dzieje w oskarżeniu o idealizację racjonalną. Bez wątpienia Sidgwick ma rację, kiedy zauważa osobliwy sposób używania przez Kanta pojęcia celu we frazie „cel sam w sobie”. Nie dostrzega jednak, że Kant sam daje do zrozumienia, iż nie posługuje się tym pojęciem w sposób zastany w języku potocznym czy choćby w filozofii (*Gr*, 428; 60). Wprowadzając zasadę celu samego w sobie, najpierw zachęca czytelnika do przyjęcia, że „istniałoby coś, czego *istnienie samo w sobie* posiada bezwzględną wartość, a co jako *cel sam w sobie* mogłoby tworzyć podstawę pewnych praw” (tamże), a następnie za pomocą osobliwego zwrotu „twierdzą oto” stawia tezę, że „człowiek i w ogóle każda istota rozumna *istnieje* jako cel sam w sobie, *nie tylko jako środek*, którego ta lub owa wola może używać wedle swego upodobania, lecz musi być uważany *zarazem za cel*” (*Gr*, 428; 60/61).

Otwieranie zdania zwrotami podobnymi do „twierdzą oto” należy u Kanta do rzadkości i występuje tam, gdzie przedstawia tezę, za którą zamierza argumentować w kolejnych akapitach (por. np. *Gr*, 448; 88–89), a nie, jak zdaje się sądzić Sidgwick, jako założenie, na którym opiera dalsze wywody. Teza, że człowiek jest celem samym w sobie, jest tezą do okazania sformułowaną na podstawie zaproponowanego wcześniej pojęcia celu. Następnie Kant poddaje zasadę wymagającą traktowania istot rozumnych i ludzi jako cele same w sobie sprawdzianowi imperatywu kategorycznego i dopiero wtedy (*Gr*, 429, 62) formułuje nową wersję imperatywu kategorycznego³³. Ignorując tę sekwencję argumentacyjną Kanta, Sidgwick znowu izoluje imperatyw kategoryczny od reszty teorii królewieckiego filozofa.

Sidgwicka krytyki imperatywu kategorycznego łączy pomijanie Kantowskiej koncepcji racjonalności praktycznej i tez o działaniu istot rozumnych a jednocześnie skończonych. Kant niejednokrotnie podkreślał — zwłaszcza w uwagach na temat woli świętej (*Gr*, 414; 40; *KpV*, 81/82; 135/136; por. *KpV*, 32/3; 56/57) — że charakterystyczną cechą istot skończonych i skończenie rozumnych jest przywiązanie do dobra połączone z wciąż obecną możliwością postępowania innego niż dobre. To właśnie ta obserwacja wymagała, zdaniem Kanta, sformułowania rozbudowanej (i przez to niewątpliwie trudnej) filozofii moralności, w której imperatyw kategoryczny jest jednym z elementów większej całości³⁴. Tymczasem Sidgwick

³³ Dokładniej sprawę tę omówiłem w *Wolność i autorytet rozumu. Racjonalność w filozofii moralnej Kanta*, Warszawa 1997, WFiS, s. 194–197.

³⁴ Więcej na temat rozróżnienia na wolę świętą i wolę niedoskonałej czy skończonej istoty rozumnej oraz o znaczeniu rozpoznania możliwości odstąpienia od dobra przez istoty niedoskonałe

zdaje się tej złożoności nie dostrzegać i interpretuje teorię Kanta w sposób obcy królewieckiemu filozofowi, pomijając tezy, bez których nie można sobie wyrobić rozsądnej opinii o jego etyce.

Kant bez wątpienia ponosi część odpowiedzialności za taką recepcję swej teorii moralnej. Niejednokrotnie przedstawia np. imperatyw kategoryczny tak, jakby był on zasadą, do której da się sprowadzić całą moralność i że namysł moralny nie stawia przed jednostką szczególnych wymagań (np. *Gr*, 403; 25). Takie wypowiedzi Kanta nie usprawiedliwiają jednak pomijania ważnych tez jego teorii przez Sidgwicka. Odwołując się do zespołu idei przedstawionego w części pierwszej tego artykułu, przedstawię obecnie drogę, która zawiodła Sidgwicka do jego interpretacji imperatywu kategorycznego oraz obecność tej interpretacji lub jej zasadniczych części u etyków anglosaskich następných pokoleń.

III. Kantowskie dziedzictwo Sidgwicka

Jak wskazałem, klucz do wyjaśnienia izolowania przez Sidgwicka imperatywu kategorycznego od reszty filozofii moralnej Kanta tkwi w zespole idei funkcjonujących w brytyjskiej etyce filozoficznej XVIII i pierwszej połowy XIX w. i w obecnym w nim rozumieniu zasad moralnych oraz w wiązanych z tym rozumieniem oczekiwaniach. Ów zespół idei może nie dostarczać pełnego wyjaśnienia postępowania Sidgwicka, lecz wydaje się, że musi odgrywać w takim wyjaśnieniu główną rolę.

Wskazówkę w poszukiwaniu wyjaśnienia korzystania przez Sidgwicka z ubogiej interpretacji etyki Kanta zawiera, jak się wydaje, już sama nazwa metody, którą jego zdaniem stosuje królewiecki filozof. Sidgwick nazywa tę metodę intuicyjną zgodnie ze znaczeniem tego określenia w Cambridge pierwszej połowy XIX w. Intuicja nie jest tu czymś tajemniczym, lecz zwykłym sposobem ujawniania się potocznych przekonań moralnych. Tak też widział etykę Whewella J. S. Mill, kiedy w 1852 roku krytykował nauczyciela Sidgwicka za intuicjonizm. W swojej energicznej krytyce Mill twierdzi, że Whewell traktuje intuicję jako źródło oczywistych, dominujących w danej społeczności, domniemyanych prawd moralnych. Według Milla intuicja faktycznie służy Whewellowi do sankcjonowania moralnego *status quo*. Tam zaś, gdzie intuicja ma być rzeczywistym źródłem pouczenia moralnego, pouczenie to jest niewielkiej wartości bez dołączenia do niego materialnych tez o dobru³⁵.

W Milla krytyce Whewella widać wyraźnie obydwa wątki obecne w Sidgwicka krytyce Kanta: przekonanie o pustym formalizmie zasady, z którą utożsamia imperatyw kategoryczny (*ME*, 386), i teza o przemycaniu weń fałszywych uogólnień empirycznych o człowieku (*ME*, 389). Inaczej niż Mill, Sidgwick nie porzuca jednak intuicji jako źródła znajomości zasad. Jak sam wyznaje, pod wpływem

dla kształtu teorii moralnej Kanta piszę w artykule pt. *Kanta odkrycie normatywności* „Przegląd Filozoficzny — Nowa Seria”, 2004, R. 13, nr 4 (52), s.147–165.

³⁵ J. S. Mill, *Dr. Whewell on Moral Philosophy*, „Westminster Review” 1852, przedrukowane [w:] *Mill's Ethical Writings*, ed. J. B. Schneewind, New York 1965, Collier Books.

lektury Clarke'a (*ME*, xix), włącza metodę intuicyjną do swojego stanowiska, lecz twierdzi też, że zasady te nie wystarczają do wyprowadzenia konkretnych pouczeń moralnych (*ME*, 379). Dlatego jest przekonany, że należy je uzupełnić treścią, którą można zaczerpnąć z motywacji ludzkich i na którą składają się tezy egoizmu i utilitaryzmu. Sidgwick wyznaje, że nie potrafi podać argumentów, które by ostatecznie rozstrzygały, która z tych dwóch metod jest nieodparcie bardziej przekonująca, sam jednak (jak twierdzi, pod wpływem lektury Kanta i Butlera) przychyliła się do wyboru metody utilitarystycznej (*ME*, xvii — xxi).

Sidgwick wykazuje ambiwalencję wobec intuicji. Uważa ją za źródło wiążących prawd, lecz zwolennikom stanowisk, które zalicza do metody intuicyjnej zarzuca brak wyraźnych konkluzji moralnych. Wydaje się, że metoda intuicyjna jest jego zdaniem atrakcyjna, lecz nie chce jej zaakceptować wraz z bagażem skojarzeń i tez, które jej zwykle towarzyszyły w brytyjskiej filozofii moralnej XIX w. Ta ambiwalencja wydaje się być przyczyną tego, że Sidgwick ma niejasne oczekiwania wobec intuicji, rozmaicie interpretuje treść intuicyjnych danych, a odmienne zasady utożsamia z imperatywem categorycznym.

Korzystając z pojęcia intuicji obecnego w brytyjskiej filozofii moralnej XIX w., Sidgwick traktuje więc Kanta tak, jakby filozof z Królewca głosił, że rozum praktyczny odkrywa imperatyw categoryczny w następstwie jakiegoś oglądu. Konkludując, że Kant błędnie wyprowadza zasadę dobroczynności z imperatywem categorycznego, Sidgwick twierdzi, że królewiecki filozof „uważa za rzecz oczywistą (*evident*) a priori, że jako istotny podmiot działający każdy człowiek musi dążyć do szczęścia innych ludzi” (*ME*, 386). Sidgwick zdaje się tu zakładać, że treścią imperatywu categorycznego jest złota reguła, a Kant nadał dziwną nazwę tej dobrze znanej, lecz mało użytecznej w namyśle praktycznym, regule. Takie postępowanie Sidgwicka jest zrozumiałe z punktu widzenia pokazanej w pierwszej części tego artykułu tradycji brytyjskiej, w której za przedmiot rozumowego oglądu uznawano złotą regułę, czasem nazywając ją imperatywem categorycznym.

Ponieważ w tradycji tej rozum utożsamiany z sumieniem miał być źródłem zasad, których rolę porównywano do roli aksjomatu myślenia moralnego, Sidgwick oczekuje, że i imperatyw categoryczny będzie aksjomatem namysłu praktycznego (*ME*, 386p), z którego to aksjomatu dedukuje się zasady działania. To pojęcie imperatywu categorycznego nie należy jednak do Kanta, lecz do osiemnastowiecznych moralistów brytyjskich oraz do ich dziewiętnastowiecznych krytyków. Sidgwick sam wskazuje, że J. S. Mill wiele razy zbiorczo odnosił się do Coleridge'a i etyki niemieckiej za pomocą zwrotów *Coleridge and the Germans* i *Germano-Coleridgean doctrine*³⁶. Określenia te wyraźnie pokazują, że tezy tych jakże odmiennych tradycji interpretowano jako różne wersje jednego stanowiska. Kiedy Sidgwick przedstawia

³⁶ *Outlines*, dz. cyt., s. 277. Tekst Milla, o którym mowa to *Coleridge* opublikowany w „The London and Westminster Review” 33 (March 1840), s. 257–302 i przedrukowany w *Autobiography and Other Writings*, ed. J. Stillinger, Boston 1969, Houghton Mifflin.

imperatyw kategoryczny jako zasadę sumienia³⁷, to przejmuje właśnie tę *niekantowską* interpretację etyki Kanta, traktując imperatyw kategoryczny jak aksjomat w systemie dedukcyjnym. Korzystając z tej interpretacji, może zarzucać Kantowi niemożliwość przeprowadzenia „bezpośredniej dedukcji” zasad działania z imperatywu kategorycznego (*ME*, 389).

Teza Sidgwicka, że z imperatywu kategorycznego nie można bezpośrednio wydedukować zasady dobroczynności (i, dodajmy, żadnej innej zasady moralnej), jest w oczywisty sposób prawdziwa. Nie dzieje się tak jednak z powodów, które przytacza Sidgwick. Z imperatywu kategorycznego nic nie wynika dlatego, że w filozofii Kanta zasada ta nie jest żadnym aksjomatem ani niczym, co odgrywałoby podobną rolę. Imperatyw kategoryczny jest elementem sprawdzianu, rodzaju kryterium, które służy do ustalania moralnej akceptowalności pewnych propozycji działania. Kant twierdzi, że dozwolone jest postępowanie na podstawie maksym, których można bez sprzeczności chcieć jako praw powszechnych; postępowanie na podstawie maksym, które tego sprawdzianu nie przetrwają, jest zakazane, a nakazane jest postępowanie przeciwne do zdyskwalifikowanego (*Gr*, 424; 55).

Sława niezwykle rzetelnego, ostrożnego i intelektualnie uczciwego autora sprawiła, że Sidgwick w znacznym stopniu kształtował kolejne pokolenia anglosaskich filozofów moralności, w tym ich interpretację Kanta. Przez następne pięćdziesiąt lat naśladowano jego styl pisarski i założenia dotyczące niektórych filozofów. W 1903 roku, tj. w roku ostatniego wydania *Methods*, ukazała się jedna z najbardziej wpływowych książek etyki XX w., *Principia ethica* G. E. Moore'a. Stylistycznie *Principia* nawiązują do *Methods* i powielają obecny w nich obraz etyki Kanta, choć sam Moore nie omawia imperatywu kategorycznego z tego względu, że koncentruje się na badaniu relacji między dobrem a jego znajomością, a nie na treści tej domniemanej wiedzy.

Moore zalicza etykę Kanta do etyki metafizycznej, twierdząc, że opiera się ona na domniemanym wglądzie w ponadmysłową rzeczywistość³⁸, w której znajduje się wiedzę o tym, co z konieczności robi czysta czy wolna wola, w postaci prawa, którego wola ta musi przestrzegać. Prawo to, twierdzi Moore, Kant przedstawia jako imperatyw dla ludzi.³⁹ Takie postępowanie Kanta ma być błędne, ponieważ prawo moralne, które w odniesieniu do woli czystej zostało przedstawione na wzór prawa przyrody teraz zostaje potraktowane jak „prawo w znaczeniu legislacyjnym”⁴⁰, Kant popełnia więc oczywisty błąd naturalistyczny, z tą różnicą, że podstawą jego błędu są tezy nie o przyrodzie, lecz o rzeczywistości ponadmysłowej.

Zbieżności między postępowaniem Moore'a a podejściem Sidgwicka są jasne. Moore traktuje Kanta jak intuicjonistę w sensie zbliżonym do Sidgwickowskiego,

³⁷ W artykule pt. *Reply to Barratt on The Suppression of Egoism*, „Mind” 1877, nr. 2 (7).

³⁸ G. E. Moore, *Zasady etyki*, przeł. Cz. Znamierowski, De Agostini, Warszawa 2003, Alaya, s. 181.

³⁹ Gdyby nie nazwisko Kanta, można by sądzić, że referując poglądy filozofa z Królewca Moore wiernie przedstawia stanowisko Samuela Clarke'a z *A Discourse of Natural Religion*, dz. cyt., s.192, 198–199.

⁴⁰ Tamże, s. 203–204 i 206.

na co sam wskazuje w *Przedmowie*⁴¹. Zgodnie z tym pojęciem intuicyjności uważa, że Kant korzysta z oglądu, którego treścią są niedowodliwe „pierwsze zasady”. W rekonstrukcji Moore’a przedmiotem Kantowskiego oglądu nie jest wprawdzie rozum czy sumienie, lecz jakaś rzeczywistość ponadzmysłowa, tym niemniej wymaga on refleksji i nie ma charakteru oczywistości. Daje za to informacje pewne o dobru. Jak u Sidgwicka, treść tego oglądu jest „podstawowym (*fundamental*) *datum*”, z którego wyprowadza się tezy o obowiązkach człowieka. Moore przypisuje Kantowi ten „geometryczny” model etyki, pomimo że Kant sam nigdzie takiego modelu nie proponuje. Bardziej prawdopodobne wydaje się natomiast podejrzenie, że tak jak jego poprzednicy, w tym i Sidgwick, Moore wierzył w etykę *more geometrico*, proponując swą etykę, „która mogłaby mieć uroszczenia do naukowości”⁴². I właśnie te *swoje* ambicje przypisywał Kantowi.

Pojęcie intuicji odziedziczone po dziewiętnastowiecznej etyce brytyjskiej wraz ze związanymi z nimi interpretacjami imperatywu kategorycznego występuje zarówno u anglosaskich krytyków, jak i u obrońców Kanta. O ile przejmowanie Sidgwickowskich wyobrażeń o etyce Kanta przez krytyków można zrozumieć jako podyktowane wygodą, to zaskakuje ono u autorów przyznających się do kantowskich inspiracji. Lista autorów, którzy w różnym stopniu powielają sidgwickowską interpretację Kanta jest długa, a już mały wyciąg z niej ilustruje rozmiary wpływu tej interpretacji na kolejnych autorów. Zarówno niechętny Kantowi Peter Singer⁴³, jak i powołujący się na kantowskie inspiracje Richard M. Hare⁴⁴ utożsamiają imperatyw kategoryczny ze złotą regułą. Z kolei krytykujący Kanta Alasdair MacIntyre⁴⁵ zgłasza zarzut, że z imperatywu kategorycznego nie można wywieść konkretnych norm, a przychylny Kantowi W. David Ross⁴⁶ powtarza oskarżenie o idealizację racjonalną, komentując użycie przez Kanta słowa cel.

To widzenie etyki Kanta przyjęło się nie tylko w pracach niektórych krytyków i naśladowców Kanta, ale też w części anglosaskich podręczników do historii filozofii i etyki. Np. w jednym z bardziej cenionych podręczników do etyki teoretycznej Richard Brandt twierdzi, że „z grubsza biorąc” imperatyw kategoryczny to złota reguła, a tłumacząc uogólnialność, odwołuje się do psychologicznej możliwości zaakceptowania zasady.⁴⁷ Natomiast Tom L. Beauchamp, wyjaśnia w swoim podręczniku do etyki, że imperatyw kategoryczny jest pustą zasadą formalną, której

⁴¹ Tamże, s. IX.

⁴² Tamże, s. VIII, por też np. s. 4.

⁴³ *Etyka praktyczna*, przeł. A. Sagan, Warszawa 2003, Książka i Wiedza, s. 26.

⁴⁴ Np. *Uniwersalny preskryptywizm* [w:] *Przewodnik po etyce*, red. P. Singer, Warszawa 1998, KiW, s. 509.

⁴⁵ *Dziedzictwo cnoty: studium z teorii moralności*, przeł., wstępem i przypisami opatrzył A. Chmielewski, Warszawa 1996, Wydawnictwo Naukowe PWN, s. 99–100.

⁴⁶ *Kant's Ethical Theory: A Commentary on the Grundlegung zur Metaphysik der Sitten*, Oxford 1954, Clarendon Press, s. 51.

⁴⁷ Richard B. Brandt, *Etyka: zagadnienia etyki normatywnej i metaetyki*, przeł. B. Stanosz, wstępem opatrzył J. Hołówka, Warszawa 1996, PWN, odpowiednio: s. 55 i s. 59.

rzeczywista wartość polega na tym, że wyraża rozpowszechnioną niechęć do czynienia wyjątków dla siebie⁴⁸. Wraz z Jamesem Childressem podobne stanowisko zajmuje w podręczniku etyki medycznej, gdzie pisze, iż imperatyw kategoryczny nie może podać konkretnych przepisów postępowania, a jego wartość sprowadza się do wymogu konsekwencji w ocenianiu⁴⁹.

Zakończenie

Sidgwicka krytyka etyki Kanta to kulminacja uprawiania etyki polegającego na poszukiwaniu naczelných zasad moralności w oczywistości intelektualnego oglądu. W tradycji tej sumienie pojawia się zamiennie z rozumem, a argumentację opartą na pochodzących z nich danych pojmowanych na kształt aksjomatów wzoruje się na dedukcji. Do aksjomatów tych zaliczano zasadę podobnego traktowania przypadków podobnych i dość często utożsamiano ją ze złotą regułą. W tradycji tej korzystano z koncepcji rozumu praktycznego, który stanowił źródło zasad i służył do dedukcyjnego wyprowadzania ich konsekwencji, nie zawierając żadnych istotnych tez na temat racjonalnego kontekstu użycia tych zasad.

Tak pojmowany rozum praktyczny nie zawierał zatem istotnych danych o dobru, a jedynie odkrywane zasady. Ponieważ były to tylko zasady, tj. reguły nie mówiące, *co* czynić w konkretnych sytuacjach, lecz jak należy porządkować bądź uzgadniać swoje przekonania moralne, a o zasadach tych zakładano, że one same mają wystarczyć do dostarczania szczegółowych pouczeń moralnych, jako niemal oczywistość pojawiał się zarzut braku konkretnych pouczeń o postępowaniu: jeśli rozum to tylko zasady, to jak można z nich wydedukować cokolwiek o wyraźnej treści moralnej?

W etyce Sidgwicka tradycja ta zatoczyła pełne koło. Okazało się, że jest jałowa i należało ją odrzucić. Sidgwick ostatecznie wyartykułował krytykę tej tradycji, jednak będąc głęboko przywiązany do właściwego tej tradycji instrumentarium, czytając prace zawierające tezy zbliżone do tez tej tradycji, nie dostrzegł zasadniczych różnic między nią a etyką Kanta. Obie traktował jak to samo stanowisko wyrażone na nieco odmienne sposoby. Pomiął przez to całą złożoność koncepcji rozumu praktycznego w filozofii Kanta. Nie zauważył na przykład, że imperatyw kategoryczny to jeden (choć centralny) element całości, jaką jest sprawdzian uogólnialności. Niemożliwość, którą Kantowski sprawdzian ma wykrywać, traktował jako psychologiczną, a nie jako konsekwencję logicznych zależności między elementami planu działania (tj. maksymy) po opatrzeniu go wielkim kwantyfikatorem.

Zaniedbując te subtelności etyki Kanta i podstawiając pod idee Kantowskie idee właściwe brytyjskiej myśli etycznej XVIII i pierwszej połowy XIX w., Sidgwick

⁴⁸ Tom L. Beauchamp, *Philosophical Ethics: An Introduction to Moral Philosophy*, New York, NY : McGraw-Hill, Inc. 1991, s. 186–187.

⁴⁹ Tom L. Beauchamp, James F. Childress, *Zasady etyki medycznej*, przeł. W. Jacórzyński, Warszawa 1996, KiW, s. 72.

w istocie nie krytykuje Kanta, lecz tradycję, z której sam wyrósł i której słabości wyraźnie dostrzegał. W paradoksalny sposób kulminacja tej tradycji przyjęła u Sidgwicka postać odrzucenia jej połączonego z zachowaniem znacznej części terminologii i podaniem „standardowej” interpretacji etyki Kanta. Wpływ tej terminologii i interpretacji okazał się tak wielki, że musiał upłynąć niemal wiek, zanim anglojęzyczni filozofowie zaczęli rzetelnie badać etykę Kanta.

The Categorical Imperative and Sidgwick's Intuitionist Legacy

The paper presents H. Sidgwick's interpretation of the categorical imperative found in *The Methods of Ethics* and the influence this interpretation had on next generations of English-speaking philosophers. In his reading of Kant Sidgwick relied on a complex of ideas which had been present in British moral philosophy before the Victorian era and which differed fundamentally from Kant's. The authority of Sidgwick led many later philosophers to accept his interpretation and criticisms of Kant's views without thorough study of the texts of the Königsberg thinker. The interpretation had its own life in many history of ethics textbooks and indirectly has been shaping thinking of many contemporary moral philosophers