

Słaba wola – w stronę eksternalizmu motywacyjnego

Karol Milczarek

(Uniwersytet Warszawski, Wydział Filozofii)

1. Wprowadzenie

Zasada, o której będzie tutaj mowa, znana jest pod wieloma nazwami: zasady intelektualizmu lub racjonalizmu etycznego¹, zasady sokratejskiej², starej formuły szkół³, zasady umiarkowanego internalizmu⁴ czy – już bardziej precyzyjnie – tezy internalizmu motywacyjnego⁵. Jest to druga przesłanka Davidsonowskiego trylematu, konstytuującego paradoks słabej woli:

(Z1) Jeżeli sprawca chce bardziej zrobić x , niż zrobić y , i jeżeli przekonany jest, że może wybrać pomiędzy x i y , wówczas w sposób zamierzony

¹ J. Hołówka, *Etyka w działaniu*, Warszawa 2002, s. 56–62.

² R. Audi, *Słabość woli i osąd praktyczny*, tłum. W.J. Popowski, w: *Filozofia moralności: postanowienie i odpowiedzialność moralna*, red. J. Hołówka, Warszawa 1997.

³ S. Tenenbaum, *The Judgement of a Weak Will*, „Philosophy and Phenomenological Research” 1999, t. 59, nr 4, s. 876.

⁴ D. Davidson, *Jak jest możliwa słabość woli?*, tłum. W. J. Popowski, w: *Filozofia moralności...*, dz. cyt., s. 87. Podkreślam, że sens, w jakim używa się tu pojęć internalizmu i eksternalizmu, jest inny niż ten, który niekiedy pojawia się w kontekście dyskusji nad działaniem niepowściągliwym, to jest w sporze o zakres racji, potencjalnie uwzględnianych przy formułowaniu najlepszego sądu (zob. np. A. McIntyre, *Is Acratic Action Always Irrational*, w: *Identity, Character and Morality*, red. O. Flanagan, A.O. Rorty, Cambridge 1993). Spór ten jest niezależny od rozstrzygnięć w kwestii (Z2).

⁵ R. Shafer-Landau, *A Defense of Motivational Externalism*, „Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition” 2000, t. 97, nr 3.

zrobi x , jeżeli w sposób zamierzony podejmie którekolwiek z tych działań.

(Z2) Jeżeli sprawca osądza, że lepiej byłoby zrobić x , niż zrobić y , wówczas bardziej chce zrobić x , niż zrobić y .

(Z3) Zdarzają się działania niepowściągliwe⁶.

Wydaje się, że każda z tych zasad osobno jest przekonująca, a jednocześnie, że łącznie nie mogą być one prawdziwe, ponieważ koniunkcja (Z1) i (Z2) implikuje negację (Z3). Aby zatem rozwiązać problem słabej woli, należy zanegować co najmniej jedną zasadę lub wyjaśnić, dlaczego między wszystkimi trzema nie ma sprzeczności.

Przykład

Załóżmy, że na przyjęciu zastanawiam się, czy dołożyć sobie jeszcze jeden kawałek tortu. Wiem, że sprawiłoby mi to przyjemność, ale jestem na diecie, więc formułuję (najlepszy) sąd, że lepiej będzie się od tego powstrzymać. Mimo to, z lekkim poczuciem winy, sięgam po kolejny kawałek. Paradoks bierze się stąd, że na mocy (Z2) sformułowanie sądu, że nie powinienem już dziś jeść słodkości, jest równoznaczne z ograniczeniem chęci wzięcia dodatkowej. Chęć ta, zgodnie z (Z1), wyrażającą naturalny związek między pragnieniami a działaniami⁷, powinna być przyczyną, która sprawi, że będę trzymał ręce z dala od tortu. W tej sytuacji, jak wskazuje (Z3), biorę jednak kolejny kawałek, czyli działam niepowściągliwie.

Działanie niepowściągliwe, nazywane także działaniem akratycznym czy działaniem będącym przejawem słabej woli⁸, to jedna z form irracjonalności

⁶ D. Davidson, *Jak jest możliwa słabość woli?*, dz. cyt., s. 83–84.

⁷ Tamże, s. 83.

⁸ Dwie uwagi techniczne. Po pierwsze, niekiedy odróżnia się akrazję od słabej woli. W tej pracy nie stosuję takiego podziału, a wymienionych pojęć używam synonimicznie. Po drugie, zajmuję się tutaj przede wszystkim działaniami niepowściągliwymi, a nie ludźmi o słabej woli. Tych drugich można określić jako osoby, które mają dyspozycję do działań niepowściągliwych, ale nie należy przy tym zakładać, że ludziom o silnej woli – cechującym się enkrateją – nie zdarza się działać akratycznie. Więcej o tych dystynkcjach zob. A. Mele, *Backsliding: Understanding Weakness of Will*, New York 2012, s. 1–11.

człowieka i jako taka jest zagadkowa. Wyraźnie widać to na przykładzie rekonstrukcji struktury działania intencjonalnego, zaproponowanej przez Donalda Davidsona – jeśli racje, z którymi działamy, są tożsame z przyczynami działań, trudno zrozumieć, jak sąd stanowiący najlepszą rację dla danego działania (związek logiczny) może różnić się od pragnienia, przyczynowo związanego z działaniem (związek kauzalny)⁹. Chcąc zatem przedstawić adekwatne ujęcie działania niepowściągliwego, źródłowo czerpiące z koncepcji Davidsona, należy uwzględnić obydwa wskazane aspekty działania.

Takie są też moje zamiary. Prawdziwość (Z1) i (Z3) biorę jako dane, ponieważ w niniejszej pracy chcę skupić się wyłącznie na omówieniu (Z2), czyli zasady internalizmu motywacyjnego. Jak zauważa Robert Audi, Davidson nie przedstawia satysfakcjonującej argumentacji na rzecz (Z2), chociaż narzuca się ona jako cel krytyki¹⁰. Wynika to z faktu, że według autora tekstu *Jak jest możliwa słabość woli?* sformułowanie jej tak, aby była fałszywa, jest łatwe, a mimo to wciąż mamy „skłonność do stwierdzenia, że jeśli ktoś naprawdę (szczerze) jest przekonany, iż powinien coś zrobić, wówczas jego zachowanie musi znaleźć odbicie w jego zachowaniu (a zatem, co oczywiste, i w jego inklinacji do działania lub w jego pragnieniu)”¹¹. Uwzględniając to, jak ten amerykański filozof analityczny rozumiał działanie intencjonalne, nietrudno dostrzec, dlaczego przyjął tę zasadę. Chcę jednak pokazać, że są dobre powody, aby nie iść w jego ślady.

To wprowadzenie miało na celu przygotowanie gruntu pod dalszą analizę (Z2). W części pierwszej artykułu (sekcja 2) przedstawiam mocną interpretację tezy internalizmu motywacyjnego i przywołuję kontrprzykłady, wskazujące, że jest to teza zbyt silna. Szkicuję także zasadniczą ideę stanowiska przeciwnego na przykładzie poglądów Alfreda Melego, aby w kolejnych częściach przejść do omówienia zarzutów przeciwko eksternalizmowi. Krytyka trzech argumentów prowadzi do sformułowania warunków eksternalistycznej analizy działania niepowściągliwego, ostatecznie sformułowanej w podsumowaniu.

⁹ D. Davidson, *Actions, Reasons, and Causes*, w: tegoż, *Essays on Actions and Events*, Oxford 1980, s. 3–20; tenże, *Paradoxes of Irrationality*, w: tegoż, *Problems of Rationality*, Oxford 2004, s. 169–187.

¹⁰ R. Audi, *Słabość woli...*, dz. cyt., s. 127.

¹¹ D. Davidson, *Jak jest możliwa słabość woli?*, dz. cyt., s. 88.

2. Mocna interpretacja

Sokrates, prekursor internalizmu motywacyjnego, miał powiedzieć, że „nie ma nic mocniejszego niż wiedza i że ona zawsze, wszędzie, gdzie jest, panuje nad rozkoszą i nad wszystkimi innymi skłonnościami”¹². Opanowanie człowieka przez namiętności, czyli to, co dziś nazywa się słabą wolą, jest więc tylko pozorne, ponieważ błąd zawsze leży po stronie rozumu. Jeśli człowiek **naprawdę** wie, że jedno działanie jest najlepsze, nigdy nie wybierze innego. Przy takim stanowisku obstawał również Richard Mervyn Hare, który twierdził, że sąd wyrażony w (Z2) jest analitycznie prawdziwy i co więcej, jest to sąd na temat uniwersalnie obowiązujących norm moralnych. Poprzednik implikacji, dotyczący osądu sprawy, Hare rozumie jako przypadek rozpoznania powinności. Rozpoznanie to jest bezosobowe – w danych okolicznościach każdy racjonalny podmiot dostrzeże te same normy moralne i się im podporządkowuje¹³.

W silnej interpretacji (Z2) opisuje logiczną zależność między ewaluacją danego działania a pragnieniem jego zrealizowania. Do takiej koncepcji zdaje się przychylić Davidson, gdy pisze, że za propozycjonalny wyraz pragnienia dowiedzenia się, która jest godzina, można uznać zdanie: „Dobrze by było dowiedzieć się, która jest godzina”¹⁴. Sama (Z2) nie wystarczy do uzasadnienia takiego rozumowania, ponieważ wyraża ona tylko jednostronną implikację. Tu zaś prowadzi ona także w drugą stronę, toteż silną wersję zasady internalizmu motywacyjnego można sformułować jako:

(Z2*) Sprawca osądza, że lepiej byłoby zrobić x niż y , wtedy i tylko wtedy, gdy bardziej chce zrobić x , niż zrobić y .

Należy jednak zachować ostrożność w przypisywaniu tak kategorycznej tezy komuś, kto wzbrania się przed jednoznacznym sformułowaniem (Z2), a na dodatek pisze, że nie jest ona wyrazem mocnego internalizmu, ponieważ „nic nie mówi na przykład o związku pomiędzy rzeczywistą wartością rzeczy (lub obligatoryjnym charakterem działań) a pragnieniami lub motywami”¹⁵. Wbrew Hare’owi Davidson traktuje osąd jako wyraz subiektywnej oceny działań, a nie

¹² Platon, *Protagoras* (357c), tłum. L. Regner, Warszawa 1995, s. 82.

¹³ R.M. Hare, *Assenting to a Value Judgement*, w: *Weakness of Will*, red. G. Mortimore, London 1971, za: J. Hołówka, *Etyka...*, dz. cyt., s. 62–67.

¹⁴ D. Davidson, *Jak jest możliwa słabość woli?*, dz. cyt., s. 94.

¹⁵ Tamże, s. 87.

jako rozpoznanie ogólnej normy moralnej, w związku z czym unika niektórych problemów – wspomnę tylko o skrajnej nieintuicyjności zanegowania istnienia akrazji i licznych przypadkach, w których działamy raczej w zgodzie z naszymi pragnieniami i planami niż z tym, co uważamy za moralnie najlepsze – lecz nie wszystkich.

Przeciwko tej zasadzie, nawet w interpretacji subiektywistycznej, przemawiają liczne kontrprzykłady. Jako że (Z2*) głosi logiczną równoważność, wyklucza ona możliwość istnienia takich sprawców (niekoniecznie ludzkich), którzy dokonywaliby ewaluacji bez posiadania pragnień, i takich, którzy pragnęliby bez dokonywania ewaluacji. Wydaje się, że wyobrażenie sobie istot z pierwszej i drugiej grupy jest jak najbardziej możliwe, co – przy założeniu, że wyobrażalność implikuje możliwość – przeczy silnej interpretacji tezy internalizmu motywacyjnego.

Nie trzeba nawet wysilać wyobraźni, żeby znaleźć kontrprzykłady dla (Z2*). Na istnienie pragnień, którym nie towarzyszy pozytywna ewaluacja, wskazuje Gary Watson:

Weźmy przypadek kobiety, która odczuwa nagłe pragnienie utopienia w kąpieli swego wrzeszczącego dziecka, albo przypadek gracza w squasha, który po sromotnej przegranej odczuwa pragnienie, aby walnąć swojego przeciwnika rakietą w twarz. Nie jest prawdą, że matka przypisuje jakąś wartość utopieniu własnego dziecka albo że gracz w squasha przypisuje jakąś wartość zranieniu swego przeciwnika i zadaniu mu cierpienia. Mimo to pragną oni tych rzeczy. Nie jest tak, iż przypisują oni tym działaniom jakąś początkową wartość, nad którą przeważają potem inne względy. **Działania te nie pojawiają się nawet z jakkolwiek, choćby najmniejszą, pozytywną wartością na pierwotnej „matrycy pożądań”¹⁶.**

Watson podkreśla irracjonalny komponent pragnień, który skłania do rozluźnienia powiązania między ewaluacją a pragnieniem. W podanych przykładach jednak pragnienia te nie są wystarczająco intensywne, przez co nie prowadzą do akratycznego działania. Na gruncie eksternalizmu twierdzi się tymczasem, że część wolnych, świadomych, intencjonalnych działań ma właśnie taką strukturę¹⁷. Za flagowego eksternalistę można uznać Alfreda Melego – odróżnia on

¹⁶ G. Watson, *Free Agency*, „The Journal of Philosophy” 1975, t. 72, 1975, s. 210, za: Ch. Swanton, *Słabość woli jako rodzaj braku odwagi wykonawczej*, tłum. W.J. Popowski, w: *Filozofia moralności...*, dz. cyt., s. 147; wyróżnienie oryginalne.

¹⁷ A. Mele, *Backsliding...*, dz. cyt., s. 15–17, 20–21.

bowiem dwa rodzaje zobowiązań praktycznych: ewaluacyjne i wykonawcze, odpowiadające dwóm perspektywom (kolejno: kognitywnej – intelektualnej oraz konatywnej – motywacyjnej), które można przyjąć wobec ludzkiego działania¹⁸.

Rozróżnienie to prowadzi do odrzucenia internalizmu. Negacja (Z2), a więc teza eksternalizmu motywacyjnego, zostaje sformułowana w słowach: „Siła motywacyjna naszych pragnień nie zawsze odpowiada ewaluacji obiektów naszych pragnień”¹⁹. Mele argumentuje, że za prawdziwością tej tezy przemawiają: 1. potoczne doświadczenie; 2. badania empiryczne, wskazujące, że wraz ze zbliżaniem się do momentu realizacji określonego pragnienia zwiększa się jego siła motywacyjna (przy założeniu, że nie zmienia się ocena przedmiotu owego pragnienia)²⁰; 3. liczne eksperymenty myślowe²¹.

Mele utrzymuje, że na ogół przejście od najlepszego sądu, żeby zrobić *x*, do sformułowania intencji, żeby zrobić *x*, jest tak płynne, że wręcz niezauważalne. Jest to domyślna ścieżka deliberacji praktycznej, co po prostu wynika z tego, „jak skonstruowany jest normalny człowiek”²². Mimo to w niektórych przypadkach dochodzi do poluzowania lub zerwania tego połączenia. Mamy wówczas do czynienia z działaniem niepowściągliwym – działaniem z podstawą w formie pragnienia, którego siła motywacyjna nie odpowiada dokonanej przez sprawcę ewaluacji.

Zaproponowane przez tego filozofa rozwiązanie paradoksu słabej woli wydaje się obiecujące. Eksternalizm musi jednak sprostać wyzwaniu, jakim jest wyjaśnienie związku logicznego między działaniem a jego przyczyną. Na gruncie przywołanego stanowiska związek między najlepszym sądem a działaniem pozostaje dość enigmatyczny, ponieważ nacisk zostaje położony na kauzalną stronę działania. Skrzętnie wykorzystują to krytycy eksternalizmu.

W dalszej części tekstu omówię trzy zarzuty wysuwane przeciwko eksternalizmowi motywacyjnemu wraz z towarzyszącymi im dookreśleniami (Z2), zaproponowanymi w celu uniknięcia części problemów silnej interpretacji zasady sokratejskiej. Pierwszy (A1) dotyczy statusu ewaluacji w koncepcjach eksternalistycznych, drugi (A2) – wolności działań akratycznych, a trzeci (A3) – rozróżnie-

¹⁸ Tamże, s. 57–64.

¹⁹ Tamże, s. 73; por. A. Mele, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, New York 1987, s. 31–49.

²⁰ A. Mele, *Backsliding...*, dz. cyt., s. 73–82.

²¹ Zob. np. A. Mele, *Irrationality...*, dz. cyt., s. 37–38.

²² A. Mele, *Springs of Action*, New York 1992, s. 231.

nia między zwyczajnymi działaniami a działaniami niepowściągliwymi. Odparcie tych zarzutów pozwoli mi na zaproponowanie kolejnych warunków działania akratycznego – po jednym dla każdego. W ten sposób sformułowany zostanie szkic eksternalistycznej analizy problemu słabej woli.

3. Argument z ewaluacji

Pierwszy nazwę argumentem z ewaluacji (A1). Wysunięty został przez Michaela Bratmana, który taką oto tezę przypisuje Davidsonowi:

T2. W takim [leżącym u podstaw działania] rozumowaniu praktycznym rozważa się, co najlepiej byłoby zrobić, aby rozstrzygnąć pytanie, co robić. Powiniennem je określić **ewaluacyjnym rozumowaniem praktycznym**²³.

Trafnie wskazuje się tutaj ścisły związek ewaluacji z deliberacją praktyczną. Gdy sprawca osądza, że lepiej zrobić *x*, niż zrobić *y*, nie rozumuje w próżni, ale stara się określić tok swojego dalszego postępowania. Nie pyta: „Co najlepiej byłoby zrobić?”, ale po prostu: „Co robić?”²⁴. Wówczas racjonalny sprawca nie jest wyłącznie obserwatorem, wydającym czysto ewaluacyjne sądy, czy – jak zgrabnie ujmuje to przywołany autor – nie jest muchą na ścianie²⁵. Oznacza to, że (Z2*) zostaje zastąpiona zasadą obowiązującą w ścisłych granicach, to jest stosującą się w ramach deliberacji praktycznej.

Bratman rozważa spór między internalizmem a eksternalizmem motywacyjnym jako dyskusję o tym, czy jest jakaś zasadnicza różnica między rozważaniem, co byłoby zrobić najlepiej, a rozważaniem, co byłoby zrobić na przykład najbardziej szykownie. Twierdzi on, że internalizm przyjmuje wyróżniony status ewaluacyjnego rozumowania praktycznego, co gwarantuje ścisły związek między deliberacją a chęcią działania, natomiast eksternalizm neguje ów status, przez co dopuszcza możliwość rozbieżności oceny i pragnień²⁶. Wyznaczenie osi sporu w ten sposób pozwala na sformułowanie argumentu przeciwko eksternalizmowi:

²³ M. Bratman, *Practical Reasoning and Weakness of Will*, „Noûs” 1979, t. 13, nr 2, s. 156; wyróżnienie oryginalne.

²⁴ Na ten fakt niekiedy zwracali uwagę także inni internaliści; zob. np. R.M. Hare, *Freedom and Reason*, Oxford 1963, s. 54–55.

²⁵ M. Bratman, *Intention, Plans, and Practical Reasons*, Cambridge, MA 1987, s. 154.

²⁶ M. Bratman, *Practical Reasoning...*, dz. cyt., s. 157–160.

- (A1) Działanie niepowściągliwe także jest poprzedzone rozważaniem ewaluacyjnym. Sprawca ocenia, które działanie jest *prima facie* najlepsze, a następnie postępuje wbrew powziętemu osądowi. T2 stosuje się zatem również odnośnie do działań wynikających ze słabej woli²⁷.

Warto mieć na uwadze to, jak autorzy rekonstruują dyskusje filozoficzne, ponieważ zmiana stawianego pytania często rodzi argumenty, które nie trafiają w swój cel. Tak jest w przypadku (A1). Wniosek z (A1) może być poprawny, ale istnieje zasadnicza, a przemilczana różnica między (Z2) a (T2), która sprawia, że nie jest to dobry zarzut przeciwko eksternalizmowi motywacyjnemu. Pierwsza z nich wiąże ewaluację z pragnieniem, natomiast druga – z rozumowaniem praktycznym. Bratman może posługiwać się (T2) jak zasadą internalizmu tylko dlatego, że przyjmuje arystotelejską przesłankę głoszącą, że „działanie albo intencja, albo wolicja do działania będzie odpowiadać konkluzji argumentu, odpowiadającemu mojemu rozumowaniu praktycznemu”²⁸.

Argument Bratmana w zasadniczo poprawnej, a jednocześnie wciąż bardzo uproszczonej formie prezentuje się zatem następująco:

1. w rozumowaniu praktycznym sprawca dokonuje ewaluacji;
 2. jeśli w rozumowaniu praktycznym sprawca dokonuje ewaluacji, to eksternalizm motywacyjny jest fałszywy;
- ergo*: eksternalizm jest fałszywy.

Druga przesłanka jest skrótowym wyrazem założenia równoważności między wnioskiem praktycznej deliberacji a intencją, wolicją czy działaniem. Jednak, żeby taka równoważność zachodziła, należy uprzednio założyć (Z2), ponieważ tylko sama teza internalizmu eliminuje możliwość rozluźnienia związku między najlepszym sądem, uzyskanym w procesie deliberacji, a pragnieniem prowadzącym do działania. Dlatego też fakt, że akrates dokonuje ewaluacji, nie stanowi zagrożenia dla eksternalizmu, ponieważ pogląd ten neguje założenie drugie, a nie pierwsze. Wprost odrzuca się tam utożsamienie wniosku rozumowania praktycznego z działaniem, wolicją czy intencją²⁹.

Na gruncie eksternalizmu mamy pełne prawo zgodzić się, że działanie niepowściągliwe jest poprzedzone ewaluacyjnym rozumowaniem praktycznym, a jed-

²⁷ Tamże, s. 159–160.

²⁸ Tamże, s. 155.

²⁹ A. Mele, *Autonomous Agents: From Self-Control to Autonomy*, New York 1995, s. 15–32.

nocześnie zaprzeczyć twierdzeniu, że wnioskowi tego rozumowania odpowiada korespondujące pragnienie. Wracając do przykładu z wprowadzenia, należałoby wskazać, iż: eksternalista nie twierdzi, że nie oceniam wzięcia kolejnego kawałka tortu jako dobrego lub złego, wartego lub niewartego realizacji, ale że mogę uznać, iż nie warto dalej jeść, a mimo to pragnąć tego i poddać się temu pragnieniu. Niech więc będzie to pierwszy warunek rozwijanej tutaj analizy: gdy sprawca S niepowściągliwie robi x , to (W1) S osądza, że lepiej byłoby zrobić y niż x .

4. Argument z przymusowości

Drugi jest argument z przymusowości działań akratycznych:

- (A2) Faktem jest, że niekiedy siła motywacyjna pragnienia nie koresponduje z ewaluacją jej obiektu. Działanie spowodowane takim pragnieniem nie jest jednak wolne. Na gruncie eksternalizmu nie da się zatem odróżnić działań niepowściągliwych od działań pod przymusem³⁰.

W formie, w jakiej ten argument pojawia się w tekście Sergia Tenenbauma, towarzyszy mu uściślenie (Z2). Autor zgadza się z klasycznym stanowiskiem w tym względzie, ponieważ uznaje zasadę sokratejską za wyraz związku pojęciowego, ale uściśla ją, dodając, że ewaluacja dokonywana jest zawsze z określonej perspektywy. Ta zaś może ulegać zmianie, toteż to, co uważam dziś za warte realizacji, jutro mogę uznać za całkowicie bezwartościowe³¹. Uwzględniając tę sugestię, precyzyjne sformułowanie (Z2) wygląda następująco:

- (Z2**) Osąd z danego punktu widzenia sprawcy S, że lepiej byłoby zrobić x , niż zrobić y , jest równoważny z pragnieniem, by zrobić x raczej, niż zrobić y .

³⁰ W różnych formach ten lub podobny argument pojawia się między innymi w: S. Tenenbaum, *The Judgement...*, dz. cyt., s. 875–911; S. Buss, *Weakness of Will*, „Pacific Philosophical Quarterly” 1997, t. 78, nr 1, s. 13–44; J. Raz, *Engaging Reason: On the Theory of Value and Action*, Oxford 1999, s. 71; G. Watson, *Skepticism about Weakness of Will*, „The Philosophical Review” 1977, t. 86, nr 3, s. 316–339.

³¹ S. Tenenbaum, *The Judgement...*, dz. cyt., s. 882.

Dzięki temu internalista może odrzucić twierdzenie, jakoby ewaluacja przedmiotów pragnień była niezmienna w eksperymentach, które wskazują na wzrost pragnienia wraz ze zbliżającą się gratyfikacją³². Może on twierdzić, że w opisanych tam przypadkach dochodzi także do zmiany w ewaluacji.

Podstawowy argument Tenenbauma przeciwko eksternalizmowi jest następujący: jeśli sprawca czegoś pragnie, to ocenia to jako pod jakimś względem dobre. Ta ewaluacja jest powodem, dla którego sprawca działa. W przypadku nieistnienia lub zbyt małej wagi takiego powodu działanie nie jest, w pełnym tego słowa znaczeniu, działaniem sprawcy. Jeśli stoi za nim tylko siła motywacyjna pozbawiona korespondującej oceny, podobne jest ono raczej do kompulsywnego wykonywania pewnej czynności, które nie jest pod kontrolą sprawcy. Jeśli więc siła motywacyjna pragnienia, żeby zrobić *y*, przeważająca nad pozytywną ewaluacją przypisaną zrobieniu *x*, prowadzi sprawcę do *y*-owania, to nie podejmuje on wolnego wyboru, żeby zrobić *y*, ale zostaje do tego zmuszony przez ogarniające go pragnienie³³.

W tym miejscu eksternalista włącza się do dyskusji, podnosząc sprzeciw, że istnieje przecież coś takiego jak samokontrola. Należy jednak rozróżnić dwa rodzaje sytuacji: w jednych możemy powstrzymać się od działania, a w innych nie mamy takiej możliwości. Do określenia lwiej części działań niepowściągliwych³⁴ wystarczy z pierwszej grupy wyróżnić te sytuacje, w których sprawca nie kontroluje się – pobłaża sobie – i decyduje się na działanie wbrew najlepszemu sądowi³⁵.

Tenenbaum kontrargumentuje, że takie rozwiązanie prowadzi do regresu *ad infinitum*. Niepodjęcie samokontroli byłoby również działaniem akratycznym, czyli niepowściągliwością wyższego rzędu. Gdyby sprawca niepowściągliwie *y*-ował, ale miał możliwość oparcia się pragnieniu, żeby zrobić *y*, i wbrew najlepszemu sądowi, żeby nie *y*-ować, nie oparł się mu, byłby to kolejny przejaw słabej woli. Jako taki musiałby znów zostać wyjaśniony przez brak samokontroli wyż-

³² Por. A. Mele, *Backsliding...*, dz. cyt., s. 77–82.

³³ S. Tenenbaum, *The Judgement...*, dz. cyt., s. 881–892.

³⁴ Mele wskazuje również, że istnieją przypadki akrazji, kiedy sprawca nie pobłaża sobie, ale nie jest świadomy możliwości lub stopnia trudności samokontroli (tenże, *Irrationality...*, dz. cyt., s. 22). Tenenbaum odpowiada, że 1. znacząco zawężyłoby to zakres działań niepowściągliwych; 2. wciąż można byłoby twierdzić, że nie byłyby one wolne; 3. nie byłyby irracjonalne (tenże, *The Judgement...*, dz. cyt., s. 889–891). Są to nieklasyczne przypadki akrazji, których omówienie zostawiam na inną okazję.

³⁵ Zob. np. A. Mele, *Irrationality...*, dz. cyt., s. 28.

szego rządu – nieoparcia się pragnieniu, żeby nie oprzeć się pragnieniu, żeby y-ować. I tak w nieskończoność³⁶.

Czy jest to argument na tyle silny, by pogрузić eksternalizm motywacyjny? Nie, a przynajmniej nie w takiej formie, jaką chciałbym tu zaproponować. Znowu rozważmy przykład kuszącego tortu. Czy jeśli decyduję się zjeść więcej niż powinienem, działam niepowściągliwie, a dodatkowo niepowściągliwie nie podejmuję kontroli nad swoim pragnieniem? Wydaje się, że nie. Raczej trzeba tu powiedzieć, że działam niepowściągliwie właśnie dlatego, że nie ograniczam swojego pragnienia. Jediną instancją słabej woli jest ta na poziomie pragnień i to od niej zależy charakter mojego działania – to, czy będzie akratyczne, czy nie. Jeśli niepowściągliwość działań zależy od braku kontroli nad pragnieniami, (A2) zostaje rozbrojony.

Sądzę, że tak właśnie jest. Człowiek nie kontroluje swoich działań bezpośrednio. Prawdziwość (Z1) wyklucza taką opcję. Odpowiednia chęć, wraz z adekwatnym zestawem przekonań, w umożliwiających działaniu okolicznościach, jest kauzalnie wystarczająca do podjęcia go. Tym, na co osoba ma bezpośredni wpływ, są jej pragnienia. Załóżmy, że sprawca chce podjąć działanie x , a jednocześnie uważa, że x nie jest warte realizacji. Wówczas, jak twierdzę, (W2) może on ograniczyć swoje pragnienie zrobienia x .

Oczywiście nie jestem zdania, że człowiek ma pełną kontrolę nad swoimi pragnieniami – to absurd. Najpewniej część działań pozornie niepowściągliwych wcale takimi nie jest, ponieważ pragnienia do nich prowadzące są zbyt silne, aby się im oprzeć (na przykład gdy mowa o uzależnieniach). Także wprost nie definiuję pojęcia ograniczania pragnień, ponieważ domyślam się, że może ono przybierać różne formy (zainteresowanych odsyłam do działu „Rozwój osobisty” w najbliższej księgarni). Nie jestem również przywiązany do samego terminu – równie dobrze można mówić tu o kontrolowaniu, hamowaniu albo trzymaniu pragnień na wodzy czy nawet o wzmaganiu opozycyjnych chęci. Sądzę jednak, że człowiek ma wpływ na to, które z jego pragnień jest dominujące.

Co więc ujmuję tutaj jako ograniczanie pragnień? To, czego istnienie powinien uznać każdy, kto zgadza się, że (Z1) jest prawdziwa, a zatem że za działania kauzalnie odpowiedzialne są określone pragnienia oraz że przynajmniej w niektórych przypadkach sprawca może dobrowolnie wybrać, które działanie podejmie. Nie zakładam tu żadnego rozstrzygnięcia problemu wolnej woli, to-

³⁶ S. Tenenbaum, *The Judgement...*, dz. cyt., s. 888–889.

też argumentacja jest warunkowa: jeśli istnieją wolne działania, to istnieją wolne działania niepowściągliwe.

(W2) przedstawia taki oto obraz podejmowania wyboru: sprawca S może zrobić x lub y . Aby doszło do x -owania, musi zaistnieć odpowiednio silne (dominujące) pragnienie S, by x -ować, analogicznie z y -owaniem. Jeśli zatem S ma dowolnie wybrać między zrobieniem x a zrobieniem y , konieczne jest, aby miał wpływ na to, które pragnienie ostatecznie stanie się dominujące. Wywieranie tego wpływu nazywam ograniczaniem czy kontrolowaniem pragnień.

5. Argument z racjonalności

Trzeci argument dotyczy irracjonalności działań niepowściągliwych i wynikającej z niej konieczności wyjaśnienia zagadkowości słabej woli. Jak pisze Hare:

W rozważanym przez nas stanowisku [eksternalizmie motywacyjnym] nie ma nic dziwniejszego w myśleniu, że coś jest najlepszą rzeczą do zrobienia w danych okolicznościach, i nieczynienia jej niż tym, że ktoś uważa pewien kamień za najbardziej okrągły w okolicy i nie podnosi go, ale podnosi jakiś inny kamień. Nie będzie niczego, co wymagałoby wyjaśnienia, jeśli zdecyduję się zrobić, powiedzmy, najgorszą możliwą rzecz, a to, co najlepsze, pozostawię niezrealizowanym³⁷.

Zarzut inspirowany przywołaną uwagą głosi rzecz następującą:

(A3) Na gruncie eksternalizmu motywacyjnego najlepszy sąd traci swoją uprzywilejowaną rolę. Sprawca może działać w zgodzie z nim, ale może także działać przeciwko niemu. Nie istnieje kryterium, które pozwalałoby na odróżnienie tych dwóch rodzajów działań. Różnica między działaniami niepowściągliwymi a działaniami, które są zgodne z najlepszym sądem sprawcy, przestaje być zatem inteligibilna³⁸.

³⁷ R.M. Hare, *Freedom and Reason...*, dz. cyt., s. 68–69.

³⁸ S. Stroud, L. Svirsky, *Weakness of Will*, w: *The Stanford Encyclopedia of Philosophy*, red. E.N. Zalta, 2021, <https://plato.stanford.edu/archives/win2021/entries/weakness-will/> (dostęp: 1.07.2022).

Wyrażona zostaje tu obawa przed całkowitym zerwaniem związku logicznego między pragnieniem a działaniem i w konsekwencji – utratą odpowiedzi na pytania, dlaczego działania niepowściągliwe są irracjonalne oraz co odróżnia je od innych. W odpowiedzi na ten argument można zaprzeczyć twierdzeniu, że nie ma kryterium, które pozwoliłoby wyznaczyć ścisłą granicę między działaniami niepowściągliwymi a tymi zgodnymi z osądem sprawcy. Należy więc wskazać taki wymóg racjonalności, który byłby naruszany w sytuacji działania akratycznego.

Na trop poszukiwanej reguły naprowadza nas z jednej strony Davidson, a z drugiej Mele. Pierwszy wskazuje tezę, którą określa zasadą powściągliwości: „należy podejmować działanie, które osądza się jako najlepsze na podstawie wszystkich dostępnych i istotnych w danej sytuacji racji”³⁹. Drugi natomiast twierdzi, że sprawca, który ocenia, że najlepiej byłoby zrobić *x*, ma racjonalne zobowiązanie (jest to zobowiązanie ewaluacyjne) do *x*-owania⁴⁰. Czy jednak owo zobowiązanie należy rozumieć jako bezpośrednią relację między ewaluacją a działaniem?

Jak sugerowałem w poprzedniej części, ustalanie kierunku działania jest za pośredniczone przez kontrolę pragnień. Z tego powodu Davidsonowska norma racjonalności powinna zostać przeformułowana. Moja propozycja jest następująca. Na gruncie eksternalizmu motywacyjnego funkcję kryterium odróżniającego działania niepowściągliwe od tych zgodnych z najlepszym sądem może adekwatnie pełnić omawiana tutaj (Z2). Nie może być ona jednak rozumiana jako zasada deskryptywna, opisująca ludzkie działanie, ale jako zasada preskryptywna – norma racjonalności postępowania. Nie twierdzę zatem, że „jeżeli sprawca osądza, że lepiej byłoby zrobić *x*, niż zrobić *y*, wówczas bardziej chce zrobić *x*, niż zrobić *y*”, ale że:

(NR) Jeżeli sprawca osądza, że lepiej byłoby zrobić *x*, niż zrobić *y*, wówczas **powinien chcieć** zrobić *x* bardziej, niż zrobić *y*.

Uwzględniając (Z1), norma ta nie może być gorsza od oryginalnej zasady powściągliwości. Jest tak, ponieważ wyróżnia ona te same działania jako akratyczne. Czyni to jednak z uwzględnieniem pragnienia, będącego ogniwem pośrednim między ewaluacją a działaniem. Nie określa bezpośrednio, jak powinno się

³⁹ D. Davidson, *Jak jest możliwa słabość woli?*, dz. cyt., s. 105.

⁴⁰ A. Mele, *Backsliding...*, dz. cyt., s. 16.

działać, ale jak powinno się pragnąć, aby działać racjonalnie. Do słów Melego, że zazwyczaj przejście od sądu do intencji jest płynne, należy dodać komponent normatywny – jest płynne i powinno takie być.

Warto zwrócić uwagę, w jaki sposób (NR) pozwala ona na wyprowadzenie trzeciego warunku eksternalistycznej analizy działania niepowściągliwego. Powróćmy na przyjęcie, gdzie zastanawiam się, czy dołożyć sobie jeszcze kawałek tortu. Przypuśćmy, że (W1) oceniam, że najlepiej będzie już sobie odpuścić, i jednocześnie (W2) mam możliwość ograniczenia pragnienia, aby sięgnąć po następną porcję. Działam niepowściągliwie, gdy nie kontroluję się, czyli (W3) nie ograniczam swojego pragnienia.

Na mocy zaproponowanej normy (NR) działanie przyczynowo związane z pragnieniem, które nie odpowiada ewaluacji i które nie zostało ograniczone, mimo że istniała taka możliwość, jest działaniem irracjonalnym. W ten sposób odrzucony zostaje (A3). Również zagadkowość słabej woli nie zostaje utracona. Idealnie racjonalny podmiot w dowolnych okolicznościach stosowałby się do (NR). Zawsze gdy miałby możliwość ograniczenia chęci, której nie uważałby za godną realizowania, czyniłby to, a logiczny związek między racjami a działaniem pozostawałby niezachwiany. Człowiek natomiast takim podmiotem nie jest, toteż niekiedy pozwala takim chęciom kierować swoim działaniem. Wówczas związek logiczny zostaje zerwany, co skutkuje irracjonalnością w postaci słabej woli.

6. Podsumowanie

Sprawca o słabej woli nie sprawuje kontroli nad własnymi pragnieniami. Gdy chce coś zrobić, ma kontrolę nad swoim działaniem, ale nie nad pragnieniem prowadzącym go do działania. Brakuje mu wolności woli w rozumieniu zaproponowanym przez Harry'ego Frankfurta: wola tego człowieka nie jest wolna, bo nie „dysponuje on wolnością pragnienia tego, czego chce pragnąć”⁴¹, a przy tym jego działanie jest wolne, ponieważ „pytanie o wolność jakiegoś działania podejmowanego przez podmiot zależy od tego, czy jest to działanie, które chce on podjąć”⁴².

⁴¹ H. Frankfurt, *Wolność woli i pojęcie osoby*, tłum. J. Nowotniak, w: *Filozofia moralności...*, dz. cyt., s. 33.

⁴² Tamże; por. J. Hołówka, *Etyka...*, dz. cyt., s. 67–71.

Korzenie działania niepowściągliwego dają się znaleźć w sferze pragnień, a dokładniej w niedoborze wysiłku wkładanego w ich ograniczanie. Naszkicowana tu koncepcja pozwala na uznanie, że w rozważaniu praktycznym ma miejsce ewaluacja, a także umożliwia odparcie argumentów z regresu i z braku kryterium wyróżnienia działań akratycznych. Rozwiązanie to wpisuje się w nurt eksternalizmu motywacyjnego.

Eksternalizm ma proste wyjaśnienie zjawiska słabej woli – człowiek działa niepowściągliwie, gdy jego pragnienia nie odpowiadają ewaluacji obiektu tego pragnienia – ale wymaga odrzucenia zasady sokratejskiej. Celem tej pracy było pokazanie, w jaki sposób można tego dokonać, przyjmując tezę o kauzalnej relacji między pragnieniami a działaniem, wyrażoną w (Z1), ale także nie ignorując logicznej struktury działania opisywanej przez Davidsona. Rozważenie argumentów przeciwko eksternalizmowi doprowadziło do wskazania trzech warunków działania niepowściągliwego. Składałam zatem fragmenty analizy w całość.

Sprawca S x -uje niepowściągliwie wtedy, gdy:

- (W0) S robi x w sposób zamierzony, ponieważ pragnie bardziej zrobić x , niż zrobić y , oraz jest przekonany, że może wybrać pomiędzy x i y .
- (W1) S osądza, że lepiej byłoby zrobić y niż x .
- (W2) S może ograniczyć swoje pragnienie zrobienia x .
- (W3) S nie ogranicza swojego pragnienia zrobienia x .

Warunkom towarzyszy norma racjonalności:

- (NR) Jeżeli sprawca osądza, że lepiej byłoby zrobić y , niż zrobić x , wówczas powinien chcieć zrobić y bardziej, niż zrobić x .

Oczywiście norma ta nie daje się pogodzić z twierdzeniem, że wszystkie działania są racjonalne, toteż należy zgodzić się z Davidsonem, że działanie człowieka niepowściągliwego jest irracjonalnie i jako takie niezrozumiałe nawet dla niego samego⁴³. Wskazana norma pozwala eksternaliście wytłumaczyć, dlaczego tak jest – racjonalność wymaga od nas kontrolowania własnych pragnień. A jeśli możemy kontrolować nasze pragnienia, możemy ograniczać również te, które prowadzą do działań niepowściągliwych.

⁴³ D. Davidson, *Jak jest możliwa słabość woli?*, dz. cyt., s. 105–106.

Już po raz ostatni pozwolę sobie wrócić do przykładu niepowściągliwego dokładania sobie tortu i przepiszę go zgodnie z tym, co zostało już powiedziane. Wygląda to tak: rozważając, czy wziąć następny kawałek, osądzam, że najlepiej będzie powstrzymać się od jedzenia (W1). Mam ochotę dalej jeść, ale jestem w stanie – odwracając wzrok, intensywnie myśląc o chorobach układu krążenia czy skupiając się na rozmowie z kimś obok – ograniczyć to pragnienie (W2). Nie robię jednak tego (W3), a chęć ta prowadzi mnie do wzięcia jeszcze jednego kawałka (W0). Nie kontroluję się, więc łamię normę (NR), głoszącą, że powinienem był zahamować pragnienie, które doprowadziło mnie do działania niepowściągliwego. W ten sposób wyrzekam się racjonalności, a „w swym własnym zamierzonym postępowaniu odkrywa[m] coś zasadniczo bezrozumnego”⁴⁴.

Bibliografia

- Audi R., *Słabość woli i osąd praktyczny*, tłum. W.J. Popowski, w: *Filozofia moralności: postanowienie i odpowiedzialność moralna*, red. J. Hołówka, Wydawnictwo SPACJA–Fundacja Aletheia, Warszawa 1997.
- Bratman M., *Intention, Plans, and Practical Reasons*, Harvard University Press, Cambridge, MA 1987, <https://doi.org/10.2307/2214395>.
- Bratman M., *Practical Reasoning and Weakness of Will*, „Noûs” 1979, t. 13, nr 2.
- Buss S., *Weakness of Will*, „Pacific Philosophical Quarterly” 1997, t. 78, nr 1, <https://doi.org/10.1111/1468-0114.00026>.
- Davidson D., *Actions, Reasons, and Causes*, w: tegoż, *Essays on Actions and Events*, Clarendon Press, Oxford 1980.
- Davidson D., *Jak jest możliwa słabość woli?*, tłum. W.J. Popowski, w: *Filozofia moralności: postanowienie i odpowiedzialność moralna*, red. J. Hołówka, Wydawnictwo SPACJA–Fundacja Aletheia, Warszawa 1997.
- Davidson D., *Paradoxes of Irrationality*, w: tegoż, *Problems of Rationality*, Clarendon Press, Oxford 2004, <https://doi.org/10.1093/0198237545.003.0011>.
- Frankfurt H., *Wolność woli i pojęcie osoby*, tłum. J. Nowotniak, w: *Filozofia moralności: postanowienie i odpowiedzialność moralna*, red. J. Hołówka, Wydawnictwo SPACJA–Fundacja Aletheia, Warszawa 1997.

⁴⁴ Tamże, s. 106.

- Hare R.M., *Assenting to a Value Judgement*, w: *Weakness of Will*, red. G. Mortimore, MacMillan, London 1971.
- Hare R.M., *Freedom and Reason*, Clarendon Press, Oxford 1963.
- Hołówka J., *Etyka w działaniu*, Prószyński i S-ka, Warszawa 2002.
- McIntyre A., *Is Acratic Action Always Irrational*, w: *Identity, Character and Morality*, red. O. Flanagan, A.O. Rorty, MIT Press, Cambridge 1993.
- Mele A., *Autonomous Agents: From Self-Control to Autonomy*, Oxford University Press, New York 1995.
- Mele A., *Backsliding: Understanding Weakness of Will*, Oxford University Press, New York 2012.
- Mele A., *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, Oxford University Press, New York 1987.
- Mele A., *Springs of Action*, Oxford University Press, New York 1992.
- Platon, *Protagoras*, tłum. L. Regner, Wydawnictwo Naukowe PWN, Warszawa 1995.
- Raz J., *Engaging Reason: On the Theory of Value and Action*, Oxford University Press UK, Oxford 1999.
- Shafer-Landau R., *A Defense of Motivational Externalism*, „Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition” 2000, t. 97, nr 3.
- Stroud S., Svirsky L., *Weakness of Will*, w: *The Stanford Encyclopedia of Philosophy*, red. E.N. Zalta, 2021, <https://plato.stanford.edu/archives/win2021/entries/weakness-will/>.
- Swanton Ch., *Słabość woli jako rodzaj braku odwagi wykonawczej*, tłum. W.J. Popowski, w: *Filozofia moralności: postanowienie i odpowiedzialność moralna*, red. J. Hołówka, Wydawnictwo SPACJA–Fundacja Aletheia, Warszawa 1997.
- Tenenbaum S., *The Judgement of a Weak Will*, „Philosophy and Phenomenological Research” 1999, t. 59, nr 4, <https://doi.org/10.2307/2653561>.
- Watson G., *Free Agency*, „The Journal of Philosophy” 1975, t. 72, <https://doi.org/10.2307/2024703>.
- Watson G., *Skepticism about Weakness of Will*, „The Philosophical Review” 1977, t. 86, nr 3, <https://doi.org/10.2307/2183785>.

Streszczenie

Jedną z fundamentalnych kwestii podnoszonych w dyskusji nad problemem słabej woli jest spór o zasadę internalizmu motywacyjnego. Zgodnie z jej treścią osąd sprawcy, że lepiej wykonać pewne działanie, pociąga chęć wykonania tego działania. W artykule omawiam argumenty wysuwane w obronie tej zasady, aby dowieść, że nie dostarczają one wystarczająco dobrych racji na rzecz jej wiarygodności. Na podstawie tych argumentów formułuję warunki działania akratycznego, które składają się na eksternalistyczną analizę słabej woli. Zaproponowana koncepcja ma pogodzić dwa twierdzenia: po pierwsze, że praktyczny osąd nie jest tożsamy z pragnieniem, oraz po drugie, że w przypadku zwyczajnych działań istnieje związek logiczny między przyczynami działania a samym działaniem. W tym celu wskazana zostaje norma racjonalności, która jest naruszana przy działaniu akratycznym.

Słowa kluczowe: słaba wola, akrazja, eksternalizm motywacyjny, zasada sokratejska, Donald Davidson

Summary

Weak Will – Towards a Motivational Externalism

One of the fundamental issues raised in the discussion of the problem of weak will is the dispute over the principle of motivational internalism. According to the principle, an agent's judgment that it is better to perform a certain action entails a desire to perform that action. In this paper, I discuss the arguments for the principle in order to show that they do not provide sufficient reasons for its plausibility. On the basis of these arguments, I formulate the conditions for acratik action in order to provide an externalist analysis of weak will. The aim is to reconcile two claims: that practical judgment is not equivalent to desire, and that in the case of ordinary actions there is a logical connection between the causes of the action and the action itself. For this purpose, the norm of rationality that is violated in acratik action is identified.

Key words: weakness of will, akrasia, motivational externalism, Socratic principle, Donald Davidson