

Bernard Williams

Wewnętrzne i zewnętrzne racje do działania

Etyka 43, 7-18

2010

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

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

Wewnętrzne i zewnętrzne racje do działania¹

Zdania o postaci „A ma powód, żeby φ ” albo „Istnieje racja dla A, żeby φ ” (gdzie φ zastępuje jakiś czasownik oznaczający działanie) można na pierwszy rzut oka interpretować dwojako². Przy pierwszej interpretacji prawdziwość takiego zdania oznacza, najogólniej rzecz biorąc, że A ma pewien motyw, któremu odpowiada czy służy podjęcie działania φ , a jeśli się okaże, że tak nie jest, to zdanie jest fałszywe; istnieje pewien warunek związany z celami podmiotu, i jeśli nie jest on spełniony, to zgodnie z tą interpretacją nie jest prawdą, że podmiot ma powód, żeby φ . Przy drugiej możliwej interpretacji nie ma takiego warunku, a brak odpowiedniego motywu nie falsyfikuje zdania o istnieniu racji do działania. Interpretacje te będę nazywał odpowiednio „wewnętrzną” i „zewnętrzną”. (Mając do dyspozycji te dwie interpretacje i dwie przytoczone na początku formy zdania, można rozsądnie założyć, że dla pierwszego zdania bardziej naturalna jest interpretacja wewnętrzna, dla drugiego zaś zewnętrzna, ale nie jest prawdą, jakoby którekolwiek z tych sformułowań dopuszczało tylko jedną interpretację). Będę też czasem dla wygody odwoływał się do „racji wewnętrznych” i „racji zewnętrznych”, jak to czynię w tytule, ale tak też, jako zwykłe udogodnienie, należy to traktować. Trzeba dopiero zbadać, czy istnieją dwa rodzaje racji do działania, odpowiadające dwóm rodzajom zdań na temat posiadanych przez ludzi racji do działania; w istocie, jak się przekonamy, kłopotliwe jest zinterpretowanie choćby jednego z tych przypadków. Zajmę się na początek interpretacją wewnętrzną, by zobaczyć, jak daleko można ją posunąć. Następnie, z większą dozą sceptycyzmu, rozważę możliwe implikacje interpretacji zewnętrznej. Zakończę kilkoma uwagami na temat związku między tym problemem a kwestią dóbr publicznych i tzw. problemem pasażera na gapę.

Najprostszy model interpretacji wewnętrznej wyglądałby tak: A ma powód, żeby φ , wtedy i tylko wtedy, gdy A ma jakieś pragnienie, którego zaspokojeniu się przysłuży, jeśli podejmie działanie φ . Można by też powiedzieć: jakieś pragnienie,

¹ Na podstawie przedruku „Internal and External Reasons”, w: *Varieties of Practical Reasoning*, ed. E. Millgram, Cambridge, Mass., MIT Press 2001, s. 77–89. Pierwodruk w: *Rational Action*, ed. R. Harrison, Cambridge University Press 1979. Tłumaczenie za zgodą spadkobierców autora i CUP. Published with permission.

² Ponieważ „racja do działania” (*reason to act*) jest terminem technicznym i razi w potocznej polszczyźnie, zwrot ten zachowano tylko w kontekstach technicznych. W pozostałych miejscach, zwłaszcza tam, gdzie Williams pisze, że dany sposób mówienia jest naturalny lub nienaturalny, konsekwentnie tłumaczono „powód/powody” (przyp. tłum.).

co do którego A jest przekonany, że ϕ posłuży jego zaspokojeniu; różnicą między tymi sformułowaniami zajmiemy się później. Taki model przypisuje się czasem Hume'owi, ale ponieważ poglądy Hume'a na tę kwestię są w rzeczywistości bardziej złożone, możemy to nazwać *modelem sub-Humowskim*. Model sub-Humowski jest z pewnością zbyt prosty. Moim celem będzie zrewidowanie go i rozbudowanie w coś bardziej adekwatnego; zacznę od zestawienia ze sobą trzech twierdzeń na temat zdań o racjach wewnętrznych.

Zasadniczo, i niejako z definicji, każdy model interpretacji wewnętrznej musi ukazywać zależność stwierdzenia istnienia racji do działania od *subiektywnego układu motywacyjnego* podmiotu, który będę nazywał jego S. Do zawartości S jeszcze dojdziemy, ale już teraz możemy powiedzieć, że:

(i) Stwierdzenie istnienia racji do działania zostaje sfalsyfikowane przez brak odpowiedniego elementu w S podmiotu.

Zgodnie z najprostszym modelem sub-Humowskim, dowolny element S dostarcza podmiotowi wewnętrznej racji do działania. Są jednak podstawy, by temu zaprzeczyć, nie ze względu na godne pożałowania, nieroztropne czy odbiegające od normy elementy w S — te sprawiają problemy innego rodzaju — ale z powodu elementów S opartych na błędnych przekonaniach. Ktoś jest przekonany, że to, co ma przed sobą, to gin, gdy w rzeczywistości to benzyna. Chce się napić ginu z tonikiem. Czy zrobiłby słusznie, gdyby zmieszał to coś z tonikiem i wypił? Czy ma powód, żeby to zrobić? Można podejść do tego problemu na dwa sposoby (jak sugerują już alternatywne sposoby formułowania modelu sub-Humowskiego). Z jednej strony powiedzenie, że ma on powód, żeby wypić tę mieszankę, brzmi po prostu bardzo dziwnie, naturalnie jest natomiast powiedzieć, że nie ma powodu, by to wypić, chociaż sądzi, że ma. Z drugiej strony, jeśli mimo wszystko podmiot to wypije, mamy nie tylko wytłumaczenie, dlaczego to zrobił (powód, dla którego to zrobił), lecz także wytłumaczenie w kategoriach racji do działania. Ich wymiar wyjaśniający jest bardzo istotny i nieraz jeszcze do niego wrócimy; jeśli istnieją racje do działania w pewien sposób, to ludzie muszą czasami działać ze względu na te racje, a jeśli to robią, ich racje muszą figurować w jakimś poprawnym wyjaśnieniu ich działania (z czego nie wynika, że muszą się znaleźć we wszystkich poprawnych wyjaśnieniach ich działania). Różnica pomiędzy fałszywymi a prawdziwymi przekonaniem po stronie podmiotu nie może zmieniać *formy* wyjaśnienia, które będzie odpowiednie dla jego działania. Ten fakt może nas skłonić do zignorowania intuicji, na którą poprzednio zwróciliśmy uwagę i doprowadzić nas wprost do uznania, że w przypadku podmiotu, który chce się napić ginu, ma on powód, by napić się stojącej przed nim benzyny.

Sądzę jednak, że nie powinniśmy tego robić. To krok w złym kierunku, sugerujący ostatecznie, że koncepcja racji wewnętrznych dotyczy tylko wyjaśnienia działania, a nie racjonalności podmiotu; co z kolei może skłaniać do poszukiwania innego rodzaju racji, związanych z jego racjonalnością. Tymczasem ta koncepcja

jak najbardziej dotyczy racjonalności podmiotu: to, co poprawnie możemy mu przypisać w trzecioosobowym stwierdzeniu istnienia racji wewnętrznej, to także, jak zobaczymy, on sam może sobie przypisać w wyniku namysłu. Dlatego sądzę, że powinniśmy raczej powiedzieć:

(ii) Element D, należący do S, nie da A powodu, żeby φ , jeśli istnienie D zależy od błędnego przekonania A lub jeśli błędne jest przekonanie A o znaczeniu φ dla zaspokojenia D.

(To podwójne sformułowanie możemy zobrazować na przykładzie sprawy ginu/benzyny: D może być potraktowane jako pragnienie napicia się tego, co znajduje się w butelce, albo jako pragnienie napicia się ginu). Mimo wszystko pozostanie prawdą, że jeśli podmiot w tych okolicznościach podejmie działanie φ , to nie tylko był po temu powód, ale, biorąc pod uwagę jego błędne przekonanie, podmiot zachowa się racjonalnie.

Możemy zanotować epistemiczną konsekwencję powyższego:

(iii) (a) A może być błędnie przekonany o prawdziwości jakiegoś zdania przypisującego mu rację wewnętrzną — a także (możemy dodać):

(b) A może nie znać jakiegoś prawdziwego zdania o posiadanej przez siebie racji wewnętrznej.

Punkt (b) pochodzi z dwóch różnych źródeł. Po pierwsze A może nie znać jakiegoś faktu tego rodzaju, że gdyby go znał, to ze względu na jakiś element swojego S byłby skłonny podjąć działanie φ — możemy powiedzieć, że ma on powód, żeby φ , chociaż o tym nie wie. Wydaje się jednak, że aby rzeczywiście miał jakiś powód, związek pomiędzy nieznanym faktem a jego działaniem musi być względnie bliski i bezpośredni; inaczej powiedzielibyśmy po prostu, że A miałby powód, żeby φ , gdyby znał ten fakt. Nie będę dalej badał kwestii, od czego zależy wyrażenie się w jeden lub drugi sposób, ale musi ona być ściśle związana z pytaniem o to, kiedy niewiedza jest częścią wyjaśnienia faktycznego działania A.

Drugim źródłem (iii) jest fakt, że A może sobie nie zdawać sprawy z istnienia jakiegoś elementu w swoim S. Musimy jednak zauważyć, że nieznaną element S, D, da A powód, żeby φ , tylko pod warunkiem, że φ jest racjonalnie związane z D; to znaczy, najogólniej rzecz biorąc, że podjęcie działania φ mogłoby być odpowiedzią na wymagające namysłu pytanie, jakie powstaje po części w następstwie istnienia D. Jeżeli D nie jest A znane, ponieważ należy do jego podświadomości, może nie spełniać tego warunku, chociaż oczywiście może stanowić powód, dla którego A podejmuje działanie φ , to jest może wyjaśniać lub pomagać wyjaśnić jego zachowanie. W takich przypadkach φ może być związane z D jedynie symbolicznie.

Jak już powiedziałem:

(iv) stwierdzenia o istnieniu racji wewnętrznych mogą być odkrywane w procesie namysłu.

Warto w tym miejscu zwrócić uwagę na coś, co już milcząco zakładaliśmy — stwierdzenie istnienia racji wewnętrznej odnosi się nie tylko do tego jedyne-

preferowanego działania, które jest ostatecznym wynikiem namysłu. „A ma powód, żeby φ ” nie oznacza „ φ to działanie, dla którego A ma ostateczną, uwzględniającą wszystkie okoliczności rację”. Może on mieć powód zrobić mnóstwo rzeczy, a zarazem mieć inny, silniejszy powód, żeby ich nie robić.

Model sub-Humowski zakłada, że φ musi się odnosić do jakiegoś elementu S jako przyczynowy środek do celu (chyba że samo realizowanie pragnień jest wprost tym elementem S). Ale to jest tylko jedna z możliwości: w istocie zwykle odkrycie, że podjęcie jakiegoś działania jest przyczynowym środkiem do celu, nie jest samo w sobie przypadkiem myślenia praktycznego³. Dobrym przykładem myślenia praktycznego jest namysł prowadzący do konkluzji, że ma się powód, aby φ , ponieważ byłby to najwygodniejszy, najprostszy, najprzyjemniejszy itp. sposób zaspokojenia jakiegoś elementu S; a o tym decyduje oczywiście jakiś inny element S, choć niekoniecznie w bardzo jasny i wyraźny sposób. Ale namysł to także m.in. zastanawianie się, jak zaspokajanie elementów S uporządkować np. w czasie; gdy zachodzi jakiś nieprzewidywalny konflikt pomiędzy elementami S, rozważanie, który z nich jest dla nas ważniejszy (nie oznacza to jednak, co istotne, że istnieje jakiś jeden rodzaj dobra, którego one w różnych ilościach dostarczają); czy wreszcie poszukiwanie konstruktywnych rozwiązań, na przykład zdecydowanie, jak przyjemnie spędzić wieczór, o ile właśnie na to mamy ochotę.

W wyniku tego rodzaju procesów podmiot może dojść do wniosku, że ma powód zrobić coś, dla czego wcześniej nie widział żadnego powodu. W ten sposób proces namyslenia się może podsunąć nam nowe działania, dla których istnieją racje wewnętrzne, a także nowe racje wewnętrzne dla danych działań. Proces namyslenia się może również usuwać elementy z samego S. Refleksja może ukazać podmiotowi, że jakieś przekonanie jest błędne, tak że uświadomi on sobie, iż w istocie nie ma żadnego powodu robić tego, o czym do tej pory sądził, że ma powód to zrobić. Dokładniej, podmiot może myśleć, że ma powód, by nadać wydarzeniom określony kierunek, ponieważ w niewystarczającym stopniu wysilił wyobraźnię, jeśli chodzi o przewidzenie możliwych rezultatów tego procesu. W samodzielnym namyśle, albo zachęcony perswazją ze strony innych, może zyskać bardziej konkretne wyobrażenie o implikacjach tej nowej sytuacji i stracić zainteresowanie jej urzeczywistnieniem; podobnie, w wymiarze pozytywnym, wyobraźnia może stwarzać nowe możliwości i nowe pragnienia. (To ważne możliwości zarówno dla polityki, jak i dla indywidualnych działań jednostek). Nie powinniśmy zatem myśleć o S jako o statycznej strukturze. Proces deliberacji może mieć najróżniejszy wpływ na S, i jest to fakt, który teoria racji wewnętrznych powinna z radością uwzględnić. Powinna też być o wiele bardziej liberalna, niż była w wydaniu niektórych teoretyków, jeśli chodzi o elementy, jakie mogą składać się na S. Omawiałem S przede wszystkim

³ Por. A. Kolnai, *Deliberation is of Ends*, [w:] *Ethics, Value and Reality* (London–Indianapolis 1978); D. Wiggins, *Deliberation and Practical Reason*, PAS 76 (1975–1976), przedrukowane częściowo w: J. Raz (ed.), *Practical Reasoning* (Oxford 1978).

w kategoriach pragnień, i jako termin formalny pragnienie nadaje się na określenie dowolnego elementu S. Ale ta terminologia może sprawić, że zapomnimy, iż S może zawierać takie rzeczy, jak dyspozycje do wydawania pewnego rodzaju ocen, wzorce reakcji emocjonalnych, osobiste zobowiązania i rozmaite projekty, jak możemy je abstrakcyjnie określić, wyrażające wymiary osobistego zaangażowania podmiotu. Nade wszystko trzeba pamiętać, że nie zakłada się tu oczywiście, jakoby pragnienia czy projekty podmiotu musiały być egoistyczne; można mieć nadzieję, że będzie on miał różnego rodzaju nieegoistyczne projekty, w nie mniejszym stopniu dostarczające mu wewnętrznych racji do działania.

W związku z zawartością S można jednak postawić kolejne pytanie: czy zgodnie z ogólną ideą racji wewnętrznych powinniśmy zakładać, że S obejmuje *potrzeby*. Z pewnością zupełnie naturalnie jest powiedzieć, że A ma powód, by zabiegać o X, po prostu na tej podstawie, że A potrzebuje X; ale czy ten wniosek wypływa naturalnie z teorii racji wewnętrznych? Osobny problem pojawia się tu tylko w przypadku, gdy potrzeby podmiotu nie znajdują wyrazu w jego pragnieniach czy motywacjach: jeśli podmiot może nie mieć motywacji do zabiegania o to, czego potrzebuje. Nie będę tu próbował rozwiązać problemu natury potrzeb, ale zakładam, że o ile istnieją możliwe do jasnego wyodrębnienia potrzeby, o tyle zachodzi taka możliwość i możliwy jest podmiot niezainteresowany zabieganiem o to, czego rzeczywiście potrzebuje. Zakładam dalej, iż ten brak zainteresowania może przetrwać namysł; i że błędem byłoby utrzymywanie, że taki brak zainteresowania musi zawsze opierać się na jakimś błędnym przekonaniu. (O ile opiera się on na błędnym przekonaniu, możemy to podciągnąć pod (ii) w omówiony wyżej sposób). Jeżeli podmiot naprawdę nie jest zainteresowany zabieganiem o to, czego potrzebuje, i nie jest to skutek błędnych przekonań na temat faktów, i nie był w stanie wyprowadzić takiego motywu z motywów już posiadanych w opisanym wyżej procesie namysłu, wówczas, moim zdaniem, musimy powiedzieć, że w sensie wewnętrznym rzeczywiście nie ma on żadnego powodu, żeby zabiegać o te rzeczy. Mówiąc to, musimy jednak pamiętać, jak silne są to założenia i jak rzadko skłonni bylibyśmy uznać, że wiemy, iż są one prawdziwe. Kiedy mówimy, że jakaś osoba ma powód, by zażyć lekarstwo, którego potrzebuje, chociaż konsekwentnie i w przekonujący sposób zaprzecza, by zależało jej na zachowaniu zdrowia, możemy nadal mówić to w sensie wewnętrznym, z myślą, że tak naprawdę na jakimś poziomie ta osoba *musi* chcieć być zdrowa.

Jeśli jednak jesteśmy pewni, że niczego takiego nie mamy na myśli, a nadal chcemy twierdzić, że ta osoba ma taki powód, wtedy z pewnością mówimy to w innym sensie, w sensie zewnętrznym. Ludzie rzeczywiście mówią rzeczy, które aż się proszą o zewnętrzną interpretację. W opowieści Jamesa o Owenie Wingrave'ie, na podstawie której Britten napisał operę, ojciec Owena próbuje go przekonać, jako do rzeczy najwyższej wagi i koniecznej, że powinien wstąpić do wojska, skoro wszyscy jego męscy przodkowie byli żołnierzami i rodowa duma wymaga od niego tego

samego. Owen Wingrave nie ma żadnej motywacji, by się zaciągnąć, przeciwnie, wszystkie jego pragnienia zdążają w innym kierunku: nienawidzi wojskowego życia i wszystkiego, z czym się ono wiąże. Jego ojciec mógłby się wyrazić w ten sposób, że *Owen ma powód, by pójść do wojska*. Świadomość, że w S Owena nie ma niczego, co poprzez rozsądny namysł doprowadziłoby go do takiej decyzji, nie skłoniłaby jego ojca do cofnięcia tych słów czy przyznania, że się pomylił. Jego zdanie należy rozumieć w sensie zewnętrznym. Czyli jakim?

Na początek trzeba zaznaczyć, że nie jest to ten sam problem, co ten, jaki niesie ze sobą pytanie o status domniemanego kategorycznego imperatywu, w kantowskim sensie „powinności”, która dotyczy podmiotu niezależnie od tego, czego akurat zdarzyło mu się chcieć: a raczej, nie można mieć pewności, że to ten sam problem. Po pierwsze, imperatyw kategoryczny rozumie się często, i tak rozumiał go Kant, jako imperatyw ściśle moralny, a tymczasem zdania na temat racji zewnętrznych niekoniecznie dotyczą moralności. Po drugie, niejasna pozostaje relacja pomiędzy „istnieje racja dla A, żeby...” a „A powinien...” Niektórzy filozofowie uważają je za równoważne i z tej perspektywy kwestia racji zewnętrznych okazuje się oczywiście o wiele bliższa problemowi statusu imperatywu kategorycznego. Nie będę tu jednak wypowiadał się na temat tej rzekomej równoważności ani poświęcał więcej miejsca kwestii „powinności”.

Rozważając, co miałyby oznaczać przypisanie komuś racji zewnętrznej, musimy następnie pamiętać o funkcji możliwego wyjaśnienia, właściwej każdej racji do działania: jeśli coś może być racją do działania, to mogłoby być powodem, dla którego ktoś w określonej sytuacji postąpił w określony sposób, a wówczas byłoby częścią wyjaśnienia tego działania. Tymczasem żadna racja zewnętrzna nie może *sama w sobie* dostarczyć wyjaśnienia czyjegokolwiek działania. Nawet gdyby było prawdą (cokolwiek może to oznaczać), że Owen miał powód, żeby pójść do wojska, sam ten fakt nie mógłby nigdy wyjaśnić żadnego jego działania, nawet pójścia do wojska. Ponieważ, jeśli w ogóle jest to prawdą, było prawdą także wtedy, gdy Owen nie miał żadnej motywacji, żeby się zaciągnąć. W stwierdzeniach istnienia racji zewnętrznych chodzi o to, że mogą być prawdziwe niezależnie od motywacji podmiotu. Ale nic nie może wyjaśnić (intencjonalnego) działania podmiotu, z wyjątkiem czegoś, co zmotywowało go do takiego działania. Zatem, by wyjaśnić działanie, potrzeba jeszcze czegoś poza prawdziwością stwierdzenia istnienia racji zewnętrznej, jakiegoś psychologicznego łącznika; a tym psychologicznym łącznikiem wydaje się przekonanie. Fakt, że A jest przekonany o prawdziwości odnoszącego się do niego stwierdzenia istnienia racji zewnętrznej, może pomóc wyjaśnić jego działanie.

Zdania o racjach zewnętrznych wprowadziliśmy po prostu w ogólnej formie „istnieje racja dla A, żeby...”; teraz jednak musimy pójść dalej, ku konkretnym stwierdzeniom istnienia racji. Bez wątplenia zdarza się, że podmiot podejmuje działanie ϕ , ponieważ sądzi, że ma powód, żeby ϕ , a zarazem nie ma żadnych

przekonań co do natury tego powodu. Tak będzie w przypadku, gdy polega na aurytacie osoby, do której ma zaufanie, albo pamięta, że znał jakiś powód, żeby φ , ale w tej chwili nie może go sobie przypomnieć. Pod tym względem racje do działania zachowują się jak racje do przyjmowania pewnych twierdzeń. Ale w przypadku jednych i drugich jest to ewidentnie sytuacja nietypowa: za podstawowy przypadek musimy uznać sytuację, gdy A podejmuje działanie φ nie tylko dlatego, że istnieje jakiś powód, by to zrobić, ale dlatego, że jakieś konkretne względy uznaje za rację przemawiającą za tym, by to zrobić. Na tej zasadzie Owen Wingrave może dojrzeć do decyzji o zaciągnięciu się, ponieważ (teraz) jest przekonany, iż wojskowe tradycje rodzinne stanowią dla niego rację do zachowania się w ten sposób.

Czy przekonanie, iż jakiś konkretny wzgląd stanowi rację do działania w pewien sposób, przynosi ze sobą, czy wręcz konstytuuje, motywację do działania? Jeśli nie, to dotąd nie ruszyliśmy z miejsca. Załóżmy, że tak. To zresztą wiarygodne założenie, przynajmniej jak długo nie zacieśnimy niepotrzebnie związku pomiędzy takimi przekonaniem a dyspozycją do działania w takim stopniu, by wykluczał akraję. W istocie jest to założenie *tak* wiarygodne, że ktoś żyjący takie przekonanie wydaje się podmiotem, któremu możemy słusznie przypisać *wewnętrzną* rację do działania: w swoim S ma teraz odpowiednią motywację. Człowiek, który jest przekonany, że wzgląd na honor rodziny jest racją do działania w pewien sposób, jest człowiekiem z pewną dyspozycją do działania, jak również do wydawania ocen, odczuwania, reakcji emocjonalnych i tak dalej.

Nie wynika stąd, że zdania o racjach zewnętrznych są pozbawione treści. Wynika natomiast, że ich treści nie odkryjemy poprzez analizę stanu kogoś, kto jest przekonany o prawdziwości takiego zdania, i tego, w jaki sposób ten stan wyjaśnia jego działanie, ponieważ jest to po prostu stan osoby, której słusznie można by przypisać *wewnętrzną* rację do działania. Treść stwierdzeń istnienia racji zewnętrznych trzeba będzie ujawnić raczej przez rozważenie, na czym polega dochodzenie do przekonania o prawdziwości takiego zdania — to tu, jeśli gdziekolwiek, ukaże się ich specyfika. Zajmiemy się przypadkiem (w rzeczywistości robimy to już od jakiegoś czasu), w którym rację zewnętrzną przypisuje się osobie, która, jak Owen Wingrave, nie jest jeszcze odpowiednio zmotywowana, a zatem jest kimś, komu nie można słusznie przypisać adekwatnej racji wewnętrznej. (Jako że różnica pomiędzy zdaniem wewnętrznymi a zewnętrznymi polega na odmiennych implikacjach dla tego, kto takie zdanie wygłasza, racje zewnętrzne można oczywiście przypisywać podmiotom, które już są odpowiednio zmotywowane; ale nie jest to interesujący przypadek). Nasz podmiot nie jest jeszcze przekonany o prawdziwości zdania o istnieniu racji zewnętrznej. Jeśli zyska takie przekonanie, zyska też motywację do działania; a zatem nabywanie takiego przekonania musi być ściśle powiązane z uzyskaniem nowej motywacji. Jak to możliwe?

Jesteśmy tu bardzo blisko starego pytania, czy, i w jaki sposób, „rozum może wzbudzić motywację”, pytania, na które, jak wiadomo, negatywnej odpowiedzi

udzielił Hume. Ale w tej formie pytanie jest samo w sobie niejasne i niejasny jest jego związek z dotychczasową argumentacją — bo oczywiście rozum, to znaczy racjonalne procesy, mogą wzbudzić nową motywację, jak to widzieliśmy w omówieniu procesu deliberacji. Co więcej, tradycyjny sposób ujmowania tego problemu także (jak postaram się pokazać) bierze na siebie ciężar wykazania, co należy uznać za „czysto racjonalny proces”, czym nie tylko nie powinien się obarczać, ale co wprost jest sprawą krytyków, którzy chcieliby się sprzeciwić ogólnemu wnioskowi Hume’a i nadać zdaniom o racjach zewnętrznych znaczącą treść — tych, których będę nazywał „teoretykami racji zewnętrznych”.

Podstawowym krokiem jest rozpoznanie faktu, że teoretyk racji zewnętrznych musi w szczególny sposób rozumieć związek pomiędzy nabywaniem nowej motywacji a uzyskiwaniem przekonania o prawdziwości zdania o istnieniu racji do działania. Bo oczywiście podmiot może na wiele różnych sposobów dojść do uzyskania motywacji, jak również do przekonania o prawdziwości zdania o istnieniu racji do działania, ale mogą to być sposoby, które nie zainteresują teoretyka racji zewnętrznych. Na przykład Owen mógłby do tego stopnia ulec motywującej retoryce swojego ojca, że uzyskałby zarazem motywację i przekonanie. Ale to wyklucza element, na którym teoretykowi racji zewnętrznych zależy najbardziej: podmiot powinien uzyskać motywację, ponieważ zyskał przekonanie o prawdziwości stwierdzenia istnienia racji do działania, a co więcej, to ostatnie powinno się stać dlatego, że w jakimś sensie podmiot właściwie pojmuje tę sprawę. Jeśli nasz teoretyk chce się upierać przy tych warunkach, to będzie musiał, jak sądzę, sformułować warunek odpowiedniego uzyskania motywacji przez podmiot mniej więcej w ten sposób, że podmiot powinien przeprowadzić poprawny namysł; a samo stwierdzenie istnienia racji zewnętrznej można by uznać za z grubsza równoważne twierdzeniu, że gdyby podmiot dokonał racjonalnego namysłu, to jakiegokolwiek byłyby jego wyjściowe motywacje, ostatecznie uzyskałby motywację, żeby ϕ .

Ale jeżeli to rozumowanie jest poprawne, to rzeczywiście wydaje się, że podstawowa teza Hume’a jest bardzo silna i mamy bardzo dobre podstawy, by uznać, że wszystkie stwierdzenia istnienia racji zewnętrznych są fałszywe. A to dlatego, że, *ex hypothesi*, nie istnieje motywacja, od której podmiot mógłby zacząć swój namysł, aby uzyskać tę nową motywację. Biorąc pod uwagę wcześniejsze motywacje podmiotu i tę jego nową motywację — jeśli stwierdzenie istnienia racji zewnętrznej ma być prawdziwe, to przy tej interpretacji do uzyskania nowej motywacji musiała prowadzić jakaś racjonalna droga, ale zarazem ta nowa motywacja nie powinna pozostawać z wcześniejszymi motywacjami w tego rodzaju racjonalnym związku, który ukazaliśmy, omawiając proces namysłu — bo w takim przypadku przede wszystkim prawdziwe byłoby stwierdzenie istnienia racji wewnętrznej. Nie widzę powodu, by zakładać, że te warunki kiedykolwiek mogłyby zostać spełnione. Można by powiedzieć, że wymowę stwierdzenia istnienia racji zewnętrznej da się wyjaśnić w następujący sposób. Takie stwierdzenie implikuje, że racjonalny podmiot byłby

motywowany do zgodnego z nim działania; i może mieć tę implikację, ponieważ racjonalny podmiot to właśnie taki, który w swoim S ma ogólną dyspozycję, by czynić to, do czego (jak jest przekonany) istnieje dla niego racja. Dlatego jeśli dojdzie do przekonania, że istnieje dla niego racja, żeby φ , będzie zmotywowany, żeby φ , chociaż wcześniej nie miał ani motywu, żeby φ , ani żadnego motywu powiązanego z φ na jeden ze sposobów, które wymieniliśmy, omawiając deliberację.

Ale ta odpowiedź tylko odsuwa problem. Stosuje na nowo przekonaniowo –motywacyjny model (z grubsza rzecz biorąc) wyjaśniania do rozpatrywanego działania, ale posługując się pragnieniem i przekonaniem, których treść należy dopiero wyjaśnić. Co takiego podmiot uznaje, kiedy uznaje, że istnieje dla niego racja, żeby φ , jeśli nie zdanie, albo coś pociągającego za sobą zdanie, że gdyby przeprowadził racjonalny namysł, zyskałby motywację do działania zgodnie z tą racją? Pytaliśmy, w jaki sposób jakiegokolwiek prawdziwe zdanie mogłoby mieć tę treść; nie można odpowiedzieć na to pytanie poprzez odwołanie się do rzekomego pragnienia aktywowanego przez przekonanie o dokładnie takiej treści. Za tymi argumentami w kwestii, co to znaczy akceptować stwierdzenie istnienia racji zewnętrznej, kryje się pewna koncepcja tego, co jest możliwe w ramach zaprezentowanego ujęcia namysłu, a co to ujęcie wyklucza. Można by jednak wysunąć zastrzeżenie, iż takie ujęcie procesu deliberacji jest bardzo niejasne i dopuszcza np. używanie wyobraźni do rozszerzania albo zawężania zawartości S podmiotu. Tymczasem, jeśli to jest możliwe, to nie jest jasne, gdzie znajduje się granica, której podmiot nie może przekroczyć, dokonując racjonalnego namysłu na podstawie swojego aktualnego S. Istotnie, jest to niejasne, lecz zachowanie i wyjaśnienie pochodzenia tej niejasności jest, moim zdaniem, w zasadzie pożądaną cechą teorii namysłu praktycznego. Jeśli chodzi o to, co można uznać za racjonalny proces deliberacji, to występuje tu nieusuwalne niezdeterminowanie. Namysł praktyczny to proces heurystyczny, wymagający wyobraźni, i nie ma sztywnych granic na kontinuum rozciągającym się od racjonalnej myśli po natchnienie i nawrócenie. Dla kogoś, kto sądzi, że racje do działania powinno się zasadniczo rozumieć w kategoriach modelu racji wewnętrznych, nie stanowi to problemu. Rzeczywiście nie do końca jasne jest, co znaczy „A ma powód, żeby φ ”, rozumiane w sensie wewnętrznym, jako że deliberatywne procesy, które mogłyby prowadzić od istniejącego S podmiotu A do uzyskania przezeń motywacji, żeby φ , można pojmować mniej lub bardziej ambitnie. Ale nie jest to żaden kłopot dla tych, którzy za podstawową uważają wewnętrzną koncepcję racji do działania: pokazuje tylko, że liczba stanów, które można by uznać za sytuację, w której A ma powód, żeby φ , jest znacznie większa i mniej określona, niż można byłoby się spodziewać.

To raczej teoretyk racji zewnętrznych ma tu problem. Jest oczywiście mnóstwo rzeczy, które możemy powiedzieć komuś, kto nie chce zrobić φ , kiedy sądzimy, że powinien chcieć: na przykład, że jest niedelikatny, okrutny, samolubny czy nierozważny; albo że sprawy, i on sam, miałyby się dużo lepiej, gdyby miał stosowne motywacje.

To wszystko mogą być bardzo rozsądne uwagi. Ale ktoś, komu szczególnie zależy na nadaniu swojej krytyce formy stwierdzenia istnienia racji zewnętrznej, chce chyba powiedzieć, że tym, co szczególnie zasługuje tu na naganę, jest *irracjonalność* krytykowanej osoby. To właśnie ten teoretyk musi precyzyjnie sformułować swój zarzut: zwłaszcza dlatego, że wymaga, by każdy racjonalny podmiot jako taki uznał wymóg zrobienia ϕ .

Ojciec Owena Wingrave'a w rzeczywistości nie użył słowa „racja”. Ale, jak sobie wyobraziliśmy, mógł posłużyć się formułą racji zewnętrznych; i sam ten fakt przysparza kłopotów teoretykowi racji zewnętrznych. Taki teoretyk, który postrzega prawdziwość stwierdzenia istnienia racji zewnętrznej jako możliwą podstawę zarzucania irracjonalności podmiotowi, który ją ignoruje, mógłby przecież powiedzieć, że gdyby Wingrave *père* nadał swoim skargom na Owena tę formę, to najprawdopodobniej twierdziłby coś, co w tym szczególnym przypadku było fałszywe. Tym jednak, co sprawiłoby teoretykowi racji zewnętrznych jeszcze większą trudność, byłoby wykazanie, że te słowa znaczą coś innego w ustach Wingrave'a, niż znaczą, gdy są, jak zakłada, wypowiedzane trafnie. Tymczasem wypowiedzane przez Wingrave'a prawie na pewno nie znaczą, że racjonalny namysł zmotywowałby Owena do pójścia do wojska — co jest (w dużym uproszczeniu) sensem czy implikacją, jaką w nich odkryliśmy, jeśli mają mieć taką wagę, jaką teoretycy racji zewnętrznych chcieliby im przypisać.

Wydaje mi się, że rozważania, które tu przedstawiłem, silnie przemawiają za tym, że stwierdzenia istnienia racji zewnętrznych, potraktowane odrębnie, są fałszywe, niespójne albo w mylący sposób mówią o czymś innym. W rzeczywistości trudniej napotkać je w rozmowach, niż to sugeruje wprowadzenie ich na początku tego artykułu; ludzie posługujący się takimi zwrotami często wydają się raczej wysuwać optymistyczne tezy na temat posiadanych przez podmiot racji wewnętrznych. Czasami jednak zdecydowanie odnoszą się do czynników spoza S podmiotu, różnych od wszystkiego, co mógłby on wywieść ze swojego S w procesie racjonalnego namysłu, i wówczas, moim zdaniem, powstaje wielka niejasność, co miałyby to znaczyć. Czasem chodzi o niewiele więcej niż o to, że sprawy miałyby się lepiej, gdyby podmiot postąpił tak, a nie inaczej. Ale ujęcie tego w formie zdania o racjach sprawia, zwłaszcza ze względu na sugestię, że podmiot zachowuje się irracjonalnie, że ta sugestia w sytuacji, gdy nie mamy wątpliwości, że nie ma podstaw do stwierdzenia istnienia racji wewnętrznej, to zwyczajny bluff.

Jeśli tak, to jedynymi prawdziwymi twierdzeniami na temat racji do działania będą twierdzenia wewnętrzne.

Problemem, o którym sądzono, że jest bardzo bliski obecnemu tematowi, jest temat dóbr publicznych i pasażerów na gapę, co dotyczy sytuacji (z grubsza rzecz ujmując), w których każdy ma egoistyczny powód, by chcieć, aby pewne dobro zostało zapewnione, ale jednocześnie każdy ma też egoistyczny powód, by nie brać udziału w zapewnieniu tego dobra. Nie będę podejmował próby omówienia

tego problemu, ale być może będzie pomocne — po prostu po to, by rozjaśnić mój pogląd na racje do działania i aby uwydatnić kontrast z innymi poglądami — jeżeli zakończę listą pytań dotyczących tego problemu wraz z odpowiedziami, których udzieliłby ktoś, kto uważa (ujmując rzecz pobieżnie), że jedyną racjonalnością działania jest racjonalność oparta na racjach wewnętrznych.

1. Czy jesteśmy w stanie zdefiniować koncepcje racjonalności, które nie są czysto egoistyczne?

Tak.

2. Czy możemy zdefiniować koncepcje racjonalności w kategoriach innych niż poszukiwanie środków do celu?

Tak.

3. Czy możemy zdefiniować koncepcję racjonalności, w ramach której działanie racjonalne w przypadku A może w żaden sposób nie być związane z jego istniejącymi motywacjami?

Nie.

4. Czy możemy wykazać, że człowiek, który ma jedynie egoistyczne motywacje, postępuje irracjonalnie, nie dążąc do nieegoistycznych celów?

Niekoniecznie, choć w szczególnych przypadkach jest to zapewne możliwe. (Irracjonalność nie jest tym, co nas w postawie egoisty najbardziej niepokoi).

Przyjmijmy, że istnieje jakieś dobro G oraz zbiór osób P taki, że każdy z członków P ma egoistyczny powód, by chcieć uzyskania G, ale osiągnięcie G wymaga podjęcia działania C, co pociąga za sobą koszty dla każdej jednostki z pewnego podzbioru P; i niech A będzie członkiem P. Wówczas:

5. Czy A ma egoistyczny powód, by zrobić C, jeśli ma rozsądną pewność albo że zbyt niewielu członków P zrobi C, by G zostało osiągnięte, albo że wystarczająco wielu innych członków P zrobi C, by G zostało osiągnięte bez jego starań?

Nie.

6. Czy można sobie wyobrazić jakieś okoliczności tego rodzaju, w których A miałby egoistyczny powód, żeby zrobić C?

Tak, w tych przypadkach, w których osiągnięcie krytycznej liczby osób, które zrobią C, zależy od jego zaangażowania albo kiedy A ma powód, by tak sądzić.

7. Czy istnieją takie motywacje, które sprawiłyby, że byłoby racjonalne dla A zrobić C, choć nie znajduje się on w opisanej wyżej sytuacji?

Tak, a jeśli nie jest on czystym egoistą, to wiele. Na przykład istnieją motywacje ekspresywne, stosowne choćby w słynnym przykładzie z głosowaniem⁴. Istnieją

⁴ Dobrze znane opracowanie problemu przedstawił M. Olson Jr. w: *The Logic of Collective Action*, Cambridge, Mass. 1965. O ekspresywnych motywacjach w tym kontekście por. również: S.I. Benn, *Rationality and Political Behaviour*, [w:] *Rationality and the Social Sciences*, ed. S.I. Benn, G.W. Mortimore, London 1976. Na temat poruszonego w tekście problemu bezstronności jest oczywiście znacznie więcej do powiedzenia, na przykład, w jaki sposób członkowie grupy są w stanie, zgodnie z wymogiem bezstronności, zgodzić się na strategię skuteczniejszą niż robienie C przez każdego (tj. przykładowo na robienie C po kolei).

także motywacje, które wypływają z poczucia sprawiedliwości. Właśnie tutaj można wyjść poza dylemat „albo daremne, albo niekonieczne” przez odwołanie się do argumentu o postaci „ktoś — a skoro nie ma powodu, by pominąć kogokolwiek konkretnego, to każdy”.

8. Czy posiadanie tego rodzaju motywacji jest irracjonalne?

W każdym sensie, w jakim to pytanie jest zrozumiałe, nie.

9. Czy jest racjonalne, z punktu widzenia społeczeństwa, wychowywanie ludzi, którzy będą mieli tego rodzaju motywacje?

Na ile to pytanie jest zrozumiałe, tak. I z pewnością mamy rację, zachęcając ludzi do nabywania tych dyspozycji — np. dlatego, że sami je posiadamy.

Muszę przyznać, że nie przychodzi mi na myśl żadne inne znaczące zagadnienie, które na tym poziomie ogólności mogłoby pomóc wyjaśnić te kwestie. Na wszystkie powyższe pytania udało się udzielić jasnych odpowiedzi, zgodnych z koncepcją racjonalności praktycznej opartej na wewnętrznych racjach do działania, i, moim zdaniem, są to odpowiedzi zupełnie rozsądne.

Przekład Agata Łukomska