

Włodek Rabinowicz

Odwracanie ról : odpowiedź na artykuł Krzysztofa Saji

Analiza i Egzystencja 12, 61-68

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.

WŁODEK RABINOWICZ

ODWRACANIE RÓL
ODPOWIEDŹ NA ARTYKUŁ KRZYSZTOFA SAJI
EKSPERYMENTY MYŚLOWE A UTYLITARYZM

Słowa kluczowe: Hare, Saja, Vendler, utilitaryzm preferencji, zasada refleksji, empatia

Keywords: Hare, Saja, Vendler, preference utilitarianism, principle of reflection, empathy

Dziękuję Krzysztofowi Saji za przetłumaczenie mojego artykułu i poddanie go w niniejszym tomie ciekawej krytyce. Nie jestem pewien, czy potrafię sobie dać z nią radę. Dobrze byłoby zapewne, gdyby mi się to nie udało – dyskutowany przeze mnie argument Hare’a okazałby się wówczas lepszy, niż myślałem. Aby czytelnik mógł to sam ocenić, skomentuję poniżej propozycję Saji.

Przedstawia on „trzy sposoby interpretowania eksperymentów myślowych Hare’a, polegających na «wejściu w skórę innych»” (Saja 2010). Według Saji Hare mógłby rozwiązać problem pozornego braku wewnętrznego konfliktu preferencji na każdy z tych sposobów. Każdy z nich pozwoliłby wykazać, że namysł moralny, jeśli przeprowadzam go poprawnie, przekształca międzyosobowy konflikt preferencji w konflikt wewnętrzny pomiędzy moimi własnymi preferencjami. Sposoby opisane przez Sają różnią się od rozwiązań, które przedstawiłem w moim artykule (por. Rabinowicz 2010 [2009]).

Pierwszy sposób: W namyśle moralnym powinienem przyjąć, że hipotetyczne sytuacje, w których byłbym w skórze innych osób, są rzeczywiste – równie rzeczywiste jak aktualna sytuacja, w której się znajduję. Innymi słowami winienem założyć, że wszystkie te sytuacje współlistnieją jako składniki aktualnego świata.

Drugi sposób: Winienem wyobrazić sobie pewien możliwy świat, niekoniecznie aktualny, w którym aktualna sytuacja i sytuacje z odwróconymi rolami współlistnieją ze sobą nawzajem.

Trzeci sposób: Jak sugeruje Vendler (1988), winienem traktować te rzekomo ‘hipotetyczne’ sytuacje odwróconych ról jako nic innego niż aktualną sytuację, tyle że widzianą z innych subiektywnych perspektyw.

Szczerze powiedziawszy, nie bardzo rozumiem, jak pierwsze dwa sposoby na tej liście mogłyby pomóc w rozwiązaniu problemu braku konfliktu. Nawet jeśli rozważane sytuacje współlistnieją w ramach tego samego świata, możliwego czy aktualnego, to moje preferencje odnośnie do każdej z nich nadal nie mają tego samego przedmiotu. W konsekwencji, preferencje te nie wchodzą we wzajemny konflikt: problem braku konfliktu nadal istnieje. Mogę mieć preferencję, aby jechać rowerem odnośnie do sytuacji, w której jestem w skórze profesora, i preferencję przeciw jeździe rowerem odnośnie do sytuacji, w której jestem w skórze czekającego studenta. Zaspokojenie jednej z tych preferencji nie przeszkadza zaspokojeniu drugiej. W tej mierze, w jakiej wydaję zalecenie moralne, zasada uniwersalizacji żąda wprawdzie, abym miał te same preferencje odnośnie do sytuacji różniących się jedynie co do roli, jaką w nich zajmuję, ale takie odwołanie się do wymogów uniwersalizacji nie wymaga przecież, aby rozważane sytuacje współlistniały w jednym świecie.

Skądinąd próba umieszczenia w tym samym świecie wszystkich sytuacji, w których zajmuję różne role, prowadzi Hare’a do dość desperackich manewrów. Jeśli te role są od siebie tak odmienne, że niemożliwe jest, aby jedna i ta sama osoba po kolei je zajmowała, Hare proponuje, że możemy traktować rozważane sytuacje jako „żywe sny” lub halucynacje, które po kolei przeżywam w danym świecie (por. Hare 1989, s. 174–177). Jednakże, przy tej interpretacji, zasada warunkowej refleksji przestaje obowiązywać. Weźmy tylko jeden przykład: We śnie, w którym jestem więźniem, intensywnie pragnę wyjść na wolność. Ale teraz, na jawie, gdy wyobrażam sobie ten sen, moje pragnienie, aby być w nim uwolnionym, nie musi być równie intensywne: W końcu, jeśli to tylko sen, to nie ma to tak wielkiego znacze-

nia, co mi się w nim przydarzy. Hare nie zgodziłby się ze mną; uważa on, że jedyne, co ma znaczenie, to subiektywne przeżycia, a te są *ex hypothesi* równie silne w realistycznym śnie jak na jawie (por. tamże, s. 177). Ale ten jego subiektywizm wyraża tylko jego prywatny pogląd, co do którego nie ma powszechnej zgody; nie można go przyjąć jako podstawy generalnego argumentu odnośnie do zakresu stosowalności zasady refleksji.

Jeśli idzie o pierwszy sposób przedstawiony przez Saja, tzn. założenie, że hipotetyczne sytuacje są rzeczywiste, nadal chcę bronić mojego pierwotnego zarzutu wobec tego rozwiązania (por. Rabinowicz, Strömberg 1996): jeśli zakładam, że wszystkie sytuacje, w których zajmuję role innych osób są rzeczywiste, to przyjmuję coś, o czym wiem, że jest nieprawdą. Wnioski, do których mogę dotrzeć wychodząc od takiej fikcji, są niewiele warte; mogę je przyjąć tylko na niby – pod warunkiem, że prawdą jest to, o czym wiem, że nie jest prawdą! Saja ma wprawdzie rację, gdy pisze: „uniwersalizacja wymaga [...] od nas, abyśmy nie rozróżniali sytuacji rzeczywistych od hipotetycznych: te drugie są równie istotne jak pierwsze”. Różnicę między rzeczywistą sytuacją i sytuacjami hipotetycznymi, w których role są odwrócone, winniśmy traktować jako moralnie nieistotną. Ale jeśli ta różnica jest nieistotna, to po co w takim razie zakładać, że wszystkie sytuacje z odwróconymi rolami są rzeczywiste, skoro to i tak nie ma znaczenia? Nie pojmuję tego.

Saja uważa, że Hare słusznie odrzuca użycie „zasłony niewiedzy” w stylu Rawlsa-Harsanyięgo:

Ponieważ [...] Hare uważał, iż poprawne ujęcie racjonalnej decyzji musi być czynione w duchu teorii racjonalności zakładającej pełną wiedzę (*full informational theory of rationality*), nie zgadzał się na interpretowanie myślenia moralnego jako procedury symulowania niewiedzy. [...] Z tego też względu Hare stwierdził, iż w naszych eksperymentach myślowych sytuacje hipotetyczne powinniśmy traktować tak, jakby miały stać się rzeczywistymi [...] Wymóg uniwersalizowalności oraz pełnowiedzowej teorii racjonalności mógłby więc spełnić eksperyment, w którym wyobrażamy sobie, że jesteśmy w skórze wszystkich osób zaangażowanych w naszą decyzję oraz że wystąpienie każdej z sytuacji $s_1 - s_3$ jest w 100% prawdopodobne. W ten sposób, abstrahując od aktualności czy hipotetyczności sytuacji s_2 i s_3 , w konsekwencji traktujemy je jako równie istotne jak sytuację rzeczywistą s_1 , gdyż zakładamy pewność, że zostaną one urzeczywistnione (Saja 2010).

Saja widzi zatem zasadniczą różnicę między udawaniem, że się nie wie, która z rozważanych sytuacji jest rzeczywista, a udawaniem, że wszystkie one są rzeczywiste. Tego też nie rozumiem. Jeśli symulowanie niewiedzy o prawdzie jest niedopuszczalne, to dlaczego dopuszczalne jest symulowanie wiedzy o nieprawdzie?

Jeśli idzie z kolei o drugi sposób na powyższej liście, to nie jest dla mnie jasne, jak moralny sąd dotyczący możliwego świata, w którym wszystkie rozważane sytuacje współlistnieją, daje się zastosować do świata aktualnego, w którym tylko jedna z tych sytuacji zachodzi. Taka ekstrapolacja nie jest uzasadniona, jeśli nie wykaże się, że oba te światy są do siebie podobne w moralnie relewantnych względach. A to wcale nie wydaje mi się oczywiste, że w danym świecie osoby zamieniają się rolami, może mieć istotne moralne znaczenie. Dla przykładu, w świecie, w którym każda osoba na przemian zajmuje rolę dającą jej okazję do uczynienia ofiary na rzecz grupy jako całości, taka ofiara może być od każdej z tych osób moralnie wymagana, z powołaniem się na fakt, że inne osoby ofiary takie też winne są uczynić. Wątpliwe jest jednak, czy można tej ofiary wymagać w świecie, w którym tylko jeden z członków grupy pełni rolę, w której może poświęcić się dla innych.

Trzeci sposób – propozycja Vendlera – jest niewątpliwie bardziej obiecujący. Jeśli Vendler ma rację, to, obiektywnie rzecz biorąc, sytuacja nie ulega zmianie, gdy w wyobraźni przyjmują role innych osób: ‘ja’ przyjmujące różne role to nie konkretna osoba (Wlodek Rabinowicz), a jedynie czysta rama świadomości. Wówczas problem braku konfliktu znika, *pod warunkiem*, że zasada warunkowej refleksji stosuje się nawet do takich eksperymentów myślowych. Wtedy bowiem, po przeprowadzeniu owej serii eksperymentów, sprzeczne ze sobą preferencje, które ‘ja’ miałbym w różnych rolach, muszą się odzwierciedlić w moich preferencjach odnośnie do rozważanej obiektywnej sytuacji, tzn. w moich preferencjach w roli eksperymentatora. Te ostatnie będą zatem wchodzić ze sobą w konflikt, dając się rozwiązać przez standardowy zabieg równoważenia.

Saja słusznie wskazuje, że w tym wypadku ‘warunkowa refleksja’ nie jest, ściśle mówiąc, ‘warunkowa’. Rozważana sytuacja jest sytuacją aktualną, a zatem moje preferencje odnoszące się do niej nie są opatrzone warunkami typu: ‘gdyby taka sytuacja zaistniała, to ...’. W tym oczywiście się zgadzamy (por. Rabinowicz 1989). Zgadzamy się też, że – ogólnie rzecz biorąc – zasada (warunkowej lub bezwarunkowej) refleksji jest wyrazem

troski własnej. Kością niezgody między nami jest *zasięg ważności* zasady refleksji. Ja uważam, że nie stosuje się ona do radykalnych eksperymentów myślowych w interpretacji Vendlera, gdyż w tych eksperymentach 'ja' nie odnosi się już do mnie samego jako do konkretnej osoby – Włodka Rabinowicza. Według mnie troska własna jest troską o własną osobę. Natomiast Saja sugeruje, że zasięg zasady refleksji nie jest w ten sposób ograniczony. Sięga ona tak daleko, jak samo 'ja' zdolne jest sięgać:

Pojęcie „ja” ma [...] dwie główne cechy. Po pierwsze, zdaniem Hare’a w pojęciu „ja” zawiera się owa możliwość transcendencji poza moje aktualne cechy i choć „ja”, w moim potocznym użyciu, wypełnione jest pewną treścią (nie jest puste), treść ta nie determinuje poprawnego użycia pojęcia „ja”. Po drugie, pojęcie „ja” posiada preskryptywny charakter: na mocy znaczeń, tam gdzie jestem „ja”, tam jest troska własna” (Saja 2010).

Nie neguję, że Saja ma rację w tej charakteryzacji stanowiska Hare’a. Ale czy stanowisko to jest słuszne? Temu chciałbym zaprzeczyć. Nie mam kłopotów z przyjęciem zasady refleksji jako wymagania praktycznej racjonalności, w tej interpretacji jednak troska własna nie sięga dalej niż do różnych hipotetycznych sytuacji, w których ja – Włodek Rabinowicz – mógłbym się znaleźć. Dopuszcza to oczywiście „możliwość transcendencji poza moje aktualne cechy”, albowiem wiele z tych cech jest takich, że mógłbym ich nie posiadać – np. mógłbym mieć inny zawód, inną żonę, innych przyjaciół, inne nazwisko, lepszą kondycję fizyczną, odmienny charakter, inne upodobania itd. Mogę sobie wyobrazić sytuację, w której – powiedzmy – złamałbym nogę i potrzebowałbym pomocy lekarskiej. Nakaz troski własnej wymaga, abym już teraz chciał, żeby w takich okolicznościach otrzymać pomoc lekarską. Mogę tego dopilnować np. przez zakup telefonu komórkowego i stałe noszenie go ze sobą. Nie ma w tym nastawieniu nic dziwnego: musimy być przygotowani na zmiany w naszym życiu i dlatego winniśmy już teraz się troszczyć o to, jakbyśmy sobie dali radę w różnych możliwych sytuacjach. Że ewolucja naturalna premiuje taką postawę, jest zrozumiałe. Można też dać ewolucyjne wytłumaczenie, dlaczego troska własna rozciąga się nawet na hipotetyczne sytuacje, o których jesteśmy przekonani, że nie mogą się przydarzyć. W końcu możemy się mylić, a zatem lepiej być przygotowa-

nym nawet na to, co uważamy za wykluczone¹. Pogląd Hare'a jest jednak bardziej radykalny: troska własna sięga tak daleko, jak moje użycie terminu 'ja' – nawet gdy termin ten przestaje się odnosić do mojej osoby. Wynika to stąd, że według Hare'a element preskryptywny zawarty jest w samym znaczeniu 'ja' jako takim. Zarazem przyjmuje on sensowność wyobrażenia sobie, że *ja* mógłbym być inną osobą, niż jestem: Jonesem, Napoleonem itd. (por. odpowiedź Hare'a na artykuł Vendlera: Hare 1988)². Innymi słowy, jest analitycznie prawdziwe, że *ja* jestem przedmiotem mojej troski, niezależnie od tego, do kogo owo 'ja' się odnosi.

Wydaje mi się to zbyt łatwe, aby było prawdziwe. Ostrożniejsza pozycja przekonuje mnie bardziej. Ponieważ 'ja' normalnie referuje do osoby, którą jestem, stwarza to naturalną psychologiczną asocjację między użyciem 'ja' a troską. Przejmuję się tym, co 'mnie' dotyczy. Ale ten związek jest przyczynowy, a nie pojęciowy. Na dodatek nie jest on ścisły: możemy się opierać skłonności do tej zgeneralizowanej troski względem kogokolwiek, do kogo nasze użycie 'ja' się odnosi³. Niektórzy z nas mają szczególną zdolność do opierania się takim skłonnościom. Osoby takie zdolne są do empatii bez sympatii.

Saja nie zgadza się z tą diagnozą: jego interpretacja empatii jest odmienna. Twierdzi on, że empatia bez sympatii jest możliwa tylko dlatego, iż empatia nie jest wyobrażeniem sobie, co *ja* bym odczuwał, będąc w skórze

¹ Innym możliwym wytłumaczeniem jest, że mechanizmy ewolucji nie są na tyle subtelne, aby brać pod uwagę różnicę między tym, czego nie jesteśmy zupełnie pewni, a tym, czego pewni jesteśmy.

² Dla znawców Hare'a ten jego tekst jest zaskakujący. We wcześniejszych pracach, o ile wiem, nie był on gotów pójść tak daleko. Przyjmował za sensowne wyobrażenie sobie, że jestem w dokładnej *pozycji* innej osoby, ale unikał sformułowań takich jak: „*gdybym był Jonesem, to ...*”

³ Hare mógłby odpowiedzieć, że ewolucyjnie ugruntowany psychologiczny związek między użyciem 'ja' a troską musiał wpłynąć na znaczenie tego terminu. W rezultacie, w potocznym użyciu, 'ja' uzyskało preskryptywne znaczenie: wypowiedzi zawierające 'ja', 'mnie' itp., jeśli są szczerze, wyrażają pewien stan motywacyjny. Być może tak jest, ale ta ewentualna preskryptywność daje się neutralizować. Tak jak nie ma sprzeczności w negacji tzw. implikatur językowych (por. „*Jones nie jest dzisiaj pijany. [...] Nie był on też pijany przy innych okazjach*”), tak też nie ma sprzeczności w neutralizacji preskryptywnego charakteru 'ja' (por. „*Zamierzam okraść Jonesa. Gdybym to ja był Jonesem, bardzo bym tę kradzież przeżył. [...] Ale to mnie ani ziębi, ani grzeje; Jones mnie nie obchodzi*”).

innej osoby. Nie jest ona wczuciem się w ‘w pierwszej osobie’, jak to wymaga myślowy eksperyment Hare’a, lecz tylko wczuciem się ‘w trzeciej osobie’. Empatia to wiedza, „co w rzeczywistej sytuacji preferuje *ktoś inny*”, a nie wiedza „co w rzeczywistej sytuacji *ja* bym preferował z innej perspektywy (np. perspektywy, w której realnie znajduje się kogoś inny)” (podkreślenia Saji). Ta interpretacja empatii jako ‘trzeciosobowej’ formy poznania nie wydaje mi się przekonująca: wątpię, czy możliwe jest wczucie się w kogoś innego bez postawienia siebie samego na jego miejscu. Wskazują na to w każdym razie współczesne teorie empatii, według których ta forma poznania ma charakter naśladowczy: wymaga ona swoistej mentalnej *symulacji* przeżyć przedmiotu przez podmiot. W tym kontekście symulację zwykło się charakteryzować w fenomenologicznych terminach jako „przyjmowanie roli” lub „stawianie siebie na miejscu innej osoby”, czyli tak, jak Hare zwykle opisuje swoje eksperymenty myślowe (por. np. Gordon 2009). W bardziej technicznych terminach mentalna symulacja opisywana jest jako proces, w którym obserwator śledzi własny system kontroli zachowania, funkcjonujący ‘off-line’ (tzn. nie na serio, ‘bez podłączenia do sieci’, jeśli tak można powiedzieć), traktując go jako model systemu zachowania obserwowanej osoby. Obserwacja funkcjonowania własnego systemu pozwala mu na sformułowanie wniosków o funkcjonowaniu systemu modelowanego (tamże).

Być może moje wątpliwości wobec tego trzeciego sposobu rozwiązania problemu braku konfliktu dają się jakoś przezwyciężyć. Nie jestem do końca pewien, kto z nas ma w tej sprawie rację. Wdzięczny jestem w każdym razie Krzysztofowi Saji i redakcji „Analizy i Egzystencji” za okazję powrotu do tego fascynującego problemu⁴.

Literatura

Gordon R.M. (2009), *Folk Psychology as Mental Simulation*, [w:] *Stanford Encyclopedia of Philosophy*, <http://www.seop.leeds.ac.uk/entries/folkpsych-simulation/>.

⁴ Chciałbym także podziękować za cenne uwagi mojemu koledze z Lundu, Toniemu Rønnow-Rasmussenowi. Szwedzka Rada Badań Naukowych (VR) dała mi finansowe wsparcie przy pracy nad tym artykułem, a Swedish Collegium for Advanced Study (SCAS) w Uppsali udzieliło mi gościny.

- Hare R.M. (1988), *Comments on Vendler*, [w:] D. Seanor, N. Fotion (red.), *Hare and Critics: Essays on Moral Thinking*, Oxford: Clarendon Press, s. 280–287.
- Rabinowicz W. (1989), *Hare on Prudence*, „Theoria” 55, s. 145–151.
- Rabinowicz W. (2010 [2009]), *Utylitaryzm preferencji poprzez zmianę preferencji?*, tłum K. Saja, „Analiza i Egzystencja” 12, s. 7–36 [*Preference Utilitarianism by Way of Preference Change?*, [w:] T. Grüne-Yanoff, S.O. Hansson (red.), *Preference Change*, Netherlands: Springer. Theory and Decision Library, vol. 42, s. 185–206].
- Rabinowicz W., Strömberg B. (1996), *What if I Were in His Shoes? On Hare’s Argument for Preference Utilitarianism*, „Theoria” 62, s. 95–23.
- Saja K. (2010), *Eksperymenty myślowe a utylitaryzm*, „Analiza i Egzystencja” 12.
- Vendler Z. (1988), *Changing Places*, [w:] D. Seanor, N. Fotion (red.), *Hare and Critics: Essays on Moral Thinking*, Oxford: Clarendon Press, s. 171–184.

ROLE REVERSALS

A REPLAY TO KRZYSZTOF SAJA’S

“THOUGHT EXPERIMENTS AND UTILITARIANISM”

Summary

In *Thought Experiments and Utilitarianism*, which is a comment on my *Utilitarianism by Way of Preference Change?*, Krzysztof Saja delineates three alternative ways in which one might interpret Richard Hare’s famous thought experiments involving role reversals. He suggests that each of them would underwrite Hare’s claim that moral deliberation transforms an interpersonal conflict of preferences into an intrapersonal one, which obtains between the deliberator’s own preferences. In this reply to Saja (*Role Reversals*), I discuss the three proposals in turn and argue that the first two do not solve the problem. The third one, which was already considered in my original paper, is based on Zeno Vendler’s suggestion that role reversals in moral thought-experiments are merely different ‘takes’ on one and the same real situation, as seen from different subjective perspectives. I argue that, in order to succeed in transforming interpersonal preference conflicts into intrapersonal ones, this proposal requires equating empathy, i.e. putting oneself in someone else’s shoes, with sympathy, which Hare would not be prepared to do.