

R o b e r t P i ł a t

Preferencje i wartości

Słowa kluczowe: *wartość, preferencja, decyzja, podmiot, racjonalność*

W artykule¹ podejmuję problem relacji pomiędzy porządkiem wartości i porządkiem preferencji. Pytanie brzmi: czy wolimy A od B, ponieważ bardziej cenimy A? Powołując się na uznawane przez siebie wartości i wynikające z nich preferencje uzasadniamy swoje wybory, lecz często uzasadnienia te nie są dobre, o czym świadczą zarówno potoczne obserwacje, jak i badania psychologiczne. Wielu czynów dokonujemy impulsywnie i nieracjonalnie. Wydaje nam się, że znamy swoje preferencje i że są one oparte na naszych wartościach, lecz często postępujemy niezgodnie z jednymi i drugimi. To, że postępujemy niezgodnie z wartościami nie jest dziwne, zważywszy że niektóre wartości odzwierciedlają trudne do realizacji ideały postępowania. Bardziej intrygujące jest to, że działamy często niezgodnie z własnymi aktualnymi preferencjami. Wolimy ciemne swetry wełniane niż jasne bawełniane, lecz wychodzimy ze sklepu z tym drugim, a nie tym pierwszym. Najwyraźniej zbyt wiele różnorodnych czynników wpływa na nasze decyzje i nie potrafimy ich zagregować w jeden osąd: „wolę A od B”. Ta nieumiejętność kłóci się z racjonalnymi założeniami na temat własnego działania, przy czym wcale nie chodzi o wygórowane normy racjonalności, lecz o minimalną racjonalność, która jest potrzebna do podejmowania wspólnych z innymi ludźmi przedsięwzięć, planowanie przyszłości itd. Zgodnie z tymi niewygórowanymi oczekiwaniami, jeśli wartość A leży w naszej osobistej hierarchii wartości wyżej niż B, to mając do wyboru A lub B i pragnąc zrealizować A lub B, wolimy A i jeśli nie zachodzą dodatkowe okoliczności, wybieramy A. Zdaniem Donalda

¹ Niniejszy tekst jest poprawioną wersją wykładu wygłoszonego w ramach X Polskiego Zjazdu Filozoficznego w Poznaniu 16 września 2015 r.

Davidsona², nie można podać racji zachowania niezgodnego z tą zasadą. Człowiek, który postępuje niezgodnie z nią, nie rozumie sam siebie – nie da się przecież ogółu przesłanek danego działania podzielić arbitralnie na dwie grupy, z których jedna wspiera preferencję $A > B$, a druga preferencję $B > A$, i nigdy nie zastanawiać się nad *relacjami* pomiędzy tymi dwiema grupami przesłanek. Davidson ma rację domagając się, by racjonalna osoba uwzględniała *całość* dostępnych jej danych, nie izolując jednych od drugich³. Nie kłóci się to ze zdolnością i skłonnością ludzkiego umysłu do takiego rozdzielania – tworzenia, jak mówią niektórzy⁴, wewnętrznych partycji, jednak tego rodzaju działania umysłu nie można nazwać racjonalnym⁵. W pewnych sytuacjach wybieramy to, co cenimy mniej. Istnieją czynniki podmiotowe i sytuacyjne, które stoją pomiędzy wartościowaniem na ogół i chwilową preferencją towarzyszącą działaniu. Zagadka owej chwilowej preferencji jest potrójna: 1) Które informacje przyczyniają się do jej kształtowania? 2) Czy rzeczywiście kształtuje się ona przed wykonaniem działania, czy raczej jest następczą racjonalizacją? 3) Jakie światło rzuca działanie impulsywne na tożsamość podmiotu? Odpowiedź na te trzy pytania byłaby bardzo ważna dla filozofii, rzuciłaby bowiem światło na źródła i ograniczenia normatywności, a także na status wartości. Wynika to z następujących zależności: (1) Preferencje są wprawdzie relacjami pomiędzy dobrami, lecz domagają się pewnego dodatkowego wglądu w niezmienniki w zbiorze dóbr, czyli wartości; samo oddziaływanie dóbr na osoby nie wystarczy, ponieważ dobra są tak różnej natury, że często nie da się ich porównać. Wartości powinny służyć jako standardy umożliwiające takie porównanie, a tym samym wyznaczenie relacji preferencji. Wydaje się, że powyższe stwierdzenie jest niezależne od znanych sporów o naturę wartości (np. obiektywizm *versus* subiektywizm, intelektualizm *versus* emocjonalizm). Niemniej, jak postaram się pokazać dalej, tradycyjne koncepcje wartości, takie jak koncepcja Maxa Schelera czy Nicolai Hartmanna, wnoszą wiele do rozumienia tego najbardziej kłopotliwego pojęcia we współczesnej teorii decyzji, jakim jest pojęcie preferencji. (2) Uznanie stabilnych preferencji wydaje się koniecznym warunkiem

² D. Davidson, *Jak możliwa jest słabość woli*, w: J. Hołówka (red.) *Filozofia moralności*, Altheia, Warszawa 1997, s. 83–84, 93.

³ D. Davidson, *Knowing One's Own Mind*, w: Q. Cassam (red.), *Self-Knowledge*, Oxford University Press, New York 1994, s. 43–64.

⁴ Najbardziej typowy jest tu pogląd Amelie Rorty, wyrażony w rozprawie o samooszukiwaniu się. Zob. A. Rorty, *The Deceptive Self: Liars, Layers, and Lairs*, w: A.O. Rorty, B.P. McLaughlin (eds.), *Perspectives on Self-Deception*, University of California Press, Berkeley 1988, s. 11–28.

⁵ Korzystanie z „partycji” jedynie oddala problem, bez widoków rozwiązania. Osoba racjonalna będzie bowiem chciała znać rację zastosowania danej partycji (dlaczego jest lepsza niż inne) i w ten sposób pierwotny problem powróci na innym poziomie.

formułowania jakichkolwiek norm; bez uzasadnionych preferencji nie możemy podać racji działania, a tych domagają się normy, jeśli mają się czymś różnić od przymusów i rozkazów. Normy są nakazami tego, co uzasadnione.

1. Niestabilność preferencji

Przekonanie, że preferencje poprzedzają akt deliberacji i decyzji, jest nie tylko częścią potocznej wizji działania, lecz również aksjomatem klasycznej teorii decyzji. W myśl tej teorii, w celu nadania możliwym wyborom użyteczności oczekiwanej (niezbędnej do zastosowania dalszych strategii teorio-decyzyjnych) musimy umieć podjąć decyzję w sytuacji loterii („Czy wolisz wartość A z prawdopodobieństwem p , czy wartość B z prawdopodobieństwem q ”). Na podstawie tych odpowiedzi można zbudować skalę interwałową cenionych przez siebie dóbr. Odpowiedzi na pytania o loterie muszą być *stabilne*, są one właściwie wszystkim, czego wymaga się tu od podmiotu decyzji. Tymczasem liczne badania pokazują, że preferencje są niestabilne. Konstruujemy je *ad hoc* w danej sytuacji, w trakcie decyzji lub nawet po jej dokonaniu. Do wyjaśnienia tego efektu nie wystarczy powołać się na jakiś jeden czynnik; prawdopodobnie wchodzi tu w grę zarówno słaba wola, jak ograniczona świadomość aksjologiczna, siła nawyku, zła ocena prawdopodobieństwa (heurystyki niezgodne z rachunkiem prawdopodobieństwa) itd. Czynników tych jest wiele i nie można ich uporządkować *a priori*. W każdym przypadku mogą pojawiać się w różnych kombinacjach i z różną siłą. Przyjrzyjmy się kilku przykładom:⁶

1. *Dodanie nowej wartości*. W powieści Marka Twaina *Przygody Tomka Sawyera* bohater wykonuje nielubianą czynność malowania płotu, nakazaną przez ciotkę, i z tego powodu jest wyśmiewany przez kolegów. Przekonuje ich jednak, że ta praca jest przywilejem, w związku z czym zmieniają oni swoją preferencję: zamiast iść się bawić, chcą teraz malować płot razem z Tomkiem.
2. *Wpływ czynników arbitralnych*. Badani mają zdecydować, za jaką zapłatę zgodziliby się słuchać przez jakiś czas bardzo nieprzyjemnego dźwięku przygotowanego przez eksperymentatora. Najpierw jednak przygotowana jest przykładowa wycena 300 sekund słuchania w sposób jawnie arbitralny, czyli dla każdej osoby oddzielnie za pomocą numeru jej dokumentu tożsamości. Następnie badani mają powiedzieć, za ile zgodziliby się rzeczywiście wysłuchać tego dźwięku. Okazuje się, że osoby badane o niższych

⁶ Przykłady zaczerpnięte są ze studiów zawartych w: P. Slovic, S. Liechtenstein (red.), *The construction of Preference*, Cambridge University Press, Cambridge 2006.

- liczbach wygenerowanych z ich numerów dokumentów decydowały się słuchać za mniejszą zapłatę⁷.
3. *Wpływ czasu*. Badani decydują, że skłonni są poczekać na większą nagrodę z pewnym prawdopodobieństwem w przyszłości niż od razu przyjąć mniejszą. Mogą jednak później zmienić zdanie. W miarę przybliżania się momentu ujawnienia, czy wygrali swoją większą nagrodę, zmieniają decyzję i przyjmują mniejszą nagrodę.
 4. *Przeskok do innego kryterium oceny*. Efekt ten widać w podwójnych zakładach. Dane są dwa zakłady, każdy z innym rozkładem wygranych i prawdopodobieństwem ich uzyskania. Pierwsze pytanie brzmi: który zakład jest atrakcyjniejszy? Drugie pytanie brzmi: za ile warto zagrać w ten zakład (jaka powinna być jego cena)? Otóż badani odpowiadają na pierwsze pytanie opierając się na prawdopodobieństwie wygranej, zaś na drugie kierując się wysokością wygranej. Dlatego w wielu przypadkach dochodzi tu do odwrócenia preferencji pomiędzy pierwszym a drugim pytaniem. Wydaje się, że chociaż wiemy, co wolimy w odniesieniu do poszczególnych cech zakładu, to jednak nie dostrzegamy zależności pomiędzy tymi cechami, a co za tym idzie, pomiędzy preferencjami.

Sarah Liechtenstein i Paul Slovic⁸ sugerują, że „...tak naprawdę nie wiemy, co wolimy; musimy konstruować swoje preferencje, w miarę jak pojawiają się sytuacje decyzyjne (...); często znamy swoje preferencje dla poszczególnych aspektów, lecz nie znamy zależności (*tradeoff*) pomiędzy nimi”. W tym samym tomie studiów Daniel Kahneman formułuje jeszcze ogólniejszy wniosek: „...utrzymanie koherentnych przekonań i preferencji jest zbyt trudne dla ograniczonego umysłu. (...) Być może przyszedł czas, by odłożyć na bok pytanie, czy ludzie są, czy nie są racjonalni, i skupić [się na tym] pod jakimi warunkami założenie o racjonalności można utrzymać jako użyteczne przybliżenie”⁹. Jeśli jednak badacze ci mają rację, to znaczy, że nasze intuicje aksjologiczne (rozumienie czy doświadczenie wartości) nie są w stanie trwale uporządkować zbioru dóbr. Wprawdzie ta niemoc aksjologiczna była już w filozofii dostrzegana – tak odmienni pod innymi względami filozofowie jak Kant i Nietzsche dalecy byli od przypisywania ludziom wiarygodnych intuicji aksjologicznych. Niemniej jednak, pomimo wątpliwości, język argumentacji odwołujący się do wartości tak się rozpowszechnił (dyskurs polityczny odwołuje się dziś głównie

⁷ D. Ariely, G. Loewenstein, D. Prelec, *Tom Sawyer and the Construction of Value*, w: P. Slovic, S. Liechtenstein (red.), *The Construction of Preference*, dz. cyt., s. 261.

⁸ S. Liechtenstein, P. Slovic, *The Construction of Preference*, dz. cyt., s. 1.

⁹ D. Kahneman, *New Challenges to the Rationality Assumption*, w: S. Liechtenstein, P. Slovic, *The Construction of Preference*, dz. cyt., s. 503.

do pojęcia wartości), że trudno nam wypracować alternatywę¹⁰. Z kolei niemoc aksjologiczna ma inną poważną konsekwencję: preferencje są podstawą działania racjonalnego, jeśli jednak one same pozbawione są gruntu aksjologicznego umożliwiającego formułowanie racji, to nawet minimalne postulaty racjonalności działania mogą się okazać, paradoksalnie, nieracjonalne. Ma to poważne konsekwencje, które krótko wskażę w następnej części.

2. Brak niezawodnego przepisu na racjonalne preferencje

Wskazane powyżej problemy są częścią pewnego szerszego zagadnienia: czy potrafimy uzgodnić swoje przekonania z pobudkami naszego działania? Odpowiedź negatywna opiera się na założeniu, że działania mają zupełnie inną strukturę logiczną i poznawczą niż przekonania – łączą się bowiem ze skomplikowaną strukturą społecznych praktyk i nie sposób zbudować apriorycznej teorii tej zależności. Odpowiedź pozytywna opiera się na przekonaniu, że kiedy uczynimy rozmyślnie i z poczuciem słuszności coś, co jest niezgodne z własnymi przekonaniem, to po zbadaniu sytuacji stwierdzimy, że tak naprawdę nasze przekonania były inne, niż sądziliśmy – wniosek ten będzie wprawdzie *ad hoc*, co nigdy nie jest zaletą, lecz raz sformułowany będzie odtąd obowiązywał daną osobę w jej działaniach: jako nowy stan jej przekonań. Kiedy natomiast nie jesteśmy w stanie dokonać takiego uzgodnienia, musimy stwierdzić irracjonalność swojego zachowania i zastanowić się, które z leżących u jego podstaw przekonań powinniśmy odrzucić. Wydaje się, że ta druga strategia jest charakterystyczna dla popularnej dziś literatury z zakresu poradnictwa, mówiącej o unikaniu irracjonalności, korygowaniu własnego działania, poprawnym wnioskowaniu praktycznym. Poniżej poddam krytyce typowy dla tej literatury sposób rozumowania.

Poczytna książka Stuarta Sutherlanda *Irrationality*¹¹ zawiera diagnozy szeregu rozpowszechnionych zachowań irracjonalnych, kładąc szczególny nacisk na błędy decyzyjne popełniane w warunkach niepewności. Chodzi głównie o błędy w szacowaniu prawdopodobieństwa i korzystaniu z informacji statystycznych. Książka ta ma charakter popularny i perswazyjny. Omawia najlepiej zbadane przypadki nieracjonalnych zachowań i podaje sposoby ich unikania. Głównym źródłem błędów wymienionych przez tego autora są heurystyki –

¹⁰ Czyni się takie próby, odwołując się do pojęcia godności, praw czy odpowiedzialności. Dobrym przykładem jest niedawna książka E. Strauba *Zur Tyrranei der Werte*, Klett-Cotta, Stuttgart 2010.

¹¹ S. Sutherland, *Irrationality*, Printer & Martin, London 2013.

skrótowe rozumowania, do których przywiązujemy większą wagę poznawczą, niż na to zasługują; są więc lepsze i gorsze heurystyki, i jest rzeczą istotną, by uzyskać dobre kryteria ich oceny. Rozważmy to na przykładach błędów omówionych w rozdziale „The Wrong Impression”.

- (1) *Wpływ emocji*. W eksperymencie z Dylematem Więźnia, osoby badane, po usłyszeniu jakiejś poruszającej, wywołującej współczucie historii, przyjmowały w tej grze strategię kooperatywną znacznie częściej niż osoby z grupy kontrolnej.
- (2) *Błąd dostępności*. Postrzeganie całej klasy zdarzeń zależy od stosunku do zdarzenia, które akurat jest dostępne poznawczo: lęk ludzi przed lataniem wzrasta się po usłyszeniu o jakimś wypadku, lecz po pewnym czasie wraca do średnich wartości.
- (3) *Błąd pierwszeństwa*. Cechy, które pojawiają się jako pierwsze, determinują w dużej mierze całościową percepcję przedmiotu, osoby czy sytuacji. Osoba spotkana po raz pierwszy w złym nastroju i w konsekwencji oceniona jako niesympatyczna, jest do jakiegoś stopnia trwale nacechowana tą oceną – tego rodzaju osądy są bardzo odporne na dalsze doświadczenia.
- (4) *Błąd nieuzasadnionego powiązania cech* (efekt aury). Osoby postrzegane jako przystojne są też często uważane za inteligentniejsze i posiadające lepszy charakter.

A oto przykładowe „dobre” reguły, sformułowane w kontekście omówionych przypadków:

- (1) Nigdy nie opieraj sądu lub decyzji na pojedynczym przypadku, niezależnie od tego, jak jest wyraźny i oczywisty (*striking*).
- (2) Wypracowując sobie opinię o jakiejś osobie, rozważaj i oceniaj każdą z jej cech osobno, nie pozwalając, by oceny jednych cech wpływały na oceny innych.
- (3) Jeśli rozważasz przedmiot składający się z pojawiających się po kolei cech (faz, części), wstrzymaj się z oceną aż do poznania ostatniej.
- (4) Unikaj potencjalnie uprzedzających informacji, na przykład o autorze ocenianego artykułu.

Niektóre z takich reguł są konkretne i łatwe do zastosowania, na przykład (4) – to, że nie zawsze dają pożądany rezultat, nie jest zarzutem, ponieważ nie oczekuje się tu, że będą niezawodne. Ale już reguły (1)–(3) są nie tylko zawodne, lecz także bardzo niejasne:

- (1) Są sytuacje, kiedy do dyspozycji jest tylko jeden przypadek: jak wówczas postąpić? Jeśli zaś zgodzimy się, że jeden przypadek, to za mało, to ile przypadków będzie już właściwą liczbą? Ponadto, co znaczy ów przymiotnik „*striking*” i czy zawsze można tę cechę ignorować? Są pewne przyczyny, dla których coś staje się uderzające, wydatne, wyraźne czy oczywiste.

- (2) Niektórych cech nie można od siebie oddzielić, są powiązane przyczynowo. Jeśli chodzi o osoby, to całość w oczywisty sposób nie jest sumą części, zatem rozważanie powiązań pomiędzy częściami jest istotne. Problem przenosi się więc dalej: jak wyosobnić interesujące i istotne powiązania.
- (3) Procedura ta wymaga przypisywania wag poznawczych kolejno pojawiającym się informacjom o danym przedmiocie. Jednak wagi takie nadaje się biorąc pod uwagę całe doświadczenie. Pewien wgląd w całość poprzedza rozważanie danego doświadczenia faza po fazie. Oczywiście ten wgląd jest tajemniczy, może się okazać iluzoryczny i prowadzić do uprzedzeń, lecz wyjściem z impasu nie jest na pewno zignorowanie jego obecności. Podobne uwagi można sformułować w stosunku do każdego z 23 rozdziałów książki Sautherlanda. Intuicyjnie czuje się, że krytykowane przez autora heurystyki są błędne i że proponowane przezeń korekty są słuszne (w tym sensie słuchanie rad autora przypuszczalnie poprawia naszą racjonalność), lecz nie wiemy, w jakiej mierze i dlaczego są słuszne – podstawa odczuwanego ulepszenia nie jest jasna. Dzieje się tak dlatego, że *pojęcia*, w jakich zostają wyrażone ulepszające postulaty, są dalekie od jasności. W kolejnych krokach tych rozważań zmierzam do wskazania solidniejszej podstawy dla krytyki irracjonalnego zachowania i postulatów poprawiania rozumowań normatywnych.

3. Preferencje i hierarchia wartości

Przytoczone wcześniej obserwacje mówiące o niestabilności preferencji nie prowadzą aż tak daleko, by w ogóle zrezygnować z pojęcia preferencji. Trzeba raczej rozwiązać dwa problemy: (1) Czy istniejące teorie wartości dają wiarygodne pojęcie standardu pozwalającego na porównywanie dóbr, a tym samym kształtowanie preferencji? (2) Czy preferencje – o ile mają dostatecznie pewny aksjologiczny grunt – posiadają logiczną strukturę pozwalającą łączyć je z sądami o stanach świata: o przyczynowości, prawdopodobieństwie zachodzenia zdarzeń itd.?

Sądzę, że w klasycznych koncepcjach wartości (Max Scheler, Nicolai Hartmann, Georg Simmel) tkwi dostateczny potencjał pojęciowy, by przynajmniej częściowo uporać się z tymi problemami. Scheler uznał hierarchiczny porządek za istotną własność każdego zbioru wartości. Nie sposób, jego zdaniem, odczuwać wartości, nie odczuwając zarazem ich hierarchii. Dlaczego jednak mielibyśmy zaakceptować jego tezę? Czy nie można sobie wyobrazić porównywania dóbr bez odwołania się do hierarchii wartości? Wydaje się, że można. Różnice pomiędzy dobrocią różnych stanów rzeczy wyjaśnialibyśmy wtedy nie przez odwołanie się do zhierarchizowanych wartości, lecz np. przez siłę pożądania odpowiednich przedmiotów. Jednak w takim przypadku eliminowalibyśmy całkiem pojęcie *wyboru*. Nikt nie dokonywałby wyboru danego dobra

względem innego, tylko szedł za silniejszym pragnieniem. Nie zachodziłoby tu w ogólne porównywanie czy ocenianie dóbr. Co więcej, nie zachodziłoby też porównywanie pragnień, a jedynie zastępowanie jednego pragnienia przez drugie – wchodzenie jednego na miejsce drugiego. Wyrażenie „silniejsze pragnienie” byłoby tylko potocznym skrótem myślowym, bez samodzielnego sensu, ponieważ relacja pomiędzy pragnieniami nie byłaby przedmiotem żadnej reprezentacji. Można sobie wyobrazić ten proces jako zdeterminowany lub nie, na przykład w ten sposób, że pragnienie *A* eliminowałoby pragnienie *B* w pewnych przypadkach, a w innych działałoby się odwrotnie, lub też znikałoby oba pragnienia – wciąż jednak nie implikowałoby to istnienia reprezentacji relacji pomiędzy pragnieniami. Jeśli więc zgodzimy się, że należy zachować pojęcie decyzji, to rozwiązanie to jest nie do przyjęcia.

Inne rozwiązanie problemu porównywania dóbr, nieodwołujące się do hierarchii wartości, polegałoby na tym, że dobra, a ściśle biorąc stany rzeczy będące ich ontyczną podstawą, dzieliłoby się na tyle stanów składowych, by wszystkie były równe pod względem dobroci. Taki zabieg byłby jednak tylko teoretyczny; kiedy z owych teoretycznych stanów rzeczy odtworzyłoby się naturalne stany rzeczy, cały problem wróciłby w pierwotnej postaci. Poza tym wątpliwe jest, czy zabieg taki dałby się, nawet w teorii, przeprowadzić. Dobra nie są dowolnie podzielne: jedne w sposób oczywisty, ponieważ odpowiednie stany rzeczy nie są podzielne, inne zaś dlatego, że po podziale znika dobroć stanu rzeczy – niektórych dóbr nie można doświadczać częściowo¹². Jeśli zatem oba wymienione sposoby „spłaszczania” zbioru dóbr są nie do przyjęcia, to trzeba przyznać, że muszą również istnieć dobra *a priori* różne pod względem miary od innych. Wartości są jednak pewnymi niezmiennikami w zbiorach dóbr – służą mianowicie do grupowania dóbr różniących się materialnie, lecz jednakowych pod względem owej miary. Jeśli zatem istnieją różne dobra różniące się miarą, to istnieje też wiele klas dóbr zespojonych z odpowiednimi wartościami, a zatem wiele wartości różniących się postulowaną w nich miarą dobra. Jeśli teraz przypomnimy sobie powyżej sformułowaną przesłankę, a mianowicie potrzebę zachowania pojęcia wyboru, to trzeba powiedzieć, że nie tylko różne co do miary wartości istnieją, lecz także są reprezentowane (przeżywane) w hierarchicznej relacji jedne do drugich. Zgadzałoby się to z założeniem Schelera, że doświadczamy bezpośrednio hierarchii wartości – jest ono częścią samego doświadczenia, że coś jest dobre¹³.

¹² Takich podziałów dokonuje się powszechnie w świecie finansów; można mieć podzielną udział w niepodzielnym dobru. To jednak inne zagadnienie, instrumenty finansowe zakładają integralność dóbr, lecz dokonują operacji podziału na czymś innym: relacjach własności, wielkości popytu i podaży, relacji do innych dóbr.

¹³ Należy zaznaczyć, że ta część dyskusji jest całkowicie niezależna od sporu o obiektywność czy subiektywność wartości. O doświadczeniu hierarchii można mówić w jednym i dru-

To, że doświadczamy hierarchicznych relacji pomiędzy wartościami, nie znaczy, że doświadczamy ich prawidłowo czy adekwatnie. Słowa „prawidłowo” i „adekwatnie” wprowadzają oczywiście nowy zestaw problemów. Nie jest wcale oczywiste, że należy podnosić tego rodzaju problemy. Doświadczanie wartości mogłoby być zarazem środkiem i kryterium. Kiedy pytamy o wartość tego doświadczenia, powołujemy do życia kłopotliwe założenie o istnieniu niezależnego porządku wartości; każda próba jego scharakteryzowania może się łatwo zakończyć filozoficznymi fantazjami. A jednak istnieje bardzo przyziemny powód, by pomimo tych trudności nie pozbywać się oceniającej perspektywy, mianowicie błędy w dziedzinie rozpoznawania relacji pomiędzy wartościami. Zdarza się często, że jesteśmy zmuszeni zmienić dostrzeganą początkowo relację pomiędzy wartościami. Uzasadnienie tej zmiany ma samo charakter aksjologiczny, to znaczy postrzegane jest jako poprawienie postrzegania wartości. Oczywiście i to poprawienie może być złudne i może wymagać następnego poprawienia, lecz znów sankcją będzie tu żywe doświadczenie wartości tego nowego postrzegania wartości. W tym duchu przedstawiał problem Nicolai Hartmann, kiedy twierdził, że wartości mają charakter idealny i tworzą idealny porządek. Mamy zdolność do wglądu w wartości, ponieważ istotną jego częścią jest aksjologiczna emocja – emocjonalnie ugruntowana aprobata. Wgląd ten jest omylny, ponieważ, jak podkreśla Hartmann, trudno nam za pomocą tych środków dostrzec *relacje* pomiędzy wartościami. Zauważmy, że sytuacja jest tu bardzo podobna do trudności, jakie mamy z preferencjami. Świadome sformułowanie preferencji jest ustawicznie rozbijane przez ingerencję pragnienia. Pragnienie jest w dużej mierze ślepe na relacje pomiędzy dobrami, tak samo jak emocja aksjologiczna jest w dużej mierze ślepa na relacje pomiędzy wartościami. Piszę jednak z rozmyślną ostrożnością: „w dużej mierze”, ponieważ, jak podkreśliłem nieco wyżej, wykreowanie i modyfikowanie wglądów w wartości (dodałbym także: preferencji) ma również komponent przeżyciowy i emocjonalny. Inaczej mówiąc, emocjonalnie przeżywamy błędy popełnione przez własne emocjonalne wglądy¹⁴.

gim przypadku. To samo można powiedzieć o metafizycznej opozycji realne-idealne. Według Nicolai Hartmanna wartości są bytami idealnymi i tworzą idealne uniwersum. Nie przeszkadza to w założeniu hierarchicznego porządku wartości, wręcz przeciwnie, jest ono raczej naturalne w tej koncepcji, jeszcze bardziej niż w ujęciu Schelera, ponieważ wybór pomiędzy dobrami o różnej mierze musi się tu dokonywać przez sam wgląd w wartości.

¹⁴ Połączenie idealnego statusu wartości z emocjonalnym ich przeżywaniem i częstymi błędami w rozpoznawaniu relacji jest bardzo intrygujące. Można powiedzieć, że uniwersum wartości ujmuje u Hartmanna pewną ogólną siłę afirmacji tkwiącą w doświadczeniu i mowie. Co dokładnie jest przedmiotem tej afirmacji, zależy od naszych (często błędnych) wglądów w stany rzeczy, lecz dynamika przeżycia wartości przeprowadza nas od jednych stanów afirmacji do innych, przy czym samo to przejście postrzegane jest emocjonalnie jako afirmacja samego potencjału i – można powiedzieć – wezwania do afirmacji.

Trudności związane z percepcją (emocjonalnym przeżyciem) wartości są analogiczne do trudności z ustalaniem preferencji. Ponieważ jednak, pomimo niedoskonałości, radzimy sobie jakoś z oboma tymi zadaniami, trzeba się zastanowić nad mechanizmami stabilizującymi nasze przeżycia aksjologiczne i preferencje, a następnie zapytać, czy mechanizmy te mają jakiś wspólny obszar, którego wskazanie pozwoliłoby lepiej ugruntować ludzkie roszczenia do racjonalności decyzji i działania.

4. Stabilizacja preferencji i wartości

Całe dotychczasowe rozumowanie prowadzi do wniosku, że zarówno zmienność w zakresie preferencji, jak zmienność w zakresie wartości domagają się pewnych środków stabilizujących. Wymagania dotyczące takiej stabilizacji są jednak dość odmienne w obu przypadkach, zatem i same środki są zupełnie różnej natury. W przypadku preferencji odwołam się do subiektywistycznej teorii decyzji i będę bronił tezy, że możliwa jest jedynie ograniczona i refleksyjna stabilizacja własnych preferencji. Jest ona ograniczona, ponieważ wstępne warunki, których wymaga, wykluczają wiele preferencji z możliwego uporządkowania. Jest refleksyjna, ponieważ posługuje się środkami logicznymi, które nie mogą należeć do efektywnego zasobu poznawczego każdego decydenta, a tylko do możliwej *teorii*, jaką można skonstruować w celu wyjaśnienia własnych (i cudzych, jako że teoria jest zawsze ogólna) preferencji. Z kolei stabilizacja wartości jest pewną transcendentalną konstrukcją – refleksyjną w bardziej jeszcze radykalnym sensie – w której specjalną rolę odgrywa podmiot, rozumiany jako zmienna przebiegająca przez *zbiór klas doświadczeń aksjologicznych*. Sens tego ostatniego wyrażenia stanie się jasny w punkcie B bieżącego rozdziału.

A. Stabilizacja preferencji

Niestabilność preferencji byłaby mniej kłopotliwa, gdyby dało się przedstawić preferencję jako relację pomiędzy przekonaniem na temat dóbr, zaś same przekonania zinterpretować jako stopniowalne, czyli możliwe do uporządkowania na skali interwałowej obejmującej wszystkie wartości rzeczywiste z jakiegoś przedziału. W potocznym sensie byłoby to trudne do wyobrażenia, choć idea pewnego rodzaju stopniowalności przekonań jest intuicyjnie zrozumiała, jak bowiem czytamy u Szekspira: „*So I heard and do in part believe it*”¹⁵. Pre-

¹⁵ Motto w książce Donalda Gillesa, *Philosophical Theories of Probability*, Routledge, London 2000.

czyjną i bardziej radykalną (wartości rzeczywiste, dobre uporządkowanie relacji) interpretacja owego „*in part*” byłaby już znacznie mniej intuicyjna. W istocie wymaga ona zupełnie nieintuicyjnych środków polegających na wprowadzeniu odwzorowania preferencji na skalę wartości rzeczywistych od 0 do 1, odpowiadających prawdopodobieństwom. Prawdopodobieństwa te są częścią złożonych sądów na temat loterii, czyli wyborów pomiędzy parami wielkości: wielkością pewnego dobra i prawdopodobieństwem jego uzyskania. Oczywiście nie potrafimy *wydedukować* własnej preferencji z (subiektywnego) pomiaru tych dwóch wielkości – pod tym względem nie jesteśmy dość kompetentni, co pokazują liczne badania uprzedzeń i zawodnych heurystyk, jakie stosujemy w ocenianiu prawdopodobieństwa stanów rzeczy na podstawie posiadanej wiedzy czy relacji pomiędzy dobrami. Klasyczna teoria decyzji von Neumanna i Morgensterna do pewnego stopnia czyniła nierealistyczne założenie, że decydując, kierujemy się własną funkcją użyteczności, która wymaga oceny prawdopodobieństwa. Można jednak te funkcje traktować jako teoretyczny i syntetyczny opis pewnej kompetencji, która przejawia się w wielu formach i nie ma jednej kanonicznej reprezentacji pojęciowej ani matematycznej. Teoria decyzji rozwijała się w krytycznym odniesieniu do wyników von Neumanna i Morgensterna, szukając alternatywnych reprezentacji owej tajemniczej kompetencji. Szczególny wkład wniosły tu ujęcia subiektywistyczne, które wyzbywają się kłopotliwego założenia, że do kształtowania, czy jak się wyraziłem, stabilizowania własnych preferencji potrzebna byłaby ocena prawdopodobieństwa. Stabilizacja preferencji zachodzi tu inaczej. Według teorii subiektywnego prawdopodobieństwa, opracowanej między innymi przez Leonarda Savage’a, można założyć, że podmiot decyzji wie zawsze, co woli, kiedy ma do czynienia z loterią. Zakłada się, że zbiór preferencji osoby racjonalnej daje się uporządkować zgodnie z zasadami, które można wyrazić za pomocą systemu formalnego. Savage udowodnił, że uporządkowany w ten sposób zbiór preferencji jest zgodny z rachunkiem prawdopodobieństwa, co znaczy, że siła relacji preferencji łącząca dobra zakłada takie różnice pomiędzy tymi dobrami, jakby wynikały one z różnego prawdopodobieństwa ich uzyskania. U Savage’a funkcja prawdopodobieństwa nie jest więc arbitralnie zakładana, lecz wyprowadzona z preferencji.

Aksjomaty Savage’a w nieformalnym ujęciu są następujące: **(A1)** Funkcja preferencji (dokładnie: jest lepsze lub równie dobre) jest relacją zupełną i przechodnią. **(A2) Definicja pomocnicza:** działania zgadzają się w zbiorze stanów B, jeśli powodują w tym zbiorze dokładnie te same skutki. **Aksjomat:** Wyobraźmy sobie dwie pary działań, czyli funkcji: f, f' oraz g, g' , które spełniają takie warunki, że dla zdarzenia kontrastowego w stosunku do pewnego zdarzenia (zdarzenia polegającego na tym, że dane zdarzenie nie zachodzi) funkcje te są parami zgodne, zaś dla zdarzenia zgodne są funkcje f z g i f'

z g' , to zachodzi zgodność preferencji pomiędzy tymi funkcjami: jeśli f jest preferowana przed g , to $i f'$ jest preferowana przed g' . **(A3)** Działania statyczne, czyli takie, których wartość skutków jest taka sama niezależnie od tego, jakie zdarzenie nastąpi, są preferowane jedne w stosunku do drugich wyłącznie na podstawie ich skutków. **(A4)** *Definicja pomocnicza*: B nie jest bardziej prawdopodobne niż A wtw gdy f przy założeniu A jest przynajmniej tak mocno preferowane jak f przy założeniu, że B , a także przy założeniu, że skutek fA jest preferowany w stosunku do $f\neg A$, oraz przy założeniu, że $fA = fB$ oraz $f\neg A = f\neg B$. *Aksjomat*: Dla każdego stanu rzeczy A i B , albo A nie jest bardziej prawdopodobne niż B , albo B nie jest bardziej prawdopodobne niż A . **(A5)** Nie jest tak, że każdy stan jest słabo preferowany w stosunku do wszystkich innych stanów. Musi istnieć stan, który jest gorszy lub lepszy w stosunku do jakiegoś innego stanu. **(A6)** Można przypisać działaniom liczby, jeśli podzielimy zbiór stanów na coraz drobniejsze podzbiory tak, że dla każdego, z wyjątkiem pewnego arbitralnie wybranego, będzie zachodziła zgoda pomiędzy odpowiednimi funkcjami, czyli działaniami.

Odkrywcze twierdzenie udowodnione przez Savage'a głosi, że liczby, które na mocy tych aksjomatów przypiszemy działaniom, mają takie własności, jak liczby, które przypisujemy niepewnościom, czyli miary prawdopodobieństwa. Inaczej mówiąc, wartości te są zgodne z aksjomatami Andrieja Kołmogorowa dla rachunku prawdopodobieństwa. Przerzucony zostaje most pomiędzy stopniowalnością ludzkich przekonań a sądami na temat prawdopodobieństwa odnoszącymi się do świata.

Widać jednak, że wynik Savage'a osiągnięty został za cenę wielu racjonalistycznych założeń. Założenia te czyni teoretyk, obserwator własnych procesów decyzyjnych. Razem z nimi wnosi normę racjonalności, ponieważ to w świetle normy nasze preferencje okazują się stabilne, spójne, pozwalające na ścisłe rozumowanie. Jako osoby refleksyjne szukamy w naszym systemie preferencji takiego regionu, do którego da się zastosować tego rodzaju normy. Savage i inni badacze stopni przekonania i preferencji w paradygmacie subiektywnej spójności dostarczają nam sposobu *testowania*, czy dla danego obszaru preferencji zachodzi ten związek – czy jesteśmy w tym zakresie racjonalni. Założona tu norma racjonalności jest jednak bardzo mocna i trudno się spodziewać, by te zakresy były szczególnie obszerne. W istocie, proponowane aksjomatyki (istnieją różne próby aksjomatyzacji, klasyczne osiągnięcie Savage'a to tylko przykład) narzucają silne restrykcje, a zatem wyznaczają bardzo wąskie klasy racjonalnych preferencji. Poza tym zakresem zachodzą w porządkowaniu preferencji liczne niespójności, co oczywiście wiemy już w chwili formułowania owych założeń – właśnie dlatego je formułujemy. Pozostaje zatem problem zobowiązania podmiotu decyzji do utrzymania swoich preferencji w stanie postulowanym przez te czy inne aksjomaty. Czy aksjomatyczne warunki narzucone na preferencje

mają zdolność stabilizowania pozostałych preferencji niemających oczekiwanych własności? Czy zaakceptowanie norm subiektywnej racjonalności sprawi, że mniej będziemy skłonni do niezrozumiałego dla samych siebie odwracania preferencji? Być może, lecz odpowiednia praca teoretyczna nie została jeszcze wykonana. Należałoby dać racjonalnemu podmiotowi nie tylko system, który w pełni odpowiada racjonalistycznym postulatom, lecz także pewien zbiór uporządkowanych systemów oddalonych w sposób kontrolowany od owego racjonalnego rdzenia. Jeśli wiemy, że nasze preferencje są dalekie od stabilności, to przynajmniej chcielibyśmy wiedzieć, jak dalekie.

B. Stabilizacja wartości

Celem tych rozważań jest wyjaśnienie związku pomiędzy wartościami i preferencjami. Wydaje się, że preferencje powinny wyrażać wartości, a wartości służyć do uzasadnienia preferencji. Lecz z powodów wyłożonych wyżej korelacja taka jest trudna do uzyskania. Wydaje się, że relacje pomiędzy dobrami będziemy postrzegać w świetle uznawanych wartości, lecz nader często postrzeganie tych relacji jest zdeterminowane zupełnie innymi czynnikami. Moja hipoteza głosi, że dla pewnego obszaru dobrze uporządkowanych preferencji korelacja z wartościami powinna być wyraźniejsza. Problem w tym, że również wartości odznaczają się niestabilnością. Analogicznie do poszukiwania stabilizacji preferencji trzeba więc rozważyć warunki stabilizacji wartości. W tym celu nawiążę do klasycznej koncepcji wartości Maxa Schelera. Wartości dane są, według Schelera, w przeżyciu emocjonalnym. Trudno sobie wyobrazić, by emocje miały utrzymać stałe relacje pomiędzy wartościami, potencjalnie wspierając dobrze uporządkowane preferencje. Ale Scheler oczywiście nie poprzestawał na wskazywaniu emocji, lecz rozwinął teorię ich działania w międzyludzkich relacjach. Otóż emocje rozwijają się w polu empatii. Wprawdzie relacje empatii są też notorycznie niestabilne i nie można zbudować klasy aksjologicznej (klasy doświadczeń empatii równoważnych z uwagi na pewną wartość) na zbiorze doświadczeń empatycznych, jednak Scheler pokazał pewne rozwiązanie problemu. Potraktował mianowicie empatię jako nakierowaną na przedmioty transcendentne – wartości. Empatia bowiem powstaje na bardziej podstawowym gruncie miłości, która prócz swego powszedniego sensu, czyli skierowania na osoby, ma i ten sens, że kieruje się na idealne związki o charakterze aksjologicznym: „...współodczuwamy tylko w takiej mierze i tak głęboko, jak kochamy. (...) Jednak nasze twierdzenie bynajmniej nie ma oznaczać, iż przedmiot, z którym współodczuwamy, musielibyśmy również kochać. Chodzi tu jedynie o ufundowanie aktów *wziętych jako istności*”¹⁶.

¹⁶ M. Scheler, *Istota i formy sympatii*, przeł. A. Węgrzecki, Warszawa, s. 224.

Być może literalna podstawa mojej interpretacji Schelera nie jest bardzo mocna, lecz sądzę, że na gruncie jego koncepcji decydująca dla empatii jest zdolność nakierowania się na pewien *wspólny przedmiot abstrakcyjny*. Jest to przedmiot aksjologiczny o charakterze aksjologicznym. Pamiętając o roli uczuć w poznawaniu wartości u Schelera, należałoby go nazwać przedmiotem ogniskującym emocje, czy po prostu przedmiotem emocjonalnym. Widzimy ów przedmiot jako osadzony w innej osobie (w przeżyciu empatycznym), lecz zarazem rozumiemy istotowy związek tej zobiektywizowanej cudzej emocji z własnymi stanami. Ten istotowy związek nie jest jednak ścisłym przypisaniem sobie samemu (jak w koncepcji wczucia, z którą Scheler polemizował), lecz co najwyżej przeżywaniem pewnej możliwości przypisania go sobie samemu. Zatem ontologicznie rzecz biorąc, wartości są wspólnymi przedmiotami możliwymi. Przedmioty te wyprowadzają poza immanencję przeżycia – implikują, zdaniem Schelera, wspólnotę. Analizowany przez filozofa przykład Robinsona na bezludnej wyspie ma pokazać, że na mocy istotnościowych związków nawet człowiek przebywając w osamotnieniu będzie wiedział, że jego życie jest odniesione do życia wspólnoty. Nie przyjdzie mu na myśl, że nie istnieje żadna wspólnota, choć żadnej nie doświadcza¹⁷. Przedmioty intencjonalne powołują do życia idealne związki pomiędzy osobami, ponieważ ich cechą jest nienasycenie. Towarzyszy im świadomość pustki, niespełnienia. Emocje wymagają dopełnienia przedmiotem, na który się intencjonalnie kierują.

Jeśli do stabilizacji własnego postrzegania wartości potrzebne jest odniesienie zarówno do wspólnoty, jak i do przedmiotów abstrakcyjnych (przedmiotów modalnych wyznaczających możliwość przeżycia pewnej emocji związanej z wartością), to dla szerszego ugruntowania wartości potrzebne jest rozumienie pewnej zmiennej w zbiorze, na który składają się różne osoby i ich przeżycia wartości. Sądzę, że taką zmienną jest racjonalny i empatyczny podmiot. Przechodząc od jednego doświadczenia aksjologicznego do drugiego, idąc śladem istotnych zależności, odkrywamy uogólniony podmiot aksjologiczny. Doświadczenia są podobne, lecz ich *podmioty* są różne. Jednak różnice pomiędzy podmiotami zbudowane są na gruncie istotowych podobieństw doświadczeń aksjologicznych. Chociaż związki fundujące konstrukcję podmiotu zachodzą pomiędzy przeżyciami i w szczególności emocjami, to jednak trzeba też założyć racjonalność uogólnionego podmiotu – a to z uwagi na konieczność rozumienia odpowiednich przedmiotów abstrakcyjnych w ich logicznych związkach. Trzeba też założyć istnienie empatycznych związków pomiędzy ludźmi, ponieważ przedmioty te prezentują się wielu osobom podobnie, o czym osoby te wiedzą na mocy emocjonalnej więzi.

¹⁷ Tamże, s. 351.

Jest jeszcze jedna cecha, która nakazuje przejście od realnych, konkretnych osób do pewnej normatywnej konstrukcji: refleksyjność. Podmiot nie jest tylko pojęciem teoretycznym, lecz sposobem, w jaki każda osoba może ujmować swoją relację zarazem z wartościami i wspólnotą osób do tych wartości odniesioną. H. Ginsborg¹⁸ słusznie podkreśla ważność sądów refleksyjnych w ustanawianiu wartości. Polegają one na uogólnianiu doświadczenia na inne osoby. Typujemy, że inni odczują to samo wobec danego doświadczenia. Innymi słowy, odwołujemy się do *emocjonalnego a priori*. Sensem tej konstrukcji jest zapewnienie nieempirycznej syntezy naszego życia emocjonalnego, jego zasadniczej zbieżności.

Istnieją oczywiście alternatywne wyjaśnienia tych samych fenomenów. Można odnieść wrażenie, że występuje dziś w filozofii szczególnie popularny nurt naturalistyczny, który ową normatywną i refleksyjną wspólnotę chce zastąpić jednością ugruntowaną w obiektywnych własnościach biologicznych i socjobiologicznych: geny, prawidłowości życia w stadach, mechanizmy związane z wychowaniem potomstwa. W optyce naturalistycznej znika refleksyjny podmiot – zobowiązania pochodzą skądinąd, wyznaczone są strukturą więzi osobistych czy grupowych. Nie jest to właściwe miejsce, by wchodzić w ogólną dyskusję z koncepcjami naturalistycznymi. Z punktu widzenia obecnych rozważań ważny jest tylko jeden argument: zobowiązania możliwe do wygenerowania na gruncie interpretacji naturalistycznej nie są *samo-zobowiązaniami*. A tymczasem bliski związek pomiędzy tymi dwoma zobowiązaniami jest kluczowy dla przeżycia wartości czy powinności. Nie wystarczy uznać, że pewna wartość ma moc wobec wielu osób; muszę założyć, że te osoby zobowiązują same siebie (formułują powinność i konstatają, że jej podlegają) na gruncie swojego własnego przeżycia wartości i że czynią to tak, jak ja sam to czynię na gruncie mojego przeżycia wartości. Idea bycia podmiotem wyraża to oczekiwanie, staje się zmienną przebiegającą zbiór doświadczeń i zobowiązań związanych z wartościami. Bycie sobą jako podmiotem byłoby natomiast pewną stałą – wartością tej zmiennej. Wspólnota aksjologiczna w sensie Schelera nie jest zbiorem osób w ich bezpośrednich relacjach, lecz zbiorem podmiotów wyodrębnionych ze struktur i więzi osobowych. To pojęcie lepiej, moim zdaniem, ujmuje refleksyjne samozobowiązanie w empatycznej wspólnotcie. Z powodu swojej abstrakcyjności pojęcie podmiotu jest łatwo przenoszone na ciała kolektywne (podmiot prawny), lecz w obecnych rozważaniach interesuje mnie wyłącznie podmiot jako pojęcie nałożone na indywidualne osoby i ich relacje z innymi indywidualnymi osobami.

¹⁸ H. Ginsborg, *Normativity in Nature. Essays on Kant's Critique of Judgment*, Oxford University Press 2014.

5. Ulepszające zmiany w zbiorach preferencji i wartości

W ostatnim kroku rozumowania chcę zaproponować pewną ogólną ideę połączenia obu mechanizmów stabilizujących: porządkowanie preferencji i porządkowanie wartości. Łączy je, jak sadzę, idea racjonalności. Decydujące i działające osoby myślą o sobie jako o podmiotach; ich przeżycia wartości podniesione są do pewnej ogólności przez uznanie, że istnieje podmiot – pewna zmienna, której prawidłowy przebieg przez zbiór doświadczeń aksjologicznych pokazuje jego spójność. Lecz stabilizacja przeżycia wartości nie daje jeszcze klucza do utrzymania spójności w obrębie konkretnych dóbr, które mają egemplifikować czy realizować te wartości. Wciąż może być wiele powodów, dla których preferowane przez samego siebie dobra rozmiągają się z własnym przeżyciem wartości. Jeśli jednak rozważymy tylko sam racjonalny rdzeń preferencji – podzbiór własnych preferencji spełniający rygorystyczne warunki logiczne, sprawa wygląda lepiej. W obrębie tego rdzenia wartości i preferencje zgadzają się ze sobą.

Spełniony jest tu również warunek samo-normowania: Otóż ani idealny podmiot empatyczny i refleksyjny, ani uporządkowany system preferencji nie są opisami *faktów*; są wytycznymi mówiącymi, jak myśleć o samym sobie. W tym myśleniu o samym sobie poszukujemy wzorca, którym powinniśmy się kierować w uzgadnianiu relacji pomiędzy tym, co sądzimy o dobru (jakie wartości uznajemy), a tym, co bezpośrednio wyjaśnia nasze działania, czyli preferencjami, które ujawniamy w konkretnej decyzji. Nasze sądy o wartościach są kontrolowane emocjonalnie i społecznie, lecz mają one generować racjonalne, *indywidualne* preferencje. Wzorcowe preferencje są uporządkowane tak, jak tego wymaga subiektywistyczna teoria decyzji w dowolnej jej postaci (np. aksjomatyki Savage'a). Zatem nasza praca nad formowaniem własnych sądów wartościujących musi wziąć pod uwagę pewien punkt docelowy: ład w dziedzinie własnych preferencji. Inaczej mówiąc, cokolwiek człowiek uzna za dobre, powinno generować spójne preferencje. Jeśli się tak nie dzieje, możemy spróbować uzyskać lepszy wgląd w uznawane przez siebie wartości, a także przemyśleć swoje preferencje w tym celu, by lepiej oddawały nasz na nowo przemyślany (zrewidowany) system wartości. *Racjonalny podmiot* jest tą sub-strukturą osoby, w której obydwa procesy się spotykają.

Powstaje zatem pytanie, w jaki sposób można racjonalnie *doskonalić* przekonania aksjologiczne i preferencje, by maksymalnie zgadzały się ze sobą i generowały spójny system przekonań i preferencji. Wydaje się, że chodzi o model, który rodzi się na przecięciu obu modeli: (1) modelu refleksyjnego podmiotu, który jest zmienną w funkcji odwzorowującej sytuację na wartości; (2) modelu racjonalnego podmiotu porządkującego swoje preferencje. Interakcja pomiędzy tymi modelami polegałaby na narzuceniu sobie granic

w postrzeganiu samego siebie jako podmiotu ze względu na racjonalny łańd własnych preferencji.

Nasze społeczne ja podlega racjonalnej restrykcji; gdyby tak nie było, szukalibyśmy takiego samookreślenia jako podmiotu, które zapewniałoby nam jak najszersze uczestnictwo we wspólnocie ludzkich doświadczeń niezależnie od ich nielogiczności, chaotyczności i dramatów, jakie muszą wówczas wywoływać konflikty wartości. Tego rodzaju podmiotowość można by nazwać *szekspirowską*, pamiętając, jak mistrzowsko Szekspir sportretował ów chaos przeciwstawnych dążeń, wobec którego zdaje się nie istnieć jakikolwiek zewnętrzny punkt widzenia mogący pełnić rolę arbitra. Z kolei zewnętrzny racjonalny punkt widzenia, promowany jest przez podmiot, który nazwijmy dla kontrastu *kantowskim*. Roszczenia tego podmiotu muszą być ograniczane przez wymagania życiowego realizmu. Musimy zaakceptować fakt, że dobre uporządkowanie własnych preferencji jest raczej normą niż faktem. Widać, że mamy tu do czynienia z dwoma zobowiązaniami: z jednej strony jest to wysiłek racjonalnego kształtowania życia, z drugiej strony realizmu w traktowaniu samych siebie i w kształtowaniu związków międzyludzkich.

Powstaje jednak trudne pytanie: w jaki sposób te dwa wzorce mają na siebie oddziaływać? Co dzieje się w ich polu przecięcia? Jak działa „podmiot szekspirowski” na „podmiot kantowski” i odwrotnie? Wydaje mi się, że oddziaływanie racjonalnego (normatywnego) rdzenia podmiotu na spontaniczne przeżywanie wartości można sobie wyobrazić jako uogólniony bayezjanizm. Jest to sposób, w jaki zmieniamy siłę swoich przekonań pod wpływem nowych doświadczeń (informacji). Podmiot szekspirowski ma skłonność, by iść za przychodzącym doświadczeniem, kolejne przeżycia przysyłają mu poprzednie, czy nawet całkiem usuwają ich wpływ, impulsywność, brak wglądu w konsekwencje swych działań, brak dystansu do siebie są jego charakterystycznymi cechami. Bayezjanizm mówi jednak, by nawet najbardziej wyraziste nowe doświadczenie modyfikowało posiadany korpus przekonań tylko w ograniczonej mierze wyznaczonej przez regułę Bayesa¹⁹ – oczywiście tam, gdzie dysponujemy danymi, które można do takiej formuły podstawić; w innych przypadkach reguła Bayesa dostarcza tylko ogólnej heurystyki w zakresie postępowania z własnymi przekonaniem. Subiektywistyczna teoria decyzji, która posłużyła mi w tych rozważaniach za model owego rdzeniowego racjonalizmu, jest silnie związana z bayezjanizmem. Na przykład kiedy otrzymujemy pozytywny test stwierdzający pewną chorobę i znając trafność i swoistość (specyficzność) tego testu podaną w procentach błędów pozytywnych i negatywnych, chcemy przekonać się, z jakim prawdopodobieństwem jesteśmy chorzy, to musimy

¹⁹ Można powiedzieć, oczywiście *cum grano salis*, że postępowanie Otella czy Leontesa stanowi przykład sprzeczności z tą regułą.

uwzględnić: (1) procent pozytywnych wyników testu, za którymi stoi choroba; (2) wszystkie przypadki chorób, które nie zostały wykryte; (3) procent chorych na tę chorobę w populacji. Właśnie ta ostatnia wielkość jest kontrowersyjna i notorycznie pomijana. A tymczasem łatwo wykazać jej nieodzowność: test wybiera przypadki chorobowe właśnie z tej puli, z całej zaś pozostałej populacji tylko wtedy, gdy się myli biorąc osobę zdrową za chorą²⁰. Tego rodzaju szacunki są ważne, jeśli na przykład mam odpowiedzieć, czy wolę operację, która leczy z prawdopodobieństwem p i zabija z prawdopodobieństwem q , od nierobienia nic w sprawie choroby, która powstaje z prawdopodobieństwem k i kończy się śmiercią z prawdopodobieństwem g .

Niemniej jednak w dzisiejszej teorii decyzji trwa dyskusja na temat stosowności reguły Bayesa do realnych procesów decyzyjnych. Zdaniem Martina Petersona²¹, wadą podejścia bayesowskiego jest to, że nie wyjaśnia preferencji, tylko się na nich opiera. Kiedy decydent wie już, co woli, i tym samym wie, co zrobić, może skonstruować funkcję bayesowską mówiącą, *o ile to woli* od określonej alternatywnej możliwości. Inaczej mówiąc, trzeba już na wstępie wiedzieć, co chce się uczynić – bayesowskie rozumowanie może pomóc zrozumieć, jak bardzo chce się to zrobić (w stosunku do innych możliwości). Ponieważ w moich obecnych rozważaniach chodzi o same podstawy stabilnych preferencji, wydaje się, że bayezjanizm niewiele tu pomoże. Rozważmy przykład. Opieram swoją dzisiejszą aktywność na dość silnym przekonaniu, że jestem stosunkowo zdrowy: snuję odległe plany, nie mam w rozkładzie najbliższego miesiąca zaplanowanych badań lekarskich itd. Ale w jakiej mierze moje dzisiejsze przekonanie jest silniejsze niż analogiczne przekonanie z zeszłego miesiąca, jeśli właśnie w ciągu ostatniego miesiąca wykonałem kilka podstawowych analiz, które (z określonym stopniem niezawodności) pozwalają sądzić, że jestem zdrowy? Otóż mój stopień przekonania o własnym zdrowiu ma prawo wzrosnąć na skutek ostatnich badań, lecz tylko w niewielkim stopniu dopuszczonym przez regułę Bayesa, zgodnie z którą mogę obliczyć, jakie jest prawdopodobieństwo tego, że otrzymane wyniki spowodowane są zadowalającym stanem mojego zdrowia. Respektowanie bayesowskiej procedury jest bardzo ważne, lecz działa ona w obrębie już ukształtowanego przekonania. Przekonanie to, razem ze związaną z nim wartością (zdrowia), wyjaśnia moje bieżące preferencje: wolę dziś popracować i pójść potem do kina, niż myśleć o kolejnych badaniach medycznych, skądinąd rozsądnych i leżących w moich planach, lecz na odległym miejscu. To, że są na odległym miejscu, wynika z moich przekonań i wartości, natomiast to, na *jak odległym* (relatywnie)

²⁰ G. Gigerenzer, *Reconing With Risk. Learning to Live With Uncertainty*, Penguin Books, London 2002, s. 170–172.

²¹ M. Peterson, *Introduction to Decision Theory*, Cambridge University Press 2009, s. 205.

są miejscu, zależy od bayesowskich restrykcji, które, jako osoba racjonalna, nakładam na zmienianie własnych przekonań pod wpływem nowych informacji.

Być może krytycy mają rację: bayesowski proces modyfikowania wiedzy (porządkowania preferencji) jest zbyt powolny, niekonkluzywny i zbyt wiele zakłada. Lecz z filozoficznego punktu widzenia, podejście bayesowskie nie musi wcale służyć przewidywaniu i wspomaganie decyzji. Nie polepsza scenariuszy działania i nie nakazuje działania. Przyczynia się raczej do zrozumienia *podmiotu decyzji*. Inaczej mówiąc, opisuje *kompetencję* do podejmowania decyzji i czyni to przez opisanie *zdolności* podmiotu do porządkowania jego preferencji w granicach jego wiedzy. Czasem korzystamy z tej zdolności w pełni, głównie w działalności naukowej, czasem jedynie częściowo, używając jej jako ogólnej heurystyki, czasem wcale nie korzystamy – jesteśmy po prostu podmiotami szekspirowskimi.

Jest równie trudno wyjaśnić, jak podmiot szekspirowski oddziałuje na podmiot kantowski. Ten pierwszy, jeśli pójść za wnikliwą interpretacją Stanleya Cavella²², cechuje głębokie doświadczenie sceptyczne. Penetracja i modyfikacja wiedzy i postaw przez postawę sceptyczną jest czymś odwiecznym i centralnym w kulturze Zachodu. I nie jest całkiem jasne, w jaki sposób oddziałuje ona sama i jej fundujące doświadczenie różnorodności i niewspółmierności ludzkich wartości. Podmiot kantowski, jeśli poważnie rozumie swoje normatywne zadanie, musi ustosunkować się do samego siebie jako empirycznego podmiotu, jego reguły racjonalności muszą być zakorzenione w cechach i skłonnościach empirycznych podmiotów. W szczególności musi respektować tę okoliczność, że niezwykle trudno utrzymać porządek własnych preferencji wobec nowych pragnień, spostrzeżeń, wpływów itd. Napięcie, jakie tu powstaje, musi być tonowane przez pewne wstępne porządkowanie doświadczenia z jednej strony, oraz przez wbudowanie pewnych ograniczeń kontekstowych w racjonalne zasady postępowania.

Wzór rozumowania w takich kwestiach znajdujemy w filozofii praktycznej Kanta oraz w jego uwagach pedagogicznych. Zdawał on sobie sprawę z tego, że istnieje luka pomiędzy akceptacją przez daną osobę powinności uformowanej zgodnie z imperatywem kategorycznym a skłonnością do działania w określony sposób. Postulował pewien pośredni poziom zbudowany na uczuciu szacunku do tych, „którzy stawiają nam przed oczy prawo”.

Analogiczna dyskusja toczy się dziś w obrębie teorii decyzji. Sugeruje się z jednej strony, że normy racjonalnej teorii decyzji są przesadne i że nasze błędne heurystyki w istocie są optymalne, jeśli chodzi o dostosowanie do normalnych warunków życia i wydolności naszych systemów poznawczych.

²² Zob. S. Cavell, *Disowning Knowledge in Seven Plays of Shakespeare*, Cambridge University Press, New York 2003.

Z drugiej strony mówi się o możliwości poprawienia własnych kompetencji w zakresach ważnych dla podejmowania racjonalnych decyzji – w tym duchu ukazuje się na przykład wiele praktycznych poradników. Nie ma tu miejsca, by wchodzić w dyskusję z pomysłami i tendencjami dotyczącymi pola przecięcia obu modeli podmiotu. Celem tych zwięzłych uwag było tylko pokazanie, że pole to nie jest puste: lepiej lub gorzej oba modele – szekspirowski i kantowski podmiot – pozostają w ciągłej interakcji. Pozostaje wszakże pytanie, którego w tych rozważaniach nie podejmuje: dlaczego interakcja pomiędzy obu modelami podmiotu miałaby prowadzić do zmiany *ulepszającej*? Nie jest wykluczone, że owo pole przecięcia generuje jedynie dalszą konfuzję, zaś intuicja polepszenia jest zarazem niejasna i uwikłana w zbyt wiele założeń. Zbadanie tej kwestii wymagałoby ogólniejszej teorii dobra, w ramach której podmiot – oderwany w tych rozważaniach od struktury osoby, zostałby jej niejako przywrócony.

Streszczenie

Artykuł przedstawia problemy wynikające z rozbieżności pomiędzy uznawanymi i przeżywanymi wartościami a preferencjami w odniesieniu do dóbr przejawianymi w sytuacjach decyzyjnych. W analizie połączone zostały wybrane postulaty klasycznych koncepcji wartości (Max Scheler, Nicolai Hartmann) z subiektywistycznym nurtem teorii decyzji (Leonard Savage). Próbuję opisać mechanizmy stabilizujące postrzeganie wartości i kształtowanie preferencji, a następnie zastanawiam się nad wzajemnym oddziaływaniem obu tych stabilizujących mechanizmów. Trwałe utrzymanie harmonii pomiędzy nimi wydaje się niemożliwe, lecz możliwe jest jej doskonalenie. To ostatnie stanowi istotę racjonalnego podmiotu.