

Katarzyna Paprzycka

## **Falszywa świadomość psychologii intencjonalnej O pewnym dogmacie współczesnej analitycznej filozofii umysłu\***

Zasadniczym celem analitycznej filozofii umysłu jest — rzecz jasna — analiza pojęcia umysłu. Jeden z ważnych sposobów podejścia do tego zadania to analiza naszego potocznego rozumienia umysłu.<sup>1</sup> Tak się — dość szczęśliwie — składa, że panuje zgoda co do zasadniczego kształtu tego pojmowania umysłu. Właściwie wszyscy

---

\* Głównej tezy miałam już okazję bronić w dwóch artykułach (1998, w druku). Bezpośrednią inspiracją do napisania poniższego tekstu były seminaria Zakładu Epistemologii UAM Prof. A. Klawitera. Jestem wdzięczna za uwagi wszystkim uczestnikom tych seminariów, a w szczególności chciałabym podziękować Andrzejowi Klawiterowi, Krzysztofowi Łastowskiemu, Agnieszce Pałac oraz Piotrowi Przybyszowi. Myśli tu przedstawione rozwijane były także dzięki dyskusjom z wieloma osobami, wśród których szczególnie ważne role odegrali Robert Almeder, Annette Baier, Robert Brandom, Felmon Davis, Peter Klein, Mark Lance, Ken Manders, Leszek Nowak oraz John McDowell. Pragnę też podziękować Prof. Jackowi Jadackiemu za skrupulatną lekturę oraz szczególne uwagi, które niejednokrotnie pozwoliły mi uniknąć niedorzeczności.

<sup>1</sup> Innym sposobem byłaby analiza naukowego rozumienia umysłu. Szybko się jednak okazuje, że nie za bardzo wiadomo, na którą analizę naukową należałoby wskazać. Ci filozofowie, którzy *explicitie* uznają, iż mówią o naukowym sposobie podejścia do umysłu, a mianowicie eliminatywiści (np. P. M. Churchland [1979], [1989a]; P. S. Churchland [1986]; Rorty [1965]), zakładają, iż należy się powoływać na zdobycze neuronauki. Inni (np. Fodor [1975]; Dennett [1991]) wskazują z kolei na zdobycze kognitywistyki (a wcale przecież nie jest jasne, że ta redukuje się do neuronauki). Dalej w tym artykule sugeruję — choć nie objaśniam tej tezy szczegółowo — że zasadniczą rolę w rozumieniu umysłu może dać (wbrew pozorom) np. psychologia społeczna. Słowem, nie istnieje żaden konsensus między tymi, którzy uznają, że należy odwołać się do refleksji naukowej, aby zrozumieć umysł. Wydaje się zatem, że nasze preteoretyczne pojęcie umysłu jest nie tylko nieprecyzyjne, ale miesza w sobie wiele poziomów myślenia o psychice. Refleksja naukowa rozwija niektóre z tych poziomów, ale kosztem odsunięcia innych — a każdy filozof umysłu odwołując się do nauki o umyśle przejmując rozstrzygnięcia pojęciowe przez tę naukę *implicitie* obrane.

analityczni filozofowie umysłu są zdania, że nasza potoczna psychologia jest psychologią pragnień i przekonań (psychologią intencjonalną).<sup>2</sup> Nie zgadzają się między sobą co do tego, czy psychologia intencjonalna jest trafna czy nie (por. P. M. Churchland [1979], [1984], [1989a]; P. S. Churchland [1986]; Stich [1983]). Nie zgadzają się co do tego, czy należy ją interpretować dosłownie (Fodor [1975], [1987]) czy tylko instrumentalnie (Dennett [1981a], [1987]). Nie zgadzają się też, czy należy ją pojmować jako teorię (P. M. Churchland [1989b]; Moran [1994]; Stich i Nichols [1992], [1995]) czy jako szczególną formę empatii (Gordon [1986], [1992]; Goldman [1989], [1992]). Jednakże wszystkie te spory opierają się na założeniu, iż psychologia potoczna *jest* w zasadzie psychologią pragnień i przekonań sprawcy.<sup>3</sup>

W poniższym tekście założenie to zakwestionuję. Jeśli przyjrzeć się wyjaśnieniom działań, które w praktyce podajemy, od razu widać, iż nie zawsze odwołują się one do pragnień i przekonań — a jeśli już, to często nie sprawcy a innej osoby. „Dlaczego Zuza myje zęby? Bo kazała jej to zrobić jej mama.” „Dlaczego Kazik podał sól? Bo poprosił o nią współbiednik.” Na takich (przynajmniej zewnętrznie) „nieintencjonalnych” wyjaśnieniach skoncentrujemy się w poniższym tekście. Pomimo tego, iż wyjaśnienia te jawią się nam jako zupełnie naturalne, wielu badaczy uważa, że w każdym wypadku *musimy* mieć do czynienia z pragnieniem jednostki. Ich zdaniem nie jest tak — jak można by naturalnie przypuszczać — iż to, jakie są najlepsze wyjaśnienia działań, jest kwestią empiryczną. Pewne wyjaśnienia wykluczają oni *a priori*, odwołując się li tylko do argumentacji filozoficznych. Okazuje się więc, że dziedzina potocznych wyjaśnień naszych działań objęta jest pewnym pojęciowym przymusem.

Temu właśnie pogładowi chcę się tutaj przeciwstawić. Utożsamić go można z pewną dyrektywą metodologiczną, a mianowicie dyrektywą indywidualizmu eksplanacyjnego. Pomimo tego, iż dyrektywa ta jest powszechnie przyjmowana (opiera się na niej w szczególności pogląd, iż psychologia potoczna jest psychologią pragnień i przekonań) — nie jest uzasadniona. Z drugiej strony myśl indywidualistyczna

<sup>2</sup> Dotyczy to głównie filozofii anglosaskiej. Oczywiście podejście intencjonalne do rozumienia działań można znaleźć już w Szkole Lwowsko-Warszawskiej u autorów dosyć skądinąd różnych, jak Ajdukiewicz [1938] czy Czeżowski [1959], choć nie jest to już tak jasne w wypadku prac Kotarbińskiego [1955]. Niewątpliwie też silna jest orientacja akcentująca wagę działań racjonalnych (np. Kmita i Nowak [1968]; Kmita [1971]; Koziński [1975]; Szaniawski [1994]) zalecająca w szczególności rekonstrukcję wyjaśnień w naukach społecznych w duchu indywidualizmu metodologicznego — tej wszakże orientacji nie zamierzam dalej podważać.

<sup>3</sup> Istnieją także głosy przeciwne — jedne bardziej, inne mniej wyraźne, ale tak rzadkie, iż można je wyliczyć: Baier [1985]; Collins [1987]; Fleming [1981]; Nowak [1987, 1991]; von Wright [1983]; Wilson [1989]. Należałoby tu także wymienić *Dociekania filozoficzne* Wittgensteina [1958/2000] — już choćby z tego względu, iż paradygmatycznym przykładem działania jest tam wykonanie rozkazu. Nie przeszkadza to jednak przyswojeniu myśli Wittgensteina przez wielu indywidualistów. W związku z tym argumentację na rzecz nieindywidualizmu Wittgensteina trzeba odłożyć na inną okazję.

znajduje swoje naturalne miejsce jako myśl normatywna (niekoniecznie jest tak, iż *zawsze* działamy powodowani swoimi własnymi pragnieniami, ale jest tak, iż działają tak *powinniśmy*). Można więc przypuszczać, że intuicyjność indywidualizmu eksplanacyjnego — który jest, moim zdaniem, nietrafny — bierze się stąd, iż trafne jest stanowisko pokrewne, mianowicie indywidualizm normatywny. Jeśli indywidualizm eksplanacyjny wydaje się nam naturalny, to znaczy, że mylimy go po prostu z indywidualizmem normatywnym.<sup>4</sup>

W rozważaniach poniższych chodzi oczywiście nie tylko o sprawę, która może się wydawać banalna, a mianowicie o to, jaki jest kształt psychologii potocznej jako takiej. Chodzi również o podstawę dla rozumienia umysłu w ogóle, a doktryna indywidualizmu eksplanacyjnego rości sobie prawa do ograniczania teorii zarówno psychologicznych, jak i filozoficznych.

## 1. PSYCHOLOGIA INTENCJONALNA

Jedną z głównych tez psychologii intencjonalnej jest twierdzenie, że jeśli jednostka  $\alpha$  pragnie tego, że  $p$ , oraz jest przekonana, że w danej sytuacji najlepszym sposobem osiągnięcia tego, że  $p$ , jest sprawienie tego, że  $q$ , wówczas *ceteris paribus*  $\alpha$  sprawi, że  $q$ .

---

<sup>4</sup> Termin 'indywidualizm' jest już tak bardzo nadużywany w literaturze fachowej, że dodanie kolejnego indywidualizmu — tym razem eksplanacyjnego, nie może być mile widziane. Niemniej jednak, jest to termin, który najlepiej oddaje intencje teoretyczne za nim stojące, pozostaje więc tylko *explicite* odróżnić go od niektórych innych stanowisk również określanych tym mianem. (1) Indywidualizm eksplanacyjny różni się od indywidualizmu metodologicznego tym, że ten drugi dotyczy kwestii, czy zjawiska społeczne należy analizować za pomocą działań jednostek, podczas gdy ten pierwszy dotyczy tego, jak należy rozumieć (wyjaśniać) działania jednostek. (2) Spór między indywidualizmem i nieindywidualizmem eksplanacyjnym nie ma nic wspólnego ze sporem między indywidualizmem i antyindywidualizmem dotyczącym treści umysłowych, którego dotyczą klasyczne prace m.in. Putnama (1975) czy Burge'a (1979) po stronie antyindywidualistów, a Searle'a (1983), Dennetta (1987) czy Fodora (1994) po stronie indywidualistów. (3) Filozofowie działania (por. np. French 1985; Tuomela 1984; Topolski 1973ab) spierają się o to, czy podmiotem działania musi być jednostka ludzka, czy może nią być kolektyw (np. firma, państwo, grupa ludzka). Ta dyskusja jest również niezależna od kwestii tu dyskutowanej. (4) Można się zastanawiać nad tym, czy możliwym jest, by podmiot działający funkcjonował bez społeczeństwa. Atomści (niekiedy też określane jako indywidualiści) twierdzą, że nie jest to supozycja logicznie sprzeczna, holiści (niekiedy nazywani nieindywidualistami) natomiast uznają nasze relacje z innymi za konstytuujące naszą naturę jako sprawców. (5) Można rozważać problem, czy działanie rozmaitych prawidłowości społecznych w jakikolwiek sposób zagraża naszemu obrazowi nas samych jako sprawców intencjonalnych. Indywidualiści (np. Pettit 1986) negują, a kolektywiści uznają tezę, że istnienie prawideł społecznych zagraża uniwersalności psychologii intencjonalnej. Choć rozróżnienie między w ten sposób rozumianym indywidualizmem i kolektywizmem nie pokrywa się z rozróżnieniem między indywidualizmem a nieindywidualizmem eksplanacyjnym, to zachodzi między tymi stanowiskami pewien związek (por. Paprzycka 1998).

Sformułowanie to wymaga od razu pewnych wyjaśnień. Po pierwsze, należy zwrócić uwagę na fakt, iż psychologia intencjonalna dotyczy działań, a nie zachowań. Obrazowo rzecz ujmując, wyjaśnieniom intencjonalnym podlega to, dlaczego ktoś schodzi ze schodów, a nie to, dlaczego z nich spada (uprzędno się potknąwszy). Po drugie, zwraca uwagę fakt, iż odwołuje się tylko do pragnień (*desires*) i przekonań (*beliefs*), a jest to odwołanie z pewnością niepełne. Wszakże potoczne wyjaśnienia działań nie ograniczają się tylko do stanów intencjonalnych<sup>5</sup> (a więc takich, które charakteryzuje pewna treść), a jeśli już takie ograniczenie wprowadzić, to oczywiście stany intencjonalne przywoływane w wyjaśnieniach obejmą też chęci, żądze, intencje, nadzieje, obawy, przypuszczenia, sugestie, etc. Pomimo tego, iż ta wielorakość rzeczywiście charakteryzuje potoczne wyjaśnienia, są argumenty za tym, by skoncentrować się tylko na pragnieniach i przekonaniach.<sup>6</sup>

Najbardziej znany z tych argumentów, który zaczerpnięty jest z głośnej pracy G. E. M. Anscombe [1963], został niedawno wykorzystany przez M. Smitha [1987, 1994] w nieco odmiennym, choć pokrewnym kontekście.<sup>7</sup> Można go przedstawić w sposób następujący. Zastanówmy się, w jakie minimum stanów intencjonalnych musiałaby być wyposażona jednostka, by mogła działać. Aby działanie było możliwe jednostka musi być w takim stanie, który będzie ją skłaniał do tego, by zmienić świat, a więc który będzie odzwierciedlał jej cel działania. Można obrazowo powiedzieć, że ten stan ma kierunek „świat-do-umysłu” (*world-to-mind fit*).<sup>8</sup> Jednocześnie, aby działanie było sensowne potrzebny jest też taki stan, który będzie informować jednostkę o aktualnym stanie świata, a więc stan, który ma kierunek „umysł-do-świata” (*mind-to-world fit*).

Pragnienie w najszerszym sensie (albo dążenie) można rozumieć jako taki stan intencjonalny, który ma kierunek „świat-do-umysłu”.<sup>9</sup> Przekonanie natomiast można rozumieć jako taki stan intencjonalny, którego kierunek jest odwrotny — „umysł-do-świata”. O zasadności tego rozumienia różnicy między pragnieniem a przekonaniem można się przekonać myśląc o tym, co się dzieje w sytuacji, gdy dany stan intencjo-

<sup>5</sup> Trudno w ten sposób interpretować np. wyjaśnienia odwołujące się do cech charakteru, szczególnie takich jak spóźniałstwo czy niedbalstwo.

<sup>6</sup> W rzeczy samej istnieją też argumenty przeciwne — za wyróżnieniem dalszych postaw. Prace m.in. G. E. M. Anscombe [1963], D. Davidsona [1978], a w szczególności M. Bratmana [1987] przyczyniły się do tego, że intencje uznaje się dziś dość powszechnie za niesprowadzalne do konstelacji pragnień i przekonań.

<sup>7</sup> Celem Smitha jest obrona poglądu neohume'owskiego przed poglądem neokantowskim. Tym niemniej argument ten daje się wykorzystać również do poparcia omawianej tezy.

<sup>8</sup> Wyrażenie 'world-to-mind' jest skrótem od 'the world to fit the mind'. Odpowiednio skrót „świat-do-umysłu” należy czytać „świat-(dopasować-)do-umysłu”.

<sup>9</sup> Takie jest rozumienie terminu 'pro-attitude' wprowadzonego przez Davidsona [1963] właśnie dla objęcia wielorakich stanów pokrewnych pragnieniom w tym sensie, że motywujących jednostkę do działania. Można domniemywać, iż chodzi tu po prostu o zakres pojęcia 'dążenie' w rozumieniu Ajdukiewicza [1938], który obejmować ma właśnie pragnienia i chcenia z jednej strony, a postanowienia i zamierzenia — z drugiej.

nalny nie jest spełniony. Załóżmy, że pewna osoba żywi przekonanie, że *p*. Jeśli zdarzy się tak, że świat akurat jest taki, że nie-*p*, wówczas osoba ta musi jakoś skorygować swój umysł (np. przez odrzucenie przekonania, że *p*). Jeśli natomiast mamy do czynienia ze stanem intencjonalnym o kierunku „świat-do-umysłu”, takim jak pragnienie, żeby *p*, wówczas jeśli zdarzy się tak, że nie-*p*, to wcale nie umysł, lecz właśnie świat należy skorygować (np. przez sprawienie, że *p*).

Przy takim rozumieniu rozróżnienia między pragnieniem a przekonaniem widać wyraźnie, że pojęcia te tak odbiegają od swojej potocznej konotacji, że wolno zasadnie mniemać, iż mogą one pełnić funkcję pojęć nadrzędnych określających dwa zasadniczo różne zbiory stanów intencjonalnych. Konkludując — zarzut, że utożsamienie psychologii potocznej z psychologią intencjonalną pomija całą wielorakość stanów intencjonalnych, jest nieporozumieniem. Wcale jej nie pomija, lecz ujmuje ją na nieco bardziej abstrakcyjnym poziomie.

Wreszcie — dla potrzeb niniejszego tekstu — psychologia intencjonalna rozumiana jest na osobowym (a nie sub-osobowym — por. Dennett [1969/1995]) poziomie analizy zjawisk psychicznych. To znaczy w szczególności, iż mowa tutaj o takim poziomie analizy umysłu, dla którego to, co się dzieje literalnie wewnątrz jednostki (w szczególności w jej mózgu) jest tak nieprzejrzyste, jak czarna skrzynka. Gdyby psychologia intencjonalna była trafną formą opisu i rozumienia naszych działań na poziomie osobowym (jak chce np. Dennett, por. [1971], a szczególnie [1978]), to jest zupełnie jeszcze otwartą kwestią, czy jest ona również trafną formą rozumienia naszego funkcjonowania na poziomie sub-osobowym (jak chce np. Fodor [1987]) — czy nie (jak sugerują niektórzy konekcjoniści, por. Smolensky [1995ab]).

## 2. INDYWIDUALIZM I NIEINDYWIDUALIZM EKSPLANACYJNY

### 2.1. Wyjaśnienia odwołujące się do własnych i cudzych pragnień

O trafności i powszechności psychologii pragnień i przekonań świadczyć mają również badania empiryczne prowadzone przez psychologów. W jednym z takich badań (Gopnik i Wellman [1992]) pokazywano trzy- i czterolatkom obrazek dziecka zagląającego pod tapczan, pod którym znajdował się pluszowy kotek, odnotowując przy tym, jak będą wyjaśniać zachowanie dziecka. Dzieci z łatwością przytaczały wyjaśnienia w rodzaju „dziecko chce się pobawić z kotkiem” czy „dziecko myśli, że kotek jest pod tapczanem”. Badania te wydają się w istocie świadczyć o tym, iż wyjaśnienia przez odwołanie do pragnień i przekonań (w szerokim sensie) opanowane są już przez małe dzieci.

Nie przytaczam tu tych badań, aby kwestionować ich rzetelność. Chcę natomiast zaoponować przeciwko pewnemu — jak sądzę — zafałszowaniu rzeczywistego obrazu naszej psychologii potocznej, któremu sprzyjają. Rzeczywiście sporą część psychologii «ludowej» daje się zrekonstruować w terminach psychologii pragnień i prze-

konań. Niemały jej fragment jednakże stanowią wyjaśnienia innego typu, a w szczególności te, które odwołują się nie do pragnień sprawcy, lecz do pragnień innej osoby. Dla zilustrowania, wyobraźmy sobie — z konieczności spekulując — że dzieciom pokazuje się dokładnie taki sam obrazek z tym tylko wyjątkiem, że obok dziecka stoi młoda kobieta.<sup>10</sup> Otóż byłoby zupełną rewelacją empiryczną gdyby dzieci oprócz wyżej wspomnianych wyjaśnień nie podawały czasem i innych w rodzaju: „mama kazała dziecku poszukać kotka” albo „ciocia prosiła, aby dziecko spojrzało pod tapczan”. Wyjaśnienia te nie mają jednak charakteru wyjaśnień w sensie psychologii intencjonalnej. O ile można je rozumieć jako odwołujące się do pragnień, to w każdym razie nie pragnień dziecka (jednostki podejmującej działania), tylko innej osoby (matki czy opiekunki).

## 2.2. Indywidualizm i nieindywidualizm eksplanacyjny

W rzeczy samej, myśl o tym, że nasze potoczne wyjaśnienia działań odwołują się do pragnień innych ludzi, ich życzeń, rozkazów czy prośb, powinna nam się wydać naturalna. Jesteśmy wszakże istotami społecznymi. Gdy współbiednik prosi nas o sól, po prostu ją podajemy — często nie podnosząc oczu znad książki; odpowiadamy wtedy na czyjeś życzenie. Gdy członek dobrze wyszkolonej załogi ratowniczej w gnieniu oka skacze do wody, gdy tylko padnie komenda — odpowiada na rozkaz. Wielu z nas doświadczyło mocy pragnień innych ludzi (szczególnie tych bardziej charyzmatycznych czy władczych), gdy podążaliśmy za nimi — często wbrew samym sobie.<sup>11</sup>

Wszystkie te zjawiska są całkowicie naturalne. To, co w nich interesujące z punktu widzenia filozofii umysłu, to fakt, iż można je interpretować w dwojaki sposób.

Z jednej strony można próbować twierdzić, iż właściwa struktura tych wyjaśnień jest zafałszowana przez przekaz słowny, a dokładniej, że jest przezeń skrócona. Wyjaśnienia te są entymematyczne: nie powołują się wprawdzie *explicite* na pragnienia sprawcy, tym niemniej ich istnienie (a także moc przyczynową) zakładają. Wszakże gdyby zacytowany biesiadnik był poproszony o sól, ale jej *nie chciał* podać, to zapewne *nie* podałyby jej. Skoro podał, to znaczy, że chciał podać — nawet jeśli sobie z tego

<sup>10</sup> Przeprowadzenie badań tego typu byłoby ciekawe. W szczególności interesującą byłaby zależność dystrybucji różnego rodzaju wyjaśnień od takich czynników, jak odpowiedź dziecka na pierwszy obrazek, doświadczenia z rodzicami, osobowość (niezależna albo uległa), to co znajduje się pod łóżkiem (prawdziwe zwierzątko czy zabawka), wyraz twarzy kobiety (przyjazny czy niechętny), wiek i płeć osoby drugiej, etc. Wyniki tych badań nie są jednak rozstrzygające dla celu, który stawiam sobie w tym tekście. Wystarczy mianowicie, że wszyscy uznajemy za zasadne przypuszczenie, iż dzieci takie wyjaśnienia mogłyby podać.

<sup>11</sup> Najbardziej dobitną ilustracją tego zjawiska są głośne eksperymenty Stanleya Milgrama [1962/2001], [1969], gdzie *dwie trzecie* badanych wykazało rzeczywistą skłonność zadawania horrendalnych wstrząsów elektrycznych (450V) niewinnej osobie pod dyktando eksperymentatora i — w dużej mierze — wbrew sobie.

nie zdawał sprawy (choćby dlatego, że był zacytany). Przytoczone rozumowanie jest zasadniczym argumentem na rzecz *indywidualizmu eksplanacyjnego*.

(eI<sub>1</sub>) Każde wyjaśnienie działania  $\alpha$ -y musi się odwoływać do pragnień (i przekonań)  $\alpha$ -y.

(eI<sub>2</sub>) Tylko pragnienia (i przekonania)  $\alpha$ -y mogą wyjaśnić działania  $\alpha$ -y.

Z drugiej strony dopuścić należy możliwość, iż przynajmniej niektóre z podawanych wyjaśnień odwołujących się do pragnień innych osób, nie są entymematyczne w tym sensie, że muszą być uzupełnione jakimś pragnieniem sprawcy. Na tym polega nieindywidualizm eksplanacyjny:

(eN<sub>1</sub>) Pewne działania  $\alpha$ -y są wyjaśnione przez pragnienia (i przekonania)  $\alpha$ -y.

(eN<sub>2</sub>) Pewne działania  $\alpha$ -y są wyjaśniane przez pragnienia (i przekonania)  $\beta$ -y.

### 2.3. Racjonalizacja a wyjaśnianie działań

Należy tu podkreślić, że stanowiska indywidualizmu i nieindywidualizmu eksplanacyjnego scharakteryzowane są w odniesieniu do pojęcia wyjaśniania działań a nie tylko tłumaczenia czy racjonalizacji działań przez racje (por. Davidson 1963). Rozróżnienie między wyjaśnianiem a racjonalizacją działań jest w zasadzie rozróżnieniem między rzeczywistym a jedynie możliwym powodem, dla którego podjęte może być — lub też jest — działanie. Załóżmy, że Agnieszka decyduje się na wykonanie ostatniej woli umierającego wujka i daje mu śmiertelny zastrzyk kończąc w ten sposób jego agonię. W grę wchodzi przynajmniej dwie racje. Po pierwsze, kocha i szanuje swego wujka, a obiecała mu w ten właśnie sposób pomóc zakończyć jego życie. Po drugie, leży w jej interesie by wuj umarł, gdyż dziedziczy w ten sposób całą jego fortunę. Pomimo faktu, iż te dwie racje stanowią *możliwe powody* dla zadania śmiertelnego zastrzyku, to wcale nie znaczy, że obydwie muszą być *rzeczywistymi powodami* jej działania. Obie racje tłumaczą („racjonalizują”) działanie, ale niekoniecznie je wyjaśniają. Agnieszka może np. dać ostateczny zastrzyk tylko dlatego, że to swemu wujowi obiecała — bez względu na to, co ma po nim odziedziczyć. W takiej sytuacji tylko pierwsza racja *wyjaśnia*, dlaczego to zrobiła. Zwrotów takich jak ‘racja (powód) *r* racjonalizuje (tłumaczy) działanie *d*’ i ‘racja (powód) *r* wyjaśnia działanie *d*’ będą zawsze używać w powyższym sensie.

Rozróżnienie to jest ważne z punktu widzenia sporu między indywidualizmem a nieindywidualizmem już choćby dlatego, iż nad wyraz łatwo jest doszukać się takiego czy innego pragnienia w zasadzie dla każdego działania podejmowanego przez sprawcę. Należy tu mianowicie podkreślić, że łatwość z jaką przypisujemy innym różnorakie pragnienia nie stanowi *żadnego* uzasadnienia dla tez indywidualizmu eksplanacyjnego. Sam fakt bowiem, że łatwo przychodzi nam *wymyślanie* tego, jakie to pragnienia sprawca *mógłby* mieć (a więc jakie pragnienia *racjonalizują* jego działa-

nie) w żaden sposób nie prowadzi do ustalenia tych pragnień, które *rzeczywiście* się do tego działania *przyczyniły* (które to działanie *wyjaśniają*). Indywidualista musi dowieść, że każde działanie musi być wyjaśnione (nie tylko zrjonalizowane) przez jakieś pragnienie sprawcy. A tego — przez powołanie się na łatwość, z jaką wymyślamy pragnienia i przekonania — zrobić się nie da.

#### 2.4. Metodologiczny charakter indywidualizmu i nieindywidualizmu eksplanacyjnego

Tak rozumiany nieindywidualizm (a także indywidualizm) dotyczy ograniczeń, jakie należy — *resp.* nie należy — nakładać zarówno na psychologiczne, jak i na filozoficzne teorie umysłu i działania. Nieindywidualizm ma więc wyłącznie wymiar metodologiczny. Nie należy go mylić np. z teorią empiryczną mającą na celu wyjaśnianie i przewidywanie zachowań. W rzeczy samej, byłoby zgoła zabawne, gdyby tezy (eN) rościły sobie jakiekolwiek pretensje w tym względzie.

Należy w związku z tym określić, jaki jest stosunek indywidualizmu i nieindywidualizmu eksplanacyjnego do psychologii intencjonalnej. Otóż, psychologia intencjonalna jest pewną formą wyjaśnień działań ludzkich. Jako taka nie jest ani indywidualistyczna ani nieindywidualistyczna w tym sensie, iż zarówno indywidualista, jak i nieindywidualista może się zgodzić, że działania ludzi są wyjaśniane przez odwołanie się do pragnień i przekonań sprawców. Spór między indywidualizmem a nieindywidualizmem rozpoczyna się wtedy, gdy indywidualista utrzymuje przekonanie o hegemonii psychologii intencjonalnej jako wyłącznej — i jedynej możliwej — formie rozumienia działań ludzkich (na poziomie osobowym, rzecz jasna), nieindywidualista zaś wskazując na prace badaczy zachowań społecznych — takich jak, np. Goffman [1986], Milgram [1962/2001], [1969], Nowak [1985], [1991], Haney, Banks, Zimbardo [1973], [1973/2001] — twierdzi, iż nie ma potrzeby reinterpretacji wyników ich badań w języku psychologii intencjonalnej.<sup>12</sup>

#### 2.5. Redukcja indywidualistyczna

Należałoby też zwrócić uwagę na pewien — pozorny — atut indywidualizmu, a mianowicie na fakt, iż indywidualista skonfrontowany z jakimkolwiek działaniem realizującym cudze pragnienia, może zawsze takie działanie zreinterpretować jako działanie realizujące własne pragnienie sprawcy, które to pragnienie jest odpowiednio skierowane na pragnienie innej osoby.

Rozważmy, jako przykład, pracownika pewnej firmy, który myje samochód szejowej na jej życzenie. Działają powodowany cudzym życzeniem, a więc i pragnieniem.

<sup>12</sup> Argumentację na rzecz tej tezy podejmuję na przykładzie studium badań Milgrama w oddzielnym artykule („O rzekomej niespójności teorii Milgrama”, w przygotowaniu).

Indywidualista nie musi wcale takiego przypadku uznać za kontrprzykład. Przeciwnie, dokona on reinterpretacji tego działania. Zapyta mianowicie: a dlaczego to niby spełnia to życzenie? Widać obawia się, że mógłby stracić pracę — jeśli tak to podejmuje mniej lub bardziej świadomie pewien wybór na podstawie swoich preferencji. Tak się składa, że w tym wypadku wygrywa pragnienie spełnienia życzenia szefowej. W innych wypadkach wygrać może pragnienie bycia wobec niej usłużnym. I tak dalej. Sam fakt, że ktoś realizuje czyjeś pragnienie nie świadczy jeszcze o tym, że indywidualizm jest fałszywy — utrzymuje.

Nie można się z tym nie zgodzić. W rzeczy samej, byłoby wysoce podejrzane, gdyby nieindywidualista z kolei miał twierdzić, że *nigdy* nie jest tak, iż spełniamy czyjeś pragnienie dlatego, że pragniemy je spełnić. Zwróćmy jednak ponownie uwagę na modalność zobowiązań teoretycznych. Indywidualista eksplanacyjny zobowiązany jest twierdzić, iż tak jest zawsze i że *nie może* być inaczej. Jest on zobowiązany przyjąć strategię redukcyjną wobec wszelkich wyjaśnień nieintencjonalnych. Nieindywidualista z kolei może przyjąć, iż *mogą* występować takie przypadki, kiedy taka redukcja jest właściwa, ale *mogą* też mieć miejsce przypadki, kiedy tak nie jest. Co innego, jeśli pracownik podejmuje pewien wybór: po rozważeniu możliwych opcji w końcu decyduje się, mniej lub bardziej świadomie, spełnić życzenie szefowej — a co innego, jeśli po mniej lub bardziej świadomym rozważeniu możliwych opcji decyduje się iść za głosem własnego sumienia i nie posłuchać jej życzenia, po czym — wbrew swej decyzji — jednak go słucha, jak to uczyniło np. wielu uczestników eksperymentów Milgrama. A jeszcze co innego, jeśli biesiadnik niemalże odruchowo odpowiada na prośbę o podanie soli odpowiednim ruchem ręki — bez wejrzenia w swe pragnienia, bez rozważenia pragnień współbiesiadników, bez zastanawiania się nad normami rządzącymi zachowaniami przy stole — po prostu sięga i podaje, tak jak został tego przez lata wyuczony.

Niech więc nie bałamuci nas myśl, że indywidualista może się podjąć takiej strategii redukcyjnej, podczas gdy nieindywidualista nie jest w stanie tego uczynić, gdyż jest ona jawnym fałszem. Załóżmy, że Zosia oddaje Jasiowi zabawkę, bo pragnie spełnić prośbę Jasia, który ją właśnie o to poprosił. Nieindywidualista również może dociekać głębszych powodów, dla których Zosia oddaje Jasiowi zabawkę realizując swe pragnienie, by spełnić prośbę Jasia. Dlaczego tak czyni? Może dlatego, że jej mama chce, by stała się mniej zaborcza i uczy ją zwracać zabawki ich właścicielom, kiedy o to poproszą. Nie jest więc jasne, dlaczego niby miałyby być tak, że ostateczną podstawą, na której opierają się wszelkie inne wyjaśnienia, muszą być wyjaśnienia intencjonalne. W szczególności nie jest jasne, dlaczego nie można byłoby przynajmniej części tychże wyjaśnień rozumieć w kategoriach (z grubsza) Wittgensteinowskich, a mianowicie w odniesieniu do szeroko rozumianych praktyk społecznych. Dlaczego Jakub poproszony o sól, podaje ją? — Dlatego, że nauczył się manier. Dlaczego Bogumiła wykonuje rozkazy kapitana załogi ratunkowej? — Tego właśnie została wyuczona. Dlaczego Zosia robi to, co każe jej mama? — Jest to część

tego, co znaczy być dzieckiem: jako dzieci, jesteśmy wsłuchani w to, co mówią, radzą czy nakazują nam robić nasi rodzice — młodzieńczy bunt zakłada posłuch.

## 2.6. Indywidualizm i nieindywidualizm a ewolucja

Jeśli czytelnik jest już z góry sceptycznie nastawiony do słabych ostatecznie — pod względem siły logicznej — tez nieindywidualizmu, warto na zakończenie rzucić je na tło naszej odległej historii. Szeroko rozpowszechniona jest hipoteza, iż ludzka ewolucja faworyzowała i selekcionowała zachowania racjonalne (por. np. Dennett [1971], [1981b]). Ponieważ działanie racjonalne (przynajmniej częściowo) polega na spełnianiu własnych pragnień (w szerokim sensie), daje to podstawy do tego, by przypuszczać, iż ludzie w normalnych warunkach będą mieli tendencję do spełniania właśnie *swoich własnych* pragnień. A to przecież kluczowa teza psychologii intencjonalnej.

Warto jednakże zaznaczyć, że są przesłanki po temu, by przypuszczać, iż nasze dziedzictwo biologiczne obejmuje również i szersze gamy zachowań. Znani teoretycy ewolucji kultury Boyd i Richerson [1985] argumentują, że podlegaliśmy ewolucyjnej presji skłaniającej ku konformizmowi. Zakładają, iż wzorce zachowań charakterystyczne dla danej społeczności są przystosowane do określonego typu sytuacji występujących w otoczeniu tejże społeczności. Dla jednostki, która do takiej grupy się dołącza, jest zatem korzystne (w sensie ewolucyjno-selekcyjnym), aby wykorzystać przetestowane wzorce zachowań (respektując w ten sposób mądrość przeszłości), zamiast ryzykować, że własne oryginalne wzorce ulegną selekcji negatywnej. Na tym właśnie polega selekcyjna korzyść płynąca z konformizmu — naszego działania zgodnie nie z naszym własnym zdaniem, ale właśnie ze zdaniem innych.<sup>13</sup>

Trzeba też zauważyć, iż te dwie części refleksji ewolucyjnych wcale się nie wykluczają. Przeciwnie, dopełniają się wzajemnie, ilustrując po prostu obecność sił prowadzących do rozwinięcia tendencji do bycia niezależnym — działania według własnych pragnień i przekonań — z jednej strony, a z drugiej do rozwinięcia tendencji zgoła przeciwnej — do bezpośredniego wczuwania się w pragnienia innych. O ile tym dwóm grupom presji ewolucyjnych rzeczywiście podlegaliśmy, o tyle można by oczekiwać, iż nasze życie będzie areną walki pomiędzy odpowiadającymi im tendencjami. I ta konstatacja zdaje się zawierać w sobie sporo prawdy — rozpoznawanej na intuicyjnym poziomie. Z tego punktu widzenia to właśnie nieindywidualizm zdaje się być bliższy prawdzie. To nieindywidualizm przecież dopuszcza myśl, iż czasem działamy wedle swych własnych pragnień, a czasem w swych działaniach podlegamy pragnieniom innych. I podobnie jak „indywidualistyczny” człon naszego ewolucyjnego-

<sup>13</sup> Należy zauważyć, iż nie jest to równoznaczne z twierdzeniem, że istnieją ewolucyjne podstawy do tego, byśmy działali *wbrew* naszemu zdaniu (jeśli takie pozory zostały stworzone to są one z pewnością podyktowane przez pewną dwuznaczność, o której będzie mowa dalej).

go dziedzictwa mógłby być uzasadnieniem dla indywidualizmu, tak jego całość może być uzasadnieniem właśnie dla nieindywidualizmu.

W rzeczy samej, ta dwuczłonowość narzuca raczej sceptyczną postawę co do strategii redukcyjnej wobec wyjaśnień odwołujących się do przekonań innych osób przyjmowanej przez indywidualistów. Strategia ta byłaby zrozumiała, gdyby te dwie tendencje ewolucyjne były ze sobą niezgodne. Wtedy rzeczywiście można by przypuszczać, iż należy być np. indywidualistą, preferując „człon indywidualistyczny”, konceptualizować wszystkie zachowania jako realizacje pragnień i przekonań jednostki, albo też być antyindywidualistą i — na odwrót — konceptualizować wszystkie zachowania jako realizacje pragnień i przekonań osób, pod których wpływem jednostka się znajduje lub znajdowała. Ale takiej niezgodności pomiędzy tymi tendencjami nie ma. Jeśli zachodzi jakakolwiek konkurencja między nimi, to dotyczy ona jedynie tego, która z tendencji przeważa w określonych sytuacjach. A jeśli tak, to nie ma powodów przypuszczać, że strategia redukcyjna — czy to w formie indywidualistycznej, czy to w formie antyindywidualistycznej — jest bardziej obiecująca. Przeciwnie, mamy wszelkie podstawy sądzić, iż oba człony naszego ewolucyjnego dziedzictwa jakoś się zaznaczyły w tym, jak zostaliśmy „zaprogramowani”. A to stanowi argument dla dyrektywy nieindywidualizmu, w obrębie którego znajduje się miejsce zarówno na działania powodowane własnymi, jak i cudzymi pragnieniami i przekonaniem.

## **2.7. Indywidualizm a idealizacyjna rekonstrukcja zasady racjonalności**

Na zakończenie tych wstępnych omówień trzeba jeszcze zwrócić uwagę na możliwość pewnego nieporozumienia. Można by mianowicie przypuszczać, że jeśli właściwie spojrzeć na zasadę racjonalności — a mianowicie jeśli rekonstruować ją jako swoiste twierdzenie idealizacyjne — wówczas można uznać spór między indywidualizmem a nieindywidualizmem za chybiony: zasada racjonalności jest przyjęta w oparciu o pewne założenia idealizujące, w szczególności o założenia, których efektem jest tymczasowe zawieszenie wpływu pragnień innych osób; po uchyleniu tych założeń zasada racjonalności musiałaby zostać skonkretyzowana w taki sposób, aby pozwalała uwzględniać wpływy pragnień innych osób na jednostkę.

Takie podejście do problemu jest bliskie ujęciu W. Patryasa [1979], który proponując idealizacyjną rekonstrukcję zasady racjonalności, opiera ją na dziesięciu założeniach idealizujących, w tym także takich, które zawężają wiedzę, z której podmiot korzysta, do wiedzy własnej podmiotu, oraz takich, które zawężają pragnienia, które podmiotem kierują, do jego własnych pragnień. Ujęcie Patryasa jest — według powyższych ustaleń — nieindywidualistyczne<sup>14</sup>, ponieważ dopuszcza możliwość tego,

<sup>14</sup> Rozróżnienie między indywidualizmem i nieindywidualizmem eksplanacyjnym zostało wprowadzone bez odwołania się do struktur idealizacyjnych. Być może warto byłoby wprowadzić jeszcze rozróżnienie między indywidualizmem i nieindywidualizmem esencjalnym. Indywidualizm esen-

iż działając realizujemy cudze pragnienia niekoniecznie tym samym realizując własne pragnienia. Charakterystyczne dla indywidualizmu eksplanacyjnego jest to, iż takiej możliwości nie dopuszcza, a nalega na to, by wszystkie działania realizujące cudze pragnienia interpretować jako działania realizujące pragnienia sprawcy.

### 3. PRÓBY UZASADNIENIA INDYWIDUALIZMU EKSPLANACYJNEGO

Tak skrojone rozróżnienie między indywidualizmem a nieindywidualizmem może się wydawać dość dziwne już choćby z tego względu, iż tezy nieindywidualizmu są — pod względem siły logicznej — niezwykle słabe. Wskazuje na to podwójne użycie kwantyfikatorów szczegółowych. Zdawać by się mogło, iż właśnie logicznie słabsza teza ma większe szanse znaleźć zwolenników. W przypadku sporu między indywidualizmem a nieindywidualizmem eksplanacyjnym jest jednak inaczej. Oczywiście zdarza się (i to nie tak rzadko), że istnieją dobre racje uzasadniające tezy logicznie silniejsze. W rozważanym wypadku tak jednak nie jest. Mało kto poczuwa się do konieczności uzasadnienia indywidualizmu — a te uzasadnienia, które można wyszukać w rozmaitych fragmentach prac filozofów umysłu, są — jak zobaczymy — niewystarczające. Poniżej zrekonstruujemy zasadniczy argument na rzecz indywidualizmu eksplanacyjnego, który — jak się okaże — wcale nie stanowi uzasadnienia dla tegoż stanowiska. Dalej zobaczymy dlaczego nie należy szukać uzasadnienia dla indywidualizmu eksplanacyjnego w jego pozornym pokrewieństwie z teorią decyzji. Następnie rozważymy argumentację za indywidualizmem eksplanacyjnym z punktu widzenia interpretacyjnych teorii stanów psychicznych.

#### 3.1. Argument zasadniczy

Warto przypomnieć, że sceptycyzm wobec nieindywidualizmu wiąże się z faktem, że jest nam nader łatwo znaleźć taką konstelację pragnień i przekonań, które mogłyby wytłumaczyć, dlaczego sprawca działa tak-a-tak. Co więcej, jest to możliwe zawsze, nawet w sytuacjach, gdy uznajemy skądinąd, iż sprawca realizuje pragnienie innej osoby. Nawet jeśli dziecku nie chce się iść do sklepu (co nakazuje mu mama), to może chcieć — z takiego czy innego powodu — usłuchać jej nakazu. I to pragnienie — choćby słabiutkie — zdaje się wystarczać dla utrzymania tez indywidualizmu eksplanacyjnego.

Rzecz w tym, że wcale nie wystarcza. Sam przez się fakt, że takie pragnienie *możemy* dziecku przypisać, nie daje nam jeszcze żadnych podstaw do przypuszczenia, iż

---

esencjalny polega na tym, że utrzymuje się, iż pierwsze twierdzenie idealizacyjne modelujące działania jednostki *musi* odwoływać się do pragnień tejże jednostki. Nieindywidualizm esencjalny natomiast dopuszcza możliwość tego, aby i w pierwszym twierdzeniu idealizacyjnym uwzględniane były cudze preferencje. Nie jest jasne, czy Patryas jest zwolennikiem tak rozumianego indywidualizmu esencjalnego (nie uzasadnia on nigdzie, iż jego rekonstrukcja jest konieczna).

to pragnienie *rzeczywiście* przyczynia się do danego działania. Niezwykła — a raczej nadmierna — łatwość *racjonalizacji* działań w terminach pragnień i przekonań nie wystarcza, by uzasadnić tezę, że pragnienia i przekonania sprawcy *rzeczywiście wyjaśniają* wszystkie jego działania. Istnieje jednakże całkiem ogólna argumentacja mająca na celu uzasadnienie właśnie tej mocniejszej tezy — koniecznej dla uzasadnienia indywidualizmu eksplanacyjnego. Twierdzi się mianowicie, że wyjaśnienia nieindywidualistyczne w ogóle nie są w stanie działań jednostki wyjaśnić — aby działanie takie wyjaśnić, konieczne jest odwołanie się do jej pragnień. Argumentację tę określam mianem zasadniczego argumentu za indywidualizmem eksplanacyjnym.

Rozważmy przykład działania, które pozornie zdaje się być wyjaśnione przez odwołanie do cudzego pragnienia. Załóżmy, że pewien kierowca zatrzymuje się by zapytać nas o drogę do Sierakowa. Wyjaśniamy mu, jak jechać i idziemy dalej. Dlaczego powiedzieliśmy mu, jak dojechać do Sierakowa? Bo o to prosił — bo pragnął (w szerokim sensie tego słowa) dowiedzieć się, jak tam dojechać. Indywidualista twierdzi, że w tej sytuacji nie tylko możliwe jest przypisanie nam samym jakiegoś pragnienia, ale że musimy to uczynić, gdyż inaczej nasze działanie pozostanie *niewyjaśnione*. Aby się o tym przekonać, rozważmy, co by się stało, gdybyśmy *nie chcieli* zrobić tego, o co kierowca prosi. Rzecz jasna, gdybyśmy *rzeczywiście* nie chcieli pomóc temu kierowcy, to byśmy tego *ceteris paribus* nie zrobili. Ponieważ pomogliśmy kierowcy, to znaczy, że *musieliśmy* tego chcieć.<sup>15</sup> Słowem, indywidualista stara się wykazać w ten sposób, iż wyjaśnienia nieintencjonalne są *zawsze* wyjaśnieniami entymematycznymi.

Rozumowanie to wygląda na prawidłowe, a i jego przesłanki mogą się wydawać nie do podważenia. Zrekonstruujmy jego ogólną formę:

- (1)  $\alpha$  wykonała działanie typu  $A$  ( $a$   $d$  jest działaniem typu  $A$ ), ponieważ  $\beta$  chciał, by  $\alpha$  wykonała działanie  $d$ .
- (2) Gdyby — dla pewnego  $X$  —  $\alpha$  nie chciała wykonać działania typu  $X$  (gdzie  $d$  jest działaniem typu  $X$ ), to *ceteris paribus*  $\alpha$  nie wykonałaby działania  $d$ .
- (3)  $\alpha$  wykonała działanie  $d$ .

Stąd, dla pewnego  $X$ ,  $\alpha$  musiała chcieć wykonać działanie typu  $X$  (gdzie  $d$  jest działaniem typu  $X$ ).

(1) nie wymaga w kontekście tej argumentacji uzasadnienia. Jest to przesłanka, którą indywidualista przyjmuje za nieindywidualistą w argumentacji przeciw nieindywidualizmowi. (3) jest przesłanką empiryczną. Jedyna przesłanka, która się w ogóle na-

<sup>15</sup> Zauważyć tu należy, iż sama treść naszego pragnienia nie jest tu przesadzona. Być może chcieliśmy pomóc kierowcy, a być może chcieliśmy spełnić to, co uznajemy za nasz obowiązek, a jest także możliwe, że wcale nie chcieliśmy pomóc kierowcy jako takiemu tylko chcieliśmy nie być uznani za nieuczynnych. Możliwości jest tu wiele. Indywidualista nie twierdzi, że chodzi o jakiś konkretne pragnienie.

daje do zakwestionowania, to przesłanka (2). Ale i ta może się wydawać zbyt oczywista by ją podważać. W końcu — można argumentować — gdybyśmy rzeczywiście nie chcieli objaśnić kierowcy, jak dojechać do Sierakowa — na przykład dlatego, że kierowca wydał się nam dość podejrzanym typem, od którego czym prędzej chcieliśmy odejść, albo choćby dlatego, że byliśmy akurat wtedy źli na całą ludzkość i nie chcieliśmy pomagać nikomu — to byśmy tego po prostu nie zrobili. Spróbuję wykazać, że uznajemy prawdziwość (2) zbyt pochopnie. Dokładniejszy wgląd w strukturę logiczną tego zdania wykaże, iż przesłanka ta jest w rzeczywistości dwuznaczna, a całe rozumowanie jest obarczone błędem ekwiwokacji.

Zanim jednak do tych rozważań przejdziemy, należy zwrócić uwagę na pewne niejasności, które wiążą się z użyciem klauzuli *ceteris paribus*. Jest to klauzula, której używa się niekiedy nawet w sformułowaniach praw naukowych (por. np. Patryas [1975], [1976]), a jest ona zgoła konieczna w sformułowaniach prawdziwości wiążących postawy intencjonalne jednostek i ich działania. Jeśli pominąć tę klauzulę, tezę (2) można sfalsyfikować nadzwyczaj łatwo. Gdyby, na przykład, spełniony był poprzednik (2), ale  $\alpha$  również zmieniła zdanie, albo gdyby oprócz tego, że  $\alpha$  nie chce objaśnić  $\beta$ -cie jak dojechać do Sierakowa,  $\alpha$  również chciała popisać się swą wiedzą przed innymi przechodniami, to wcale nie jest jasne, czy następnik (2) byłby prawdziwy. Takie przypadki można mnożyć w nieskończoność, stąd zasadność zastosowania klauzuli *ceteris paribus*. Kwestionowanie (2) ze względu na jawne niejasności, jakie klauzula ta ze sobą niesie, byłoby jednak bezcelowe i niezyciliwe (z punktu widzenia nieindywidualizmu). Przyjmijmy tu zatem, iż wszelkie warunki, które mogłyby być ujęte w tej klauzuli, sprzyjają indywidualistom.

Zasadniczy problem polega na tym, że przesłanka (2) jest dwuznaczna. Zdanie ' $\alpha$  nie chce sprawić tego, że  $p$ '<sup>16</sup> może być użyte dla oznaczenia dwóch różnych stanów rzeczy:

- (a)  $\alpha$  chce nie sprawić tego, że  $p$  (kontrpragnienie)
- (b) nie jest tak, że  $\alpha$  chce sprawić to, że  $p$  (brak pragnienia).

---

<sup>16</sup> Zakładam tu, że każde zdanie orzekające o tym, że  $\alpha$  podejmuje jakieś działanie  $d$  można w języku polskim sparafrazować jako ' $\alpha$  sprawia to, że  $p$ ' dla pewnego  $p$ , gdzie ' $p$ ' jest opisem pewnego stanu rzeczy. Wykorzystuję tu ustalenia logiki sprawstwa zaproponowanej przez N. Belnapa i M. Perloffą [1990], [2001]. Autorzy utrzymują, iż każde zdanie sprawcze (tj. orzekające o tym, iż podmiot podjął pewne działanie) można sparafrazować przy użyciu formuły [ $\alpha$  stit:  $p$ ], gdzie 'stit' to skrót angielskiego 'sees to it that', którego najbliższym odpowiednikiem wydaje mi się zwrot użyty powyżej. Milcząco założona jest tu więc skutkowa koncepcja działania opisana przez A. Klawitera [1978, s. 58—61]. Autor ten rozróżnia skutkowe, racjonalistyczne oraz behawioralne koncepcje działania, kwalifikujące działania w zależności od ich skutku, celu lub pewnych cech czynności na te działania się składających. Można jednakże bronić tezy, iż skutkowe koncepcje działania są w pewnym sensie podstawowe — pozostałe można bowiem rozumieć jako ich szczególne przypadki. Tak mianowicie, że skutkiem jest np. osiągnięcie pewnego celu lub wykonanie ruchu o takich-a-takich cechach.

W przypadku pierwszym,  $\alpha$  żywi pewne *pragnienie*, a mianowicie pragnienie niesprawienia tego, że  $p$ ; można też powiedzieć, że  $\alpha$  posiada kontrpragnienie sprawienia tego, że  $p$ . W przypadku drugim natomiast, orzekamy wcale nie o tym, że  $\alpha$  jakiegokolwiek pragnienie czy kontrpragnienie posiada, ale *zaprzeczamy* po prostu temu, jakoby posiadała *pragnienie* sprawienia tego, że  $p$ . Oczywiście — o ile pragnienia  $\alpha$ -y nie są sprzeczne — (b) wynika z (a). Nie jest jednak tak, że (a) wynika z (b). W przypadku (b) — ale nie w przypadku (a) —  $\alpha$  może być zupełnie obojętna wobec sprawienia (a także niesprawienia) tego, że  $p$ .

Fakt, iż obydwie te sposoby odczytania wyjściowego zdania ' $\alpha$  nie chce sprawić, że  $p$ ' są możliwe, nie podlega chyba kwestii. Ponadto, jest to dwuznaczność rozpozszeczona w stosunku do większości postaw intencjonalnych. Wyrażenie ' $\alpha$  nie sądzi, że  $p$ ' może znaczyć albo, że (a')  $\alpha$  sądzi, że nie- $p$  albo, że (b') nie jest tak, że  $\alpha$  sądzi, że  $p$ . Podobnie wyrażenie ' $\alpha$  nie zamierza sprawić tego, że  $p$ ' może znaczyć albo, że (a'')  $\alpha$  zamierza nie sprawić tego, że  $p$ , albo, że (b'') nie jest tak, że  $\alpha$  zamierza sprawić to, że  $p$ . Warto jednocześnie zauważyć, iż w mowie potocznej notorycznie te dwa znaczenia ze sobą mieszamy. Formy słowne, których gramatyka powierzchniowa sugeruje drugie znaczenie, stosujemy do wyrażenia pierwszego. I tak orzeczenie przed sądem „Nie mam zamiaru się podporządkowywać wyrokowi Sądu!” niedwuznacznie odczytujemy w sensie (a) jako oznajmiające posiadanie *zamiaru* — i to silnego — właśnie *niepodporządkowania* się wyrokom Sądu, a nie w sensie (b) sugerowanym przez gramatykę powierzchniową. Podobnie oznajmienie „Nie chcę się z nią bawić!” wcale nie oznacza braku pragnienia, ale wręcz przeciwnie, ogłasza posiadanie pewnego kontrpragnienia.

Jeśli tak, to przesłankę (2) można odczytać na dwa różne sposoby:

- (2a) Gdyby  $\alpha$  chciała nie sprawić tego, że  $p$  to *ceteris paribus*  $\alpha$  nie sprawiłaby tego, że  $p$
- (2b) Gdyby nie było tak, że  $\alpha$  chce sprawić to, że  $p$  to *ceteris paribus*  $\alpha$  nie sprawiłaby tego, że  $p$ .

Aby wykazać, że pragnienie  $\alpha$ -y jest *konieczne* do wyjaśnienia jej działania, indywidualista musi wykazać prawdziwość (2b) — nie wystarczy prawdziwość (2a). Tymczasem powody, które skłaniały nas do utrzymania dwuznacznej przesłanki (2) są powodami do utrzymania tylko (2a). Odwołajmy się najpierw do powyższych przykładów. Jeśli zatrzymujący się kierowca wydaje się nam podejrzany i chcemy się od jego obecności czym prędzej uwolnić, to wcale nie jest tak, iż brak nam po prostu pragnienia by wskazać mu drogę — przeciwnie, jest tak, iż *pragniemy* mu drogi *nie* wskazywać. Podobnie jest wtedy, gdy jesteśmy „źli na cały świat”. Nie jest wówczas tak, że brakuje nam pragnienia, aby komuś udzielić pomocy, *lecz posiadamy* właśnie *pragnienie* by nikomu (w tym i owemu kierowcy) *nie* pomagać (choćby przez wskazanie drogi).

W istocie rzeczy krótka refleksja nad naturą psychologii intencjonalnej wystarczy, by zrozumieć, iż nie jest ona w stanie zapewnić żadnego dodatkowego uzasadnienia dla tezy (2b). Dzieje się tak po prostu dlatego, że prawidłowości psychologii intencjonalnej dotyczą tylko tego, co się dzieje jeśli osoby *posiadają* pewne pragnienia i przekonania — nie mówi nic na temat tego, co się dzieje, jeśli takich pragnień i przekonań nie posiadają. Jeśli jednak przypomnimy sobie, że to (2b) jest konieczne, by wykazać niezbędność wyjaśnień intencjonalnych, to w ogóle upada całość argumentacji indywidualisty. Nieindywidualista mógłby się nawet zgodzić z twierdzeniem (2a)<sup>17</sup> i twierdzić tylko, że skoro sprawca jest sam obojętny wobec sprawienia czy niesprawienia, że *p*, to cudze pragnienie, ażeby sprawił on, że *p*, może przyczynić się do tego, że tak właśnie uczyni. Jest to przypadek zgodny z (2a), ale zakazany przez dyrektywę indywidualizmu eksplanacyjnego. Zmuszeni jesteśmy zatem uznać, iż tzw. argument zasadniczy za indywidualizmem eksplanacyjnym nie stanowi żadnego uzasadnienia dla tego stanowiska.

### 3.2. Teoria decyzji a indywidualizm eksplanacyjny

Nieco zniecierpliwiony czytelnik mógłby rozumować tak: „O co tu właściwie chodzi? Z jednej strony mamy indywidualizm eksplanacyjny — może nie dość ugruntowany filozoficznie, ale przecież zakładany przez solidne koncepcje choćby w postaci teorii decyzji, nie wspominając o poparciu empirycznym. A cóż może zaferować nieindywidualizm? Jeśli niewiele, to wybór jest jasny.”

Zarzut ten byłby trafny, gdyby nieindywidualizm uzurpował sobie prawo do bycia traktowanym na równi z formalną teorią decyzji i działań ludzkich. Tak jednak nie jest. Po pierwsze, spór, który jest przedmiotem tych rozważań, jest sporem pomiędzy stanowiskami metodologicznymi określającymi granice możliwych koncepcji filozoficznych i psychologicznych. Nieindywidualizm (tak jak i indywidualizm) stanowi pewną dyrektywę metodologiczną, a nie teorię — a tym bardziej nie teorię formalną. Dalej — w przeciwieństwie do indywidualizmu — na nieindywidualizm składa się właśnie dyrektywa bardziej permissywna, nie narzucająca kategorii teoretycznych. Nieindywidualista nie twierdzi w szczególności, że działania nasze nie mogą być dyktowane naszymi preferencjami ani też, że działania nasze muszą być — choćby czasem — dyktowane pragnieniami innych. Wszelkie konkretne teorie, które odwołują się do działań powodowanych przez pragnienia sprawcy można uznać za uwiarygodniające nieindywidualizm, jeśli nie opatrzone są twierdzeniem, że tylko pragnienia sprawcy (a nie innych osób) mogą przyczynić się do tego, że jednostki podejmują działania.

<sup>17</sup> Nie utrzymuję, że jest to stanowisko, do którego nieindywidualista jest jakoś przymuszony. Dużo bardziej trafnym wydaje się również zakwestionować (2a) wskazując choćby na eksperymenty Milgrama. Powyższy wywód pokazuje tylko niekonkluzywność argumentacji indywidualistycznej.

Po drugie, należy zauważyć, iż w teorii decyzji nie istnieje nic, co przesądzałoby, że nasze działania nie mogą być powodowane przez cudze pragnienia. Nieindywidualista mógłby zatem mieć nadzieję, że być może kiedyś i nieindywidualistyczny członek naszej natury doczeka się formalizacji.

Nie jest jednakże wykluczone, że tak się nie stanie — z zupełnie prostego powodu. Może się mianowicie okazać, że wyjaśnienia działań odwołujące się do cudzych pragnień są — z formalnego punktu widzenia — zupełnie nieciekawe. Każde takie działanie można przecież — za pomocą manewru formalnego określonego wyżej mianem redukcji indywidualistycznej (por. § 2.5) — przedstawić jako działanie powodowane pewnym pragnieniem jednostki skierowanym na cudze pragnienie. Jeśli więc nie jest tak, że istnieją jakiegokolwiek *formalne* różnice pomiędzy działaniem powodowanym przez cudze pragnienie (pragnienie mamy, aby synek przyniósł mąkę), a działaniem powodowanym przez pragnienie jednostki skierowane na cudze pragnienie (pragnienie synka, aby spełnić życzenie mamy, żeby przyniósł mąkę), to nie należy się spodziewać, że rozróżnienie, o którym tu mowa, zostanie jakkolwiek w formalnej teorii decyzji uwzględnione.

Po trzecie wreszcie, należy zauważyć, że teoria decyzji jest pewnym rachunkiem formalnym. Nic nie stoi więc na przeszkodzie temu, by sformułować „antyindywidualistyczną interpretację” teorii decyzji,<sup>18</sup> która uznawałaby, że zawsze spełniamy pragnienia innej jednostki; nazwijmy tę jednostkę  $\Omega$ .<sup>19</sup> Wyrażenie ‘ $u(x)$ ’, które przy interpretacji indywidualistycznej czytamy jako ‘subiektywna użyteczność dla podmiotu  $\alpha$  skutku  $x$  działania  $\alpha$ -y’, w antyindywidualistycznej interpretacji należałoby odczytać ‘subiektywna użyteczność dla  $\Omega$ -i skutku  $x$  działania  $\alpha$ -y’. (Możemy założyć, że wyrażenie ‘ $p(x)$ ’ jest odczytywane tak samo w indywidualistycznej, jak i w antyindywidualistycznej interpretacji.<sup>20</sup>) Podobnie ‘ $x \succeq y$ ’, które w indywidualistycznej interpretacji czytamy jako ‘ $\alpha$  słabo preferuje skutek  $x$  wobec skutku  $y$ ’, w antyindywidualistycznej interpretacji czytamy jako ‘ $\Omega$  słabo preferuje skutek  $x$  działania  $\alpha$ -y wobec skutku  $y$  działania  $\alpha$ -y’. Z założenia różnice między stanowiskiem indywidualistycznym a antyindywidualistycznym zachodzą tylko w interpretacji rachunku formalnego. Aksjomaty teorii decyzji natomiast zdają się być równie intuicyjne zarówno przy interpretacji indywidualistycznej, jak i antyindywidualistycznej. Na przykład aksjomat o przechodniości preferencji brzmiałby mniej więcej tak: jeśli  $\Omega$  słabo preferuje skutek  $x$  działania  $\alpha$ -y wobec skutku  $y$  działania  $\alpha$ -y oraz  $\Omega$  słabo preferuje skutek  $y$  działania  $\alpha$ -y wobec skutku  $z$  działania  $\alpha$ -y, to  $\Omega$  słabo preferuje skutek  $x$  działania  $\alpha$ -y wobec skutku  $z$  działania  $\alpha$ -y.

<sup>18</sup> Interpretacja ta jest *antyindywidualistyczna* (nie: nieindywidualistyczna) ponieważ zakłada, iż zawsze (nie: niekiedy) kierujemy się pragnieniami innej jednostki.

<sup>19</sup> Należy tu zwrócić uwagę, że ‘ $\Omega$ ’ odnosi się do jednej konkretnej jednostki, tej samej dla wszystkich podmiotów.

<sup>20</sup> Należy pamiętać, że spór pomiędzy nieindywidualizmem a indywidualizmem dotyczy tego, czy cudze pragnienia (nie przekonania) mogą wpływać na jednostkę.

Z formalnego punktu widzenia, antyindywidualistyczna interpretacja teorii decyzji jest równoważna interpretacji indywidualistycznej, a jest tak po prostu dlatego, że owego odniesienia do  $\Omega$  w żaden sposób nie wykorzystuje się w rachunku formalnym. Jednocześnie przedwczesna wydaje się obrona interpretacji indywidualistycznej czyniona w duchu jej rzekomo wyższej intuicyjności. Po pierwsze, nie można twierdzić, że antyindywidualistyczna interpretacja zupełnie nie uwzględnia preferencji  $\alpha$ -y. Oczywiście, że je uwzględnia — dzięki manewrowi redukcji antyindywidualistycznej, tj. przy założeniu, że preferencje  $\Omega$ -i obejmują spełnienie preferencji  $\alpha$ -y. (Proszę zauważyć, że sam manewr redukcji antyindywidualistycznej — jak i rzecz jasna jego odpowiednik, a mianowicie manewr redukcji indywidualistycznej — stanowią już element interpretacji teorii decyzji.) Po wtóre, to, co mogłoby wywołać opory, to fakt, iż wyżej zarysowana antyindywidualistyczna interpretacja zakłada, że preferencje jakiejś jednostki  $\Omega$  mogą mieć roszczenia do czynienia racjonalnymi działaniami podejmowanymi przez  $\alpha$ -ę. Istotnie, gdyby odniesieniem 'Ω' był Jan Kowalski, taka interpretacja mogłaby się wydać śmieszna, ale jeśli byłby nim Bóg, to śmieszna być przestaje — a bywały okresy w historii, kiedy interpretacja taka zapewne spotkałaby się z jednomyślną akceptacją.

Dalekie od moich intencji jest forsowanie teologicznej interpretacji teorii decyzji. Powyższe rozważania pokazują jedynie, że zbyt pochopnie dokonujemy jednoznacznie indywidualistycznej interpretacji ustaleń tejże teorii. Tak jak bezzasadne byłoby twierdzenie zwolenników antyindywidualistycznej interpretacji teorii decyzji, że nigdy nie kierujemy się własnymi preferencjami, tak bezzasadne jest twierdzenie zwolenników indywidualistycznej interpretacji teorii decyzji, że nigdy nie działamy (i nie możemy działać) powodowani wyłącznie cudzymi pragnieniami. Samą teorię decyzji należy wyraźnie odgraniczyć od jej interpretacji semantycznej, jak również właściwie rozumieć formalny sens redukcji — czy to indywidualistycznej, czy to antyindywidualistycznej. Nie należy (bez dalszych argumentów) rozumieć jej jako redukcji zjawisk; należy ją rozumieć jako formę ujęcia pewnych zjawisk, które — z *formalnego* punktu widzenia — niczym szczególnym się nie wyróżniają. Należy dalej zdać sobie sprawę z faktu, że nawet jeśli działania powodowane przez cudze pragnienia są z formalnego punktu widzenia nieciekawe, to nie stanowi to wcale o tym, czy są ciekawe z innych punktów widzenia — w tym i z punktu widzenia psychologa starającego się zrozumieć, co się faktycznie dzieje.

### 3.3. Interpretacyjne teorie stanów psychicznych

Należałoby teraz rozpatrzeć pewną niedorzeczność, o którą nieindywidualizm mógłby zostać posądzony. Przyjęło się mianowicie rozumieć postawy intencjonalne jako wewnętrzne stany sprawcy. Zwykle też rozumie się działania jako te zdarzenia, które są spowodowane m.in. przez postawy intencjonalne sprawcy. Skoro jednak tak, to jeśli nawet cudze postawy intencjonalne są w pewnych sytuacjach istotne dla działania

sprawcy, to przecież własne postawy intencjonalne sprawcy są konieczne — to one w końcu powodują zdarzenie, które jest działaniem! Wydawać by się mogło, że negacja tezy, iż postawy intencjonalne sprawcy przyczyniają się do jego działania, jest negacją tezy, iż stany wewnętrzne sprawcy się do tego przyczyniają — a to przecież absurd.

Argumentacja ta nie jest jednak przekonująca. Pomijam już fakt, iż nie ma jedno-myślności co do założenia, że postawy intencjonalne (które identyfikujemy na poziomie osobowym) należy utożsamiać z podklasą stanów wewnętrznych (które identyfikowane są na poziomie subosobowym), por. np. prace Baker [1995], Collinsa [1987], Dennetta [1981a], [1987] czy Hornsby [1980—1], [1997]. Zasadniczym problemem jest to, iż ta argumentacja za indywidualizmem skrycie indywidualizm zakłada i z tego powodu nie może być użyta jako argument przeciw nieindywidualizmowi.

Można mianowicie zadać pytanie, które to stany wewnętrzne należy utożsamiać z określonymi postawami intencjonalnymi. Jedna z bardziej rozpowszechnionych odpowiedzi na to pytanie głosi, iż przypisanie postaw intencjonalnych jest fragmentem holistycznego zrozumienia całości działań sprawcy (Davidson [1980], [1984]). Według tego poglądu dany stan wewnętrzny osoby  $\alpha$  można uznać za pragnienie  $\alpha$ -y, żeby  $p$ , tylko o tyle, o ile przypisując jej pragnienie, żeby  $p$ , maksymalizujemy zrozumienie całości jej działań. Taka atrybucja jest zatem zapośredniczona przez fakt, iż przyjmujemy pewne ogólne tezy dotyczące ludzkich działań, a w szczególności tezę, że ludzie działają zgodnie z własnymi pragnieniami i przekonaniami. Innymi słowy, utożsamienie naszych postaw intencjonalnych zakłada już pewne rozumienie psychologii potocznej.

To właśnie jest moment, w którym nieindywidualista musi przeciąć powyższy tok rozumowania. Winien zauważyć mianowicie, że przyjęcie nieindywidualizmu wpływa na naturę tego, jak identyfikujemy nasze postawy intencjonalne. Jeśli od początku uznajemy, że ludzie działają nie tylko zgodnie z własnymi, ale też zgodnie z cudzymi pragnieniami i przekonaniami, to możemy maksymalizować zrozumienie tego, dlaczego jednostka działa tak, a nie inaczej, poprzez przypisanie pewnego pragnienia wcale nie jej, lecz właśnie innej osobie. Załóżmy np., że pewna osoba  $\alpha$  zdaje się zachowywać zgodnie z preferencjami  $\beta$ -y (o których wiemy na podstawie uprzedniego doświadczenia), ale tylko w bliskiej obecności  $\beta$ -y. Nie przesądza to oczywiście ostatecznie o konieczności uznania, iż  $\alpha$  działa tak, a nie inaczej dlatego, że tego chce  $\beta$ , niemniej jednak stanowi to wstępne uzasadnienie dla takiej supozycji.

Skoro jednak tak, to nieindywidualista nie musi odrzucić wyżej wymienionych tez. Może aprobować powszechny pogląd, że postawy intencjonalne to stany wewnętrzne. Może utrzymywać, że działania są zdarzeniami, które są powodowane m.in. przez stany wewnętrzne sprawcy. I może również utrzymywać, że atrybucja postaw intencjonalnych jest częścią holistycznego rozumienia zachowania sprawcy.<sup>21</sup>

---

<sup>21</sup> Nie znaczy to, rzecz jasna, że nieindywidualista jest do tych poglądów niejako zmuszony. Trzeba jednak podkreślić, że indywidualizm nie może usprawiedliwiać swej hegemonii wstępując w przymierze ze zwolennikami tych szeroko rozpowszechnionych poglądów.

Nieindywidualista odnotować musi jedynie, iż fakt, że działania są spowodowane przez stany wewnętrzne sprawcy, nie znaczy wcale, że działania muszą być spowodowane przez pragnienia i przekonania sprawcy. Jest tak dlatego, iż tylko te spośród stanów wewnętrznych sprawcy, co do których mamy (holistyczne) powody, aby rozumieć je jako pragnienia są identyfikowane jako pragnienia. Przez przyjęcie nieindywidualizmu otwieramy możliwość rozumienia działań sprawcy nie tylko na podstawie jego własnych, ale i cudzych pragnień.

#### 4. O DWÓCH POZIOMACH STOSOWANIA PSYCHOLOGII INTENCJONALNEJ

Podsumujmy dotychczasowe wywody. Z jednej strony, jak starałam się wykazać, nieindywidualizm eksplanacyjny jest stanowiskiem naturalnym — a to, co zdaje się go „wynaturzać” to argumenty filozoficzne, których ostrze — przy bliższej analizie — okazuje się dość tępe. Z drugiej strony jednakże okazało się, że argumentacja filozoficzna nie pozwala bynajmniej na podważenie przekonania, że indywidualizm jest równie naturalny. Stajemy w obliczu impasu.

Indywidualista godzi te dwa fakty wysuwając tezę redukcyjną, a mianowicie tezę, że każde wyjaśnienie odwołujące się tylko do cudzych pragnień jest niepełne w tym sensie, iż kryje się pod nim wyjaśnienie intencjonalne (to nie pragnienia matki powodują działaniami syna, lecz co najwyżej jego pragnienie, aby spełnić pragnienia matki). W tym paragrafie postaram się zaproponować odmienne pogodzenie tych faktów. Sądzę, mianowicie, iż psychologia potoczna operuje niejako na dwóch poziomach: normatywnym i eksplanacyjnym, przy czym nieindywidualizm jest trafny na poziomie eksplanacyjnym, indywidualizm natomiast — na poziomie normatywnym.

##### 4.1. Poziom eksplanacyjny i normatywny

Cele psychologii potocznej (na poziomie eksplanacyjnym) obejmują przedstawienie pewnego (jakkolwiek tymczasowego) sposobu wyjaśniania i przewidywania tego, co ludzie czynią. Poziom normatywny natomiast dotyczy nie tyle wyjaśniania naszych działań, ile ujmowania ich w pewien szczególny sposób. Celem jest takie ujęcie naszych działań, które utrwałaby — w nas i w innych — obraz nas samych jako jednostek silnych, niezależnych i autonomicznych. Na tak pojętym terenie normatywnym myśl indywidualistyczna znajduje swe naturalne miejsce, zarówno w tym sensie, że jest przez nas — przynajmniej w kulturze zachodniej — w sposób naturalny traktowana, jak i w tym, że jest pozbawiona poważniejszej konkurencji. Nie jest bowiem tak, jak twierdzi indywidualizm eksplanacyjny, że *rzeczywiście* działamy powodowani tylko własnymi pragnieniami i przekonaniami. Jest co najwyżej tak, jak twierdzi indywidualizm normatywny: że — w miarę możliwości — *powinniśmy*

działać tylko zgodnie z własnymi pragnieniami i przekonaniem — o ile mamy być uznawani za niezależne i autonomiczne jednostki.

Rozważmy dwa przykłady, które zdają się wspierać to rozróżnienie. Wyobraźmy sobie gospodynię domową, która — na pytanie, dlaczego sprząta, pierze, gotuje, itd. — odpowiada (zupełnie poważnie), iż jest to jej społeczną rolą jako żony i pani domu, która to rola jest częścią istniejącego porządku patriarchalnego. Odczuwamy przecież, że coś w tym wyjaśnieniu jest „nie tak”, pomimo tego, że wielu z nas sądzi, iż wyjaśnienie to jest prawidłowe, a fakty, do których się odwołuje, mają rzeczywiste miejsce. Gdyby takie wyjaśnienie jej działań podał ktoś inny, przyjęlibyśmy je przecież bez wahania. Co więc stoi na przeszkodzie naszej akceptacji jej własnego wyjaśnienia jej własnych działań? Dlaczego jej wyjaśnienie nie utwierdza nas dodatkowo w przekonaniu, iż nasze wyjaśnienie jest trafne?

Odpowiedź na te pytania jest prosta: nie takiego „wyjaśnienia” od niej oczekujemy. To, co chcemy *od niej* usłyszeć, to wyjaśnienie odwołujące się do *racji*, jakie *nią* kierowały przy wyborze takich a nie innych działań. Przyczyną tego, że jej socjologiczne wyjaśnienie niejako nam „nie pasuje”, jest po prostu to, że nie podaje *racji*, które przemawiałyby za takim działaniem. Przeciwnie, zdaje się właśnie podawać *racje* przeciw takiemu zachowaniu — kto w końcu pragnie trwać w poddaństwie?

Rozważmy drugi przykład. Po tym, jak Milgram (1969) przeprowadził swe słynne eksperymenty, prosił ich uczestników o refleksję (*post factum*) nad tym, dlaczego zadawali wstrząsy elektryczne do samego końca, słuchając poleceń eksperymentatora, a nie głosu sumienia. Jedną z tych odpowiedzi jest szczególnie pouczająca — brzmi ona: „Nie wiem”. W pewnym sensie oczywiście jest to odpowiedź po prostu fałszywa. Jeśli brać pod uwagę tylko poziom eksplanacyjny, to oczywiście jest (również dla pytanego wtedy człowieka), że działał właśnie dlatego, że eksperymentator takie działanie mu nakazał. W innym sensie (na poziomie normatywnym), jest to odpowiedź jak najbardziej prawidłowa. Ten człowiek wyraził po prostu fakt, iż nie posiada wyjaśnienia swego działania, którego od niego oczekujemy, a mianowicie wyjaśnienia odwołującego się do *racji*, które nim powodowały.

#### 4.2. Łatwość przypisywania pragnień i przekonań

Warto tu zauważyć, że rozróżnienie tych dwóch poziomów rzuca odmienne światło na fakt, iż tak łatwo przychodzi nam przypisywanie ludziom postaw intencjonalnych. Ten fakt sprawia, iż psychologia intencjonalna jest idealnym narzędziem dla przyjętego w naszej kulturze celu ukształtowania nas jako niezależnych i silnych jednostek, których działania możemy (prawie zawsze) postrzegać jako powodowane przez ich własne — jeśli nie podzielane przez nas, to w każdym razie jakieś — pragnienie czy przekonanie.

Dwuznaczność, o której mowa była w § 3.1, jest tu szczególnie pouczająca. Jak pamiętamy, łatwo mylimy brak pewnego pragnienia z obecnością odpowiedniego

kontrpragnienia. Zważmy, co się dzieje z prawem wyłączonego środka zastosowanym do pragnień:

(ZD) Dla każdego działania, albo jest tak, że sprawca chce to działanie podjąć albo nie jest tak, że sprawca chce to działanie podjąć.

Bardziej potoczycie: dla każdego działania, albo sprawca chce je podjąć, albo nie chce go podjąć. Przy wspomnianej tendencji do odczytywania wyrażeń o formie 'α chce sprawić tego, że p' jako oznaczających fakt, iż α chce nie sprawić tego, że p, łatwo — choć opacznie — przechodzimy do fałszywego twierdzenia:

(ZPP) Dla każdego działania, albo jest tak, że sprawca chce to działanie podjąć (ma pragnienie podjęcia tego działania), albo jest tak, że sprawca chce tego działania nie podejmować (ma kontrpragnienie podjęcia tego działania).

Twierdzenie (ZPP) można określić mianem *zasady polaryzacji pragnień*, gdyż pozwala ono na atrybucję jakiegoś (pro- czy kontr-) pragnienia dla *każdego* działania. Jest to istotne w obliczu ideału, do którego dążymy. Najgorsze, co mogłoby nas — jako aspirujące do samodzielności i niezależności jednostki — spotkać, to ewentualność tego, iż okazalibyśmy się obojętni. — „Jeśli zjadłeś szpinak, to znaczy, że musiał ci smakować, bo gdyby ci nie smakował, to byś go nie zjadł.” Oczywiście dopuszczalne jest podanie innego powodu, dla którego szpinak został zjedzony (np. żeby nie urazić gospodarzy). Niemniej jednak trzeba się wykazać jakimś stosunkiem do sprawy. Możliwość, że ktoś zjadł szpinak tak po prostu, bez zastanowienia, bez jakiegokolwiek pragnienia (za czy przeciw), jest wykluczona przez zasadę polaryzacji pragnień. Zasada ta stoi na straży tego, by nasze *postawy* odzwierciedlały nasze działania.

### 4.3. Normatywny nieindywidualizm?

Można by przeczyć temu, że jesteśmy w życiu codziennym indywidualistami normatywnymi. Przecież — można argumentować — uznajemy także ideały nastawione na innych, np. pomocy innym, pracy na rzecz dobra społeczności, przekładania dobra innych nad swe własne, itd. Co więcej, to, jakie ideały są propagowane, jest w dużym stopniu zależne od kultury. Nietrudno jest jednakże uzgodnić zasadę, że ludzie powinni się kierować swymi własnymi racjami, z innymi normami (pomocy innym, pracy na rzecz dobra wspólnego, etc.). Gorzej przecież traktowalibyśmy człowieka, który zawsze postępowałby z normami nastawionymi na pomoc innym, działanie na rzecz wspólnego dobra, etc., ale tylko dlatego, że jego matka uznawała to za wskazane, niżli człowieka, który by to czynił kierując się własnymi wartościami. Jeśli tak, to zasadne wydaje się uogólnienie tego przypuszczenia w taki sposób, że normatywno-nieindywidualistyczne ideały, owszem, utrzymujemy, ale jako wtórne wobec ideałów normatywno-indywidualistycznych. Jednostka, która stosuje się do ide-

ałów normatywno-nieindywidualistycznych, ale nie do ideałów normatywno-indywidualistycznych, może być co prawda pożyteczna społecznie, ale nie okaże się godna miana niezależnej i autonomicznej jednostki.

#### 4.4. Napięcia pojęciowe: altruizm, słaba wola i zniewolenie

Chcę zakończyć moje rozważania, przywołując fakty potwierdzające tezę, że główny cel psychologii intencjonalnej to nie wyjaśnianie, jak rzeczywiście postępujemy, lecz ukształtowanie nas w samodzielne jednostki. Jest to więc cel nie poznawczy, lecz normatywny. Zwracałam już niejednokrotnie uwagę na zadziwiającą łatwość, z jaką przypisujemy ludziom pragnienia i przekonania. Ta łatwość okupiona jest jednakże dającymi do myślenia „napięciami konceptualnymi”. Występują one mianowicie, gdy staramy się w terminach psychologii intencjonalnej ująć zjawiska, które zdają się przeczyć ideałowi silnej i niezależnej jednostki. Kategorie psychologii intencjonalnej samoistnie niejako zakłamują te zjawiska, przekształcając je tak, aby uzgodnić się dały z ideałem psychologii intencjonalnej. Przytoczę tu trzy przykłady: altruizmu, słabej woli i zniewolenia.

Najlepiej znanym przykładem takiego „pojęciowego napięcia” jest paradoks altruizmu. Sama próba określenia tego, czym jest działanie altruistyczne, a mianowicie działanie na rzecz preferencji innego człowieka, zdaje się niemożliwe, bo od razu zmuszeni jesteśmy odnajdować tam preferencje sprawcy — wszakże musiał *chcieć* to zrobić dla tego człowieka (w końcu, *gdyby nie chciał*, to...). Zatem jego działanie musi być rozumiane jako powodowane przez jego własny cel (nawet jeśli tym celem jest realizowanie celu innego człowieka). A jeśli tak, to działanie to wcale nie jest altruistyczne, ale właśnie egoistyczne — wywodzi indywidualistycznie zorientowany zdrowy rozsądek.

Paradoks altruizmu jest interesujący, gdyż powstaje niejako z samego zastosowania kategorii intencjonalnych. Bierze się stąd, że trudno nam w tych kategoriach wyrazić to, co chcieliśmy wyrazić. Oczywiście kwestią dalszą jest, jak ten paradoks potraktować: czy przyjąć go za wyznacznik tego, że altruizm nie istnieje, czy też próbować wyróżnić np. rodzaje pragnień, które nie świadczyłyby o tym, że jesteśmy egoistami. Interesujący jest raczej sam przez się fakt, że ten paradoks daje się sformułować, że próba ujęcia zjawiska altruizmu kolidującego w jakimś stopniu z ideałem silnej i niezależnej jednostki, jaki psychologia intencjonalna ma forsować, prowadzi do jego przekłamania. Nie twierdzę, że to przekłamanie jest nieodwracalne — zwracam tylko uwagę na to, że zachodzi.

Podobne napięcie zachodzi przy próbach konceptualizacji zjawiska słabej woli. To, że ludzie wykazują słabą wolę, nie ulega wątpliwości. Ale już próba zrozumienia tego zjawiska w języku psychologii intencjonalnej okazuje się trudna. Gdy wyobrażamy sobie akratyka, który decyduje, że nie sprawi tego, że *p*, jest w pełni umotywowany, by nie sprawić tego, że *p*, a potem sprawia to, że *p*, jesteśmy natychmiast

skłonni do przypuszczenia, że musiał jednak — w jakimś sensie — chcieć sprawić to, że *p.* (W końcu, *gdyby nie chciał*, to...) Być może zmienił zdanie, a może znalazł się pod wpływem jakiegoś chwilowego pragnienia, żeby *p.* Ale jeśli tak, to jego działanie wcale nie było przejawem słabej woli — przeciwnie, robił dokładnie to, co chciał! W istocie rzeczy, jeśli przyjrzeć się poszczególnym działaniom akratycznym, bardzo kusi taka próba ich rekonstrukcji, która przejawy słabej woli przeistacza właśnie w przejawy silnej woli. W konsekwencji jesteśmy pewniejsi istnienia zjawiska akrazji jako takiej niż poszczególnych jej przejawów, z których każdy — przy bliższym (ale intencjonalnym!) przyjrzeniu się — zdaje się być właśnie zjawiskiem dokładnie odwrotnym. Jak w wypadku altruizmu, tak i w tym wypadku, takie samoistne przekłamanie naszych intencji jest zadziwiająco spójne z ideałem jednostki silnej i niezależnej.

Niech ostatnim przykładem pojęciowych napięć, jaki tu przytaczam, będą zjawiska zniewolenia, opisane m. in. przez L. Nowaka [1985], [1987], [1991, t. III]. Większość z nas wykazuje tendencję do odpowiadania wrogością na kolejne akty wrogości ze strony drugiej osoby. Nowak zauważa, że — dla każdego z nas — nadchodzi poziom wrogości (lub moment trwania wrogości ustalonej), kiedy dalszy przyrost wrogości wobec nas nie będzie prowadzić do eskalacji odpowiedzi wrogich z naszej strony, lecz do odpowiedzi coraz słabszych, a w końcu przy kolejnym progu staje się wręcz przeciwnie: będziemy odpowiadać życzliwością na coraz to bardziej wrogie akty. Klasyczne przypadki ilustrujące ten fenomen to postać Winstona Smitha z 1984 Orwella, jak również zakładnicy, których kuriozalne zachowanie wobec porywaczy określono mianem „syndromu sztokholmskiego” i generalnie przypadki ludzi, którzy zostali „złamani”: więźniów, żołnierzy, pacjentów psychiatrycznych, kobiet, niewolników — wszystkich tych, którzy byli poddani torturze psychicznej, fizycznej czy sytuacyjnej. Gdy taka osoba zachowuje się życzliwie wobec swego oprawcy, jesteśmy skłonni interpretować jej zachowanie jako przejaw służalczości. Lecz gdy tylko spróbujemy zrozumieć jej działania intencjonalnie — czy to odwołując się do tego, że *chce* być życzliwa, czy nawet do tego, że *chce* być służalcza, samo określenie jej zachowania jako służalczego zdaje się być zagrożone. W końcu — pobrzmiewa w nas głos wewnętrzny — ona może robić to, co chce: jeśli tak się właśnie *chce* wobec swego oprawcy zachowywać, to jest to *jej* prawo. Jeśli jednakże jest to właściwy rys psychologiczny tej osoby, to jawi się ona nam przecież nie jako osoba służalcza, ale właśnie jako osoba niezależna i silna. Po raz kolejny jest więc tak, że samo scharakteryzowanie jej zachowania w języku psychologii intencjonalnej zdaje się przekłamywać nasze wyjściowe intuicje, i to przekłamywać dokładnie w kierunku ideału jednostki, który — jak twierdzą — psychologia intencjonalna ma forsować.

We wszystkich trzech wypadkach — altruizmu, akrazji i zniewolenia, widzimy skłonność do przekłamania naszych wyjściowych intuicji. Jest tak, jak gdyby samo ujęcie danego zjawiska w kategoriach intencjonalnych już zaopatrywało je w wewnętrzną inercję, którą trudno jest w trakcie interpretacji tego zjawiska przełamać. Oczywiście sam fakt, że trudno jest wyrazić pewne zjawiska przy użyciu języka intencjonalnego, nie znaczy jeszcze, że jest to niemożliwe. Tomy całe zostały poświę-

cone szczegółowym próbom ujęcia tych zjawisk w języku psychologii intencjonalnej za pomocą rozmaitych rozstrzygnięć i rozróżnień pojęciowych. A jednak samo istnienie tych napięć już na presystematycznym, czysto intuicyjnym poziomie tłumaczenia działań, wystarcza dla poparcia wysuwanej tu tezy o głównie „wychowawczej” a nie teoretycznej roli psychologii intencjonalnej.

## BIBLIOGRAFIA

- Ajdukiewicz, K. (1938) — „Postępowanie człowieka”, [w:] (1960), s. 314—364.
- Ajdukiewicz, K. (1960) — *Język i poznanie*, Tom 1, Warszawa: Państwowe Wydawnictwo Naukowe.
- Anscombe, G. E. M. (1963) — *Intention*, 2nd edition, Ithaca: Cornell University Press.
- Aronson, E., red. (2001) — *Człowiek istota społeczna: Wybór Tekstów*, Warszawa: PWN.
- Baier, A. (1985) — *Postures of the Mind. Essays on Mind and Morals*, Minneapolis: University of Minnesota Press.
- Baker, L. R. (1995) — *Explaining Attitudes. A Practical Approach to the Mind*, Cambridge: Cambridge University Press.
- Belnap, N., Perloff, M. (1990) — „Seeing to It that: A Canonical Form for Agentives”, [w:] Kyburg, Jr. et. al. (1990), s. 175—199.
- Belnap, N., Perloff, M., Xu, M. (2001) — *Facing the Future. Agents and Choices in Our Indeterminist World*, Oxford: Oxford University Press.
- Boyd, R., Richerson, P. (1985) — *Culture and the Evolutionary Process*, Chicago: Chicago University Press.
- Bratman, M.E. (1987) — *Intention, Plans, and Practical Reason*, Cambridge, MA: Harvard University Press.
- Burge, T. (1979) — „Individualism and the Mental”, [w:] French et. al. (1979), s. 73—122.
- Churchland, P. S. (1986) — *Neurophilosophy. Toward a Unified Science of the Mind/Brain*, Cambridge, MA: MIT Press.
- Churchland, P. M. (1979) — *Scientific Realism and the Plasticity of Mind*, Cambridge: Cambridge University Press.
- Churchland, P. M. (1984) — *Matter and Consciousness. A Contemporary Introduction to Philosophy of Mind*, Cambridge, MA: The MIT Press.
- Churchland, P. M. (1989a) — *A Neurocomputational Perspective. The Nature of Mind and the Structure of Science*, Cambridge, MA: MIT Press.
- Churchland, P. M. (1989b) — „Folk Psychology and the Explanation of Human Behavior”, [w:] Tomberlin (1989), s. 225—242.
- Chwedeńczuk, B., red. (1995) — *Filozofia umysłu*, Warszawa: Wydawnictwo SPACJA.
- Collins, A.W. (1987) — *The Nature of Mental Things*, Notre Dame, IN: University of Notre Dame Press.
- Czeżowski, T. (1959) — „O postępowaniu”, [w:] (1989), s. 247—282.
- Czeżowski, T. (1989) — *Pisma z etyki i teorii wartości*, Wrocław: Wydawnictwo Zakładu Narodowego im. Ossolińskich.
- Davidson, D. (1963) — „Actions, Reasons, and Causes”, [w:] (1980), s. 3—19.
- Davidson, D. (1978) — „Intending”, [w:] (1980), s. 83—102.
- Davidson, D. (1980) — *Essays on Actions and Events*, Oxford: Clarendon Press.
- Davidson, D. (1984) — *Inquiries into Truth and Interpretation*, Oxford: Clarendon Press.

- Davies, M., Stone, T., red. (1995a) — *Folk Psychology. The Theory of Mind Debate*, Oxford: Basil Blackwell.
- Davies, M., Stone, T., red. (1995b) — *Mental Simulation. Evaluations and Applications*, Oxford: Basil Blackwell.
- Dennett, D. C. (1969/1995) — „Osobowy i subosobowy poziom wyjaśniania”, [w:] Chwedeńczuk (1995), s. 103—109.
- Dennett, D. C. (1971) — „Intentional Systems”, [w:] (1981a), s. 3—22.
- Dennett, D. C. (1978) — „Three Kinds of Intentional Psychology”, [w:] (1987), s. 43—82.
- Dennett, D. C. (1981a) — *Brainstorms. Philosophical Essays on Mind and Psychology*, Cambridge, MA: Bradford Books.
- Dennett, D. C. (1981b) — „True Believers: The Intentional Strategy and Why It Works”, [w:] (1987), s. 13—35.
- Dennett, D. C. (1987) — *The Intentional Stance*, Cambridge, MA.: Bradford Books.
- Dennett, D. C. (1991) — *Consciousness Explained*, Boston: Little, Brown.
- Falkiewicz, A., Nowak, L., red. (1996) — *Przestrzenie świadomości. Studia z Filozofii Literatury*, Poznań: Zysk i S-ka Wydawnictwo.
- Fleming, N. (1981) — „Autonomy of the Will”, *Mind* 90, 201—223.
- Flew, A., red. (1951) — *Essays on Logic and Language*, Oxford: Blackwell.
- Fodor, J. A. (1975) — *The Language of Thought*, New York: Crowell.
- Fodor, J. A. (1987) — *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*, Cambridge, MA: The MIT Press.
- Fodor, J. A. (1994) — *The Elm and the Expert. Mentalese and Its Semantics*, Cambridge, MA: The MIT Press; wyd. pol.: *Eksperci od wiązów. Język myślenia i jego semantyka*, tłum. M. Gokieli, Warszawa: Fundacja Aletheia, 2001.
- French, P. A. (1985) — „Fishing the Red Herrings Out of the Sea of Moral Responsibility”, [w:] LePore i McLaughlin (1985), s. 73—87.
- French, P. A., Uehling, Jr., T.E., Wettstein, H.K., red. (1979) — *Studies in Metaphysics. Midwest Studies in Philosophy*, vol. 4, Minneapolis: University of Minnesota Press.
- Goffman, E. (1986) — *Frame Analysis. An Essay on the Organization of Experience*, Boston: Northeastern University Press.
- Goldman, A. (1989) — „Interpretation Psychologized”, *Mind and Language* 4, 161—185.
- Goldman, A. I. (1992) — „In Defence of the Simulation Theory”, [w:] Davies i Stone (1995a), s. 191—206.
- Gopnik, A., Wellman, H. M. (1992) — „Why the Child's Theory of Mind Really Is a Theory”, *Mind & Language* 7, 145—171.
- Gordon, R. M. (1986) — „Folk Psychology as Simulation”, *Mind and Language* 1, 158—171.
- Gordon, R. M. (1992) — „The Simulation Theory: Objections and Misconceptions”, [w:] Davies i Stone (1995a), s. 100—123.
- Haney, C., Banks, C., Zimbardo, P. G. (1973) — „Interpersonal Dynamics in a Simulated Prison”, *International Journal of Criminology and Penology* 1, 69—97.
- Haney, C., Banks, C., Zimbardo, P. G. (1973/2001) — „Więźniowie i strażnicy — badanie w symulowanym więzieniu”, [w:] Aronson (2001), s. 83—101.
- Hart, H. L. A. (1951) — „The Ascription of Responsibility and Rights”, [w:] Flew (1951), s. 145—166.
- Hornsby, J. (1980-1) — „Which Physical Events are Mental Events”, *Proceedings of the Aristotelian Society* 81, 73—92.

- Hornsby, J. (1997) — *Simple Mindedness. In Defense of Naive Naturalism in the Philosophy of Mind*, Cambridge, MA: Harvard University Press.
- Klawiter, A. (1978) — *Problem metodologicznego statusu materializmu historycznego*, Warszawa: Państwowe Wydawnictwo Naukowe.
- Kmita, J. (1971) — *Z metodologicznych problemów interpretacji humanistycznej*, Warszawa: PWN.
- Kmita, J., red. (1973) — *Elementy marksistowskiej metodologii humanistyki*, Poznań: Wydawnictwo Poznańskie.
- Kmita, J., Nowak, L. (1968) — *Studia nad teoretycznymi podstawami humanistyki*, Poznań: Wydawnictwo Naukowe UAM.
- Kotarbiński, T. (1955) — *Traktat o dobrej robocie*, Warszawa.
- Kozielecki, J. (1975) — *Psychologiczna teoria decyzji*, Warszawa: Wydawnictwo Naukowe PWN.
- Kyburg, Jr., H. E., Loui, R. P., Carlson, G. N., red. (1990) — *Knowledge Representation and Defeasible Reasoning*, Dordrecht: Kluwer.
- LePore, E., McLaughlin, B. P., red. (1985) — *Actions and Events. Perspectives on the Philosophy of Donald Davidson*, Oxford: Basil Blackwell.
- Macdonald, C., Macdonald, G., red. (1995) — *Connectionism. Debates on Psychological Explanation*, Oxford: Basil Blackwell.
- Milgram, S. (1962/2001) — Behavioralne badanie posłuszeństwa, [w:] Aronson (2001), s. 48—64.
- Milgram, S. (1969) — *Obedience to Authority*, New York: Harper & Row.
- Moran, R. (1994) — „Interpretation Theory and the First Person”, *The Philosophical Quarterly* 44, 154—173.
- Nowak, L. (1985) — „Człowiek i ludzie”, *Obecność* 9, 45—54.
- Nowak, L. (1987) — „Man and People”, *Social Theory and Practice* 14, 1—17.
- Nowak, L. (1991) — *Power and Civil Society. Toward a Dynamic Theory of Real Socialism*, New York: Greenwood Press.
- Nowak, L. (1991) — *U podstaw teorii socjalizmu*, Poznań: Wydawnictwo NAKOM.
- Paprzycka, K. (1996) — „Gombrowicz a analityczne teorie działania”, [w:] Falkiewicz i Nowak (1996), s. 205—218.
- Paprzycka, K. (1997) — *Social Anatomy of Action. Toward a Responsibility-Based Conception of Agency*, Ph.D. Dissertation: University of Pittsburgh.
- Paprzycka, K. (1998) — „Collectivism on the Horizon: A Challenge to Pettit's Critique of Collectivism”, *Australasian Journal of Philosophy* 76, 165—181.
- Paprzycka, K. (2002) — „False Consciousness of Intentional Psychology”, *Philosophical Psychology* 15—3, 271—295.
- Patryas, W. (1975) — „An Analysis of the Ceteris Paribus Clause”, *Poznań Studies in the Philosophy of the Sciences and the Humanities* 1, 59—64.
- Patryas, W. (1976) — *Eksperyment a idealizacja*, Warszawa-Poznań: PWN.
- Patryas, W. (1979) — *Idealizacyjny charakter interpretacji humanistycznej*, Poznań: Wydawnictwo Naukowe UAM.
- Pettit, P. (1986) — *The Common Mind. An Essay on Psychology, Society, and Politics*, Oxford: Oxford University Press.
- Pettit, P., Smith, M. (1990) — „Backgrounding Desire”, *Philosophical Review* 99, 565—592.
- Putnam, H. (1975) — *Mind, Language and Reality. Philosophical Papers, vol. II*, Cambridge: Cambridge University Press.
- Putnam, H. (1975) — „The Meaning of Meaning”, [w:] Putnam (1975), s. 215—271.

- Rorty, R. (1965) — „Mind-Body Identity, Privacy and Categories”, *Review of Metaphysics* 19, 24—54; tłum. polskie: „Identyczność ciała i umysłu, prywatność a kategorie ontologiczne”, tłum. M. Szczubiałka, [w:] Chwedeńczuk (1995), s. 215-244.
- Searle, J. R. (1983) — *Intentionality. An Essay in the Philosophy of Mind*, Cambridge: Cambridge University Press.
- Smith, M. (1987) — „The Humean Theory of Motivation”, *Mind* 96, 36—61.
- Smith, M. (1994) — *The Moral Problem*, Oxford: Blackwell.
- Smolensky, P. (1995a) — „On the Projectable Predicates of Connectionist Psychology”, [w:] Macdonald i Macdonald (1995).
- Smolensky, P. (1995b) — „On the Proper Treatment of Connectionism”, [w:] Macdonald i Macdonald (1995).
- Stich, S. (1983) — *From Folk Psychology to Cognitive Science*, Cambridge, MA.: MIT Press.
- Stich, S., Nichols, S. (1992) — „Folk Psychology: Simulation or Tacit Theory”, [w:] Davies i Stone (1995a), s. 100—122.
- Stich, S., Nichols, S. (1995) — „Second Thoughts on Simulation”, [w:] Davies i Stone (1995b), s. 87—108.
- Szaniawski, K. (1994) — *O nauce, rozumowaniu i wartościach*, Warszawa: Wydawnictwo Naukowe PWN.
- Tomberlin, J. E., red. (1989) — *Philosophical Perspectives*, Vol. 3: *Philosophy of Mind and Action Theory*, Atascadero, CA: Ridgeview Publishing Company.
- Topolski, J. (1973) — „Aktywistyczna koncepcja procesu dziejowego”, [w:] Kmita (1973), s. 255—273.
- Topolski, J. (1973) — „Dyrektywa racjonalizowania działań ludzkich”, [w:] Kmita (1973), s. 222—236.
- Tuomela, R. (1984) — *A Theory of Social Action*, Dordrecht: Reidel.
- von Wright, G. H. (1983) — „Explanation and Understanding of Action”, [w:] (1983), s. 53—66.
- von Wright, G. H. (1983) — *Practical Reason. Philosophical Papers*, vol. 1, Ithaca: Cornell University Press.
- Wilson, G. M. (1989) — *The Intentionality of Human Action*, Stanford: Stanford University Press.
- Wittgenstein, L. (1958/2000) — *Philosophical Investigations*, New York: Macmillan; wyd. pol.: *Dociekania filozoficzne*, tłum. Bogusław Wolniewicz, Warszawa: Wydawnictwo Naukowe PWN.