

STANISŁAW JUDYCKI

## INTROSPEKCJA JAKO PROBLEM FILOZOFICZNY

### 1. UWAGI OGÓLNE

Zagadnienie natury introspekcji ma przynajmniej cztery najważniejsze wymiary teoretyczne. Są w nim obecne aspekty epistemologiczne, metafizyczne (ontologiczne), psychologiczne oraz antropologiczne. Samo słowo „introspekcja” (z łacińskich *intro* – „do wewnątrz” i *specto* – „patrzeć”) oznacza spojrzenie do wewnątrz, a w związku tym, że introspekcja jest pewnym zdarzeniem w ramach świadomości, można powiedzieć, że jest to **spojrzenie od wewnątrz do wewnątrz**. Używane są także i inne słowa na określenie tej władzy poznawczej: spostrzeganie wewnętrzne, samoświadomość, samoobserwacja, refleksja. Każdemu z nich przypisywane są niekiedy nieco inne znaczenia. Część tych przesunięć znaczeniowych zostanie zasygnalizowana w dalszym ciągu rozważań. Czasami introspekcja przeciwstawiana jest retrospekcji, czyli przypomnieniu. Podstawową jednak opozycją do introspekcji jest spostrzeganie zewnętrzne (percepcja zmysłowa, raczej rzadko określana mianem ekstraspekcji). **Percepcja zmysłowa przedstawia nam świat zewnętrzny, świat przedmiotów fizycznych, natomiast w introspekcji mamy dany świat wewnętrzny, świat przedmiotów psychicznych**. Przedmiotem introspekcji są akty, procesy i zdarzenia należące do strumienia świadomości. W pewnej mierze przedmiotem introspekcji mogą być również dyspozycje (np. „bycie inteligentnym” czy „bycie odważnym”), lecz zazwyczaj zwraca się uwagę na to, że nasze własności dyspozycyjne poznajemy w dzia-

łaniu, gdyż introspekcyjna „samokontemplacja” niewiele może nam powiedzieć o tego rodzaju własnościach. Przedmiotem introspekcji może być też nasze „ja”, czyli centrum „strumienia” świadomości, z którego wypływają (i do którego się odnoszą) wszystkie akty, procesy i zdarzenia świadome. Gdy na przykład spostrzegam zamglone wybrzeże morza, to przedmiotem percepcji zmysłowej są morze, mgła itd., a przedmiotem introspekcji może być moje przeżycie, powiedzmy, melancholii. Dzięki introspekcji mogę jednak zwrócić uwagę również na to, że spostrzegam wybrzeże morza, a nie że je sobie tylko wyobrażam, że widoki, które oglądam, są niewyraźne itp. Introspekcja dotyczy więc nie tylko doznań „czysto” wewnętrznych, lecz także percepcji zewnętrznych, o ile staną się one przedmiotem obserwacji wewnętrznej.

Epistemologiczna problematyka introspekcji dotyczy pytania o jej naturę, zasięg i wartość jako źródła formowania przekonań i zdobywania wiedzy na temat naszego życia psychicznego. Metafizyczne (ontologiczne) aspekty introspekcji odnoszą się do zagadnienia natury naszego umysłu: czy introspekcja informuje o istnieniu odmiennej ontycznie od świata fizycznego dziedziny świata duchowego? Psychologiczna problematyka introspekcji to głównie pytanie, czy introspekcja może być narzędziem poznania naukowego<sup>1</sup>. Behawioryzm w XX w. był skierowany przeciwko introspekcji jako możliwemu narzędziu badawczemu naukowej psychologii. Zagadnienia z zakresu metodologii psychologii przecinają się częściowo z problematyką epistemologiczną. Istnieje też aspekt antropologiczny introspekcji: jako źródło poznania samego siebie stanowi ona niezbędne założenie możliwości autodeterminacji, podejmowania wolnych i rozumnych decyzji, poczucia winy i możliwości przebudowy moralnej itp. Ta problematyka jest bliska zagadnieniom metafizycznym (ontologicznym).

Współczesna problematyka natury introspekcji ma swoje korzenie historyczne w filozofii starożytnej i średniowiecznej. Arystoteles w traktacie *O duszy* pisał:

Rozum jest dla siebie przedmiotem myślenia, tak jak są nimi inne jego przedmioty myślenia. W rzeczach bowiem bezmaterialnych myślący i przedmiot jego myśli są czymś jednym i tym samym. Wiedza teoretyczna bowiem jest w gruncie rzeczy identyczna z tym, co się wie<sup>2</sup>.

<sup>1</sup> Por. M. K r e u t z, *Metody współczesnej psychologii*, Warszawa: PWN 1962.

<sup>2</sup> *O duszy*, 430 a 2, tłum. P. Siwek, Warszawa: PWN 1972; por. *Metafizyka*, 1072 b, tłum. K. Leśniak, Warszawa: PWN 1983.

Św. Augustyn twierdził, że duch poznaje siebie przez siebie samego, ponieważ jest niematerialny<sup>3</sup>. Św. Tomasz z Akwinu uznawał, że dusza nie poznaje siebie samej bezpośrednio co do swojej istoty (*per suam essentiam*), lecz tylko co do swojego istnienia. Poznanie siebie, według Tomasza, odbywa się zawsze poprzez poznanie rzeczy zewnętrznych i myślenie o nich<sup>4</sup>. Według Dunska Szkota dusza poznaje siebie tylko przez swój obraz, a według Ockhama mamy intuicyjne poznanie naszej duszy. W literaturze współczesnej przytacza się zazwyczaj poglądy Kartezjusza, który twierdził, że podmiot, myśląc, ujmuje siebie samego jako niematerialną substancję, której istotą jest myślenie. Ten pogląd uważa się za wzorcowy dla stanowiska podkreślającego unikalność poznania introspekcyjnego. Już jednak we wcześniejszych stanowiskach filozofów widać pewne wahanie, **czy samopoznanie dostarcza wiedzy o istocie naszego umysłu, czy też raczej prezentuje tylko jego zjawisko (obraz).**

Agnostyczny pogląd co do możliwości poznawczego uchwycenia własnej istoty metafizycznej za pomocą introspekcji wyraźnie sformułował Kant. Według niego zmysłowość charakteryzuje się receptywnością, a więc zdolnością do nabywania przedstawięń przez pobudzenie zewnętrzne. Jeśli zatem poznanie świata zewnętrznego wymaga pobudzenia przez jakieś bodźce, to w wypadku samopoznania trzeba przyjąć, że oprócz zmysłu zewnętrznego istnieje też zmysł wewnętrzny. W poznaniu introspekcyjnym nasz umysł (świadomość) nie wytwarza przedstawięń spontanicznie (samoistnie), lecz musi sam siebie pobudzać, aby uzyskać poznanie siebie. Formami zmysłu zewnętrznego są, według Kanta, przestrzeń i czas. Obie te formy zmysłowości są subiektywne. Formą zaś zmysłu wewnętrznego jest tylko czas, gdyż nasze stany mentalne nie mają żadnej rozciągłości przestrzennej. Poznając siebie w subiektywnej formie czasu, w wyniku pobudzenia siebie przez siebie samego (*Selbstaffektion*), umysł nie poznaje siebie takim, jakim jest, lecz tylko jako pewne zjawisko. Oficjalna doktryna głoszona przez Kanta stwierdzała, że na podstawie tego zjawiskowego poznania siebie nie możemy uzyskać wiedzy, czy nasz umysł jest w swojej istocie (czyli jako tzw. *noumen*)

---

<sup>3</sup> „Mens se ipsam novit per se ipsam, quoniam est incorporeal”. *De Trinitate*, eds. W. J. Mountain, F. Glorie, Turnhout: Brepols 1968, XI, 3.

<sup>4</sup> „Mens nostra non potest seipsam intelligere, ita quod seipsam immediate apprehendat, sed ex hoc, quod apprehendit alia, devenit in suam cognitionem”. *Quaestiones disputatae de veritate*, w: *Opera omnia, iussu Leonis XIII edita cura et studio Fratrum Praedicatorum*, t. XXII, Roma 1970, X, 8.

niematerialną substancją (duszą), czy też raczej jego substratem jest substancja czysto fizyczna (mózg)<sup>5</sup>. Większość stanowisk w filozofii XX w. kontynuuje tę agnostyczną linię kantowską, a tym samym dystansuje się od linii kartezjańskiej i od poprzedzających ją interpretacji bardziej metafizycznie optymistycznych. **Dzisiaj filozofowie skłaniają się raczej do tezy agnostycznej, że samopoznanie w formie introspekcyjnej jest bardzo ograniczone** i że na jego podstawie nie możemy wnosić o istnieniu niematerialnego podłoża stanów naszego umysłu. Za główny przedmiot krytyki obiera się, traktowane jako wzorcowe, poglądy Kartezjusza.

## 2. TRADYCYJNA TEORIA INTROSPEKCJI

Teorii spostrzegania wewnętrznego, obecnej w pismach Kartezjusza, przypisuje się współcześnie miano tradycyjnej teorii introspekcji i uznaje się, że funkcjonują w niej przynajmniej następujące założenia.

Dzięki introspekcji mamy **uprzywilejowany dostęp** (*privileged access*) do naszych stanów mentalnych. Według teorii kartezjańskiej dane introspekcyjne stanowią wzorcowy (i być może jedyny) przypadek nieinferencyjnej wiedzy bezpośredniej. Wystarczy, że jakiś stan mentalny zachodzi, a my od razu mamy do niego dostęp. Jeśli chcemy się czegoś dowiedzieć na temat tzw. świata zewnętrznego, to musimy dokonywać różnego rodzaju badań, wnioskowań itd., a więc dążyć do odsłonięcia zakrytej istoty danego zjawiska. Jeżeli natomiast odczuwamy np. ból, to od razu wiemy, czym on jest. Nie istnieje więc żadna zakryta istota zjawisk mentalnych. Inaczej mówiąc, dla danych introspekcyjnych nie da się sformułować opozycji zjawisko–rzeczywistość. W odróżnieniu od danych odnoszących się do świata rzeczy fizycznych, dane introspekcyjne są jakby „płaskie”, tzn. nie przejawiają się przez wyglądy, przez tzw. dane zmysłowe (*sense data*) itp. Gdy w introspekcji ujmujemy nasze „ja”, wówczas tym samym poznajemy pewną duchową real-

---

<sup>5</sup> „Nawet wewnętrzna i zmysłowa naoczność naszego umysłu (*Gemüt*), którego własności (*Bestimmung*) przedstawiamy sobie przez następstwo różnych stanów w czasie, nie jest właściwym Ja (*Selbst*), takim, jakim ono istnieje samo w sobie, czyli podmiotem transcendentnym, lecz tylko zjawiskiem, które zostało dane zmysłowości owej nieznaną nam istoty”. I. K a n t, *Krytyka czystego rozumu*, tłum. R. Ingarden, t. I, Warszawa: PWN 1957, A 492. Por. też S. J u d y c k i, *Kantowska teoria umysłu*, „Roczniki Filozoficzne”, 45(1997), z. 1, s. 35-60.

ność, która sama nie jest już przejawem czegoś innego, np. jakichś stanów nieświadomych czy fizjologicznych stanów mózgu.

Według tradycyjnej teorii introspekcji, z tezy uprzywilejowanego dostępu ma wynikać **niepowątpiewalność danych introspekcyjnych** (czy ich niekorygowalność, nieobalalność). Jeżeli odczuwam ból, to nie mogę się mylić co do tego, że go odczuwam. Podobnie jeżeli spostrzegam w ogrodzie sylwetkę człowieka, to nie mogę się mylić co do danej introspekcyjnej mówiącej, że przeżywam widzenie takiej sylwetki, chociaż w rzeczywistości może to być np. słomiany chochoł. Percepcje zewnętrzne, potraktowane jako dane introspekcyjne, byłyby zatem – podobnie jak zdarzenia czysto wewnętrzne, np. bóle – niepowątpiewalne. Niepowątpiewalność, niekorygowalność czy też nieobalalność danych introspekcyjnych bywa różnie interpretowana: albo jako psychologiczna niemożliwość pomyślenia, że coś jest w rzeczywistości inne, niż dane introspekcyjne to prezentują, albo jako logiczna niemożliwość wątplenia w te dane, a więc, że każda próba wątplenia w dane tego rodzaju musi prowadzić do sprzeczności, albo też jako niemożliwość pomyślenia doniosłych racji wątplenia w te dane.

**Introspekcja jest źródłem danych całkowicie prywatnych.** Tylko ja mogę introspekcyjnie poznawać mój umysł i ja mogę introspekcyjnie poznawać tylko mój umysł. Stany umysłów innych ludzi są mi dane tylko w wyniku zapośredniczenia przez ich cielesne zachowanie się (trzeba tu pamiętać, że zeznania werbalne na temat stanów mentalnych, które słyszę od innych ludzi, są dla mnie też tylko pewnym rodzajem ich cielesnego zachowania się). Pewność więc mogę mieć tylko co do istnienia i co do treści mojego umysłu. Istnienie innych umysłów jest wyłącznie prawdopodobne, gdyż musi być wywnioskowane z objawów cielesnych. W ten sposób teza o prywatności danych introspekcyjnych prowadzi do problematyki solipsyzmu. Jeżeli wiedza zakłada pewność, to mogę mieć pewność co do istnienia własnego umysłu, istnienie zaś innych umysłów staje się co najwyżej prawdopodobną hipotezą. Takie stanowisko nazywa się niekiedy solipsyzmem epistemologicznym. W myśl tej interpretacji solipsyzmu dysponuję wiedzą (pewną) tylko na temat istnienia jednego umysłu, tzn. mojego własnego. Solipsyzm metafizyczny jest tezą radykalniejszą, a mianowicie, że istnieje tylko jeden umysł, tzn. umysł aktualnie filozofującego podmiotu.

**Dla aktów introspekcji umysł jest czymś całkowicie transparentnym.** Byłoby czymś sprzecznym zakładanie, że mogę znajdować się w pewnym stanie mentalnym bez tego, aby był on dostępny dla introspekcji. Nie mogą więc istnieć stany mentalne, które są niedostępne dla introspekcji. Stany nie-

dostępne dla introspekcji są albo stanami fizjologicznymi mózgu, albo bliżej nieokreślonymi stanami dyspozycjonalnymi, lecz w obu tych wypadkach nie są to stany mentalne. Mentalność oznacza świadomość jakichś treści, a świadomość równa się introspekcyjnemu dostępowi.

**Introspekcja jest osobnym aktem**, w którym dokonuje się ujęcie życia mentalnego rozgrywającego się na pierwszym poziomie (spontanicznie). W akcie introspekcji umysł nasz zdolny jest do uczynienia siebie samego przedmiotem. Interpretacja introspekcji jako osobnego aktu prowadzi do trudności ciągu w nieskończoność: aby dowolny akt introspekcji mógł zostać uświadomiony, potrzebny jest następny akt introspekcji wyższego stopnia i tak w nieskończoność.

**Introspekcja jest spostrzeganiem wewnętrznym**, a więc czymś analogicznym do spostrzeżenia zewnętrznego zmysłowego (percepcji zmysłowej). Problem polega na tym, że w wypadku introspekcji nie da się znaleźć wielu istotnych elementów obecnych w normalnym procesie percepcji zmysłowej. Percepcja zmysłowa funkcjonuje dzięki określonemu mechanizmowi przyczynowemu, który doprowadza do powstania przekonań percepcyjnych. Rzeczy należące do naszego otoczenia fizycznego oddziałują na nasz aparat zmysłowy, co w efekcie prowadzi do powstania przekonań percepcyjnych. W wypadku introspekcji nie wiadomo jednak, jak w ogóle mogłoby wyglądać mechanizm przyczynowy prowadzący do powstania przekonań introspekcyjnych. W jaki sposób miałyby dochodzić do, mówiąc językiem Kanta, „samopobudzenia” się (*Selbstaffektion*) umysłu? Analogia między spostrzeżeniem zewnętrznym a spostrzeżeniem wewnętrznym jest także dlatego problematyczna, że trudno jest wskazać na taki organ spostrzegania wewnętrznego, który byłby funkcjonalnie podobny do organów spostrzegania zmysłowego. Trudno również powiedzieć, na czym polega odrębny, jakościowy (fenomenologiczny) charakter doświadczeń introspekcyjnych. Wydaje się bowiem, że nie istnieje specyficzna naoczność introspekcyjna, różna od dwóch podstawowych rodzajów naoczności, tzn. od naoczności zmysłowej i naoczności wyobrazeniowej. Analogia między tymi dwoma rodzajami spostrzegania jest także dlatego nie-trafna, że spostrzeganie zmysłowe zawiera w sobie możliwość błędu, a uzyskanie prawdy percepcyjnej można nawet nazwać pewnym sukcesem percepcyjnym. Introspekcja zaś miałaby niejako wbudowaną w siebie gwarancję

sukcesu, gdyż – zgodnie ze stanowiskiem kartezjańskim – jest ona nieomylną władzą poznawczą<sup>6</sup>.

**Introspekcja dotyczy tego, co aktualnie rozgrywa się w naszym umyśle.** Zwracano jednak uwagę, że introspekcja może mieć formę wyłącznie retrospekcji. Gdy jesteśmy zajęci spostrzeganiem jakiegoś obiektu, wówczas nie możemy jednocześnie być zajęci spełnianiem aktu introspekcji.

Zgodnie z epistemologicznym projektem Kartezjusza **niepowątpiewalne dane introspekcyjne miały stanowić podstawę wiedzy o świecie zewnętrznym.** Według tej koncepcji jakieś przekonanie jest uzasadnione wtedy, gdy albo samo jest niepowątpiewalne, albo jest w odpowiedni sposób – najlepiej dedukcyjny – wyprowadzone z przekonań niepowątpiewalnych. Tego rodzaju teorię epistemologiczną nazywa się dzisiaj fundacjonalizmem<sup>7</sup>. W odniesieniu do radykalnej (kartezjańskiej) wersji fundacjonalizmu formułuje się zazwyczaj dwa zasadnicze problemy. Po pierwsze: jak, zaczynając od niepowątpiewalnych danych doświadczenia wewnętrznego, można przejść do przekonań na temat świata zewnętrznego rzeczy fizycznych oraz istnienia innych umysłów? Takie przedsięwzięcie, jak się często zarzuca, wydaje się skazane na immanentyzm i solipsyzm. Po drugie, duża część współczesnych filozofów sądzi, że dane introspekcyjne nie są niepowątpiewalne (niekorygowalne, nieobalalne), a stąd też, że radykalny program fundacjonalizmu kartezjańskiego jest niewykonalny.

Współczesny fundacjonalizm za niepowątpiewalną podstawę ludzkiej wiedzy uznaje zdania dotyczące bezpośrednich rejestracji doświadczenia, które są wypowiedane w trybie subiektywnym (np. „Przeżywam teraz widzenie czerwonej płaszczyzny”). Nie możemy się mylić co do tego, że przeżywamy takie widzenie, ale możemy być w błędzie co do tego, iż ta płaszczyzna jest rzeczywiście czerwona (np. z powodu złego oświetlenia). Przejście od trybu subiektywnego do obiektywnego fundacjonalistów opierają na wnioskowaniu według zasady najlepszego możliwego wyjaśnienia: najlepszym wyjaśnieniem

---

<sup>6</sup> Por. V. D a l m i y a, *Introspection*, w: *A Companion to Epistemology*, eds. J. Dancy, E. Sosa, Oxford: Blackwell 1992, s. 220-221.

<sup>7</sup> Systematyczne rozważania dotyczące fundacjonalizmu, koherencjonizmu i sceptycyzmu zostały zapoczątkowane w Kole Wiedeńskim, por. dyskusję między M. Schlickiem (*Facts and Propositions*, „Analysis” 2[1934/1935], s. 65-70) a C. G. Hemplem (*On the ‘Logical Positivists’ Theory of Truth*, tamże, s. 49-59). Za klasycznych fundacjonalistów z obszaru filozofii analitycznej uchodzą współcześnie C. I. Lewis (*An Analysis of Knowledge and Valuation*, La Salle IL: Open Court 1946) i R. Chisholm (*Theory of Knowledge*, Engelwood Cliffs NJ: Prentice-Hall 1966 [1977<sup>2</sup>, 1989<sup>3</sup>]).

tego, że teraz przeżywam widzenie czerwonej płaszczyzny, jest właśnie to, iż taka płaszczyzna rzeczywiście istnieje. Zakłada się przy tym, że jest to wyjaśnienie lepsze od konkurencyjnych hipotez, a więc od hipotez sceptycznych w stylu Kartezjusza, że podlegamy uniwersalnemu złudzeniu, wytwarzanemu w nas przez złego i potężnego demona. Współczesne wersje fundacjonalizmu albo rezygnują z mocnej interpretacji zdań bazowych, przyznając im jednak status zdań jakoś wyróżnionych epistemicznie, albo uznają, że sposób derywacji zdań pochodnych ze zdań bazowych nie musi być dedukcyjny, lecz może być jakąś formą rozumowania probabilistycznego, albo jednocześnie rezygnują z obu tych postulatów kartezjańskiego fundacjonalizmu.

Jeden z głównych argumentów przeciwko fundacjonalizmowi, a na rzecz koherencjonizmu, został po raz pierwszy wyraźnie sformułowany w artykule Sellarsa<sup>8</sup>. Starał się on podważyć fundacjonalistyczną koncepcję, że nasze przekonania bazowe są uprawomocnione przez bezpośrednie doświadczenie, np. przekonanie, że teraz przeżywam widzenie czerwonej płaszczyzny, jest uprawomocnione przez widzenie czerwonej płaszczyzny w ten sposób, iż samo widzenie nie zawiera już żadnych elementów konceptualnych i przekonaniowych. Według fundacjonalistów bezpośrednie doświadczenie w tym sensie miałyby „gasić” regres uprawomocnienia epistemicznego, gdyż nie ma sensu dalej pytać, za pomocą czego samo widzenie jest uzasadnione. Sellars odpierał ten argument, twierdząc, że bezpośrednie doświadczenie jest albo „surowym czuciem” świata (np. takim, jakie występuje u bardzo małych dzieci), ale wtedy nie może odgrywać żadnej roli w uzasadnianiu przekonań, albo ma ono pełnić funkcję w uzasadnianiu epistemicznym, a wtedy musi zawierać elementy konceptualne i przekonaniowe (np. muszę mieć pojęcia kolorów i muszę być przekonany, że oświetlenie jest wystarczająco dobre, abym mógł rozróżnić kolory itd.). Stąd Sellars wyciągał wniosek, że koncepcja tego, co dane (danych bezpośrednich), jest mitem epistemologii i że każde uzasadnianie epistemiczne musi się odnieść do innych przekonań; musi je założyć jako możliwie spójne tło. Każde uprawomocnienie musi więc być koherencjonistyczne, a nie fundacjonalistyczne. Projekt epistemologii fenomenologicznej sformułowany przez Husserla uznaje się za jedną z głównych XX-wiecznych ilustracji fundacjonalistycznego podejścia do problemu poznania. Według

---

<sup>8</sup> Por. W. S e l l a r s, *Empiricism and the Philosophy of Mind*, w: t e n ż e, *Science, Perception and Reality*, London: Routledge & Kegan Paul 1963, s. 127-196 (*Empiryzm i filozofia umysłu*, w: *Empiryzm współczesny*, red. B. Stanosz, Warszawa: Wydawnictwo UW 1991, s. 173-257).



Husserla w spostrzeganiu zmysłowym przedmiotów fizycznych występuje zawsze splecenie elementu sygnitywnego i elementu naocznego wypełnienia. Pióro wyjęte z szuflady jest wprawdzie dane naocznie, lecz jedynie częściowo, gdyż jego tylna strona jest tylko pusto (sygnitywnie) zasugerowana. Można wprawdzie pióro obrócić, ale wtedy sytuacja się powtórzy. Husserl sądził, że tego rodzaju perspektywizacja spostrzegania zmysłowego dopuszcza zawsze możliwość błędu, a tym samym możliwość wątpienia: nasze sygnitywne domniemania mogą zostać przekreślone przez następne ciągi spostrzeżeń, przedmiot może się okazać zupełnie inny, niż przypuszczaliśmy, lub może wcale nie istnieć. Wątpienie i sceptycyzm są natomiast wykluczone wtedy, gdy zamiast ujmować przedmioty jako realne jednostki istniejące w świecie fizycznym, potraktujemy je wyłącznie jako dane tzw. spostrzegania immanentnego. Widziany przedmiot nie będzie wtedy realną jednostką, lecz „płaską daną” wziętą tylko tak, jak ona się naocznie prezentuje w spostrzeganiu immanentnym. W takiej sytuacji spostrzeganie i to, co w nim spostrzeżone, tworzą – zdaniem Husserla – „niczym niezapśredniczoną jedność”, w której istnienie nie można sensownie wątpić. „W wypadku immanentnie skierowanego, albo krócej mówiąc immanentnego spostrzeżenia (tzw. «wewnętrznego») spostrzeżenie i to, co spostrzeżone, tworzą niczym niezapśredniczoną jedność, jedność jednej jedynej *cogitatio*. Spostrzeżenie kryje tu w sobie swój obiekt w ten sposób, że tylko przez abstrakcję da się go od niego oddzielić, tylko jako coś istotnie niesamodzielnego”<sup>9</sup>. Terminu „spozstrzeganie immanentne” Husserl używał jako terminu technicznego i odróżniał go od spostrzeżenia wewnętrznego, w którym ujmujemy swój strumień przeżyć świadomych z całym bogactwem i głębią różnorodnych apercpcji, tzn. ujęć wykraczających poza „płaskie” dane spostrzeżenia immanentnego. W spostrzeganiu immanentnym widzę nie pióro, lecz plamę barwną o pewnych kształtach. Mogę się wprawdzie mylić co do tego, czy jest to pióro, czy jakiś inny przedmiot, ale nie mogę się mylić, że przeżywam widzenie pewnych barwnych i rozciągłych przestrzennie danych. Według Husserla zadaniem filozoficznej teorii poznania powinna być krytyka rozumu realizowana za pomocą narzędzia, jakim jest spostrzeganie immanentne. W ten sposób badania epistemologiczne byłyby autonomiczne w stosunku do wiedzy naukowej oraz miałyby walor niepowątpiewalności. Przeprowadzający je epistemolog badałby, jak z „płaskich” danych, tzn. z danych pozbawionych sygnitywnych

---

<sup>9</sup> E. H u s s e r l, *Idee czystej fenomenologii i fenomenologicznej filozofii*. Księga pierwsza, tłum. D. Gierulanka, Warszawa: PWN 1975, s. 68.

domniemań, powstają „pełne” przedmioty, jak np. „pełne” pióro powstaje w wyniku syntezy danych naocznych i elementów czysto znakowych albo jak niższe akty zmysłowe podbudowują wyższe teoretycznie konstrukcje myślowe. Ten, wydawałoby się, jeszcze nic nie przesądzający program badawczy zawierał już pewne rozstrzygnięcie dotyczące jednego z głównych problemów metafizycznych i epistemologicznych nowożytności, a mianowicie zagadnienia realizm–idealizm. W polskiej literaturze filozoficznej tezę o niepowątpiewalności danych spostrzegania immanentnego podtrzymywali R. Ingarden<sup>10</sup> i A. B. Stępień<sup>11</sup>. Trudność interpretacyjna polega na tym, że – jak się wydaje – obaj ci autorzy wydają się przypisywać niepowątpiewalność danym tzw. intuicji przeżywania, która, według nich, jest czymś o mniejszym zasięgu epistemicznym niż spostrzeganie immanentne w sensie Husserla<sup>12</sup>.

### 3. INTROSPEKCJA I BEHAWIORYZM

W potocznym przekonaniu behawioryzm kojarzony jest głównie z teorią psychologiczną badającą zachowanie ludzkie. W rzeczywistości „behawioryzm” należy raczej traktować jako hasło wywoławcze dla różnego rodzaju szkół psychologicznych i stanowisk filozoficznych, które powstały w XX w. i przez cały ten wiek podlegały różnym przemianom. W rozważaniach nad introspekcją i świadomością behawioryzm jest ważny dlatego, że przedstawiciele „ruchu behawiorystycznego” głosili zbędność pojęcia świadomości dla opisu i wyjaśniania zachowania ludzkiego. Ze względu na wielość formuł behawioryzmu odróżnia się zazwyczaj **behawioryzm metodologiczny**, nazywany także **naukowym**, od **behawioryzmu filozoficznego**, nazywanego niekiedy **logicznym**. Pierwszy dotyczył statusu metodologicznego psychologii jako nauki, drugi był stanowiskiem filozoficznym odnoszącym się do charakteru zjawisk mentalnych i ich relacji do ciała. Tego, co filozoficzne, i tego, co psychologiczne, w behawioryzmie nie da się jednak ściśle od siebie oddzielić. Behawioryzm metodologiczny w psychologii był inspirowany ideami filozoficznymi, a postulujący go psychologowie wkraczali często na teren

---

<sup>10</sup> *O niebezpieczeństwie ‘petitionis principii’ w teorii poznania*, w: t e n ż e, *U podstaw teorii poznania*, Warszawa: PWN 1971, s. 357-380.

<sup>11</sup> *O metodzie teorii poznania*, Lublin: TN KUL 1966.

<sup>12</sup> Na temat tej dyskusji por. R. Z i e m i ń s k a, *Od tłumacza*, w: R. M. C h i - s h o l m, *Teoria poznania*, tłum. R. Ziemińska, Lublin: Daimonion 1994, s. 193-208.

zasadniczych rozstrzygnięć filozoficznych, ale też, z drugiej strony, rozwój badań w psychologii prowadzonych pod etykietą behawioryzmu wpływał na nowe pomysły teoretyczne w filozofii. Dzisiaj przeważa przekonanie, że behawioryzm jest sprawą minioną. W psychologii został zastąpiony nowym podejściem, tzw. paradygmatem kognitywnym. W filozofii również przestał być uważany za stanowisko warte reprezentowania i obrony. Istnieją jednak autorzy uważający, że wszystko zależy od tego, jak zinterpretujemy słowo „behawioryzm”. Przy pewnej interpretacji można nawet powiedzieć, że behawioryzm jest dzisiaj ciągle żywy, szczególnie na terenie filozofii<sup>13</sup>. Istnieje pogląd, że „behawioryzm” (logiczny) był raczej tylko słowem sugerującym istnienie pewnego odrębnego stanowiska w kwestii natury świadomości, podczas gdy w istocie chodziło tu o pewną formę materializmu (fizykalizmu lub naturalizmu). Ta opinia wydaje się częściowo słuszna i dlatego można rzeczywiście twierdzić, że behawioryzm jest dzisiaj ciągle żywy, gdyż większość współczesnych filozofów w dziedzinie tzw. filozofii umysłu przyznaje się do tej lub innej formy materializmu.

Idee behawiorystyczne zaczęły być formułowane na początku XX w. Jeszcze w 1890 r. James, jeden z twórców nowoczesnej psychologii, pisał:

Psychologia jest nauką o życiu umysłowym, zarówno o jego fenomenach, jak i ich warunkach. Fenomenami są takie rzeczy, jak uczucia, pragnienia, akty poznania, rozumowania, decyzje itp.<sup>14</sup>

Publikacje z początku XX w., autorstwa Jenningsa i Watsona, a także późniejsze prace Hulla, Skinnera i wielu innych, zakwestionowały rozumienie psychologii jako nauki o życiu umysłowym i wysunęły nowy ideał psychologii jako nauki badającej zewnętrzne zachowania się różnych organizmów<sup>15</sup>. Patrząc z dzisiejszego punktu widzenia, można już wyraźnie powiedzieć, **jakie czynniki spowodowały przejście od nauki o życiu mentalnym,**

---

<sup>13</sup> „[...] behaviourism, in one form or another, is still alive and kicking”. A. B y r n e, *Behaviourism*, w: *A Companion to the Philosophy of Mind*, ed. S. Guttenplan, London: Blackwell 1994, s. 132.

<sup>14</sup> W. J a m e s, *Principles of Psychology*, New York: Dover 1890 (1955), s. 1.

<sup>15</sup> H. S. J e n n i n g s, *Behavior of the Lower Organisms*, Bloomington: Indiana University Press 1906 (1976); J. B. W a t s o n, *Psychology as the behaviorist views it*, „Psychological Review”, 20(1913), s. 158-177; C. L. H u l l, *A Behavior System*, New Haven CT: Yale University Press 1952; B. F. S k i n n e r, *The Behavior of Organisms*, New York: Appleton-Century-Crofts 1938; t e n ż e, *Verbal Behavior*, New York: Appleton-Century-Crofts 1957.

**posługującej się przede wszystkim narzędziem introspekcji, do nauki o zachowaniu się.** Behawioryzm w psychologii był motywowany potrzebą znalezienia metodycznego podejścia umożliwiającego weryfikację twierdzeń dotyczących zachowania się zwierząt, a także zachowania się dzieci nie potrafiących jeszcze werbalizować swoich przeżyć, do postaci raportów introspekcyjnych. Tezy na temat świadomości zwierząt, czy to negatywne, czy pozytywne, nie są podatne na weryfikację introspekcyjną. Uznano więc, że w tych sytuacjach tylko obiektywny proces zachowania może być przedmiotem badań naukowych. Tłem koncepcji behawiorystycznych była też Darwinowska teoria ewolucji biologicznej. Twórcy behawioryzmu chcieli zeweryfikować hipotezę Darwina o istnieniu ciągłości między organizmami zwierzęcymi a człowiekiem. Pytanie brzmiało, czy pomiędzy psychologią człowieka a psychologią pozaludzką istnieje przepaść jakościowa, czy też raczej ciągłość. Odpowiedzi na to pytanie, zdaniem behawiorystów, mogło dostarczyć badanie zachowania człowieka i porównanie go z zachowaniem się organizmów pozaludzkich. Innym czynnikiem motywującym behawioryzm było dążenie do obiektywizmu. Psychologię tylko wtedy będzie można uznać za naukę, gdy będzie ona w stanie formułować ogólne prawa dotyczące zachowania się, podobne do praw w innych naukach przyrodniczych. Dążenie do największego możliwego obiektywizmu doprowadziło do krytyki introspekcji jako narzędzia poznawczego psychologii. Dane otrzymywane na podstawie introspekcji uznano za nieobiektywne. Dane te bowiem nie dadzą się dowolnie powtarzać, niemożliwa jest ich intersubiektywna weryfikacja, a samoobserwacja prowadzi do zmian obiektu obserwowanego (uczucie gniewu zmienia się i gaśnie, gdy zaczynamy się mu introspekcyjnie przyglądać). Obiektywność była w tym wypadku rozumiana jako intersubiektywna dostępność badanych zjawisk. W krytyce introspekcji behawioryzm zbieżny był z ideami powstającej na początku XX w. psychoanalizy.

Psychologia behawiorystyczna miała różne fazy, odmiany i interpretacje. Jest także nazywana psychologią uczenia się, gdyż pierwszorzędym przedmiotem jej zainteresowania był proces uczenia się, oparty na doświadczeniach i połączeniach asocjacyjnych pomiędzy nimi. Nazywa się ją też psychologią bodźca i reakcji, psychologią *S – R* (*S – stimulus*, *R – reaction*). Zasadnicze podejście polegało na tym, że za determinanty różnego typu zachowań uznano dające się obserwacyjnie uchwycić oddziaływania środowiskowe, nie zaś żadne czynniki wewnętrzne, np. czynniki psychiczne czy fizjologiczne, pochodzące z ontogenezy lub wrodzone. Zazwyczaj podaje się, że **w behawioryzmie pojawiły się dwie odmienne interpretacje wyjaśniania w psycholo-**

gii<sup>16</sup>. Skrajna interpretacja uważała, że zadaniem psychologii jest wyjaśnianie fenomenalistyczne, tzn. pomiędzy *S* i *R* nie należy wprowadzać żadnych pośredniczących konstrukcji teoretycznych. Psychologia powinna dążyć wyłącznie do formułowania praw indukcyjnych w języku wyłącznie obserwacyjnym (fenomenalistycznym). Aby wyjaśnić zachowanie się jakiegoś organizmu, nie jest potrzebne wprowadzanie jakichkolwiek terminów oznaczających czynniki, których nie da się bezpośrednio obserwować, np. takie jak: pamięć, powiązanie asocjacyjne, stan motywacyjny, uwaga. Druga interpretacja stwierdzała, że tego rodzaju pośredniczące konstrukcje teoretyczne są dopuszczalne i że relacja *S* – *R* powinna mieć postać *S* (zmienna niezależna) – *O* (termin teoretyczny) – *R* (zmienna zależna). Jeśli np. szczur nauczył się naciskać dźwignię w celu otrzymania pożywienia, to skrajny behawioryzm wyjaśniał to przez odniesienie do tzw. prawa efektu (*law of effect*) stwierdzającego, że jeśli reakcja powstaje w obecności bodźca, po którym następuje wzmocnienie, wtedy prawdopodobieństwo reakcji w obecności tego bodźca także wzrasta. Behawioryzm dopuszczający konstrukcje teoretyczne uważał natomiast, że dopuszczalna jest spekulacja na temat nie dających się obserwować mechanizmów fizjologicznych, które mogą występować w procesie uczenia się. Przypuszczano więc, że pożywienie wywołuje pewne stany fizjologiczne, które zmieniają synaptyczne połączenia nerwowe, i tym można wyjaśniać zmieniony wzorzec zachowania się szczura.

Idee behawiorystyczne są nadal żywe w dzisiejszej psychologii. Behawioryzm przyczynił się istotnie do wyostrenia samoświadomości metodologicznej wśród psychologów, spowodował, że zaczęto przywiązywać większą wagę do precyzyjnego określenia warunków eksperymentowania itp. **Generalne wątpliwości w stosunku do behawioryzmu dotyczyły samej teorii uczenia się.** Stwierdzono na przykład, że u pewnych gatunków ptaków występuje czysto pasywne uczenie się, czyli takie, które nie opiera się na schemacie bodziec–reakcja. Ptaki te, zanim zaczną latać, pasywnie obserwują rotację gwiazd na ciemnym niebie i dopiero później wykorzystują tę wiedzę w lataniu. Wyzwaniem dla behawiorystycznego antynatywizmu były też teorie wskazujące, że pewne mechanizmy uczenia się należy uznać za wrodzone, gdyż inaczej nie da się wyjaśnić bezpośrednio dających się obserwować umiejętności. Trzeba np. przyjąć wrodzoność podstawowych struktur obsługujących uczenie się języka, ponieważ inaczej nie można będzie wyjaśnić szyb-

---

<sup>16</sup> Por. K. Z a m i a r a, *Behawioryzm*, w: *Filozofia a nauka. Zarys encyklopedyczny*, red. Z. Cackowski, J. Kmita, K. Szaniawski, Wrocław: PAN 1987, s. 48-57.

kiego nabywania kompetencji językowych przez bardzo małe dzieci. Podejście komputacyjne do umysłu, rozpowszechniające się przynajmniej od lat 70. XX w., także stało w opozycji do behawiorystycznej teorii uczenia się. Źródłem umiejętności oceny położenia w przestrzeni jest kalkulacja szybkości i czasu w trakcie przemieszczania się. Tego rodzaju proces, realizujący się zarówno u ludzi, jak i u zwierząt, jest w swojej naturze raczej komputacyjny niż oparty na kojarzeniu doświadczeń<sup>17</sup>.

**Behawioryzmem logicznym (filozoficznym)** nazywane są stanowiska próbujące uzasadnić tezę o językowej przekładalności terminów opisujących stany mentalne człowieka na terminy opisujące jego zachowanie się. Behawioryzm logiczny w XX w. miewał różne motywacje, ale dominujący wpływ wywarły tu empirystyczne poglądy Koła Wiedeńskiego. Jeden z jego przedstawicieli, C. G. Hempel, tak pisał na ten temat:

Wszystkie należące do psychologii zdania, które mają sens, a więc takie, które są w zasadzie weryfikowalne, są przekładalne na zdania, które nie zawierają pojęć psychologicznych, a tylko pojęcia należące do fizyki<sup>18</sup>.

Stanowisko Hempela opierało się na dwóch założeniach. Po pierwsze, na głoszonej przez Koło Wiedeńskie weryfikacjonistycznej teorii znaczenia, według której znaczenie wyrażen językowych jest określone przez warunki ich weryfikacji. Weryfikacja, o którą chodziło Kołu Wiedeńskiemu, była przeprowadzana za pomocą doświadczenia zmysłowego. Wyrażenia, które nie są bezpośrednio lub pośrednio sprawdzalne zmysłowo, uznano za wyrażenia pozbawione jakiegokolwiek znaczenia kognitywnego. Po drugie, Hempel uznawał, że zewnętrzne zachowanie się osoby jest w zasadniczym stopniu racją przypisywania jej stanów mentalnych. Stąd powstała teza, że zdania odnoszące się do stanów mentalnych danej osoby są w dużej mierze równoznaczne ze zdaniem odnoszącymi się do jej fizycznego zachowania się.

---

<sup>17</sup> Por. C. R. Gallistel, *The Organization of Learning*, Cambridge MA: MIT Press 1990. Mimo tej zmasowanej krytyki behawioryzmu można też spotkać przekonanie, że podejście behawiorystyczne może dobrze dać sobie radę ze wszystkimi podnoszonymi przez krytyków trudnościami, por. E. Wasserman, *Behaviorism*, w: *The MIT Encyclopedia of the Cognitive Sciences*, eds. R. A. Wilson, F. C. Keil, Cambridge MA: The MIT Press, s. 77-80.

<sup>18</sup> C. G. Hempel, *The Logical Analysis of Psychology*, w: *Readings in Philosophical Analysis*, eds. H. Feigl, W. Sellars, New York: Appleton-Century-Crofts 1949, s. 373-384, przedruk w: *Readings in Philosophy of Psychology*, vol. I, ed. N. Block, Cambridge MA: Harvard University Press, s. 18.

Stanowisko Hempla określa się mianem behawiorystycznego, chociaż do warunków weryfikacji zdań mentalistycznych zaliczał on nie tylko bezpośrednio obserwowalne stany zachowaniowe danej osoby, lecz także nie dające się bezpośrednio obserwować zmiany w jej ciele. Zdanie „Kazimierza boli ząb” jest weryfikowalne nie tylko za pomocą odniesienia się do jego zachowania, lecz także przez wzięcie pod uwagę jego procesów trawiennych, ciśnienia krwi, stanu jego systemu nerwowego itp.

W stanowisku Hempla uwidaczniają się motywy behawioryzmu logicznego (filozoficznego): materialistyczne (fizykalistyczne) dążenie do unifikacji słownika naukowego, tzn. do zredukowania go do języka zawierającego wyłącznie terminy fizyczne, oraz dążenie do unaukowania psychologii. **Stanowisko behawioryzmu filozoficznego w czystej postaci można zinterpretować jako tezę, że zdania opisujące stany mentalne dadzą się przełożyć bez utraty znaczenia na zdania dotyczące zewnętrznego zachowania się podmiotu.** W odróżnieniu od stanowiska Hempla, nie mówi się tu o wewnętrznych stanach fizycznych podmiotu, których nie da się bezpośrednio obserwować. Pewne elementy takiego czystego behawioryzmu można znaleźć u Ryle’a. Jest on zazwyczaj traktowany jako główny przedstawiciel behawioryzmu filozoficznego. W swojej książce z 1949 r. *The Concept of Mind*<sup>19</sup> Ryle argumentował przeciwko dualistycznej koncepcji człowieka, którą określał mianem kartezjańskiej idei „ducha w maszynie”. Według Kartezjusza człowiek miałby składać się z dwóch komponentów: fizycznego ciała (maszyny) i нефizycznego umysłu (ducha, ew. duszy). Poglądy Ryle’a były zbieżne z poglądami Wittgensteina z *Dociekań filozoficznych*. Obaj uważali, że dualizm jest stanowiskiem, które musi prowadzić do solipsyzmu. Jeżeli stany mentalne są czymś, co należy do umysłu jako нефizycznej substancji duchowej, to są stanami całkowicie prywatnymi. Gdyby tak było, wtedy nie dałoby się wytłumaczyć, w jaki sposób odnosimy się do nich, używając języka publicznego. Stąd jedynym wyjściem jest przyjęcie, że nie są to stany prywatne, lecz że należy je charakteryzować w terminach dającego się bezpośrednio obserwować zachowania się podmiotu. Jeżeli komuś przypisujemy pewien stan mentalny, to na podstawie jego zachowania się nie wnioskujemy o czymś całkowicie prywatnym i нефizycznym, lecz charakteryzujemy bezpośrednio dane stany jego ciała. Gdy mówimy np., że ktoś jest próżny, to nie chodzi nam o istnienie w nim jakiegoś wewnętrznego stanu, złożonego z ta-

---

<sup>19</sup> G. R y l e, *The Concept of Mind*, Harmondsworth: Penguin 1949 (*Czym jest umysł?*, tłum. W. Marciszewski, Warszawa: PWN 1970).

jenniczego materiału duchowego, lecz o stwierdzenie, że w jego zachowaniu da się wykryć pewien charakterystyczny wzorzec. Tym wzorcem może być np. to, że wykorzystuje on każdą okazję, aby zaprezentować się w jak najlepszym świetle, często podziwia siebie w lustrze itd. Gdy przypisujemy komuś posiadanie stanów mentalnych, nie chodzi nam tylko o stany aktualne, lecz zazwyczaj o dyspozycje do zachowania się w pewien sposób. Jeśli mówimy, że coś jest kruche, to nie mamy na myśli, że ta rzecz aktualnie pęka, lecz że pęknie, jeżeli poddamy ją odpowiedniemu naciskowi. Podobnie, charakteryzując kogoś jako próżnego, odnosimy się nie do jego aktualnego zachowania, lecz do dyspozycji do zachowania się w charakterystyczny sposób w określonych warunkach. Jeżeli ktoś odczuwa depresję, to – według behawioryzmu w tej interpretacji – będzie on wykazywał tendencję do zachowania się w sposób charakterystyczny dla depresji, a w tym również będzie zawarta tendencja do formułowania wypowiedzi typu: „Odczuwam depresję”. Behawioryzm filozoficzny Ryle’a należy więc zinterpretować jako tezę, że **zdania dotyczące stanów mentalnych dadzą się przełożyć bez utraty swego znaczenia na zdania dotyczące aktualnego lub potencjalnego zachowania się podmiotu.**

Prawie oczywisty zarzut w stosunku do behawioryzmu w tej postaci brzmi następująco: zdania odnoszące się do cudzych stanów mentalnych nie dadzą się przełożyć bez utraty znaczenia na zdania dotyczące aktualnego lub potencjalnego zachowania się podmiotu, gdyż **ktoś może symulować stany mentalne, wcale ich w danym momencie nie odczuwając.** Jego zachowanie będzie nieodróżnialne od zachowania się kogoś, kto je rzeczywiście odczuwa. Ryle starał się wówczas argumentować, że w takim wypadku będzie jednak istniała różnica w pewnych szczególnie doniosłych sytuacjach życiowych. Ktoś, kto odczuwa głęboką skruchę z powodu wyrządzenia komuś innemu krzywdy, będzie gotowy do zrekompensowania mu tej krzywdy, ponosząc znaczące ofiary i wysiłek, podczas gdy ktoś, kogo skrucha jest tylko symulacją, nie będzie zachowywał się w ten sposób.

Według behawioryzmu logicznego w duchu Ryle’a musiałyby nie istnieć żadna istotna różnica między sposobem, w jaki stwierdzamy własne stany mentalne, a sposobem, w jaki poznajemy cudze stany mentalne. Przeciwno temu należy jednak stwierdzić, że takie raporty introspekcyjne, jak np. „Jestem głodny”, nie dadzą się skonstruować jako coś, co dotyczy wyłącznie naszego zachowania się, a nie tego, co bezpośrednio odczuwamy. Widać to szczególnie wtedy, gdy zrelacjonujemy je w czasie przeszłym. W takim wypadku nie mogą one być potraktowane jako zachowaniowe manifestacje



naszego własnego stanu mentalnego, lecz muszą zostać potraktowane jako proste raporty o tym, co odczuwaliśmy w przeszłości. Stąd można wnioskować, że behawiorystyczna strategia nie może zasypać przepaści pomiędzy naszą wiedzą o nas samych a wiedzą na temat stanów mentalnych innych podmiotów. Kwestia ta ukazuje też trudność w określeniu samego pojęcia zachowania się.

**Pojęcie „zachowania” dopuszcza dwie interpretacje:** albo „zachowanie” oznacza kierowane intencją i wolą działanie, albo pewien czysto fizyczny ruch w przestrzeni. Coś jest przypadkiem rzetelnego zachowania się wtedy, gdy jest spowodowane przez akt woli. Ryle próbował zarzucić tej koncepcji, że prowadzi do nieskończonego regresu: jeżeli coś jest rzeczywistym działaniem wtedy, gdy jest spowodowane jakimś aktem woli, to przyczyną tego aktu woli musi być następny akt woli. Jeżeli tą przyczyną jest akt woli, to jego przyczyną musi być następny akt woli i tak w nieskończoność. W tak sformułowanym przez Ryle’a zarzucie nieskończonego regresu pozostaje jednak otwarte pytanie, czy akty woli trzeba opisywać jako coś chcianego lub niechcianego w innych, osobnych aktach. Gdy pominiemy kwestię nieskończonego regresu i przyznamy, że coś jest przykładem rzetelnego zachowania się, gdy jest spowodowane przez akt woli, wtedy behawiorizm logiczny nie da się utrzymać. Jeżeli nasze stany mentalne są przyczynami naszego zachowania się, to nie da się ich zidentyfikować z dyspozycjami do zachowania się w określony sposób.

Najbardziej podstawowym zarzutem w stosunku do idei logicznego behawiorizmu jest jednak to, że **behawiorystyczne podejście do mentalności w ogóle nie pozwala mówić o różnych rodzajach stanów mentalnych**. Gdy weźmiemy pod uwagę różnorakie manifestacje behawioralne, wówczas widać, że nie istnieje żaden wzorzec zachowania się wspólny im wszystkim. Nie istnieje żaden powód, aby różnorakie zdarzenia fizyczne traktować jako manifestacje takiego lub innego rodzaju stanu mentalnego. Klasyfikacja czegoś jako dumy, złości czy próżności jest możliwa tylko wtedy, gdy na podstawie wewnętrznego doświadczenia wiemy, czym się różnią od siebie duma, złość czy próżność. Widać więc, dlaczego nie jesteśmy w stanie przeprowadzić analizy stanów mentalnych w terminach zachowaniowych, w terminach dyspozycji do zachowania się w ten, a nie inny sposób. Żeby to uczynić, potrzebujemy świadomości stanów mentalnych, aby móc rozpoznać, czy behawioralne manifestacje należą do tego czy innego stanu. Nie możemy zacząć od pewnego zbioru zachowań i już jakoś uchwycić, że należą one do określonego rodzaju, bez odniesienia się do stanów mentalnych leżących u ich pod-

łoża. Duma, złość czy próżność mogą się manifestować na prawie nieskończenie wiele sposobów fizycznych, które jako takie nie mają ze sobą nic wspólnego. To samo przekonanie może zostać wyrażone w mowie, na piśmie, w wielu językach, alfabetem Morse'a itd.<sup>20</sup>

Istnieje też pewien ogólny argument przeciwko behawiorystycznej koncepcji redukcji języka mentalnego do języka behawioralnego. To, co ktoś robi, lub to, do czego ktoś ma dyspozycję, nie zależy tylko od faktu, że aktualnie ma on pewne określone, partykularne przekonanie. Zależy to też od jego pragnień i innych przekonań. Stąd też nie może być mowy o przeprowadzeniu prostej („atomowej”) analizy behawioralnej, która korelowałaby każde przekonanie ze specyficznym dla niego rodzajem zachowania. Praktycznie rzecz biorąc, jakieś określone przekonanie może prowadzić do każdego rodzaju zachowania się, zależnie od innych przekonań i stanów mentalnych podmiotu<sup>21</sup>.

#### 4. STRUKTURA INTROSPEKCJI

Przedstawione powyżej zarzuty wobec tradycyjnej teorii introspekcji charakteryzują współczesny stan myślenia filozofów na ten temat. Większość z nich krytykuje tradycyjny model, który powszechnie określa się mianem modelu kartezjańskiego. Wydaje się jednak, że część tych zarzutów można odeprzeć za pomocą odpowiedniej interpretacji struktury aktów spostrzegania wewnętrznego. Inne trudności dotyczą szerszych zagadnień, np. problemu zasięgu epistemicznego introspekcji czy problemu nieświadomości, i muszą być rozważane osobno.

Świadomość może mieć charakter prostej, pre-refleksyjnej (pre-introspekcyjnej) świadomości obiektów znajdujących się w polu percepcyjnym, ale także charakter samoświadomości introspekcyjnej (refleksyjnej)<sup>22</sup>. **Prosta**

---

<sup>20</sup> Por. G. M a d e l l, *Behaviourism*, w: *An Encyclopaedia of Philosophy*, ed. G. H. R. Parkinson, London: Routledge 1988, s. 427-440.

<sup>21</sup> Por. P. T. G e a c h, *Mental Acts: Their Content and their Objects*, London: Routledge & Kegan Paul 1957, s. 8.

<sup>22</sup> Niekiedy termin „refleksja” bywa używany zamiennie z terminami „introspekcja”, „spostreżenie wewnętrzne”, „samoobserwacja” na oznaczenie procesu zauważania własnych stanów psychicznych i ich modyfikacji. Istnieje jednak tradycja filozoficzna nie tyle rozumiejąca przez „refleksję” pewną partykularną władzę poznawczą, ile przypisująca jej specyficzne znaczenie. „Refleksja” jest wtedy interpretowana jako zasadniczy rodzaj pozna-

**pre-introspekcyjna świadomość towarzyszy wszystkim aktom, zdarzeniom, stanom i procesom mentalnym.** Spostrzegając zmysłowo jakiś przedmiot lub myśląc jakąś myśl, jesteśmy ich świadomi na sposób prostej świadomości ich obecności. Czym innym wszakże jest samoobserwacja, czyli zwrócenie się w osobnym akcie ujęcia do własnego strumienia przeżyć świadomych pierwszego stopnia. Ten akt zwrócenia się do siebie samego może się modyfikować zależnie od stopnia uwagi poświęconej prostemu przeżywaniu i może też być różnie nasycony intelektualnie. Modyfikacje ze względu na stopień uwagi dotyczą oczywiście też spostrzegania zewnętrznego: mogę mniej lub bardziej uważnie przyglądać się krajobrazom widzianym z okna pociągu. W wypadku introspekcji **modyfikacje uważnościowe** są jednak o wiele bardziej doniosłe. Mogę mniej lub bardziej uważnie kierować się do tego, jak przeżywam widzenie świata przedmiotów, mniej lub bardziej zwracać uwagę na swój sposób doświadczania czegoś, myślenia o czymś. Jeżeli ten stopień uwagi jest duży, to moja świadomość nabiera kształtu przedmiotowego, podczas gdy modyfikacje uważnościowe w ramach percepcji zmysłowej nie mogą aż w takim stopniu wpływać na przedmiotową obecność i wyartykułowanie świata zewnętrznego. Krajobraz widziany z pociągu pozostanie nadal wyar-

---

nia, odsłaniający sposób istnienia i sposób działania ducha ludzkiego, a jednocześnie jako podstawowe narzędzie filozofowania w ogóle. Dialektyka Kanta, Fichtego, Hegla, neokantyzm, a także współczesna filozofia transcendentálna były przedsięwzięciami, które za pomocą refleksyjnej analizy umysłu dążyły do określenia relacji pomiędzy kategoriami bytu a kategoriami myślenia. Charakterystyczna dla tego podejścia jest książka H. Wagnera *Philosophie und Reflexion*, München–Basel: E. Reinhardt 1959. W filozofii angielskojęzycznej po drugiej wojnie światowej upowszechniła się (głównie za sprawą książki P. F. Strawsona *Individuals. An Essay in Descriptive Metaphysics*, London 1959; wyd. pol.: *Indywidualia. Próba metafizyki opisowej*, tłum. B. Chwedeńczuk, Warszawa: IW PAX 1980) pewna wersja tej problematyki nazywana „argumentacją transcendentálną”. Ten sposób „wykorzystania” refleksji w filozofii miał charakter epistemologiczno-ontologiczny (metafizyczny). Charakter egzystencjalno-ontologiczny miały zaś te stanowiska, które w refleksji widziały zasadnicze narzędzie dostępu poznawczego do struktury egzystencjalnej człowieka. Takie podejście można odnaleźć u św. Augustyna, Pascala, Kierkegarda, Nietzschego, w tzw. filozofii życia, aż po różne współczesne odmiany personalizmu, tzw. dialogizmu, egzystencjalizmu i hermeneutyki. Za najbardziej wpływową w tym nurcie należy uznać filozofię M. Heideggera. Dys-tansował się on wprawdzie od transcendentalistycznie rozumianego pojęcia refleksji i od tradycyjnych pojęć ducha i samoświadomości, lecz całość analiz zawartych w *Sein und Zeit* (Halle 1927) jest kulminacją filozofii, która antropologię filozoficzną (a także metafizykę) opiera na refleksyjnej analizie działania ducha ludzkiego (egzystencji). Na temat radykalnej krytyki tzw. refleksyjnego sposobu filozofowania por. W. C h u d y, *Rozwój filozofowania a „pułapka refleksji”*. *Filozofia refleksji i próby jej przewyższenia*, Lublin: RW KUL 1995<sup>2</sup>.

tykułowaną daną przedmiotową, mimo że nie poświęcam mu zbyt wiele uwagi. Modyfikacje uważnościowe sprawiają, że akty introspekcji nie są czymś, co pojawia się skokowo, czymś, co istnieje albo nie istnieje, lecz czymś, co kontynualnie przechodzi od prostej, pre-introspekcyjnej świadomości do postaci wyraźnie wyartykułowanej samoświadomości introspekcyjnej.

Artikulacja i wyraźna obecność danych introspekcyjnych zależne są też od stopnia nasycenia tych danych treściami konceptualnymi. Tak jak zdolność do rozróżniania rozmaitych gatunków win czy rodzajów tytoniu jest zależna od posiadanego doświadczenia w tych dziedzinach, od treningu i od repertuaru pojęć dotyczących win i tytoniu, tak też introspekcja może się rozciągać od bieguna prostego zwrócenia uwagi na swój sposób przeżywania świata aż do bieguna wyrafinowanej pojęciowo umiejętności samoanalizy, do której zdolny jest np. wykształcony psychoterapeuta. Należy więc odróżniać prostą świadomość własnych stanów, która towarzyszy wszystkim tym stanom, od kontynualnie z nią splecionych uważnościowych i intelektualnych odmian procesu samoobserwacji.

Tego rodzaju kontynualne przechodzenie w siebie wzajemnie aktów prostej, pre-introspekcyjnej świadomości i wyraźnie wyartykułowanych aktów introspekcji nie powinno jednak zaciemniać analitycznego faktu, że prosta świadomość jest czymś strukturalnie bardziej podstawowym. Wyartykułowane uważnościowo i intelektualnie akty introspekcji same są uświadamiane dzięki prostej świadomości ich przeżywania. W ten sposób **można odeprzeć zarzut, że introspekcyjna teoria świadomości musi prowadzić do ciągu w nieskończoność**. W myśl tego zarzutu żadna rzecz nie mogłaby się stać świadoma, ponieważ każdy akt introspekcji, aby mógł zostać uświadomiony, domagałby się następnego aktu introspekcji. Akty introspekcji jednak, tak jak wszystkie inne zdarzenia świadome, samoujawniają się w trakcie swojego zachodzenia, są uświadamiane na sposób świadomości prostej, a nie na sposób świadomości refleksyjnej<sup>23</sup>.

---

<sup>23</sup> „Przeżywanie jest immanentnym momentem dokonującego się przeżycia, jest współpłynięciem świadomości ze stanem lub aktem, który jest przez nią ujmowany drugorzędnie, nie jako temat uwagi; przeżycie jest poczuciem zachodzenia bez przedmiotowego dystansowania tego, z czym zaznajamiam, czego jest poczuciem. ‘Przeżywanie’ to nazwa dla tej własności świadomości, że się sama przez swoje zachodzenie ujawnia”. A. B. S t ę p i e ń, *Rodzaje bezpośredniego poznania*, „Roczniki Filozoficzne”, 19(1971), z. 1, s. 106. Por. też W. C h u d y, *Refleksja a poznanie bytu. Refleksja ‘in actu exercito’ i jej funkcja w poznaniu metafizycznym*, Lublin: RW KUL 1984.

Płynna i kontynualna relacja pomiędzy introspekcją a pre-refleksyjną świadomością spontaniczną pozwala też częściowo przeciwstawić się obiekcji, że **introspekcja jest niemożliwa, ponieważ nie jest możliwe jednoczesne uświadamianie sobie czegoś i bycie świadomym tego uświadamiania sobie czegoś**. Według tego zarzutu nie mógłbym np. kierować uwagi na obiekt *O*, a jednocześnie kierować jej na swój sposób spostrzegania obiektu *O*. Podmiot może śledzić z uwagą tylko jeden obiekt w jednej chwili. W ten sposób introspekcja byłaby możliwa tylko jako retrospekcja. Wydaje się jednak, że tego rodzaju krytyka zbyt mocno analitycznie usztywnia samo pojęcie uwagi. „Szerokość” uwagi jest czymś elastycznym i łatwo rozciągającym się od jednego obiektu do drugiego. Tak więc, chociaż w sensie ścisłym nie możemy jednocześnie zauważać jakiegoś obiektu i naszego sposobu zauważania tego obiektu, to jednak luka czasowa jest tutaj tak mała, iż można twierdzić, że introspekcja jest aktem jednoczesnym z aktem pierwszego stopnia. Nie jest ona z pewnością przypomnieniem w sensie odrębnego aktu przywołania własnej przeszłości, chociaż oczywiście dzięki przypomnieniu możemy odnosić się do własnych minionych stanów świadomych pierwszego stopnia i do aktów introspekcji, których niegdyś dokonaliśmy. Zarzut (podnoszony już przez A. Comte’a), że introspekcja nie jest możliwa, ponieważ życie świadome człowieka nie może jednocześnie rozpaść się na dwa strumienie, należy więc uznać za nietrafny, gdyż nie jest to literalne rozejście się aktywności pierwszego i drugiego stopnia, lecz ciągle „falowanie” modyfikacji uważnościowych<sup>24</sup>.

---

<sup>24</sup> „Z jednej strony, zachodzenie takiego aktu spostrzegania wewnętrznego wymagałoby od obserwatora możliwości obejmowania uwagą dwóch rzeczy naraz. Postanawiałby on na przykład wstać wcześniej rano, a jednocześnie obserwowałby swój akt postanowienia [...]. Nie jest to może zarzut druzgocący logicznie, ponieważ można by argumentować, że niektórzy ludzie potrafią, może po pewnej praktyce, połączyć uwagę w prowadzeniu samochodu z uwagą poświęconą rozmowie [...]. A jednak, nawet gdy się głosi, że dokonując introspekcji obejmuje się uwagą jednocześnie dwie rzeczy, trzeba się zgodzić, że istnieje jakaś granica dla liczby możliwych jednocześnie aktów uwagi, z tego zaś wynika, że muszą być jakieś procesy umysłowe niedostępne introspekcji, te mianowicie jej akty, które zawierają w sobie największą liczbę równoczesnych aktów uwagi. Przed zwolennikami owej teorii stanęłoby wówczas pytanie, na jakiej drodze można wykryć zachodzenie takich aktów; bo skoro nie na drodze introspekcyjnej, to wynikałoby stąd, że wiedza człowieka o procesach zachodzących we własnym umyśle nie zawsze może opierać się na introspekcji” (R y l e, *Czym jest umysł?*, s. 269-270). Ten klasyczny już fragment rozważań Ryle’a, które zmierzały do wykazania niemożliwości introspekcyjnej samoświadomości, dobrze ilustruje powiązanie między kwestią ‘ciągu w nieskończoność’ aktów introspekcji a problemem możliwości jednoczesnego spostrzegania dwóch obiektów w ramach doświadczenia wewnętrznego.

Bardzo często podnoszona jest trudność, że introspekcja nie może być rzetelnym źródłem przekonań i wiedzy na temat tzw. świata wewnętrznego, ponieważ w trakcie jej dokonywania zmieniają się własności prezentowanych przez nią przedmiotów. Gdy np. zaczynam obserwować własną złość, wówczas bardzo często ona przygasa, zmienia swój charakter, a niekiedy nawet znika całkowicie. Zarzut ten można jednak uznać za pewną przesadę analizy. Zmiany w wewnętrznie obserwowanym przedmiocie są zależne od stopnia uwagi, z jaką dokonuję samoobserwacji, i dopiero nasycona intelektualnie samoanaliza może prowadzić do istotnego zakłócenia spostrzeganego wewnątrznie fenomenu. Inaczej mówiąc, **możliwe jest takie introspekcyjne współzauważanie, które tylko w małym stopniu, jeśli w ogóle, zmienia dane introspekcyjne.**

Unikalność pewnych własności aktów introspekcji nie pozwala na prostą analogię z aktami spostrzegania zewnętrznego (zmysłowego). Rzeczywiście, w introspekcji nie występuje osobny rodzaj naoczności, gdyż introspekcyjnie możemy zarówno uświadamiać sobie elementy, które są wyposażone w charakterystyki jakościowe, np. wyobrażenia, jak i analizować jakościową strukturę naszego doświadczenia zmysłowego, możemy również introspekcyjnie uświadomić sobie własne „abstrakcyjne myśli”, a więc struktury pozbawione wszelkich jakości naocznościowych. Tego rodzaju **naocznościowa nieokreśloność introspekcji** nie powinna jednak skłaniać do poglądu, że jeśli introspekcja jest czymś rzeczywistym, to musi być czymś innym, a więc do wniosku, że świadomość introspekcyjna może być czymś iluzorycznym w stosunku do „prawdziwej rzeczywistości”, którą stanowią procesy neuronalne w mózgu. Można (jak F. Brentano) twierdzić przeciwnie, że spostrzeganie wewnętrzne (introspekcja) jest jedynym spostrzeżeniem we właściwym sensie, gdyż to, co się w nim prezentuje, prezentuje się absolutnie bezpośrednio.

W introspekcji jakby rzeczywiście nie miało miejsca, naturalne w percepcji zewnętrznej, **zdwojenie na zjawisko i rzeczywistość**. Dane introspekcyjne bowiem nie przejawiają się przez różnorokie wyglądy (odcienie), tak jak np. przejawia się góra lodowa w słońcu, gdy płyniemy dookoła niej. Z tej obojętności introspekcji wobec dualności „zjawisko–rzeczywistość” nie musi jednak wynikać esencjalna nieomyślność introspekcji, a więc wniosek, że dane introspekcyjne nigdy nie mogą wprowadzać w błąd podmiotu, który je otrzy-

---

Ryle ma rację, że same akty introspekcyjnego zauważania muszą ostatecznie być zauważane nie-introspekcyjnie, lecz z tego nie wynika, że introspekcja na pewnym poziomie nie jest możliwa jako (prawie) jednoczesne zauważanie dwóch obiektów.

muje, podczas gdy percepcja zewnętrzna byłaby czymś z istoty skazanym na błąd ze względu na swój zjawiskowy charakter. Wydaje się, że w dyskusji dotyczącej nieomylności introspekcji nie są wyraźnie odróżniane dwie kwestie. Po pierwsze, generalne sceptyczne przypuszczenie, że introspekcja, podobnie jak cała świadomość, może być przejawem czegoś całkowicie innego, np. procesów mózgowych czy nieświadomych procesów mentalnych. Wtedy oczywiście nie tylko dane introspekcyjne, lecz także cała świadomość będzie swoistego rodzaju złudzeniem dla posiadających ją istot. Czymś jednak różnym od tego przypuszczenia jest pytanie, w jakiej mierze poszczególne dane introspekcyjne mogą być błędne. Pytania te prowadzą do problematyki epistemicznego zasięgu introspekcji.

#### 5. ZASIĘG EPISTEMICZNY INTROSPEKCJI

Sławna formuła Kartezjusza „myślę, więc jestem” stała się powodem do mówienia o kartezjańskim modelu świadomości introspekcyjnej i o jej nieomylności. Pojęcie myślenia było przez Kartezjusza rozumiane bardzo szeroko. „Myślenie” obejmowało właściwie wszystkie odmiany aktów świadomych, a więc oprócz aktów potocznie uznawanych za myślenie, np. abstrakcyjnego wnioskowania z jednych treści o innych, myśleniem były także akty percepcji zmysłowej, akty przypominania sobie czegoś, akty wyobrażania czegoś itd. Nie mogę się mylić co do tego, że myślę, gdyż wątplenie jest jedną z odmian myślenia w tym szerokim sensie. Przekładając to na język teorii introspekcji: nie mogę się mylić co do tego, że przeżywam widzenie światła – przeżywam w sensie danej introspekcyjnej – chociaż mogę się mylić co do tego, że rzeczywiście jest przede mną jakieś światło. Podobnie jest ze wszystkimi innymi świadomymi odniesieniami: nie mogę się mylić co do tego, że przeżywam pewne doświadczenia (tryb subiektywny), chociaż mogę się mylić co do tego, czy one rzeczywiście prezentują to, co mi się wydaje, że prezentują (tryb obiektywny). Introspekcja dostarcza zatem danych niepowątpiewalnych. Czy rzeczywiście tak jest?

Przeciw tezie o nieomylności introspekcji wytacza się czasami bardzo ogólny argument, że ból i (introspekcyjna) świadomość bólu są dwiema różnymi jednostkami. Jeżeli są takimi dwiema różnymi jednostkami, to **zachodzi przynajmniej logiczna możliwość, że ból jest obecny, a nie ma świadomości bólu**. Inaczej mówiąc, że ulegamy swoistej halucynacji, iż odczuwamy ból, gdy w rzeczywistości nic nas nie boli. Logiczna możliwość oznacza tu

niesprzeczność, a więc nie byłoby czymś sprzecznym logicznie pomyślenie, że może istnieć świadomość bólu bez bólu<sup>25</sup>. Pomyśleć, w sensie czystej możliwości, że ból i świadomość bólu są dwiema różnymi jednostkami, można – ale czy rzeczywiście tak jest? Wydaje się, iż raczej jest tak, że ból i świadomość bólu stanowią „niczym nie zapośredniczoną jedność”. Świadome stany mentalne refleksyjnie reprezentują same siebie, a więc ból jako zdarzenie mentalne, przez samo swoje zachodzenie, reprezentuje sam siebie. Ta sytuacja zachodzi także dla introspekcji rozumianej jako osobny akt zwrócenia się do własnego strumienia świadomości, a więc introspekcyjna świadomość bólu i ból też tworzą „niczym nie zapośredniczoną jedność”. Aby podważyć nieomyślność tego rodzaju danych świadomych, trzeba byłoby się odwołać do czegoś więcej niż tylko do logicznej możliwości pomyślenia sobie świadomości bólu bez bólu. Potrzebne byłyby empiryczne przykłady sytuacji bólu bez (rzeczywistego) bólu.

Podaje się niekiedy fikcyjne przybliżenia takich sytuacji, w których byłaby możliwa świadomość bólu bez (rzeczywistego) bólu. Możemy na przykład wyobrazić sobie, że wiedza na temat sposobu funkcjonowania ludzkiego mózgu posunęła się aż tak daleko, że jesteśmy w stanie dokładnie wskazać, jakie zdarzenia neuronalne odpowiadają określonym zdarzeniom świadomym. W takiej sytuacji lekarz mógłby powiedzieć pacjentowi, iż się myli, gdy twierdzi, że odczuwa ból, gdyż obserwując aktualnie jego mózg, wie, że nie zachodzą w nim zdarzenia, które normalnie wywołują ból. Albo nieco mniej skomplikowany przykład: eksperymentator mówi, że przyłoży do szyi osoby badanej gorący przedmiot, tak że osoba ta nie będzie widziała tego przedmiotu, a potem w rzeczywistości przykłada kawałek lodu; wtedy większości badanych wydaje się, że odczuwają ból, gdy naprawdę powinni odczuwać tylko chłodne dotknięcie jakiegoś przedmiotu. Myliliby się więc co do tego, że ich boli. Wydaje się, że wobec tych przykładów nasza intuicja się waha.

---

<sup>25</sup> „Jasne jest, że jeżeli istnieje nie podlegająca korekcie wiedza o naszych własnych stanach umysłowych, to nie może odnosić się do przeszłości, a tylko do teraźniejszości. Jeżeli bowiem stan umysłowy umieścimy w przeszłości, choćby ułamek sekundy wstecz, to błąd będzie logicznie możliwy. Dwa zdarzenia, które występują w różnym czasie, muszą być «odrębnymi bytami»: zawsze jest logicznie możliwe, że jedno zdarzenie wystąpi, a drugie nie. Jeżeli myślę, że przed ułamkiem sekundy czułem ból, to mimo krótkiego odstępu czasu myśl i ból muszą być «odrębnymi bytami». A więc musi istnieć logiczna możliwość, że nie było żadnego bólu, który odpowiadałby mojej myśli” (D. M. A r m s t r o n g, *A Materialist Theory of Mind*. London: Routledge & Kegan Paul 1968; wyd. pol.: *Materialistyczna teoria umysłu*, tłum. H. Kraheńska, Warszawa: PWN 1982, s. 147).



Można wprawdzie odpowiedzieć, że w jednym i w drugim wypadku odczuwam to, co odczuwam, i że co do tego nie mogę się mylić. Mylić mógł się natomiast ten, kto na podstawie badań empirycznych błędnie skorelował określone zdarzenia w mózgu z określonymi zdarzeniami w świadomości i uznał, że taka korelacja obowiązuje u wszystkich ludzi. W wypadku zaś przykładania lodu, mogę się mylić co do tego, jak nazwać moje doznanie, lecz nie co do tego, co faktycznie odczuwam. Z drugiej jednak strony powstaje pytanie: **czy możemy dysponować wiedzą całkowicie niezależną od pojęć, których używamy?** Czy faktycznie można odczuwać chłodne dotknięcie, a pomylić się tylko co do sposobu nazwania tego doświadczenia?

Powyższa dyskusja i podane przykłady dotyczą kwestii nieomyślności danych introspekcyjnych na poziomie prostej świadomości fenomenalnej, a więc na poziomie takich zdarzeń świadomych, jak doznawanie bólu, widzenie czerwieni pewnej płaszczyzny, słyszenie jakiś dźwięków itd. Większość jednak zarzutów w stosunku do introspekcji koncentruje się współcześnie na danych introspekcyjnych odnoszących się do bardziej skomplikowanych stanów rzeczy. Twierdzi się, że teza o nieomyślności introspekcji jest tezą bardzo radykalną, ponieważ we wszystkich dziedzinach empirycznych mechanizmy, które w nich funkcjonują, mogą przestać funkcjonować poprawnie. **Jeżeli „mentalny mechanizm”, który obsługuje dane introspekcyjne, byłby całkowicie niezawodny, to graniczyłoby to z cudem.** Introspekcja, tak jak inne mechanizmy kognitywne człowieka, musi być także niekompletna i zawodna.

Tezę o zawodności introspekcji uzasadnia się przykładami należącymi do różnych kategorii. Są to albo pewne dane eksperymentalne, albo też pewne ogólniejsze hipotezy teoretyczne. Wskazuje się np., że **samoobserwacja podatna jest na konfabulację.** Identyfikując introspekcyjnie własne przekonania i postawy, ludzie mówią zazwyczaj o takich przekonaniach i postawach, które uważają za racjonalne lub za statystycznie normalne, nawet wtedy, gdy ich w rzeczywistości nie mają. Przykładem może być sytuacja przypadkowego świadka wypadku. Większość osób zeznaje, że ich gotowość do niesienia pomocy ofiarom wypadku nie jest zależna od liczby innych osób będących świadkami wypadku. Badania wykazują jednak, że gotowość do niesienia pomocy maleje wraz z rosnącą liczbą osób towarzyszących danemu zdarzeniu<sup>26</sup>.

---

<sup>26</sup> Por. B. L a n t a n é, J. D a r l e y, *The Unconscious Bystander: Why Doesn't He Help Us?*, New York: Appleton-Century-Crofts 1970.

Raporty introspekcyjne są nie tylko wystawione na możliwość konfabulacji, lecz także uwarunkowane teoriami kulturowymi, które stosujemy do siebie i do innych. **Raporty introspekcyjne są także zależne od stopnia werbalnego wyartykułowania informacji introspekcyjnych.** Zgodnie z wynikami pewnych eksperymentów należy odróżnić raporty dokonywane w czasie wypełniania jakiegoś głównego zadania od wszystkich innych raportów, które są retrospekcyjne. Wśród raportów retrospekcyjnych największą poprawność otrzymujemy wtedy, gdy opierają się one na pamięci krótkoterminowej. W tych sytuacjach informacja werbalna jest ciągle obecna i wymaga tylko przywołania w trakcie wykonywania głównego zadania. Raporty odnoszące się do pamięci długoterminowej wymagają albo ponownej artykulacji werbalnej – np. w postaci przekształcenia posiadanych informacji z informacji wizualnych w informacje werbalne – albo odpowiedniej filtracji i znalezienia nowych informacji, na które normalnie nie zwracamy uwagi, gdyż wymagają np. pamięciowego odszukania racji pewnych dokonanych już wyborów lub motywów zapamiętanego zachowania się w jakiejś dawniej zaistniałej sytuacji. Widziana w tym świetle introspekcja byłaby w zasadzie zawsze retrospekcją i jej niezawodność malałaby wraz z oddaleniem czasowym oraz wynikającą z niego koniecznością odpowiedniego przetwarzania informacji<sup>27</sup>.

Najbardziej chyba znaną strategią metodologiczną zmierzającą do podważenia wiarygodności i kompletności danych introspekcyjnych są **teorie psychoanalityczne**. Teorie te postulują konieczność przyjęcia nieświadomych stanów mentalnych w celu wypełnienia luki eksplanacyjnej pomiędzy publicznie dającym się obserwować zachowaniem danej osoby a dostarczonymi przez nią zeznaniami introspekcyjnymi. Według tego podejścia takie fenomeny, jak pomyłki językowe, dziwne sny, pewne skojarzenia itd., dadzą się wyjaśnić tylko wtedy, gdy przyjmie się istnienie nieświadomych introspekcyjnie stanów mentalnych. Generalny zarzut wobec psychoanalitycznego podważenia wartości introspekcji polega na zwróceniu uwagi, że tego typu wyjaśnienia zawsze będą otwarte na inne, konkurencyjne wyjaśnienia, które nie muszą przyjmować istnienia czegoś tak wątpliwego, jak nieświadome i introspekcyjnie nieuświadomialne stany mentalne<sup>28</sup>.

---

<sup>27</sup> Por. K. A. E r i c s s o n, H. A. S i m o n, *Protocol Analysis: Verbal Reports as Data*, Cambridge MA: MIT Press 1993<sup>2</sup>.

<sup>28</sup> Por. A. G r ü n b a u m, *The Foundations of Psychoanalysis: A Philosophical Critique*, Berkeley CA: University of California Press 1984. Najbardziej radykalną krytyką teorii psychoanalitycznych jest twierdzenie Poppera, że psychoanaliza (podobnie jak marksizm) nie jest teorią naukową, gdyż nie jest podatna na żaden rodzaj falsyfikacji. Według

Odmianą grupę obiekcji w stosunku do tezy o nieomyślności i pełności danych introspekcyjnych tworzą dane eksperymentalne mające dowodzić, że **ludzie są w stanie przyjmować informacje poza progiem świadomości introspekcyjnej**, a potem świadomie z nich korzystać w działaniu. Z przykładem takiej sytuacji mielibyśmy do czynienia wtedy, gdy osobom poddawanym eksperymentowi podaje się w postaci niewyraźnego szumu do jednego ucha sugestie odpowiedzi na pytania podawane do drugiego ucha. Liczba poprawnych odpowiedzi wzrasta znacząco w stosunku do osób, które słyszały tylko pytania<sup>29</sup>. Innym przykładem tego rodzaju jest sławny fenomen tzw. ślepego widzenia (*blindsight*). Pacjenci z mechanicznym uszkodzeniem dróg nerwowych, które łączą siatkówkę oka z częściami mózgu odpowiedzialnymi za widzenie, mają częściowo całkowicie nieczynny pewien fragment pola spostrzeżeniowego, a jednak są w stanie poprawnie opisać własności przedmiotów umieszczonych w tym fragmencie pola. Na tej podstawie przypuszcza się, że istnieje introspekcyjnie nieświadome przetwarzanie bodźców wzrokowych, które dopiero potem, w efekcie odpowiednich pytań zadawanych przez eksperymentatora, uzyskują postać werbalną. Można powiedzieć, że osoby te mówią o tym, czego (świadomie) nie widzą<sup>30</sup>.

Jeden z najbardziej przekonujących zarzutów w stosunku do pełności i zasięgu epistemicznego introspekcji pochodzi z lingwistyki teoretycznej. Według teorii sformułowanej przez N. Chomsky'ego proces nabywania języka przez dzieci polega na tworzeniu i testowaniu hipotez na temat jego struktury, a możliwość posługiwania się językiem wymaga, żeby jego reguły były

---

Poppera pewna teoria jest teorią naukową tylko wtedy, gdy istnieją jakieś możliwe, choć niekoniecznie aktualne doświadczenia, które mogłyby ją sfalsyfikować. Jeżeli „teoria” jest zgodna z każdymi możliwymi danymi obserwacyjnymi, to taka „teoria” nie jest nauką, lecz pseudonauką (K. R. P o p p e r, *Conjectures and Refutations*, New York: Basic Books 1963). Dzisiaj wskazuje się raczej, że chociaż nieświadome motywacje nie są dostępne bezpośrednio, to jednak teorie psychoanalityczne mogą podlegać falsyfikacji. Przeprowadzono różne badania, w wyniku których bogate treściowo hipotezy Freuda – trzy fazy seksualności, powstanie superego w wyniku przezwyciężenia kompleksu Edypa czy seksualne podstawy represji nieświadomych motywacji – znalazły niewielkie, jeśli w ogóle, potwierdzenie eksperymentalne (por. H. J. E y s e n c k, G. D. W i l s o n, *The Experimental Study of Freudian Theories*, London: Methuen 1973). Istnieje też przypuszczenie, że pozytywne efekty terapii psychoanalitycznych mogą być spowodowane efektem *placebo*, a konkretnie w tym wypadku zrozumieniem i wsparciem ze strony terapeuty.

<sup>29</sup> J. L a c k n e r, M. G a r r e t t, *Resolving Ambiguity: Effects of Biasing Context in the Unattended Ear*, „Cognition”, 1(1972), s. 359-372.

<sup>30</sup> Por. L. W e i s k r a n t z, *Blindsight: A Case Study and Implications*, Oxford: Oxford University Press 1986.

jakoś reprezentowane naszym umyśle. **Ani jednak hipotezy strukturalne dotyczące języka, formułowane w trakcie jego uczenia się, ani też reguły potrzebne do posługiwania się nim nie są dostępne dla introspekcji.** Gdy mówimy jakimś językiem, wówczas nie zdajemy sobie sprawy z jego reguł ani też zwykłe introspekcyjne „zajrzenie” w stany naszej świadomości nie pozwoli nam ich odkryć. Mimo to nasz umysł musi je jakoś reprezentować i przetwarzać<sup>31</sup>.

Przedstawione powyżej przykłady miały ukazać, że dane introspekcyjne są przesiąknięte treściami teoretycznymi, elementami konceptualnymi, i że jako takie mogą podlegać warunkowaniu, które albo nie jest aktualnie uświadomione, albo też w ogóle nie da się introspekcyjnie uświadomić. W tym aspekcie przewaga argumentacyjna znajduje się po stronie tych, którzy podkreślają ułomność i niepełność zaawansowanych teoretycznie raportów na temat własnych stanów mentalnych. W przypadku zaś raportów dotyczących stosunkowo prostych danych fenomenalnych, teza o wysokim stopniu niezawodności introspekcji wydaje się bardziej uzasadniona. Niestety, ten wysoki stopień niezawodności jest okupiony małą zawartością informacyjną zeznań introspekcyjnych na poziomie świadomości fenomenalnej. Niekiedy jednak dowodzenie, że dane introspekcyjne są w podobny sposób zawodne, jak dane ekstraspekcyjne, jest łączone z wyraźną sugestią uniwersalnego sceptycyzmu w stosunku do danych introspekcji. Gdyby ten sceptycyzm trzeba było potraktować jako jedyne możliwe stanowisko, wtedy wszelkie rozpoznawanie motywów, racji działania, jak również wszelkie decyzje, należałoby uznać za czynności całkowicie pozorne. Taki jednak uniwersalny sceptycyzm w stosunku do danych introspekcji jest często współcześnie tylko wstępem do teorii całkowicie nie-sceptycznych, a mianowicie do behawioryzmu, fizykalizmu czy materializmu. W argumentacjach generalnie podważających wartość informacyjną introspekcji istotną rolę odgrywa pojęcie nieświadomości. Jeśli całe nasze życie świadome byłoby wyłącznie pewnego rodzaju epifenomenem w stosunku do głębokiego warunkowania ze strony stanów nieświadomych i z istoty swej nieuświadomialnych, to rzeczywiście trzeba byłoby powiedzieć, że nie wiemy, co myślimy i kim jesteśmy.

---

<sup>31</sup> Por. N. C h o m s k y, *Knowledge of Language*, New York: Praeger 1986.

## 6. INTROSPEKCJA I NIEŚWIADOMOŚĆ

Mimo że pojęcie nieświadomości należy do repertuaru pojęć przekazywanych na poziomie ogólnego wykształcenia i funkcjonuje w języku potocznym, to jednak bardzo trudno powiedzieć, co właściwie ono znaczy<sup>32</sup>. W każdej chwili naszego życia jesteśmy każdorazowo świadomi tylko bardzo wąskiego wycinka naszej mentalności. Większość zakresu naszej mentalności wymaga ciągłego pamięciowego przywoływania. W tym sensie wszystko, co jest przedmiotem przypomnienia, należałoby zaliczyć do sfery nieświadomości. Jednakże wyrażenie „nieświadomość” nie jest używane na określenie tego, co możemy sobie przypomnieć. To, co potencjalnie leży w zasięgu możliwości przypomnienia, nie jest tym, co nieświadome w tym sensie, który jest zazwyczaj łączony z poglądami psychoanalizy wywodzącej się ze stanowiska Freuda. **Pamięciowa nieświadomość**, jeżeli można użyć takiego wyrażenia, jest dla nas z zasady dostępna, podczas gdy to, co nieświadome w sensie Freuda, jest czymś, co z zasady nie jest dostępne dla przypomnienia w zwykłym rozumieniu tego słowa.

---

<sup>32</sup> Rozróżnienie między procesami umysłowych świadomymi a nieświadomymi sięga przynajmniej Leibnizjańskiej koncepcji *petites perceptions*, czyli tych stanów monad, które same nie są zauważane, ale których współdziałanie i natężenie konstytuuje świadomość. W formowaniu się rozróżnienia między umysłem świadomym a nieświadomym istotną rolę odegrał Kant. Cała kantowska „maszyna” transcendentna funkcjonowała nieświadomie, tzn. aplikacja form naoczności i kategorii do materiału wrażeniowego odbywała się bez udziału świadomego umysłu. Dopiero ten nieświadomy proces doprowadzał do powstania introspekcyjnie dostępnej świadomości empirycznej. W XIX w. niedoceniany „filozof nieświadomości”, E. von Hartmann, skonstruował syntezę Heglowskiej teorii absolutu, teorii Schopenhauera, że absolut jest wolą, Schellinga interpretacji absolutu jako bytu nieświadomego, Leibnizjańskiego indywidualizmu i pluralizmu oraz realizmu teoriopoznawczego współczesnego sobie przyrodoznawstwa. Wpłynęli na niego również I. H. Fichte, G. Th. Fechner i R. H. Lotze. Głównym instrumentem tej syntezy było pojęcie nieświadomości. Hartmann rozszerzył pojęcie psychiczności w ten sposób, że obejmowało ono również nieświadome procesy psychiczne. Odróżniał to, co jest relatywnie nieświadome w złożonym indywiduum, czyli to, co jest nieświadome dla centralnej świadomości (a co może być świadome dla podporządkowanej świadomości niższego stopnia), od tego, co jest absolutnie nieświadome. Absolutna nieświadomość nie indywidualizuje się i jest równoważna duchowi absolutnemu. Fenomeny psychiczne są pasywne, natomiast czynności psychiczne są aktywne, ale zawsze nieświadome. Gdy dochodzi do kolizji czynności psychicznych, wtedy powstają fenomeny świadome. Por. główne dzieło E. von Hartmanna *Philosophie des Unbewußten. Versuch einer Weltanschauung*, Berlin: W. Friedrich 1869. Przez metafizykę nieświadomości Hartmanna i teorie Freuda dyskusja nad ideą nieświadomego umysłu doprowadziła do często dzisiaj używanej dystynkcji między nieświadomym umysłem komputacyjnym a świadomym umysłem fenomenologicznym.

Do nieświadomości nie zaliczamy też całej wiedzy habitualnej, tzn. nie mówimy, że nasze umiejętności pływania, chodzenia, siadania czy pisania należą do dziedziny tego, co nieświadome, mimo że w trakcie ich dokonywania nie uświadamiamy sobie reguł, według których dokonujemy tych czynności. Tego rodzaju **latentna wiedza habitualna w zasadzie nie jest zaliczana do sfery nieświadomej**. Jest tak dlatego, że tego typu sprawności nie polegają na posiadaniu stanów mentalnych wyposażonych w treści intencjonalne. W trakcie pływania można sobie uświadamiać reguły tej czynności, lecz jest to tylko „kognitywny naddatek”, gdyż cały proces odbywa się bez świadomości rządzących nim reguł. W oczywisty sposób nie ma też sensu powiedzenie, że przedmioty pozbawione całkowicie świadomości działają według nie uświadomionych reguł czy praw. Planety nie poruszają się po swoich orbitach nieświadomie, postępując według reguł czy praw ruchu fizycznego. Inaczej mówiąc, **pojęcie nieświadomości może się stosować tylko do istot posiadających świadomość**. Jaką zatem treść należy przypisać pojęciu nieświadomości, skoro nie chodzi w nim ani o treści mentalne potencjalnie dostępne w przypomnieniu, ani o latentną wiedzę habitualną, ani też o reguły czy prawa, według których funkcjonują przedmioty całkowicie pozbawione świadomości?

Wydaje się, że pojęcie nieświadomości, spopularyzowane w kulturowym słownictwie XX w. przez teorie psychoanalityczne, ma przynajmniej dwie cechy istotne. Po pierwsze – **to, co nieświadome, nie jest dostępne świadomości na normalnej drodze introspekcyjnej** ani też na normalnej drodze przypomnienia (drodze retrospekcyjnej). Po drugie – **to, co nieświadome, w istotny sposób warunkuje (lub może warunkować) to, do czego mamy normalny dostęp introspekcyjny bądź retrospekcyjny**. Sam sposób tego warunkowania nie jest również dostępny ani introspekcyjnie, ani retrospekcyjnie. Teorie nieświadomości różnicują się w zależności od uznania zakresu tego rodzaju warunkowania. Uznanie, że warunkowanie tego, co świadome i uświadamialne introspekcyjnie, przez to, co nieświadome i introspekcyjnie nieuświadamialne, ma stosunkowo mały zakres i niewielką doniosłość dla świadomego życia podmiotu, nie niesie ze sobą żadnych poważniejszych konsekwencji. Teorie nieświadomego warunkowania stają się interesujące wtedy, gdy zaczynają twierdzić, że całe nasze życie świadome jest zależne od stanów nieświadomych, które nie mogą być introspekcyjnie uświadomione. To, co na poziomie świadomym pojawia się jako mające taki a taki sens, może być więc wytworem stanów i procesów o całkowicie odmiennym charakterze. Szczególną radykalizacją tego stanowiska jest teza, że również

nasze „ja” jest produktem procesów nieświadomych. W efekcie „ja” (jako ośrodek życia świadomego oraz jako instancja kontrolująca zachowanie) staje się produktem ubocznym (epifenomenem) głębokich procesów nieświadomych. Gdy mówimy o „nieświadomości”, nie chodzi więc tylko o to, że mogą istnieć w naszym umyśle jakieś zdarzenia czy procesy niedostępne nam na normalnej drodze introspekcyjnej, lecz że są to zdarzenia i procesy, które istotnie determinują nasze zachowanie, aż do sytuacji, kiedy przestaje mieć sens mówienie, że jestem autorem swoich czynów, że moje działanie i myślenie podlega kontroli z mojej strony (ze strony „mojego ja”). Można, nieco drastycznie, powiedzieć, że jeżeli świadome siebie „ja” jest tylko epifenomenem jakichś nieświadomych procesów, to nigdy nikt z nas nie jest sobą.

I te wyjaśnienia jednak nie rozświetlają wszystkiego na temat pojęcia nieświadomości, którym się tak powszechnie posługujemy. Czym są domniemane stany i procesy nieświadome? Nie są tymi stanami i procesami, które uświadamiamy sobie aktualnie, ani tymi wszystkimi, które, wykorzystując pamięć, jesteśmy w stanie sobie w zasadzie uświadomić na normalnej drodze introspekcyjnej. Procesy nieświadome nie byłyby dostępne za pomocą introspekcji, a jednocześnie miałyby spełniać funkcję warunkującą w stosunku do procesów świadomych. Określenia te jednak nie mówią, czym są stany i procesy nieświadome, wskazują tylko na sposób dostępu do nich i na ich rolę w stosunku do stanów i procesów świadomych. Albo procesy nieświadome są procesami czysto fizycznymi (fizjologicznymi) i wtedy nie ma sensu mówienie o nieświadomych stanach i procesach mentalnych, albo stany i procesy nieświadome są jednak w jakiś sposób mentalne, chociaż brakuje im własności bycia świadomymi, ewentualnie własności bycia potencjalnie uświadamialnymi na normalnej drodze introspekcyjnej<sup>33</sup>. W pierwszym wypadku, tzn. gdy pojęcie nieświadomości oznacza nieświadome stany i procesy mózgowo, traci sens mówienie o nieświadomości jako o dziedzinie odrębnej

---

<sup>33</sup> Analizując stanowisko J. Findlaya (filozofa znającego dobrze zarówno analityczną, jak i fenomenologiczną tradycję filozofii umysłu), E. S. Casey stwierdza: „The critical question, then is whether there is any genuinely unconscious mental activity, that is activity which is temporarily or permanently outside conscious ken and yet still psychical in nature [...]. The question remains, however, whether there are mental processes which are unconscious in the more rigorous sense of being altogether inaccessible or admissible to consciousness” (*Findley's Philosophy of Mind*, w: *Studies in the Philosophy of J. N. Findlay*, eds. R. S. Cohen, R. M. Martin, M. Westphal, Albany: State University of New York Press 1985, s. 128). „To fall back on the body or the brain as the exclusive explanatory basis of unconscious activity is thus to adopt a tactic of *obscurum per obscurius*” (tamże, s. 131).

od dziedziny czysto fizycznej. Nie ma więc sensu sugestia, występująca w pojęciu nieświadomości, że stany i procesy nieświadome zawierają w sobie pewne określone treści intencjonalne, które mogą wpływać na aktualnie uświadamiane treści intencjonalne. Mózg, tak jak jest dany w spostrzeganiu zmysłowym, nie zawiera w sobie żadnych treści intencjonalnych. Przy takiej interpretacji nieświadomości trzeba powiedzieć, że psychoanalityk nie wpływa terapeutycznie na zorganizowaną treściowo w pewien sposób, intencjonalną nieświadomość pacjenta, lecz w rzeczywistości stara się wywrzeć wpływ na jego mózg. Gdyby pojęcie nieświadomości oznaczało fizyczne stany mózgu, wtedy psychoanalizę i podobne do niej teorie należałoby uznać za wprawdzie dość heroiczny, lecz nieudolny wstęp do prawdziwej medycyny nieświadomości. Stałaby się nią w przyszłości tzw. neuronauka (*neuroscience*), która wykorzystując pełną wiedzę na temat biologicznej i chemicznej działalności mózgu, byłaby tym samym pełną wiedzą na temat świadomości i nieświadomości. Tak zwane choroby psychiczne można byłoby leczyć w ten sposób, w jaki dzisiaj leczymy zapalenie płuc<sup>34</sup>.

Czy jednak ma sens mówienie o nieświadomych stanach mentalnych, niedostępnych na normalnej drodze introspekcyjnej, które mimo to są wyposażone w określone treści intencjonalne? Tego rodzaju pojęcie nieświadomości, rozpatrywane z punktu widzenia materialistycznych teorii umysłu, wydaje się całkowicie niekonsekwentne, gdyż implikuje **dualizm dwóch rodzajów sfer: z jednej strony sfery intencjonalnych stanów nieświadomych, z drugiej zaś sfery czysto fizycznych zdarzeń mózgowych**. Dualizm nie był z pewnością stanowiskiem twórcy psychoanalizy. Freud uważał, że wszystkie stany mentalne, zarówno świadome, jak i nieświadome, są ostatecznie stanami mózgu. Można zatem twierdzić, że w ogóle pojęcie nieświadomości, jako pojęcie czegoś odrębnego od sfery fizycznych zdarzeń mózgowych, jest niespójne. Albo coś jest aktualnie świadome i potencjalnie uświadamialne,

---

<sup>34</sup> „Materializm eliminacyjny jest tezą, że nasza potoczna koncepcja fenomenów psychicznych stanowi radykalnie fałszywą teorię, teorię tak fundamentalnie złą, że zarówno jej zasady, jak i jej ontologia zostaną przez w pełni rozwiniętą neuronaukę (*neuroscience*) ostatecznie raczej odrzucone niż łagodnie zredukowane. Nasze wzajemne porozumiewanie się, a nawet nasza introspekcja, mogłyby wtedy zostać ponownie ukonstytuowane w obrębie pojęciowych ram tej pełnej neuronauki – czyli teorii, która, jak należy oczekiwać, będzie bardziej zasobna eksplanacyjnie (*powerful*) niż zastępowana przez nią nasza potoczna psychologia i, ogólnie biorąc, bardziej istotnie zintegrowana z naukami fizykalnymi”. P. M. Churchland, *Neurocomputational Perspective. The Nature of Mind and the Structure of Science*, Cambridge MA: The MIT Press 1989, s. 1.



a wtedy jest mentalne, albo coś jest nieświadome i nieuświadomiane, a wtedy nie może być mentalne (intencjonalne, wyposażone w treści). Mózg można porównać do płyty kompaktowej, która aktualnie nie jest odtwarzana. Na podstawie zawartego w niej zapisu można odtworzyć koncert albo też wygenerować pokaz świetlny lub inny rodzaj zdarzenia. Zgodnie z tą analogią nieświadome stany fizyczne mózgu nie mogą być więc stanami zawierającymi określone treści intencjonalne, gdyż dopiero gdy staną się świadome, nabywają treści intencjonalnych. Nieświadomy mózg, podobnie jak płyta kompaktowa, umożliwi dowolne inne efekty niż efekty intencjonalne<sup>35</sup>.

Jednakże siła eksplanacyjna teorii Freuda, hipotezy Chomsky'ego, siła przekonywająca przykładów dotyczących tzw. ślepego widzenia oraz innych interpretacji nieświadomych stanów mentalnych polega na tym, że są to hipotezy dobrze wyjaśniające wrażliwe na treść zachowanie się podmiotu. **Do czasu, gdy się pokaże, że tego rodzaju fenomeny mogą zostać wyjaśnione inaczej niż przez odwoływanie się do nieświadomych stanów mentalnych, pojęcie nieświadomości ciągle będzie odgrywać ważną rolę.** Zwolennicy pojęcia nieświadomości mogą więc powiedzieć, że różnica pomiędzy mózgiem a płytą kompaktową polega na tym, że własności muzyczne aktualnie nie odgrywanej płyty kompaktowej pozostają kauzalnie nieczynne, natomiast własności nieświadomego mózgu są kauzalnie aktywne w trakcie dokonywania czynności świadomych. Jest to bardzo istotna różnica dla tych, którzy argumentują na rzecz istnienia nieświadomych stanów mentalnych.

Teoria nieświadomości staje się teorią doniosłą antropologicznie wtedy, gdy przybiera formę tezy, że również nasze „ja” jest wytworem tego, co nieświadome – obojętnie, czy będzie to nieświadomość o cechach stanów i procesów intencjonalnych, czy nieświadomość w sensie „czystych” procesów mózgowych. Wtedy nie tylko introspekcję można porównać do światła księżyca, ślizgającego się po powierzchni głębokiego i nieprzezroczystego stawu, w którym pływają tajemnicze ryby (a więc stany intencjonalne, lecz nieświadome), lecz także podmiot dokonujący aktów introspekcji jest tylko cieniem i swoistego rodzaju złudzeniem dla siebie samego. Przy założeniu

---

<sup>35</sup> Strawson twierdzi, że „ściśle mówiąc, nie istnieją żadne nie-dyspozycjonalne i nie-doświadczeniowe zjawiska mentalne”, i porównuje mózg, który niczego nie doświadcza, do płyty CD, gdzie nagrany jest kwartet Beethovena (aktualnie nie odtwarzany). Zapis fizyczny na płycie mógłby zostać wykorzystany albo do wytworzenia jakiegoś rzeczywistego zdarzenia akustycznego, albo do wygenerowania pokazu świetlnego (por. G. S t r a w s o n, *Mental Reality*, Cambridge MA: MIT Press 1994, s. 167).

mocnego i całkowitego warunkowania tego, co świadome, przez to, co nieświadome, podmiot dokonujący introspekcji staje się wytworem głębokich stanów nieświadomych, które – jak należałoby chyba powiedzieć we współczesnym języku – próbują same siebie skanować. Podmiot (ja) nie jest wtedy żadnym autonomicznym ośrodkiem myślenia i działania. Gdy jednak uznamy, że nasze „ja” jest wytworem nieświadomych stanów i procesów, wtedy trzeba powiedzieć, że nie wiemy, ani o czym myślimy, ani kim jesteśmy. Jakikolwiek przypisanie komukolwiek odpowiedzialności za akty myślenia i działania traci sens, a doświadczenie posiadania wolnej woli jest tylko iluzją. Problem nieświadomości nie jest więc tylko pewnym wąskim zagadnieniem z zakresu dyskursu na temat metodologicznych podstaw psychologii. Dyskusja dotycząca pojęcia nieświadomości musi ostatecznie uwzględnić szeroką problematykę relacji między ciałem (mózgiem) a świadomym umysłem w ogóle<sup>36</sup>.

---

<sup>36</sup> Searle przytacza powiedzenie Lashleya: „No activity of mind is ever conscious”, podkreślając, że nie było ono domniemane dosłownie, lecz chodziło w nim raczej o to, że procesy, za pomocą których wywoływane są różne cechy świadomości, same nie są świadome (por. K. L a s h l e y, *Cerebral Organization and Behaviour*, w: *The Brain and Human Behaviour*, eds. H. Solomon, S. Cobb, W. Penfield, Baltimore: Williams and Wilkins Press 1956). Dystynkcję „umysł fenomenologiczny–umysł komputacyjny” miał, według Searle’a, wprowadzić J. Jackendoff w książce *Consciousness and the Computational Mind* (Cambridge MA: The MIT Press 1987). Gdy np. weźmiemy pod uwagę ilość informacji, które muszą zostać przetworzone przez nasz mózg, abyśmy mogli trafić kulą śnieżną w przejeżdżający samochód, wtedy koncepcja kalkulującego nasze ruchy i kontrolującego nasze działania, nieświadomego umysłu komputacyjnego może się stać bardzo przekonująca. Może pojawić się myśl, że świadomość jest tylko małym wycinkiem gigantycznego procesu umysłowego, który dokonuje się nieświadomie i który ma charakter neuronalnego przetwarzania informacji. W myśl interpretacji, która dominuje dzisiaj w tzw. naukach kognitywnych (*cognitive science*), przetwarzanie to ma charakter komputacyjny, co m.in. oznacza, iż odbywa się ono według reguł, które nigdy i z zasady nie mogą zostać uświadomione. Ostatnio koncepcję tę próbował podważyć Searle, formułując kryterium oddzielenia „rzetelnych” stanów umysłowych od stanów neuronalnych mózgu: „[...] pojęcie umysłowego stanu nieświadomego rozumiemy tylko jako pojęcie możliwej treści świadomości, tylko jako ten rodzaj rzeczy, która chociaż nieświadoma, i być może niemożliwa do uświadomienia z wielu różnych powodów, jest jednak czymś, co może (lub mogłoby) być uświadomione. Tę właśnie ideę, że nieświadome stany intencjonalne są w zasadzie dostępne dla świadomości, nazywam zasadą powiązania (*connection principle*)”. J. R. S e a r l e, *The Rediscovery of the Mind*, Cambridge MA: The MIT Press 1992, s. 155-156 (wyd. pol.: *Umysł na nowo odkryty*, tłum. T. Baszniak, Warszawa: PIW 1999). Tego rodzaju obrona pewnej samodzielności tego, co mentalne, w stosunku do tego, co nieświadome, jest jednak stawiana w dwuznacznym świetle przez samego Searle’a, a to ze względu na głoszony przez niego biologiczny naturalizm, uznający mózg za jedyną podstawę świadomości.

Mogłoby się wydawać, że dobrym wyjściem z tych trudności jest argumentacja na rzecz dualizmu. W myśl teorii dualistycznych nasze „ja” dysponuje pewną autonomią w stosunku do tego, co fizyczne i nieświadome, dlatego że jest aktualizacją pewnej potencjalności naszej duszy, tzn. potencjalności poza-fizycznego substratu stanów mentalnych. Dualizm – czy to w wersji kartezjańskiej, czy w sensie hylemorfizmu arystotelesowskiego – może się na pierwszy rzut oka wydać adekwatną odpowiedzią na materialistyczne pojęcie nieświadomości. Nawet jednak filozofowie dalecy od materializmu, np. św. Tomasz z Akwinu, podkreślali, że naszej duszy (substancji duchowej, którą jesteśmy) nie poznamy co do jej istoty, lecz tylko poprzez jej akty. Istota substancji duchowej, którą jesteśmy, nie jest nam, przynajmniej całkowicie, dostępna. Jako taka jest więc ona nieświadoma i może wywierać warunkujący wpływ na nasze myślenie i działanie. Pojęcie duchowej nieświadomości było od dawna obecne w różnych teoriach twórczości artystycznej i w interpretacjach dotyczących źródeł geniuszu twórczego. Nawet więc jeśli przyjmemy, że nasze „ja” jest czymś „zakotwiczonym” w poza-fizycznej (duchowej) substancji, to nadal pozostaniemy z problemem nieświadomego warunkowania ze strony tej substancji, które może pozbawić samodzielności nasze świadome „ja”. Inaczej mówiąc, pojęcie nieświadomości zagraża pojęciu autonomii świadomego siebie podmiotu niezależnie od tego, czy nieświadomość będzie zinterpretowana materialnie czy pozamaterialnie.

Poczucie autonomii naszego „ja”, którego doświadczamy introspekcyjnie, jest czymś bardzo swoistym, czymś, co nie daje się wyprowadzić nawet z pojęcia duchowości (niematerialności). Zauważenie, że autonomia naszego „ja” nie jest możliwa do zrozumienia ani materialistycznie, ani spirytualistycznie, nie powinno jednak skłaniać do wniosku, że jest ona pochodna, że jest produktem ubocznym jakichś nieświadomych stanów i procesów materialnych lub duchowych. Tego rodzaju nieeksplikowalność poczucia autonomii powinna raczej skłonić do myślenia, że „ja” dokonujące aktów introspekcji jest czymś całkowicie unikalnym metafizycznie, czymś, co „wynosi się” ponad wszelkie stany i procesy materialne, jak też ponad wszelkie nieświadome stany i procesy duchowe. To „ja” jest czymś, co może się zdystansować od wszystkich tych stanów i procesów, może poddać je osądowi i odpowiednio do tego osądu działać.

## BIBLIOGRAFIA

- A r m s t r o n g D. M., *A Materialist Theory of Mind*, London: Routledge & Kegan Paul 1968 (Materialistyczna teoria umysłu, tłum. H. Krahelska, Warszawa: PWN 1982).
- A r y s t o t e l e s, *O duszy*, tłum. P. Siwek, Warszawa: PWN 1972.
- A r y s t o t e l e s, *Metafizyka*, tłum. K. Leśniak, Warszawa: PWN 1983.
- A u g u s t y n, *De Trinitate*, ed. W. J. Mountain, F. Glorie, Turnhout: Brepols 1968.
- B y r n e A., *Behaviourism*, w: *A Companion to the Philosophy of Mind*, ed. S. Guttenplan, London: Blackwell 1994, s. 132-140.
- C a s e y E. S., *Findley's Philosophy of Mind*, w: *Studies in the Philosophy of J. N. Findlay*, eds. R. S. Cohen, R. M. Martin, M. Westphal, Albany: State University of New York Press 1985, s. 114-138.
- C h i s h o l m R., *Theory of Knowledge*, Engelwood Cliffs NJ: Prentice-Hall 1966 (1977<sup>2</sup>, 1989<sup>3</sup>).
- C h o m s k y N., *A Review of B. F. Skinner's Verbal Behavior*, „*Language*” 35(1959), nr 1, przedruk: *Readings in Philosophy of Psychology*, ed. N. Block, vol. I, London: Methuen 1980, s. 48-63.
- C h o m s k y N., *Knowledge of Language*, New York: Praeger 1986.
- C h u d y W., *Refleksja a poznanie bytu. Refleksja 'in actu exercito' i jej funkcja w poznaniu metafizycznym*, Lublin: RW KUL 1984.
- C h u d y W., *Rozwój filozofowania a „pułapka refleksji”*. *Filozofia refleksji i próby jej przewyciężenia*, Lublin: RW KUL 1995<sup>2</sup>.
- C h u r c h l a n d P. M., *Neurocomputational Perspective. The Nature of Mind and the Structure of Science*, Cambridge MA: The MIT Press 1989.
- D a l m i y a V., *Introspection*, w: *A Companion to Epistemology*, eds. J. Dancy, E. Sosa, Oxford: Blackwell 1992, s. 218-222.
- E r i c s s o n K. A., S i m o n H. A., *Protocol Analysis: Verbal Reports as Data*, Cambridge MA: MIT Press 1993<sup>2</sup>.
- E y s e n c k H. J., W i l s o n G. D., *The Experimental Study of Freudian Theories*, London: Methuen 1973.
- G a l l i s t e l C. R., *The Organization of Learning*, Cambridge MA: MIT Press 1990.
- G e a c h P. T., *Mental Acts: Their Content and their Objects*, London: Routledge & Kegan Paul 1957.
- G r ü n b a u m A., *The Foundations of Psychoanalysis: A Philosophical Critique*, Berkeley CA: University of California Press 1984.
- H a r t m a n n E. von, *Philosophie des Unbewußten. Versuch einer Weltanschauung*, Berlin: W. Friedrich 1869.
- H e i d e g g e r M., *Sein und Zeit*, Halle 1927.
- H e m p e l C. G., *On the 'Logical Positivists' Theory of Truth*, „*Analysis*”, 2(1934/1935), s. 49-59.

- H e m p e l C. G., The Logical Analysis of Psychology, w: Readings in Philosophical Analysis, eds. H. Feigl, W. Sellars, New York: Appleton-Century-Crofts 1949, s. 373-384; przedruk: Readings in Philosophy of Psychology, ed. N. Block, vol. I, Cambridge MA: Harvard University Press, s. 14-23.
- H u l l C. L., A Behavior System, New Haven CT: Yale University Press 1952.
- H u s s e r l E., Idee czystej fenomenologii i fenomenologicznej filozofii. Księga pierwsza, tłum. D. Gierulanka, Warszawa: PWN 1975.
- I n g a r d e n R., O niebezpieczeństwie 'petitionis principii' w teorii poznania, w: t e n ż e, U podstaw teorii poznania, Warszawa: PWN 1971, s. 357-380.
- J a c k e n d o f f J., Consciousness and the Computational Mind, Cambridge MA: The MIT Press 1987.
- J a m e s W., Principles of Psychology, New York: Dover 1890 (1955).
- J e n n i n g s H. S., Behavior of the Lower Organisms, Bloomington: Indiana University Press 1906 (1976).
- J u d y c k i S., Kantowska teoria umysłu, „Roczniki Filozoficzne”, 45(1997), z. 1, s. 35-60.
- K a n t I., Krytyka czystego rozumu, tłum. R. Ingarden, Warszawa: PWN 1957.
- K r e u t z M., Metody współczesnej psychologii, Warszawa: PWN 1962.
- L a c k n e r J., G a r r e t t M., Resolving Ambiguity: Effects of Biasing Context in the Unattended Ear, „Cognition”, 1 (1972), s. 359-372.
- L a n t a n é B., D a r l e y J., The Unconscious Bystander: Why Doesn't He Help Us?, New York: Appleton-Century-Crofts 1970.
- L a s h l e y K., Cerebral Organisation and Behaviour, w: The Brain and Human Behaviour, eds. H. Solomon, S. Cobb, W. Penfield, Baltimore: Williams and Wilkins Press 1956.
- L e w i s C. I., An Analysis of Knowledge and Valuation, La Salle IL: Open Court 1946.
- M a d e l l G., Behaviourism, w: An Encyclopaedia of Philosophy, ed. G. H. R. Parkinson, London: Routledge 1988, s. 427-440.
- P o p p e r K. R., Conjectures and Refutations, New York: Basic Books 1963.
- R y l e G., The Concept of Mind, Harmondsworth: Penguin 1949 (Czym jest umysł?, tłum. W. Marciszewski, Warszawa: PWN 1970).
- S c h l i c k M., Facts and Propositions, „Analysis”, 2(1934/1935), s. 65-70.
- S e a r l e J. R., The Rediscovery of the Mind, Cambridge MA: The MIT Press 1992 (Umysł na nowo odkryty, tłum. T. Baszniak, Warszawa: PIW 1999).
- S e l l a r s W., Empiricism and the Philosophy of Mind, w: t e n ż e, Science, Perception and Reality, London: Routledge & Kegan Paul 1963, s. 127-196 (Empiryzm i filozofia umysłu, w: Empiryzm współczesny, red. B. Stanosz, Warszawa: Wydawnictwo UW 1991, s. 173-257).
- S k i n n e r B. F., The Behavior of Organisms, New York: Appleton-Century-Crofts 1938.
- S k i n n e r B. F., Verbal Behavior, New York: Appleton-Century-Crofts 1957.
- S t r a w s o n P. F., Individuals. An Essay in Descriptive Metaphysics, London 1959 (Indywidualność. Próba metafizyki opisowej, tłum. B. Chwedeńczuk, Warszawa: IW PAX 1980).

- S t r a w s o n G., *Mental Reality*, Cambridge MA: MIT Press 1994.
- S t ę p i e ń A. B., *O metodzie teorii poznania*, Lublin: TN KUL 1966.
- S t ę p i e ń A. B., *Rodzaje bezpośredniego poznania*, „Roczniki Filozoficzne”, 19(1971), z. 1, s. 95-126.
- T o m a s z z A k w i n u, *Quaestiones disputatae de veritate*, w: *Opera omnia, iussu Leonis XIII edita cura et studio Fratrum Praedicatorum*, Tomus XXII, Roma 1970.
- W a g n e r H., *Philosophie und Reflexion*, München–Basel: E. Reinhardt 1959.
- W a s s e r m a n E., *Behaviorism*, w: *The MIT Encyclopedia of the Cognitive Sciences*, eds. R. A. Wilson, F. C. Keil, Cambridge MA: The MIT Press 1999, s. 77-80.
- W a t s o n J. B., *Psychology as the behaviorist views it*, „*Psychological Review*”, 20(1913), s. 158-177.
- W e i s k r a n t z L., *Blindsight: A Case Study and Implications*, Oxford: Oxford University Press 1986.
- Z a m i a r a K., *Behawioryzm*, w: *Filozofia a nauka. Zarys encyklopedyczny*, red. Z. Cackowski, J. Kmita, K. Szaniawski, Wrocław: PAN 1987, s. 48-57.
- Z i e m i ń s k a R., *Od tłumacza*, w: R. M. C h i s h o l m, *Teoria poznania*, tłum. R. Ziemińska, Lublin: Daimonion 1994, s. 193-208.

#### INTROSPECTION AS A PHILOSOPHICAL PROBLEM

##### S u m m a r y

In its first half the article presents the main assumptions of the traditional (so-called Cartesian) theory of introspection. Among others (1) privileged access, (2) undoubtfulness of introspective data, (3) privacy of introspective data, (4) proposition that introspection is a separate act of consciousness, (5) proposition that introspection is a kind of inner perception, (6) proposition that introspection is concerned with what is going on in the consciousness at the moment, (7) proposition that introspective data may be the basis for knowledge about the outer world (foundationalism) belong to them. In the part concerned with behaviourism the main assumptions of scientific (methodological) behaviourism and of philosophical (logical) behaviourism are presented. The author accepts the main point of the criticism directed at the project of G. Ryle's philosophical behaviourism. He maintains the charge that the behaviourist approach to mentality does not allow one to talk about various kinds of mental states. In the part concerning the structure of introspection stress is laid on the role of attention. With the help of the idea of modification of attention the author tries to dismiss the charge that performing a definite act and at the same time introspectively defining it is impossible. With the use of the idea of attention he also shows how it is possible to oppose **the objection of continuing acts of introspection into infinity** and the charge that introspection modifies its object. Also the issues of infallibility and of completeness of introspective data in the light of certain experimental data and theoretical hypotheses (e.g. blindsight, psychoanalysis, N. Chomsky's theory of the language) are discussed. The last part concerns the problem of

the relation between the ideas of introspection and unconsciousness. 'Unconsciousness' in the strong sense means a certain kind of conditioning to which human mind may be subjected. This conditioning is not only actually unconscious, but by its nature is unrealisable in the normal introspective way. Various ideas of unconsciousness are analysed: unconsciousness as what can be called to mind, unconscious habitual knowledge, unconsciousness in the psychoanalytical interpretation, unconsciousness as conditioning by the so-called **computative mind**. The author shows that the autonomy of the 'I' given in introspection is something peculiar and as such cannot be reduced to any kind of unconscious conditioning.

*Translated by Tadeusz Karłowicz*

**Słowa kluczowe:** epistemologia, psychologia, introspekcja, behawioryzm, nieświadomość.

**Key words:** epistemology, psychology, introspection, behaviourism, unconsciousness.