

Katarzyna Paprzycka

Monizm anomalny Davidsona a problem epifenomenalizmu w świetle Idealizacyjnej Koncepcji Nauki*

1. Pomiędzy problemem psychofizycznym a redukcjonizmem

Filozoficzne dyskusje nad umysłem ograniczone są niejako z dwóch stron. Z jednej strony, chodzi o to, by podkreślić fakt, iż to, co umysłowe, jest w pewien zasadniczy sposób różne od tego, co fizyczne. Największe wrażenie na Kartezjuszu wywiera fakt, iż nasza wiedza o tym, co myślimy, odczuwamy etc., a więc nasza wiedza o treści naszego umysłu w danej chwili, jest pewna – nie mogą jej podważyć nawet najcięższe argumenty sceptyków. Wyjaśnia on ten osobliwy fakt, postulując istnienie dwóch zupełnie odrębnych bytów: psychicznego (pozaoprzestrzennego i niematerialnego) i fizycznego (przestrzennego i materialnego). W ten sposób też wpada w sidła problemu psychofizycznego, bo jest zupełnie przecież niejasne, jak coś, co jest poza przestrzenią, może wpływać (przyczynowo-skutkowo) na coś, co jest rozciągnięte w przestrzeni.

Receptą na rozwiązanie problemu psychofizycznego jest monizm, czy to neutralny, czy spirytualistyczny, czy wreszcie – współcześnie najbardziej popularny – materialistyczny. Zarzucając ideę rozziwu pomiędzy dwiema zupełnie odrębnymi sferami bytu, kładzie podwaliny pod zrozumienie, jak to, co psychiczne, może współdziałać z tym, co fizyczne. Odpowiedź jest prosta: może, bo jest ontologicznie jednorodne.

W tym momencie docieramy do ograniczenia z drugiej strony. Jedną z najbardziej naturalnych form monizmu jest tak zwana teoria identyczności, wedle której relacja tożsamości zachodzi nie tylko pomiędzy zdarzeniami psychicznymi

* Wcześniejsza wersja tego tekstu została zreferowana na zebraniu Polskiego Towarzystwa Filozoficznego w Poznaniu. Pragnę podziękować wszystkim uczestnikom za dyskusję, a szczególnie serdecznie Profesorowi Krzysztofowi Łastowskiemu. Chciałabym również podziękować Profesorowi L. Nowakowi oraz Profesorowi J. Hołowce za pomocne wskazówki i sugestie.

i fizycznymi, ale również pomiędzy ich typami (rodzajami)¹. Jest więc konsekwencją teorii identyczności, że pojęcia dojrzałej nauki psychologicznej będzie można zdefiniować w oparciu o pojęcia dojrzałej fizyki, dając w ten sposób podstawy do redukcji praw psychologii do praw fizyki². Kryjąca się tu wizja jedności nauki ma tę wadę, że zdaje się przeczyć intuicjom wyjściowym dotyczącym natury umysłu jako czegoś swoistego. Stąd też, m.in., bierze się notorycznie powracające zagadnienia natury świadomości (Nagel 1974; Jackson 1986; Chalmers 1996).

Pomiędzy tymi dwoma krańcami mieści się monizm nieredukcyjny³. Zaproponowanych zostało wiele konkretnych propozycji od wczesnego funkcjonalizmu (Putnam 1967/1975⁴) do być może najlepiej znanej propozycji monizmu anomalnego, przedstawionej przez Donalda Davidsona (1970/1992). Davidson próbuje utrzymać zarówno intuicję monistyczną (a nawet fizykalistyczną, dając priorytet prawom fizyki, co kwestionuje np. McDowell 1985), jak i intuicję antyredukcjonistyczną, twierdząc nie tylko, że prawa psychologiczne są niesprowadzalne do praw fizyki, lecz również, że prawa psychologii w ścisłym tego słowa znaczeniu nie istnieją. Sugeruje on, że myśli te można pogodzić, jeśli zamiast wymogu utożsamienia rodzajów (typów) zdarzeń fizycznych i mentalnych wprowadzi się wymóg utożsamienia tylko jednostek (egzemplarzy) zdarzeń fizycznych i mentalnych⁵. Propozycja teorii tożsamości jednostkowej (*token-token identity theory*) spotkała się z mieszanym odbiorem. Z jednej strony stanowi podstawę dla antyredukcjonizmu. Z drugiej jednak strony propozycja Davidsona spotkała się z zarzutem epi-fenomenalizmu rodzajowego (Stoutland 1976, 1985; Honderich 1982; McLaughlin 1989; Kim 1993a).

Jest jasne choćby z powyższego zarysu, że kwestia, o której mowa, znajduje się na pograniczu filozofii umysłu, metafizyki i filozofii nauki. Większość toczonych dyskusji na ten temat pozostawała w obrębie pierwszych dwóch. Celem tego artykułu jest zarysowanie, jak dyskusję tę można byłoby wzbogacić, uzupełniając ją o pewne propozycje filozofii nauki. Zamierzam w szczególności pokazać, że przyjęcie perspektywy Idealizacyjnej Koncepcji Nauki (Nowak 1971, 1977, 1980;

¹ Mówi się niekiedy, iż typy (tudzież rodzaje) zdarzeń psychicznych są tożsame z typami (rodzajami) zdarzeń fizycznych. W niniejszym tekście ograniczymy się do omawiania tezy, że tożsamości te zachodzą pomiędzy predykatami dojrzałej teorii psychologicznej i fizycznej.

² Oczywiście „dać podstawy” nie znaczy „zagwarantować”. Nadto przywołanie fizyki w tym kontekście jest wielkim uproszczeniem. Można by żywić nadzieję, że psychologia zostanie zredukowana do neurofizjologii; z nadzieją, że zostanie zredukowana do fizyki, należałoby w każdym razie poczekać.

³ Ograniczam się dalej do rozważania monizmu materialistycznego.

⁴ Ścisły związek funkcjonalizmu i antyredukcjonizmu podważają m.in. Lewis (1972) i Jackson *et al.* (1982).

⁵ Dalej przyjmujemy za W. Strawieńskim (1997) konwencję mówienia o jednostkach i rodzajach (nie o egzemplarzach i typach) jako odpowiedników rozróżnienia między *token* a *type*. Powinno tu być jasne jednakże, że za Davidsonem nie przypisujemy rodzajom pozytywnego statusu ontycznego.

dalej będę używać skrótu IKN) pozwala z jednej strony na zrozumienie tego, na czym mogłaby polegać nieredukowalność jednej teorii do drugiej, a z drugiej strony na uniknięcie zarzutu epifenomenalizmu rodzajowego.

Należy też zaznaczyć, że pomimo tego, iż propozycja niżej przedstawiona jest sformułowana w terminach IKN, to nie jest ona do niej ograniczona. Rozróżnienie, które okaże się być zasadnicze, a mianowicie rozróżnienie między teoriami istotnościowo zgodnymi i niezgodnymi, można powielić na gruncie wszelkich koncepcji uwzględniających zróżnicowanie istotnościowe czynników. Takimi koncepcjami są np. teoria Nancy Cartwright (1983; 1989), lecz również – choć po jakimś uzupełnieniu – koncepcja Wesleya Salmona (1984) oraz L. Jonathana Cohena (1977).

2. O osobliwości rozważań nad antyredukcjonizmem

Rozważania dotyczące zarówno redukcjonizmu, jak i antyredukcjonizmu mają tę wadę, że są „palcem po wodzie pisane”. Są z konieczności pograżone w sferze spekulacji w złym tego słowa znaczeniu. Zwolenników redukcjonizmu dywagacje zwolenników antyredukcjonizmu mogą tylko irytować. I na odwrót.

Warto sobie zdać sprawę z tego, że o definitywnym rozstrzygnięciu sporu pomiędzy redukcjonizmem a antyredukcjonizmem wolno nam dzisiaj tylko marzyć. Teorie naukowe podlegają ciągłemu rozwojowi i doprawdy trudno się wypowiadać na temat tego, co naukowcy mogą – a czego nie mogą – jeszcze wymyślić. W takiej sytuacji można obrać dwa rodzaje postaw, a w obrębie drugiej można wyróżnić jeszcze dwa podrodzaje. Można przyjąć (postawa „albo jedno, albo drugie”), że jesteśmy w stanie – i powinniśmy – opowiedzieć się albo po stronie redukcjonizmu, albo po stronie antyredukcjonizmu, opierając się na racjach już dostępnych, ewentualnie pomnażając jeszcze ich liczbę. Można jednak to założenie zakwestionować, odwołując się do wyżej wyłuszczonych racji, i optować za stanowiskiem agnostycznym zakazującym opowiadania się za słusznością stanowiska redukcjonistycznego czy antyredukcjonistycznego. Postawa taka może przybrać dwie formy. Można przyjąć postawę agnozy bierniej i w ogóle się od kwestii dyskutowanej odciąć (postawa „ani jedno, ani drugie”). W obliczu tych samych argumentów można też optować za postawą agnozy czynnej i chcieć zadbać o to, aby wziąć pod uwagę jak najwięcej możliwych stanowisk (postawa „i jedno, i drugie”).

W niniejszej pracy przyjmiemy to ostatnie stanowisko. Wydaje się ono być owocne we wszelkich nierozstrzygalnych lub trudno rozstrzygalnych sporach filozoficznych (z wyjątkiem być może sporów etycznych), a ponadto pozwala lepiej zrozumieć dość osobliwą przemianę, która zaszła w angloamerykańskiej filozofii umysłu. Podczas gdy na początku XX wieku panowała „moda” na redukcjonizm, to pod koniec tego wieku została zastąpiona w zupełności przez „modę” na antyredukcjonizm. Bynajmniej nie stało się tak ze względu na to, że to ostatnie stanowisko zostało w tym czasie dobrze opracowane. Przeciwnie. Jeśli przyj-

rzeć się analizom redukcjonizmu, to są one opracowane o wiele lepiej (por. np. Lewis 1966, 1972; Hellman & Thompson 1975, 1977; Kim 1993b, 1998; na gruncie polskim koncepcję redukcjonistyczną rozwija m.in. Strawiński 1997). Nadto stanowiska takie jak np. Davidsona są podatne na liczne zarzuty (w tym problem epifenomenalizmu rodzajowego, który omówimy niżej). Można więc postawić pytanie, jakie są źródła popularności antyredukcjonizmu. Staje się ona zrozumiała przy przyjęciu postawy agnozy czynnej (oraz założeniu, że społeczność filozoficzna właśnie taką postawę przyjmuje mniej lub bardziej świadomie): to właśnie dlatego, że stanowisko redukcjonistyczne jest opracowane lepiej, a stanowisko antyredukcjonistyczne boryka się z rozmaitego typu trudnościami, należy inwestować energię w rozwój stanowiska antyredukcjonistycznego. W tym też duchu postaramy się niżej naszkicować pewną postać stanowiska antyredukcjonistycznego.

3. Redukcja ontologiczna według E. Nagela

Rozpocznijmy od powszechnie przyjętego punktu wyjścia zarówno dla koncepcji redukcjonistycznych, jak i antyredukcjonistycznych, a mianowicie od koncepcji redukcji E. Nagela (1961). Wyróżnia on dwa rodzaje redukcji: homogeniczną i heterogeniczną (ontologiczną). Redukcja homogeniczna dotyczy teorii proponowanych dla wyjaśnienia pewnych zjawisk w tej samej dziedzinie wiedzy stosujących pokrewne pojęcia. Przykładem redukcji tego rodzaju są przejścia pomiędzy teoriami spadku swobodnego Galileusza i prawami ruchu planet Keplera z jednej strony a mechaniką Newtona z drugiej oraz mechaniką Newtona a szczególnie teorią względności. Redukcja heterogeniczna dotyczy teorii mających wyjaśniać zjawiska, które mogą się wydawać z pozoru różne, używających odmiennych systemów konceptualizacji. Przykładem redukcji drugiego rodzaju są przejścia między „termodynamiką jakościową” a termodynamiką molekularną oraz między genetyką Mendla a genetyką molekularną. Można przypuszczać, że jeśli w ogóle można będzie kiedyś myśleć o redukcji rozwiniętej teorii psychologicznej do rozwiniętej teorii neurofizjologicznej (tudzież fizycznej), to w grę wchodzi tu redukcja heterogeniczna, a nie homogeniczna⁶.

E. Nagel charakteryzuje redukcję heterogeniczną, odwołując się do dwóch typów warunków: warunków *formalnych* oraz warunków *empirycznych*. Warunki empiryczne obejmują kwestie dotyczące uzasadnienia teorii, uwzględnienia rozwoju historycznego, poprawnego rozumienia natury redukcji, a także systematy-

⁶ Warto w tym miejscu odnotować fakt osobliwy. Dotychczasowe przypadki redukcji heterogenicznych podyktowane były rozwojem nauk, a utożsamienie ich pozornie różnych dziedzin i zjawisk było niejakiem przełomem – w każdym razie nie było oczekiwane. W przypadku relacji psychofizycznych jest natomiast wręcz odwrotnie. Nie ma w zasadzie śladu ani po teoriach neurofizjologicznych, ani psychologicznych – mamy do czynienia co najwyżej ze zrozumieniem lokalnym i cząstkowym działania mechanizmów neurofizjologicznych.

zacji mającej być skutkiem redukcji. Warunki formalne można dalej podzielić na warunki formalne *wstępne*, przygotowujące niejako grunt pod warunki formalne *właściwe*, a mianowicie warunek spójności oraz warunek wynikania logicznego. Na warunki wstępne składa się wymóg kodyfikacji teorii na grupy zdań pełniących określoną rolę w teorii (postulaty teoretyczne, reguły koordynujące, jednostkowe zdania obserwacyjne etc.) oraz wymóg określenia znaczenia zarówno terminów obserwacyjnych, jak i teoretycznych teorii.

Warunek spójności głosi, że terminy występujące w teorii redukowanej (wtórnej) a niewystępujące w teorii redukującej (pierwotnej) trzeba wyrazić w terminach teorii pierwotnej. Powstają w ten sposób definicje koordynujące (zwane też prawami mostowymi), których natura rodzi wiele pytań (szerzej tę kwestię omawia Strawiński 1997). Na potrzeby dalszych rozważań uczynimy założenie upraszczające, które jest standardowe w literaturze z zakresu filozofii umysłu, a mianowicie, że terminy teorii wtórnej należy utożsamiać z terminami teorii pierwotnej. Jest to niewątpliwie założenie dalece upraszczające, ale rozważenie wszystkich możliwych przypadków relacji pomiędzy terminami wtórnymi i pierwotnymi niepotrzebnie by tylko skomplikowało obraz⁷. Warunek wynikania logicznego natomiast głosi, że:

- (W) prawa eksperymentalne nauki wtórnej (oraz ewentualnie jej teorie, o ile istnieją) są logicznymi konsekwencjami postulatów teoretycznych oraz definicji koordynujących nauki pierwotnej (Nagel 1961, s. 307).

Warto tu podkreślić, że tak rozumiana redukcja w sposób zasadniczy opiera się na definicjach koordynujących, które zdają sprawę z zależności rodzajowych pomiędzy pojęciami dwu teorii.

4. Nieredukowalność I: Identyczność (wyłącznie) jednostkowa

Czy jest zatem możliwe, aby istniały dwie prawdziwe teorie traktujące o tych samych zjawiskach, z których żadna nie redukuje się do drugiej? Davidson odpowiada na to pytanie twierdząco. Aby tak było, pojęcia jednej i drugiej teorii nie

⁷ Założenie, iż terminy obu teorii można ze sobą utożsamiać, jest wysoce upraszczające przynajmniej z dwóch względów. Po pierwsze, redukcja teorii nie wymaga utożsamienia czynników, dopuszcza się niekiedy ich korelacje. Filozofowie umysłu są jednak szczególnie zainteresowani przypadkami redukcji, w których mamy do czynienia z utożsamieniem czynników, ze względu na to, iż sytuacja taka stanowi niejako najgorszy możliwy przypadek z punktu widzenia autonomii psychologii. Po drugie, nie jest wcale jasny status definicji koordynujących. W rozprawie habilitacyjnej, której niniejszy artykuł jest częścią, pokazuję, iż owoć jest spojrzeć na nie jako na zdania eksplikacyjne, których status jest dość szczególny (szerzej tę kwestię poruszałam w pracy Paprzycka 1999). Pozwala to m.in. na rozjaśnienie problemu niewspółmierności teorii (Ajdukiewicz 1934/1960; Kuhn 1962/2001).

mogłyby być do siebie sprowadzalne, a więc nie mogłyby istnieć definicje koordynujące, w tym przypadku psychofizyczne prawa mostowe. Natomiast pomiędzy jednostkami należącymi do zakresów tych pojęć zachodziłyby relacje tożsamości.

Taką propozycję przedstawia Davidson. Cztery twierdzenia składają się na jego monizm anomalny:

- (1) Zasada Interakcji Przyczynowej: „przynajmniej niektóre zdarzenia mentalne wchodzi w interakcje przyczynowe ze zdarzeniami fizycznymi” (Davidson 1970/1992, s. 164).
- (2) Zasada Anomalnego Charakteru Zdarzeń Mentalnych: „nie istnieją ścisłe, deterministyczne prawa, na podstawie których można przewidywać i wyjaśniać zdarzenia mentalne” (Davidson 1970/1992, pp. 165–166).
- (3) Zasada Nomologicznego Charakteru Przyczynowości: „zdarzenia, które są ze sobą powiązane jako przyczyna i skutek, podpadają pod ścisłe prawa deterministyczne” (Davidson 1970/1992, p. 165).
- (4) Teza Identyczności (wyłącznie) Jednostkowej (nie rodzajowej) „jednostkowe zdarzenia mentalne są tożsame z jednostkowymi zdarzeniami fizycznymi pomimo tego, iż nie zachodzi relacja tożsamości pomiędzy rodzajami mentalnymi a rodzajami fizycznymi”.

Davidson przyjmuje pierwsze trzy zasady. Jeśli dalej przyjąć za teoriami tożsamości rodzajowej, że pojęcia psychologiczne można sprowadzić do pojęć fizycznych, wówczas tezy te prowadzą do sprzeczności. Prześledźmy jego rozumowanie na schematycznym przykładzie. Według Zasady Interakcji Przyczynowej pewne zdarzenie psychiczne (np. zdarzenie d , które opisać można m.in. za pomocą predykatu psychologicznego M_1) przyczynia się do pewnego zdarzenia fizycznego (np. zdarzenia d' , które opisać można m.in. za pomocą predykatu fizycznego F_2)⁸. Przyjmując, że d można również opisać za pomocą predykatu F_1 , Zasada Nomologicznego Charakteru Przyczynowości głosi, że istnieje ścisłe prawo fizyczne (Davidson dalej utrzymuje, iż domeną ścisłych praw jest fizyka), którego konsekwencją jest zależność $(x)(y)(F_1(x) \rightarrow F_2(y))$. Jeśli więc zachodziłyby tożsamości pomiędzy predykatami fizycznymi i psychologicznymi, np. $M_1 = F_1$ i $M_2 = F_2$, wówczas zachodziłaby ścisła zależność psychologiczna, a mianowicie $(x)(y)(M_1(x) \rightarrow M_2(y))$, a temu zaprzecza Zasada Anomalnego Charakteru Zdarzeń Mentalnych.

Teza Identyczności (wyłącznie) Jednostkowej głosi, iż jednostkowe zdarzenia psychiczne są identyczne z jednostkowymi zdarzeniami fizycznymi, natomiast zaprzecza temu, że predykaty fizyczne są identyczne z predykatami psychologicz-

⁸ Według Davidsona zdarzenie jest fizyczne wtedy i tylko wtedy, jeśli można je opisać za pomocą pewnego predykatu fizycznego, a zdarzenie jest psychiczne wtedy i tylko wtedy, gdy można je opisać za pomocą pewnego predykatu psychologicznego.

nymi. Zachodzą tylko tożsamości pomiędzy jednostkowymi zdarzeniami, które można opisać zarówno w języku fizyki (podpadają pod prawa fizyki), jak i w języku psychologii (choć nie podpadają pod ścisłe prawa psychologiczne, gdyż takich – według Davidsona – nie ma).

Odrzucenie tezy o zachodzeniu tożsamości rodzajowych, a ściślej tożsamości pomiędzy predykatami psychologicznymi i fizycznymi, stanowi zatem podstawę dla antyredukcjonizmu Davidsona.

5. Zarzut epifenomenalizmu rodzajowego

Teorii identyczności jednostkowej Davidsona postawiono zarzut, iż prowadzi do pewnej wersji epifenomenalizmu, a mianowicie epifenomenalizmu rodzajowego⁹. Stanowisko Davidsona nie prowadzi do epifenomenalizmu klasycznego, wedle którego zdarzenia psychiczne nie mają mocy sprawczej. W ujęciu teorii identyczności jednostkowej zdarzenia psychiczne mają moc sprawczą – dokładnie taką jak zdarzenia fizyczne, z którymi są wszakże identyczne. W świetle tej koncepcji nie posiadają natomiast mocy sprawczej zdarzenia ze względu na własności psychiczne¹⁰. Ograniczenie to określane jest mianem epifenomenalizmu rodzajowego.

Aby się przekonać, iż powszechnie przyjmujemy, że zdarzenia mogą przyczyniać się do innych zdarzeń ze względu na swe różne własności, rozważmy następujący przykład. – Jan krzyczy wniebogłosy „Świat jest wspaniały!” tak głośno, że pękają szyby w oknach, a zdepresjonowana sąsiadka wyskakuje przez okno na grządkę i podrzuca mu parę chwastów przez nowo otwarte możliwości, sc. dziury w oknach. – Okrzyk Jana jest przyczyną przynajmniej dwóch zdarzeń: pęknięcia szyb oraz tego, że sąsiadka wyskakuje przez okno. W pierwszym przypadku jasne jest, że w grę wchodzi tylko fizyczne własności tego okrzyku – jego gwałtownie wzrastająca amplituda. W drugim przypadku pewną rolę odgrywają również inne własności, w szczególności semantyczne – jeśli zdepresjonowana sąsiadka burzliwie się nie zgadza z treścią okrzyku, a także psychiczne – jeśli ją oburza manifestowany w ten sposób stan ducha krzyczącego.

Przykład ten wskazuje na to, iż relacje przyczynowo-skutkowe zachodzą nie tyle między zdarzeniami (jak sugeruje Davidson), ile między zdarzeniami ze względu na pewne własności tych zdarzeń. W powyższym przykładzie mamy do czynienia ze zdarzeniem jednostkowym, a mianowicie krzykiem Jana, które ze względu na dwie różne własności przyczynia się do dwóch dalszych zdarzeń. Ze względu

⁹ Jako pierwszy zarzut ten wysunął F. Stoutland (1976). Później przedstawiano go w wielu wersjach, por. np. Honderich (1982), McLaughlin (1989), Kim (1993a).

¹⁰ Zarzut ten formułowany bywa bardzo różnie. Niekiedy przypisuje się ową niemoc przyczynową własnościom psychicznym, a niekiedy zdarzeniom psychicznym *qua* zdarzeniom psychicznym.

na gwałtownie wzrastającą amplitudę przyczynia się do pęknięcia szyb, a ze względu na swą treść – do zdenerwowania sąsiadki.

Davidson przyjmując fizykalistyczną Zasadę Nomologicznego Charakteru Przyczynowości, zdaje się być zobligowany twierdzić, że wszelkie relacje przyczynowe zachodzą tylko ze względu na fizyczne własności tych zdarzeń. Jeśli tak, to nie ma miejsca w teorii Davidsona na to, by zachodziły związki przyczynowe ze względu na własności mentalne zdarzeń. Ściślej rzecz ujmując, Davidson mógłby twierdzić, że pewne relacje przyczynowe mogą zachodzić ze względu na własności mentalne przyczyn, jeśli własności mentalne można by utożsamić z własnościami fizycznymi. Nie może jednakże tego uczynić, gdyż właśnie to twierdzenie teorii identyczności rodzajowej odrzuca i musi odrzucić, jeśli chce utrzymać zasady (1)–(3). W rezultacie Davidson zdaje się być skazany na twierdzenie o epifenomenalizmie rodzajowym zdarzeń psychicznych.

Stajemy zatem wobec dylematu. Albo przyjmiemy za Davidsonem antyredukcjonizm, ale wówczas wiąże się to z odrzuceniem zależności (rodzajowych) pomiędzy pojęciami psychologii i fizyki, co prowadzi do epifenomenalizmu rodzajowego, albo unikniemy epifenomenalizmu rodzajowego, przyjmując zależności rodzajowe, ale kosztem zarzucenia antyredukcjonizmu. Okazuje się, że jest to dylemat fałszywy – można przyjąć, że zachodzą zależności rodzajowe, i uniknąć redukcjonizmu.

6. Redukcja w terminach Idealizacyjnej Koncepcji Nauki

Powyższy dylemat powstaje na gruncie tzw. zastanego poglądu na strukturę nauki, którego integralną częścią jest koncepcja redukcji E. Nagela. Pomimo tego, iż większość koncepcji składających się na „pogląd zastany” była poddawana krytyce z rozmaitych perspektyw (dobry przegląd krytyczny zawarty jest np. w pracach: Suppe 1974, Salmon 1989), koncepcje te nadal stanowią niejako domyślny punkt wyjścia w szczególności dla rozważań prowadzonych w innych dziedzinach. Niżej postaram się pokazać, iż przyjmując pojęcie redukcji zaproponowane na gruncie Idealizacyjnej Koncepcji Nauki, można wybrnąć z tak postawionego dylematu.

Redukcji heterogenicznej (ew. ontologicznej) dotychczas w terminach idealizacyjnych nie rozważano, koncentrując uwagę na tym, co Nagel określa mianem redukcji homogenicznej, a co w IKN funkcjonuje pod nazwą korespondencji (por. Nowakowa 1975, 1994; Krajewski 1977; Paprzycka 1990). Ponieważ zajmujemy się tu bardzo specjalną kwestią, nie będziemy tych rozważań nadto komplikować. Wprowadzimy tylko minimalną poprawkę do ogólnej koncepcji Nagela związaną z naturą teorii idealizacyjnych. Do warunków formalnych wstępnych Nagela trzeba dołączyć warunek idealizacyjnej kodyfikacji teorii redukowanej i redukującej, który głosi tylko tyle, że są one teoriami idealizacyjnymi. Warunek

spójności pozostaje bez zmian. Dla uproszczenia nadal przyjmujemy, iż definicje koordynujące zdają sprawę z tożsamości pomiędzy pojęciami teorii redukowanej i redukującej.

Trzeba natomiast zmodyfikować warunek wynikania logicznego ze względu na idealizacyjną strukturę teorii. Z redukcją (heterogeniczną) teorii idealizacyjnej t do teorii idealizacyjnej T mamy do czynienia, jeśli

(W_i) z teorii idealizacyjnej T wzbogaconej o odpowiednie definicje koordynujące wynika tak prawo idealizacyjne teorii t , jak i jego konkretyzacje.

Tylko w takim wypadku możemy mówić o redukcji jednej teorii do drugiej. Przykładem może tu być redukcja teorii gazów doskonałych do mechaniki statystycznej. Prawo idealizacyjne Boyle'a-Charlesa dla gazów doskonałych można wyprowadzić z mechaniki statystycznej przy przyjęciu odpowiednich definicji koordynujących (w szczególności dotyczących relacji pomiędzy temperaturą rozumianą jako średnia energia kinetyczna cząsteczek oraz ciśnieniem rozumianym jako średni pęd cząsteczek gazu). Jego konkretyzacje, zaproponowane przez van der Waalsa, również da się wyprowadzić przy przyjęciu odpowiednich założeń (por. Kuipers 1982, 1985, 1990).

Należy jednak zwrócić uwagę, że dopuszczalna byłaby sytuacja taka, że z teorii T wraz z definicjami koordynującymi wynikałyby tylko prawa faktualne (ostatnie konkretyzacje) teorii t , natomiast ani nie prawo idealizacyjne teorii t , ani nie jego pośrednie konkretyzacje. W takim wypadku nie można byłoby jednak mówić o redukcji jednej teorii do drugiej, a co najwyżej o subsumpcji wyników empirycznych jednej teorii przez drugą.

7. Nieredukowalność II: Zróznicowanie istotnościowe

Jeśli przyjąć założenia Idealizacyjnej Koncepcji Nauki, niewątpliwie możliwe jest przyjęcie koncepcji nieredukowalności w oparciu o tożsamość wyłącznie jednostkową. Otwiera się jednakże przynajmniej jeszcze jedna możliwość konceptualizacji nieredukowalności, która jest o tyle godna uwagi, że nie prowadzi do zarzutu epifenomenalizmu rodzajowego.

Intuicyjnie chodzi tu o dwie teorie, których dziedziny pokrywają się przynajmniej częściowo. Może być przy tym tak, że przestrzenie (zbiory) czynników istotnych dla wielkości badanych w obu teoriach będą tożsame (przy przyjęciu definicji koordynujących), natomiast teorie te proponują zasadniczo odmienną hierarchizację istotnościową tych przestrzeni. Przyjmijmy, że to, co dla wielkości badanej w jednej teorii jest czynnikiem głównym (uwzględnionym już w jej prawie idealizacyjnym), dla wielkości badanej w drugiej teorii jest czynnikiem ubocznym, który uwzględniony jest w jakiejś konkretyzacji właściwego jej prawa idealizacyjnego.

Założmy, że rozważamy dwie teorie idealizacyjne t i T . Teoria t bada czynnik C , dla którego czynnikami istotnymi są M , m_1 i m_2 . Hierarchia czynników istotnych dla C

$$\begin{aligned} S_C: & M \\ & M, m_1 \\ & M, m_1, m_2 \end{aligned}$$

stanowi tzw. strukturę esencjalną, na której najwyższy poziom składa się czynnik główny, a której kolejne poziomy uwzględniają czynniki uboczne w kolejności ich istotności.

Przyjmujemy, że w obrębie teorii t wysunięto i sprawdzono następujące prawo idealizacyjne:

$$(t_0) (x)[m_1(x) = 0 \wedge m_2(x) = 0 \rightarrow C(x) = g_0(M(x))],$$

gdzie wyrażenie „ $m_1(x) = 0$ ” jest założeniem idealizującym, przyjęcie którego jest równoznaczne z uznaniem, iż czynnik m_1 przysługujący przedmiotowi x nie wywiera wpływu na czynnik C przysługujący x , natomiast funkcja g_0 wyraża zależność zachodzącą pomiędzy czynnikiem głównym M a czynnikiem określanym C przy założeniu, iż czynniki uboczne nie oddziałują na C – g_0 określa się też mianem prawidłowości.

Przyjmujemy dalej, że prawo idealizacyjne (t_0) poddano konkretyzacji najpierw ze względu na najbardziej istotny spośród czynników ubocznych, a mianowicie czynnik m_1 , postulując i sprawdzając pierwszą konkretyzację prawa idealizacyjnego:

$$\begin{aligned} (t_1) (x)[m_1(x) \neq 0 \wedge m_2(x) = 0 \rightarrow C(x) &= h_{k_1}(g_0(M(x)), k_1(m_1(x))) \\ &= g_1(M(x), m_1(x))], \end{aligned}$$

gdzie funkcja k_1 jest tzw. funkcją korekcyjną pokazującą, jak pomijany dotąd czynnik uboczny wpływa na czynnik badany, funkcja kierunkowa h_{k_1} pokazuje, jak funkcja korekcyjna modyfikuje prawidłowość; po złożeniu tych funkcji możemy mówić o zależności g_1 wyrażającej wpływ czynników M i m_1 na czynnik określany. Założymy też, że dokonano ostatniej konkretyzacji uwzględniającej wpływ czynnika m_2 :

$$\begin{aligned} (t_2) (x)[m_1(x) \neq 0 \wedge m_2(x) \neq 0 \rightarrow C(x) &= h_{k_2}(g_1(M(x), m_1(x)), k_2(m_2(x))) = \\ &= g_2(M(x), m_1(x), m_2(x))], \end{aligned}$$

gdzie k_2 to funkcja korekcyjna, h_{k_2} to funkcja kierunkowa, a g_2 wyraża zależność czynnika C od wszystkich czynników dlań istotnych. Ostatnia konkretyzacja prawa idealizacyjnego jest określana mianem twierdzenia faktualnego, ponieważ dotyczy już niewyidealizowanej rzeczywistości. Ciąg twierdzeń (t_0) , (t_1) , (t_2) określany jest mianem prostej teorii idealizacyjnej.

Założymy dalej, że mamy też do czynienia z teorią T badającą czynnik D , dla którego czynnikami istotnymi są N , n_1 i n_2 . Teoria T zakłada strukturę esencjalną S_D

$$\begin{aligned} S_D: & N \\ & N, n_1 \\ & N, n_1, n_2 \end{aligned}$$

Przyjmijmy, że w obrębie teorii T wysunięto i sprawdzono następujące prawo i jego konkretyzacje:

$$\begin{aligned} (T_0) (x) [n_1(x) = 0 \wedge n_2(x) = 0 \rightarrow D(x) = f_0(N(x))] \\ (T_1) (x) [n_1(x) \neq 0 \wedge n_2(x) = 0 \rightarrow D(x) = f_1(N(x), n_1(x))] \\ (T_2) (x) [n_1(x) \neq 0 \wedge n_2(x) \neq 0 \rightarrow D(x) = f_2(N(x), n_1(x), n_2(x))], \end{aligned}$$

gdzie f_0 wyraża prawidłowość teorii T , pokazującą zależność pomiędzy czynnikiem badanym D a czynnikiem głównym N przy założeniu, iż czynniki uboczne n_1 i n_2 nie wpływają na D , natomiast f_1 i f_2 wyrażają zależności pomiędzy czynnikiem badanym oraz odpowiednimi czynnikami dlań istotnymi.

Założmy dalej, że teorie t i T są do siebie sprowadzalne, tj. (a) czynniki badane C i D związane są definicją koordynującą

$$C = D,$$

(b) pozostałe czynniki obu teorii również związane są definicjami koordynującymi. Jeśli czynniki obu teorii skoordynowane są ze sobą w taki sposób, że istnieje izomorfizm między strukturami esencjalnymi obu teorii, wówczas mówimy, że teorie te są *istotnościowo zgodne*; w przeciwnym razie teorie są *istotnościowo niezgodne*.

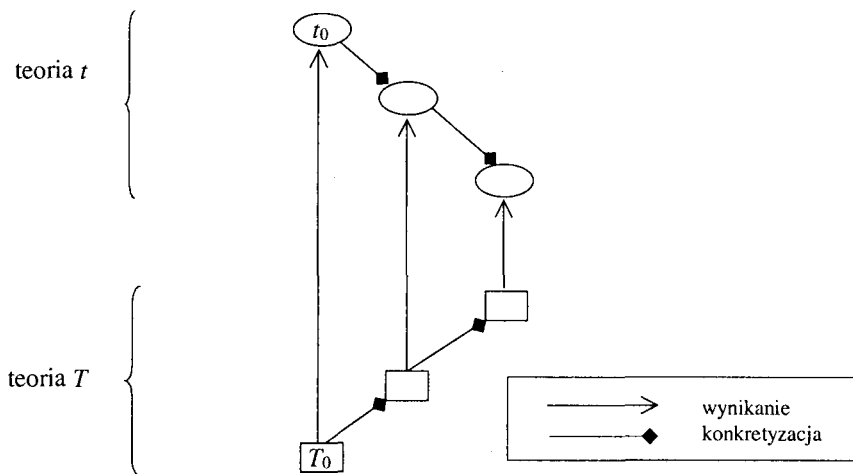
Teorie istotnościowo zgodne. Kierując się powyższym przykładem teorii T i t , możemy uznać, że są one istotnościowo zgodne, jeśli stosowne są następujące definicje koordynujące:

$$\begin{aligned} M &= N \\ m_1 &= n_1 \\ m_2 &= n_2. \end{aligned}$$

W ten sposób powielone jest uporządkowanie istotnościowe czynników. Jeśli tak, to z prawa (T_0) , można wyprowadzić (t_0) , a także jego konkretyzacje. Strukturę taką odzwierciedla rys. 1.

Przypadek teorii istotnościowo zgodnych odpowiada intuicjom redukcjonistycznym. Jeśli teorie są istotnościowo zgodne, wówczas możliwe jest wyprowadzenie prawa idealizacyjnego teorii redukowanej t z prawa idealizacyjnego teorii redukującej T , a także wyprowadzenie każdej konkretyzacji teorii t z odpowiadającej jej konkretyzacji teorii T . Za przykład może tu posłużyć wspomniana już

redukcja termodynamiki jakościowej do termodynamiki statystycznej, gdzie tak prawo idealizacyjne (prawo gazu idealnego), jak i jego konkretyzacje (twierdzenia van der Waalsa) można wyprowadzić z termodynamiki statystycznej (por. Kuipers 1982, 1985, 1990).



Rys. 1. Teorie idealizacyjne istotnościowo zgodne

Teorie istotnościowo niezgodne. W przypadku teorii istotnościowo niezgodnych natomiast czynniki teorii t i T związane są przez definicje koordynujące, które nie zachowują izomorfizmu pomiędzy porządkami istotnościowymi obu teorii (wiążąc czynniki główne jednej teorii z czynnikami ubocznymi drugiej). Weźmy na przykład następujące definicje koordynujące:

$$\begin{aligned} M &= n_1 \\ m_1 &= n_2 \\ m_2 &= N. \end{aligned}$$

Przy takich definicjach koordynujących teorii t nie da się zredukować do teorii T . Rozważmy prawa idealizacyjne. Zwróćmy najpierw uwagę na same następniki. Jeśli dane jest (T_0) , to wiemy, że w pewnych warunkach spełniona jest zależność

$$D(x) = f(N(x))$$

Z tego następnika dałoby się co najwyżej wyprowadzić zależność

$$C(x) = k_1(m_2(x)),$$

a więc jedynie funkcję korekcyjną uwzględniającą wpływ czynnika ubocznego m_2 . Nie stanowi więc ona zależności koniecznej dla dokonania redukcji t do T , a mia-

nowicie $C(x) = g(M(x))$. Tym bardziej, że musimy też uwzględnić warunki, w jakich następnik ma zachodzić. Przy przyjętych definicjach koordynujących z (T_0) wyprowadzić można twierdzenie

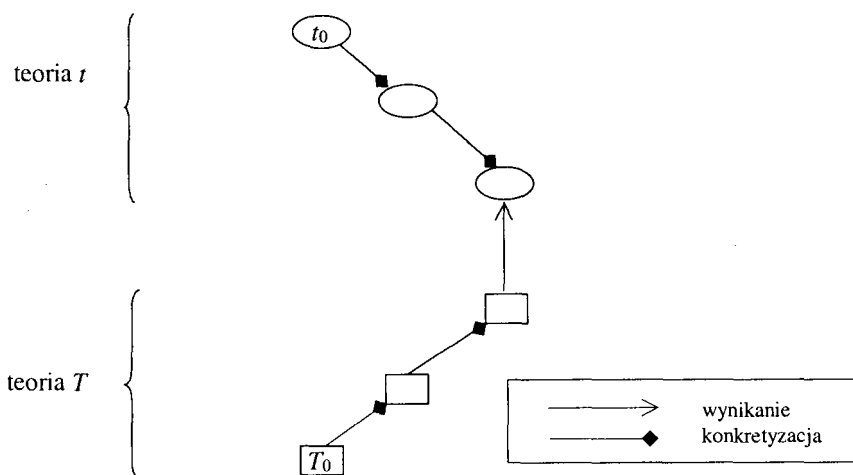
$$(T'_0)(x)[M(x) = 0 \wedge m_1(x) = 0 \rightarrow C(x) = h(m_2(x))]$$

a dalej

$$(T'_1)(x)[M(x) \neq 0 \wedge m_1(x) = 0 \rightarrow C(x) = h'(m_2(x), M(x))]$$

$$(T'_2)(x)[M(x) \neq 0 \wedge m_1(x) \neq 0 \rightarrow C(x) = h''(m_2(x), M(x), m_1(x))]$$

Nie jest tu zatem spełniony warunek wynikania dla praw idealizacyjnych. Związek pomiędzy teoriami niezgodnymi obrazuje rys. 2.



Rys. 2. Teorie idealizacyjne istotnościowo niezgodne.

Można co najwyżej wyprowadzić twierdzenie faktualne (ostatnią konkretyzację) (T'_2) z (T_2) . Natomiast nie można wyprowadzić całości struktury nomologicznej. Ponieważ to właśnie odwołanie do prawa idealizacyjnego – a więc szczytu takiej struktury – jest konieczne dla wyjaśniania (Nowak 1977), w związku z tym fakt, że nie daje się wyprowadzić prawa idealizacyjnego, jest nierozwiązalnie połączony z faktem, że nie daje się powielić struktury wyjaśniania w kategoriach starej teorii. Ponieważ wyjaśnianie jest ściśle związane z redukcją, więc nie dziwi fakt, że niezachowanie struktury wyjaśniania idzie w parze z niemożnością dokonania redukcji.

Jednakże to, iż możliwe jest wyprowadzenie twierdzeń faktualnych, jest ważne, gdyż znaczy to, iż zachodzą pewnego typu zależności nieprzypadkowe (i rodzajowe) pomiędzy pojęciami i zależnościami jednej i drugiej teorii. Pozwalają

one na ich zbieżność w detalicznych i szczególnych generalizacjach, natomiast nie na związanie prawidłowości idealizacyjnych. Innymi słowy, teorie te wyjaśniają te same zjawiska, ale na różne sposoby.

Dzięki temu związkowi można jednakże uniknąć zarzutu epifenomenalizmu rodzajowego. Davidson przyjmując fizykalistyczną Zasadę Nomologicznego Charakteru Przyczynowości, zobligowany jest twierdzić, iż wszelkie związki przyczynowe zachodzą ze względu na własności fizyczne zdarzeń. Zarzut epifenomenalizmu rodzajowego wziął się stąd, iż Davidson, aby utrzymać antyredukcjonizm, uznał, iż musi zaprzeczyć temu, jakoby pojęcia psychologii były sprowadzalne do pojęć fizyki, zarzucając tym samym podstawy dla przyjęcia tezy, że związki przyczynowe zachodzą również ze względu na własności psychiczne zdarzeń.

Przedstawiona wyżej propozycja pokazuje, jak można utrzymać antyredukcjonizm bez zarzucenia tożsamości rodzajowych. Nie ma zatem przeszkód, by przyjąć intuicyjne twierdzenie, iż związki przyczynowe pomiędzy zdarzeniami fizycznymi i psychicznymi zachodzą nie tylko ze względu na fizyczne własności zdarzeń, ale również ze względu na ich własności psychiczne. Niezgodność istotnościowa gwarantująca nieredukowalność teorii idealizacyjnych opiera się wszakże na tożsamościach (rodzajowych) pomiędzy poszczególnymi czynnikami obu teorii (tyle że aby niezgodność istotnościowa zachodziła, nie mogą być jednocześnie powielone związki istotnościowe pomiędzy czynnikami). Gwarantuje to zasadnicze różnice między prawami idealizacyjnymi oraz ich konkretyzacjami, przy zachowaniu jednak nie tylko jednostkowej, ale i rodzajowej zbieżności między tym, co teorie te wyjaśniają.

8. Zakończenie

Widzieliśmy wyżej, jak wzbogacając stanowisko Davidsona o pewne propozycje Idealizacyjnej Koncepcji Nauki, można z jednej strony utrzymać stanowisko antyredukcjonizmu, a z drugiej uniknąć zarzutu epifenomenalizmu rodzajowego. – By ująć rzecz nieco metaforycznie, Idealizacyjna Koncepcja Nauki pokazuje, że mamy do czynienia nie z jednym szeregiem, ale z hierarchią rodzajów (tu: czynników). Prawa naukowe dotyczą czynników najwyższego poziomu w hierarchii, ich konkretyzacje natomiast – poziomów coraz to niższych. Jednym sposobem zrozumienia, na czym mogłaby polegać nieredukowalność teorii do siebie, jest przypuszczenie, iż utożsamić należy tylko zdarzenia jednostkowe, do których stosują się te teorie – tak jak to czyni Davidson. Takie rozwiązanie obciążone jest jednakże kosztem w postaci zarzutu epifenomenalizmu rodzajowego. Innym sposobem zrozumienia, jak możliwa byłaby nieredukowalność jednej teorii do drugiej, jest przypuszczenie, iż utożsamione są nie tylko jednostkowe zdarzenia, ale i czynniki obu teorii, tyle że na wskroś ich uporządkowań istotnościowych. W takiej sytuacji relacja wynikania zachodzić będzie tylko pomiędzy odle-

głymi konkretyzacjami dwóch teorii istotnościowo niezgodnych, ale nie pomiędzy prawami idealizacyjnymi, gwarantując w ten sposób nieredukowalność, a zatem i autonomiczność tych teorii. Jednocześnie, ponieważ utożsamiane są również czynniki obu teorii, a nie tylko jednostkowe zdarzenia, unika się zarzutu epifenomenalizmu rodzajowego.

Należy tu podkreślić, że wydźwięk niniejszego tekstu jest tylko pozornie antyredukcyjny. Pokazaliśmy wszakże, w jakich warunkach (mianowicie niezgodności istotnościowej) redukcja dwóch teorii idealizacyjnych byłaby niemożliwa. Czy taka sytuacja ma miejsce w przypadku relacji między jakimikolwiek teoriami – a psychologią i neurofizjologią w szczególności – jest i pozostanie kwestią otwartą. Zadaniem dla filozofii jest zrozumieć, na czym redukcja i autonomia mogłyby w ogóle polegać. Czy zasadny jest redukcjonizm czy antyredukcyjnizm, rozstrzygną natomiast ci, do których to rozstrzygnięcie należy – naukowcy w bliższej lub dalszej przyszłości.

Bibliografia

- Ajdukiewicz, K. (1934/1960) – *Obraz świata i aparatura pojęciowa*, w: *Język i poznanie*, Tom 1, Warszawa: Państwowe Wydawnictwo Naukowe, 1960, s. 175–195.
- Brzeziński, J., Coniglione, F., Kuipers, T.A.F., Nowak, L., red. (1990) – *Idealization I: General Problems. Poznań Studies in the Philosophy of the Humanities*, vol. 16, Amsterdam: Rodopi.
- Cartwright, N. (1983) – *How the Laws of Physics Lie*, Oxford: Clarendon Press.
- Cartwright, N. (1989) – *Capacities and their Measurement*, Oxford: Oxford University Press.
- Chalmers, D. (1996) – *The Conscious Mind*, Oxford: Oxford University Press.
- Cohen, L.J. (1977) – *The Probable and the Provable*, Oxford: Basil Blackwell.
- Davidson, D. (1970/1992) – *Zdarzenia mentalne*, w: *Eseje o prawdzie, języku i umyśle*, Warszawa: PWN, 1992, s. 163–193.
- Hellman, G., Thompson, F. (1975) – *Physicalism: Ontology, Determination, and Reduction*, „Journal of Philosophy” 72, 551–564.
- Hellman, G., Thompson, F. (1977) – *Physicalist Materialism*, „Nous” 11, 309–345.
- Honderich, T. (1982) – *The Argument for Anomalous Monism*, „Analysis” 42, 59–64.
- Jackson, F. (1986) – *What Mary Didn't Know*, „Journal of Philosophy” 83, 291–295.
- Jackson, F., Pargetter, R., Prior, E. (1982) – *Functionalism and Type-Type Identity Theories*, „Philosophical Studies” 42, 209–225.
- Kim, J. (1993a) – *Can Supervenience and „Non-Strict Laws” Save Anomalous Monism*, w: J. Heil, A. Mele (red.), *Mental Causation*, Oxford: Clarendon Press, s. 19–26.

- Kim, J. (1993b) – *Supervenience and Mind. Selected Philosophical Essays*, Cambridge: Cambridge University Press.
- Kim, J. (1998) – *Mind in a Physical World. An Essay on the Mind-Body Problem and Mental Causation*, Cambridge, MA: The MIT Press.
- Krajewski, W. (1977) – *The Principle of Correspondence and the Growth of Knowledge*, Dordrecht-Boston: Reidel.
- Kuhn, T. (1962/2001) – *Struktura rewolucji naukowych*, Warszawa: Aletheia.
- Kuipers, T.A.F. (1982) – *The Reduction of the Phenomenological to Kinetic Thermostatics*, „Philosophy of Science” 49, 107–119.
- Kuipers, T.A.F. (1985) – *The Paradigm of Concretization: The Law of van der Waals*, w: J. Brzeziński, (red.), *Consciousness: Methodological and Psychological Approaches: Poznań Studies in the Philosophy of the Sciences and the Humanities*, vol. 8, Amsterdam: Rodopi, s. 185–199.
- Kuipers, T.A.F. (1990) – *Reduction of Laws and Concepts*, w: Brzeziński et al. (1990), s. 241–276.
- LePore, E., McLaughlin, B.P., red. (1985) – *Actions and Events. Perspectives on the Philosophy of Donald Davidson*, Oxford: Basil Blackwell.
- Lewis, D. (1966) – *An Argument for the Identity Theory*, „Journal of Philosophy” 63, 17–25.
- Lewis, D. (1972) – *Psychophysical and Theoretical Identifications*, „Australasian Journal of Philosophy” 50, 249–258.
- McDowell, J. (1985) – *Functionalism and Anomalous Monism*, w: LePore & McLaughlin (1985), s. 387–398.
- McLaughlin, B.P. (1989) – *Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical*, „Philosophical Perspectives” 3, 109–136.
- Nagel, E. (1961) – *Struktura nauki. Zagadnienia logiki wyjaśnień naukowych*, Warszawa: PWN.
- Nagel, T. (1974) – *What Is It like to Be a Bat?*, „Philosophical Review” 83, 435–450.
- Nowak, L. (1971) – *U podstaw marksowskiej metodologii nauk*, Warszawa: PWN.
- Nowak, L. (1977) – *Wstęp do idealizacyjnej teorii nauki*, Warszawa: PWN.
- Nowak, L. (1980) – *The Structure of Idealization*, Dordrecht/Boston: Reidel.
- Nowakowa, I. (1975) – *Dialektyczna korespondencja a rozwój nauki*, Warszawa–Poznań: PWN.
- Nowakowa, I. (1994) – *The Dynamics of Idealizations*, Amsterdam–Atlanta: Rodopi.
- Paprzycka, K. (1990) – *Reduction and Correspondence in the Idealizational Approach to Science*, w: Brzeziński et al. (1990), s. 277–286.
- Paprzycka, K. (1999) – *Socrates Meets Carnap: Explication in the Theaetetus*, „Logical Analysis and History of Philosophy” 2, 87–108.
- Putnam, H. (1967/1975) – *The Nature of Mental States*, w: *Mind, Language and Reality: Philosophical Papers*, vol. II, Cambridge: Cambridge University Press, 1975, s. 429–440.

- Salmon, W.C. (1984) – *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press.
- Salmon, W.C. (1989) – *Four Decades of Scientific Explanation*, w: P. Kitcher i W.C. Salmon (red.), *Scientific Explanation. Minnesota Studies in the Philosophy of Science, Vol. XIII*, Minneapolis: University of Minnesota Press, s. 3–219.
- Stoutland, F. (1976) – *The Causation of Behavior*, w: J. Hintikka (red.), *Essays on Wittgenstein in Honor of G.H. von Wright: Acta Philosophica Fennica* 28, Amsterdam: North-Holland, s. 286–325.
- Stoutland, F. (1985) – *Davidson on Intentional Behaviour*, w: LePore & McLaughlin (1985), s. 44–59.
- Strawiński, W. (1997) – *Jedność nauki, redukcja, emergencja. Z metodologicznych i ontologicznych problemów integracji wiedzy*, Warszawa: Fundacja Aletheia.
- Suppe, F., red. (1974) – *The Structure of Scientific Theories*, Urbana: University of Illinois Press.

Davidson's Anomalous Monism and the Problem of Epiphenomenalism in the Light of the Idealizational Conception of Science

Davidson's anomalous monism provides a metaphysical foundation for the autonomy of psychology. However, by eschewing all type-type identities between mental and physical events, his token-token identity theory falls prey to the problem of type phenomenalism (mental events can only be causally efficacious in virtue of their physical not in virtue of their mental properties). I argue that one can accept type-type identities and thereby avoid the charge of epiphenomenalism, while still allowing room for the autonomy of a higher-level science. Using the apparatus of the Idealizational Conception of Science it is possible to show that when two theories are incompatible in their essential orderings, one idealizational theory will be in principle irreducible to another idealizational theory even if it is possible to provide type-type identities for all relevant factors.