



Dr hab. Marcin Miłkowski, prof. IFiS PAN
Kierownik Zakładu Logiki i Kognitywistyki
Instytut Filozofii i Socjologii PAN

OPINIA O OSIĄGNIĘCIU HABILITACYJNYM DOKTORA MICHAŁA PIEKARSKIEGO
„MECHANIZMY PREDYKCYJNE I ICH NORMATYWNOŚĆ”
ORAZ O JEGO DOROBKU NAUKOWYM

Pan dr Michał Piekarski jest zatrudniony jako adiunkt w Instytucie Filozofii na Wydziale Filozofii Chrześcijańskiej UKSW w Warszawie. Zajmuje się filozofią języka (zwłaszcza Ludwiga Wittgensteina), filozofią umysłu, polską fenomenologią, etyką zwierząt oraz filozofią nauk o procesach poznawczych. Opublikował pracę „Mechanizmy predykcyjne i ich normatywność” (Liberi Libri, Warszawa 2020; poziom II ministerialnego wykazu wydawnictw naukowych), która została przedstawiona jako osiągnięcie naukowe. Należy zaznaczyć, że książka udostępniona jest również w wolnym dostępie w formacie PDF, co świadczy o dbałości autora o szerokie upowszechnianie wyników jego pracy. W mojej opinii skupię się na tej książce, gdyż pozostałe prace dołączone do dokumentacji nie zostały wskazane jako część osiągnięcia naukowego.

Praca habilitacyjna, jest pierwszą opublikowaną w języku polskim monografią na temat tradycji badawczej określanej mianem „przetwarzania predykcyjnego” bądź „kodowania predykcyjnego”. We współczesnych badaniach nad mózgiem i umysłem tradycja ta odgrywa obecnie kolosalną rolę, gdyż jej zwolennicy sądzą, iż może posłużyć ona do teoretycznej unifikacji modeli percepcji, myślenia i działania. Liczba modeli opartych na założeniach tej tradycji rośnie w bardzo dużym tempie, a jednocześnie rozwijane są podstawy teoretyczne tego podejścia. W uproszczeniu można je przedstawić jako koncepcję, zgodnie z którą umysły tworzą modele rzeczywistości, mające



wyjaśnić przyczyny ich pobudzeń zmysłowych oraz kierować ich działaniem. Modele te są aktualizowane w czasie, gdy wykryte zostaną rozbieżności między przewidywaniami modelu a rzeczywistością. Do tworzenia przewidywań oraz aktualizacji modeli służą poszczególne algorytmy stosowane przez zwolenników przetwarzania predykcyjnego (najbardziej popularne przewidują istnienie hierarchicznych modeli generatywnych). Jednocześnie trwa ożywiona debata filozoficzna na temat teoretycznych założeń stosowanych w modelowaniu.

Monografia habilitanta nawet na tle coraz bogatszej światowej literatury stanowi novum, gdyż do tej pory opublikowano bodaj jedynie dwie inne monografie na ten temat (autorstwa A. Clarka i J. Hohwy'ego), a trzecia, poświęcona obronie zasady energii swobodnej (K. Fristona) jeszcze nie ukazała się drukiem (dostępny jest jedynie preprint). W odróżnieniu jednak od tych pozycji autor skupia się nie tyle na obronie pewnych zasad przyświecających badaczom korzystającym z metodologii tej tradycji badawczej do modelowania procesów poznawczych, ile raczej na pogłębionej analizie zagadnień teoretycznych, skupionych wokół normatywności działania, modelowanej przy użyciu modeli mechanizmów predykcyjnych. Tematyka ta należy do fundamentalnych zagadnień dla tej tradycji, ale nie doczekała się jeszcze pogłębionej dyskusji (na forum międzynarodowym habilitant w pewnej mierze ją zapoczątkował, publikując krótkie artykuły polemiczne na łamach „Frontiers in Psychology”, do których nawiązuje w aktualnych pracach też J. Hohwy).

Monografia została zredagowana dosyć starannie (jedynym większym uchybieniem, które zauważyłem, jest występowanie pozycji Korbak 2019 w tekście głównym, podczas gdy w bibliografii są dwie pozycje z tą datą, 2019a i 2019b). Praca została opatrzona bardzo szczegółowym spisem treści oraz indeksami (osobowym i przedmiotowym). Byłem jednym z czytelników wstępnych wersji wybranych fragmentów manuskryptu książki i przedstawiłem trochę uwag krytycznych (za które dr Piekarski dziękuje mi na s. 33), więc w tym miejscu chciałbym zaznaczyć, że obecna wersja książki na tyle odbiega od czytanej wcześniej przeze mnie, iż w



lekturze stanowiła dla mnie całkowitą nowość.

Autor postawił sobie wyjątkowo trudne zadanie, gdyż nawiązuje do prac badawczych, które nie należą do zamierchłej historii nauki, lecz są nauką nieomal *in statu nascendi*, co dotyczy w szczególności matematycznych założeń przyjmowanych przez zwolenników koncepcji aktywnego wnioskowania, opartej na zasadzie energii swobodnej. Przedstawienie istoty tego podejścia bez epatowania czytelnika zbędnymi szczegółami matematycznymi (co prowadzi niekiedy w pracach zwolenników aktywnego wnioskowania do czegoś, co Kazimierz Twardowski określiłby z pewnością mianem „symbolomanii”) jest również zadaniem wyjątkowo niewdzięcznym. Z zadania tego habilitant wywiązał się jednak bardzo dobrze, skupiając się przede wszystkim na filozoficznych i teoretycznych założeniach opisywanej tradycji badawczej, a jednocześnie dochowując wierności zasadzie życzliwości. Zaznacza bowiem w obszernych przypisach (i odesłaniach do bogatej bibliografii), że wiele założeń (zwłaszcza idealizacyjnych) przyjmowanych obecnie w tej tradycji badawczej nadal podlega ciągłej modyfikacji.

Zanim przystąpię do krótkiego omówienia zawartości poszczególnych rozdziałów, przedstawię charakterystykę podstawowej idei, która stanowi, moim zdaniem, samo sedno przedstawionego do oceny osiągnięcia. Dr Piekarski precyzyjnie wylicza główne hipotezy postawione w monografii (s. 22-25), ale wydaje się, że można w uproszczeniu przedstawić je tak: jeśli procesy poznawcze można poprawnie wyjaśnić w kategoriach przyczynowo-skutkowych, zgodnie z założeniami tzw. nowego mechanicyzmu w filozofii nauki, to wyjaśnienia te będą miały charakter normatywny. Będą bowiem one dotyczyć tego, jak podmioty powinny modyfikować swoje przekonania bądź działania, aby pozostały podmiotami, których natura jest taka, że realnie korzystają z bayesowskiego wnioskowania probabilistycznego. Wskazana powinność, czy raczej – jak określa to autor – normatywność nie ma charakteru etycznego, lecz „konstytutywny”, tzn. dotyczy pewnego rodzaju norm lub standardów działania powiązanych z istotą samych bayesowskich podmiotów poznawczych. Dr Piekarski wiąże tę odmianę normatywności z pojęciem funkcji, wpisując się w



tradycję naturalistycznej filozofii (obecną zwłaszcza w pracach Ruth Millikan i Marka Bickharda). Są to wszystko dosyć mocne i oryginalne tezy, których obrona prowadzona jest systematycznie w kolejnych rozdziałach pracy.

Moja charakterystyka treści poszczególnych rozdziałów będzie stosunkowo krótka, gdyż sam autor bardzo dobrze przedstawia ich zawartość zarówno w autoreferacie, jak i w zakończeniu opiniowanej rozprawy. We wprowadzeniu habilitant przedstawia wstępną charakterystykę pojęcia normatywności oraz stawiane w pracy hipotezy. W przedmowie (która, moim zdaniem, powinna poprzedzać wprowadzenie, gdyż ma charakter bardziej osobisty) autor wskazuje m.in. źródła swoich inspiracji i dziękuje za uwagi.

W pierwszym rozdziale wprowadza pojęcie przetwarzania predykcyjnego; interesującym wątkiem w tym rozdziale jest pogłębiona analiza transcendentalizmu kantowskiego. Habilitant przekonująco i dogłębnie argumentuje, że wbrew panującym powszechnie opiniom przetwarzanie predykcyjne nie ma charakteru kantowskiego, gdyż nie opiera się na argumentach transcendentalnych. W rozdziale tym charakteryzowane są umiarkowane i radykalne wersje przetwarzania predykcyjnego oraz wskazywane są elementy tej koncepcji, które podlegały krytyce. W rozdziale drugim autor skupia się na powiązaniu przetwarzania predykcyjnego z bayesowskim modelem wnioskowania probabilistycznego. Krytycznie też ustosunkowuje się do argumentów dotyczących preskryptywnego bądź deskryptywnego charakteru tego wnioskowania. Rozdział trzeci poświęcony został zasadzie energii swobodnej, bronionej przez K. Fristona i jego współpracowników. Rozdział czwarty służy obronie odpowiedniego pojęcia funkcji, które powiązane zostaje z pojęciem mechanizmu (w rozumieniu nowego mechanicyzmu). W dużej mierze propozycja dr. Piekarskiego jest zbieżna z tym, czego broni J. Garson w klasycznym artykule na temat mechanizmów funkcjonalnych (2013): nie istnieją mianowicie mechanizmy (funkcjonalne) odpowiedzialne np. za schizofrenię; tego rodzaju zjawiska, jako patologiczne, należy pojmować jako dysfunkcjonalne. W odróżnieniu od Garsona Piekarski posługuje się terminem „mechanizm normatywny” (co jest



różnicą jedynie terminologiczną), ale też znacznie mocniejszym pojęciem funkcji (co już jest różnicą merytoryczną). To pojęcie funkcji opracowane zostaje w wyniku krytycznej analizy pojęcia funkcji systemowej bronionej przez Marka Bickharda, notabene bardzo zdecydowanego przeciwnika przetwarzania predykcijnego, ale podzielającego wiele założeń z jego zwolennikami (np. o pojęciowym związku działania z poznaniem).

Ostatni, piąty rozdział poświęca dr Piekarski problemowi ciemnego pokoju, który jest uznawany w literaturze przedmiotu za najpoważniejszy zarzut wobec przetwarzania predykcijnego i aktywnego wnioskowania. Argumentuje on, że rozwiązania przyjmowane wcześniej są nieadekwatne. Wskazuje wręcz, że ani przetwarzanie predykcyjne, ani aktywne wnioskowanie nie ma możliwości, aby się przed tym zarzutem wybronić. Twierdzi w związku z tym, że przetwarzanie predykcyjne nie może być kompletną teorią (pretendującą do miana ogólnej teorii wszystkiego, przynajmniej w kognitywistyce, jeśli nie we wszystkich naukach o życiu, do czego skłaniają się zwolennicy aktywnego wnioskowania tacy jak K. Friston), lecz musi być uzupełnione o elementy zewnętrzne. Musi być ono mianowicie zintegrowane z pojęciami motywacji i nastawienia, które wyznaczają ograniczenia mechanizmom normatywnym przetwarzania predykcijnego. W zakończeniu podsumowuje cały przebieg argumentacji i przypomina bronione hipotezy.

Rozdział piąty zasługuje na szczególną uwagę, gdyż tu przejawia się bardzo oryginalna i kontrowersyjna idea autora, a mianowicie, że istnieją pewne obiekty, nazywane przez niego „ograniczeniami”, które tak modyfikują działanie mechanizmów normatywnych, że nie dochodzi do powstania problemu ciemnego pokoju. Problem ten ma powstawać, zdaniem przeciwników przetwarzania predykcijnego, dlatego, że zgodnie z tą koncepcją poznanie ma polegać na minimalizacji błędu przewidywania. Minimalizacja ta może polegać na modyfikacji przekonań i pragnień bądź na działaniu (czyli dostosowywaniu świata do przekonań i pragnień). Innymi słowy, podmiot działający ma postępować tak, by rozbieżność między jego przewidywaniami a rzeczywistością była (zwykle dodaje się tu: na dłuższą metę lub przeciętnie) jak najmniejsza.



Tymczasem najprostszą metodą, aby zminimalizować predykcje, jest udać się do ciemnego pokoju, w którym odczuwa się minimalnie mało bodźców. Przewidywanie bowiem, że będzie dalej ciemno i nic nie będzie widać, czuć, słyszeć itd. jest banalnie proste i skuteczne. Tylko że ludzie tak nie działają (chyba że idą spać). Najprostsze rozwiązanie, które też krytycznie omawia Piekarski, polega na przyjęciu, że ewolucja wbudowała w podmioty działające założenia, iż nie są one w ciemnym pokoju, stąd będą one aktywnie eksplorować rzeczywistość, a nie siedzieć nieruchomo w ciemnym pokoju, by umrzeć w nim z głodu. Tymczasem jest to rozwiązanie niezadowolające, np. dlatego, że niektóre organizmy (bądź ich populacje) lubią spać w ciemnych pokojach przez większość doby, ale też z powodu zmienności międzyosobniczej. Zdaniem Piekarskiego, dopiero relacje motywacyjne z otoczeniem oraz nastawienie (termin „nastawienie” stosuje Piekarski na oznaczenie wewnętrznych relacji motywacyjnych) mogą doprowadzić do odpowiedniego wyboru, kiedy podmiot jest postawiony przed decyzją: zmieniać świat czy własny model świata. Dopiero to może zmniejszać liczbę stopni swobody mechanizmu predykcyjnego, a zatem działać jako jego ograniczenie w sensie analizowanym przez dr. Piekarskiego. Co prawda, przedstawione przez habilitanta rozwiązanie problemu jest szkicowe (w tym sensie, że nie sposób byłoby zbudować na jego podstawie modeli obliczeniowych, co jest główną metodą modelowania umysłu w przetwarzaniu predykcyjnym), ale na pewno jest godne uwagi i oryginalne. Kontrowersyjne jest ono z tego względu, że podważa założenia bardzo wielu interpretacji przetwarzania predykcyjnego, a przy tym nie jest do końca jasne, w jaki sposób te relacje mają wpływać na działanie modeli generatywnych. Czy te relacje motywacyjne są też w jakiś sposób wprost modelowane w postaci rozkładów prawdopodobieństw? Czy jakoś inaczej? Rozprawa habilitanta sugeruje chyba to drugie, ale trzeba by też się zastanowić, czy zwolennik kodowania predykcyjnego nie mógłby twierdzić, że tzw. hiperpriors w modelach generatywnych nie są właśnie odpowiednikami relacji motywacyjnych. Mówiąc inaczej, rozwiązanie habilitanta pobudza do wielu nowych pytań (w odróżnieniu od pozostałych prób obrony przetwarzania predykcyjnego, gdzie problem ciemnego pokoju zbywa się często jako problem pozorny). Wydaje się, że warto tę koncepcję rozwijać dalej i precyzować.



W pracy habilitacyjnej ujawnia się bardzo ryzykowna strategia badawcza dr. Piekarskiego. Otóż tam, gdzie typowa analiza pojęciowa wskazywałaby na wieloznaczność pewnych pojęć (np. pojęcia „ograniczenie” w rozdziale V), habilitant szuka podobieństw. Jest to o tyle ryzykowne, że naraża się w ten sposób na zarzut popełnienia błędu ekwiwokacji. Jednocześnie analizy te przypominają w pewnej mierze wyniki wariacji ejdetycznej, a więc metody fenomenologicznej: dr Piekarski skupia się raczej na wspólnym rdzeniu znaczeniowym pojęć stosowanych przez dosyć różnych autorów, takich jak np. H. Pattee, W. Bechtel czy C. Craver (w przypadku „ograniczenia”). Pozwala mu to wskazać na to, co łączy te podejścia, a nie tylko poprzestać na skonstatowaniu istnienia różnic. Można by rzec, że ta strategia jest elementem metafizologii habilitanta. Mimo początkowego sceptycyzmu stwierdzam, że istotnie prowadzi ona do ciekawych spostrzeżeń. Udaje mu się dostrzec las, gdzie inni widzą tylko różne drzewa.

Podsumuję teraz ocenę osiągnięcia habilitacyjnego. Przedstawiona do oceny monografia świadczy o dojrzałości badawczej dr. Piekarskiego, jego rozległej erudycji oraz wysokich kompetencjach filozoficznych, a także odwadze, by zajmować się problemami trudnymi – nawet gdy stanowią coś w rodzaju ruchomego celu, gdy założenia badawcze przyjmowane przez krytykowanych przezeń autorów ulegają ciągłym zmianom. Nie ulega dla mnie wątpliwości, że praca ta stanowi osiągnięcie naukowe stanowiące znaczny wkład w rozwój jego dyscypliny naukowej (a więc filozofii) w rozumieniu art. 219 Prawa o szkolnictwie wyższym i nauce z dn. 20 lipca 2018

Dr Piekarski załączył szczegółowy raport bibliometryczny, z którego wynika, że jego sumaryczny impact factor wynosi 9,366, co jest stosunkowo wysokim wynikiem dla osób na jego etapie kariery naukowej (wyniki te wynikają jednak z publikowania w czasopismach w dużej mierze niefilozoficznych, lecz publikujących sporadycznie prace filozofów). Cytowania w bazie Web of Science są niewysokie (2), podobnie jak indeks Hirscha (1). Wg Google Scholar (w Publish or Perish) jest cytowany 49 razy, a indeks Hirscha wynosi 4. To wynik dosyć dobry, a jakościowa



analiza cytowań wskazuje, że prace habilitanta na temat normatywności przetwarzania predykcijnego oddziałują już na znanych i cenionych autorów (np. J. Hohwy'ego).

Zajmę się teraz jego pozostałym dorobkiem. Praca doktorska dr. Piekarskiego również ukazała się w postaci monografii w j. polskim. Habilitant po uzyskaniu stopnia doktora opublikował 27 artykułów, 2 rozdziały w monografiach oraz wygłosił 20 referatów na konferencjach krajowych i tyleż na międzynarodowych. Tematyka tych prac jest bardzo zróżnicowana: od przetwarzania predykcijnego, przez filozofię Ludwiga Wittgensteina (tematyka rozprawy doktorskiej), etykę zwierząt po polską fenomenologię. Sama liczba artykułów świadczy o dużej pracowitości autora i znacznie przewyższa dorobek wielu innych kandydatów do stopnia doktora habilitowanego. Część z tych prac ukazała się w dobrych czasopismach międzynarodowych i spotkała się z pewnym odzewem (m.in. krytyczny artykuł o etyce zwierząt w podejściu Davida Gunkela i Marka Coeckelberga, na który krytykowani autorzy odpowiedzieli).

Kierował jednym projektem badawczym NCN: w wyniku grantu MINIATURA opublikował artykuł w języku angielskim w czasopiśmie „Frontiers in Psychology”. Od roku 2016 współpracuje naukowo z Projektem Avant, co pozwala mu współpracować z badaczami z kraju i zagranicy. Był współorganizatorem lub głównym organizatorem 9 konferencji czy sympozjów naukowych we współpracy z różnymi instytucjami naukowymi (m.in. IFiS PAN). Recenzował teksty dla wielu czasopism naukowych, w tym cenionych czasopism międzynarodowych („Philosophy & Technology” i „Axiomathes”). Był też członkiem redakcji czasopisma „Kronos” (2011-2016). Z dostarczonej dokumentacji wnioskuję, że był recenzentem prac dyplomowych, lecz nie występował w roli promotora. Był członkiem komitetu okręgowego Olimpiady Filozoficznej w Warszawie (2017), prowadził wykłady z filozofii umysłu dla Uniwersytetu III wieku w Milanówku i gościł w popularyzatorskich audycjach radiowych. Autoreferat wskazuje, że niedawno zdobył uprawnienia pedagogiczne, co zapewne będzie związane z jego aktywnością jako nauczyciela filozofii w szkolnictwie średnim.



W świetle tego, co zostało wyżej napisane, uważam, że pan dr Michał Piekarski jest w pełni dojrzałym badaczem i ma solidny dorobek naukowy, stanowiący znaczny wkład w rozwój filozofii. Opublikował już dwie monografie naukowe i kilkadziesiąt artykułów. Ma zacięcie popularyzatorskie.

Stawiam zatem wniosek o nadanie drowi Michałowi Piekarskiemu stopnia doktora habilitowanego w zakresie filozofii.

dr hab. Marcin Miłkowski

Warszawa, 08.07.2021

