

# Maszyna do ćwierkania w pojemniku z nirwaną

## Recenzja książki *Mechanizmy predykcyjne i ich normatywność*

Tytuł: *Mechanizmy predykcyjne i ich normatywność*

Autor: Michał Piekarski

Wydawnictwo: Liberi Libri

ISBN: 978-83-63487-44-7

[Wersja elektroniczna w wolnym dostępie](#)

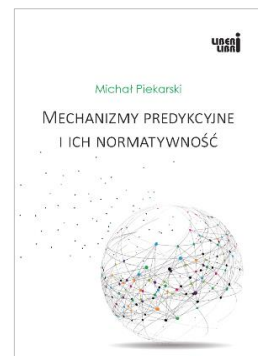
Rok wydania: 2020

Liczba stron: 362

**Witold Wachowski** 

Uniwersytet Marii Curie-Skłodowskiej w Lublinie

[witold.wachowski@mail.umcs.pl](mailto:witold.wachowski@mail.umcs.pl)



Przyjęto 4 września 2022; zaakceptowano 26 września 2022; opublikowano *Online First* 7 października 2022.

### Abstrakt

Artykuł ten jest recenzją książki *Mechanizmy predykcyjne i ich normatywność*, skoncentrowaną na trzech zakresach tej publikacji. Pierwszy zakres obejmuje kontekst wydawniczy, czyli kwestię zapotrzebowania na taką pracę, biorąc pod uwagę język publikacji, jej dostępność itd. Drugi zakres dotyczy akademickiego potencjału tej książki jako obszernego opracowania koncepcji przetwarzania predykcyjnego. Trzeci zakres natomiast wiąże się z główną ambicją Autora książki i jednocześnie najbardziej wymagającym wątkiem tej pracy, jako studium problemu normatywności w ujęciu koncepcji przetwarzania predykcyjnego i wyjaśniania (neo)mechanistycznego.

**Słowa kluczowe:** energia swobodna; mechanizm; model generatywny; mózg; normatywność; przetwarzanie predykcyjne

*Jak w biciu dzwonów można się dosłuchać wszystkich możliwych słów,  
tak i widzialność świata wyrazić można na różne rodzaje i sposoby,  
a nikt nie ma prawa twierdzić, że jedno jest prawdziwsze od drugiego.*

(Heinrich Wölfflin, 2017<sup>1</sup>)

## 1. Wprowadzenie

Kiedy przed laty brałem udział – tylko jako słuchacz – w kameralnym wydarzeniu konferencyjnym w Warszawie, które poświęcono badaniom nad przetwarzaniem predykcyjnym, mało co rozumiałem z wygłaszanych tam referatów. Oczywiście, z powodu mojego nieprzygotowania. Drażniło mnie przy tym używanie terminów, które najwyraźniej były kalkami językowymi. Referaty były polskojęzyczne, co jednak w niczym nie pomogło, może nawet przeszkodziło. Tematu predykcji jednak nie mogłem zamieść sobie pod dywan, gdyż zaczął powracać tu i ówdzie, coraz częściej i w nieoczekiwanych kontekstach, zagarniając coraz większy obszar problemowy w kognitywistyce. Trochę tak jak niegdyś temat ucieleśnienia poznania. Pomogło więc zapoznanie się z ważniejszymi publikacjami anglojęzycznymi, wysłuchanie wykładów na kolejnych konferencjach, udział w seminariach i dyskusjach. Ostatecznie, chociaż przetwarzanie predykcyjne nie stało się moim pierwszoplanowym tematem badawczym, to uwzględniam je coraz częściej we własnych pracach, poświęciłem mu również kurs dla studentów kognitywistyki.

I w obu tych kontekstach próbuję ocenić wartość książki Michała Piekarskiego *Mechanizmy predykcyjne i ich normatywność*, wydanej przeszło rok temu w Wydawnictwie Liberi Libri. Ponadto, oczywiście, omawiam pokrótce wywiązanie się Autora z realizacji głównego celu badawczego, związanego z problematyką normatywności.

Niezorientowanym w temacie zasygnalizuję, że zgodnie z przedstawianą tu koncepcją system poznawczy (mózg, umysł) można traktować jako rodzaj „maszyny” predykcyjnej, która dzięki minimalizowaniu błędu predykcyjnego (rozbieżność między danymi a oczekiwaniami) wspomaga – a może wręcz umożliwia – skuteczność percepcji i działania w świecie. Wyraziłem to jednak w wielkim, abstrakcyjnym skrócie pozbawionym problematyzowania.

Dość nietypowym odniesieniem (o nieoczywistej przydatności) będzie tu dla mnie „Maszyna do ćwierkania” Paula Klee z 1922 r. – obraz, o którym złośliwi lub zniecierpliwieni powiedzą, że pasuje do każdej interpretacji, czy to będzie triumf techniki nad naturą, czy też odwrotnie: natury nad techniką. „Maszynę” obejrzeć można sobie [tutaj](#). I z tym odniesieniem na razie się rozstaję, by powrócić do niego w ostatniej części recenzji.

---

<sup>1</sup> Motto przechwycone z portalu <http://www.isztuka.edu.pl> (dostęp 18.06.2022).

## 2. Kontekst

Książka Piekarskiego – jako recenzowana monografia – ukazała się w Polsce w wolnym dostępie i na licencji Creative Commons CC 4.0. Każdy z tych faktów wart jest odnotowania, zarówno dlatego, że polityka publikacji książek w wolnym dostępie i na jednej z wolnych licencji nie jest taka prosta i oczywista, w odróżnieniu od trendu udostępniania w ten sposób artykułów w czasopiśmie, jak i dlatego, że nadal trudniej zdobyć w wersji elektronicznej polską niż zagraniczną książkę naukową. Mowa tu, oczywiście, nie o przysłowiowym „wydawnictwie Krzak”, tylko o docenionej oficynie wydawniczej, w której ukazują się potencjalnie ważne prace naukowe.

Wydaje się to istotne, gdy brać pod uwagę ambicje czy też potencjał danej pracy, która nie powstała sobie a Muzom. Mowa przecież o pierwszej polskojęzycznej monografii na względnie świeży temat naukowy, żywo dyskutowany w nurcie kognitywistyki i nie tylko.

Poza tą książką mamy do dyspozycji kilka polskojęzycznych artykułów naukowych, gdzie tematyka przetwarzania predykcyjnego jest na pierwszym lub drugim planie. Nie należy jednak zapominać o publikacjach książkowych częściowo bliskich tej tematyce – czy to w historii nauki, czy to w wymiarze popularyzatorskim, czy technicznym zastosowaniu. Przykładem pierwszego jest monografia *Oko i mózg. Psychologia widzenia* Richarda L. Gregory’ego (1971), która nadal, choć od wydania samego przekładu minęło pół wieku, ciekawie i przejrzyście wprowadza w badania nad mechanizmami percepcji wzrokowej, odnosząc zjawiska związane z przewidywaniem, uzupełnianiem czy odgadywaniem między innymi do psychologii postaci. Jako przykład drugiego typu wspomnianych publikacji wskażę tu książkę *Mózg nie służy do myślenia (7 i ½ wywrotowych lekcji o mózgu)* Lisy Feldman-Barrett (2021), gdzie w rozdziale „Nasz mózg przewiduje (niemal) wszystko, co robimy” autorka przekazuje trochę zbyt gładko i bez problematyzowania „prawdy naukowej” o tym zjawisku (publikacja przyjemnie wprowadzająca do tego, co w nim ciekawe, osoby spoza świata akademickiego, a nawet studiujące – ale to wszystko, w tekście brak bezpośredniego osadzenia w literaturze tematu i dylematach badawczych). Preferuję jednak inny rodzaj popularyzatorskiego przedstawiania badań naukowych: taki, jaki w kontekście tematu przetwarzania predykcyjnego proponuje Paweł Gładziejewski w artykule „Percepcja jako kontrolowana halucynacja” na łamach magazynu *Filozofuj!* (2017). Czyli z perspektywy filozofii nauki. A jako przykład trzeciego rodzaju przywoływanej tu literatury wymienię książkę (również przekład) *Deep learning i modelowanie generatywne. Jak nauczyć komputer malowania, pisanie, komponowania i granie* autorstwa Davida Fostera (2021), w której modelowanie generatywne, w tej recenzji związane z przetwarzaniem predykcyjnym, odnosi się do uczenia maszynowego. Mimo mocno praktycznego, inżynierskiego zastosowania, pozycja ta we fragmentach przydać się może również studentom zapoznającym się z tematem predykcji w kognitywistycznym ujęciu (co sprawdziłem już na moich

studentkach i studentach). Mam jedynie pewne zastrzeżenia do samego tłumaczenia pracy, w tym do ostrożnościowego pomijania w przekładzie niektórych pojęć i sformułowań – co widać już w jej tytule.

W takim to – z grubsza biorąc – kontekście polskojęzycznych publikacji sięgamy dzisiaj po książkę Piekarskiego. Nie znaczy to, że nie warto zupełnie odnosić jej do publikacji anglojęzycznych – chociażby do wydanej niedawno pracy autorstwa zasłużonego w tym temacie Karla Fristona oraz Thomasa Parra i Giovanniego Pezzulo *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior* (2022). Przyglądam się jednak tutaj książce polskojęzycznej, dlatego w tym zakresie w dużej mierze ją oceniam.

### 3. Koncepcja

W sekcji tej odnoszę się zarówno do koncepcji przetwarzania predykcyjnego, jak i koncepcji samej książki *Mechanizmy predykcyjne i ich normatywność*.

Kiedy zapoznajemy się z jakąś teorią, nurtem czy tradycją badawczą, często przedstawia się nam (a) historię, (b) zestaw idei i pojęć, jak i (c) różne odmiany, a czasem też etapy. Jednak monografia Piekarskiego nie przystaje tak łatwo do tej uproszczonej struktury, ponieważ jest przede wszystkim efektem autorskiego projektu badawczego, a nie publikacją o charakterze podręcznikowym. Dlatego też najpierw zrekonstruuje zawartą tam koncepcję przetwarzania predykcyjnego, a następnie odniosę się do zawartości poszczególnych rozdziałów – nieprzesadnie je referując, ponieważ, co zaznaczam, robi to dość starannie sam Autor w Przedmowie.

Omawiana koncepcja jest stosunkowo nowa. Na tyle nowa, że trzymam się dość ostrożnościowego, ogólnego określenia „koncepcja”, jako że nie mamy tutaj jeszcze, jak wskazują krytycy, do czynienia z teorią w bardziej wymagającym sensie „teorii”, która powinna służyć przewidywaniu, wyjaśnianiu i kontroli badań, działając jako zintegrowana całość (zob. Litwin i Miłkowski, 2020).

Trzeba jednak pamiętać, ile zrobiono już wcześniej w filozofii i nauce na rzecz opisania i zrozumienia fenomenu predykcji w procesach poznawczych<sup>2</sup>. Jako pionier z odległej przeszłości jawi się tutaj Ibn al-Hajtam (X–XI w.), który analizował wpływ wnioskowania na percepcję. Z oczywistych względów należy wskazać tu na Immanuela Kanta (XVIII–XIX w.), rozważającego w tym kontekście między innymi aktywny charakter postrzegania. Ważną postacią jest też Hermann von Helmholtz (XIX w.), zresztą neokantysta, analizujący mózg jako urządzenie testujące hipotezy. Zasłużona w tej tematyce jest wspomniana już psychologia postaci (np. Koffka, 1922), jeśli chodzi o zagadnienie strukturalizowania bodźców percepcyjnych. Istotnych wskazówek dostarczyły badania nad

---

<sup>2</sup> Przykłady, których nie zaczerpnąłem z prac Piekarskiego, opatrzone zostały osobnymi przypisami.

transmisją obrazu telewizyjnego (jakie prowadził na przykład Ernest R. Kretzmer w latach 50.), w których zwrócono uwagę na ograniczenie przesyłanych informacji do sygnału zmian. Kolejni zasłużeni to Donald MacKay, Ulrich Neisser i Richard Gregory – przedstawiciele psychologii „analizy przez syntezę”; według nich mózg nie tworzy wewnętrznego modelu świata na podstawie oddolnych danych, lecz aktywnie stwarza reprezentacje, które potem dopasowuje do danych. Innego rodzaju podwalin pod koncepcję dostarczył Philip Johnson-Laird (1994), analizujący myślenie probalistyczne oraz manipulowanie modelami umysłowymi jako zbiorami (zbiorów) możliwości. Swego rodzaju wstępem do omawianej tu koncepcji była koncepcja emulacji Ricka Grusha (1995), zgodnie z którą mózg konstruuje modele ciała lub otoczenia, używane jako surogaty, i zastępuje rzeczywistą interakcją z modelami. Wreszcie – przywołać trzeba tu (również za Piekarskim) publikację uważaną za przełomową dla koncepcji przetwarzania predykcyjnego: artykuł Rajesha Rao i Dany Ballarda pt. „Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects” opublikowany w 1999 r. w *Nature Neuroscience*. W oparciu o badania empiryczne przedstawiono tam model przetwarzania informacji wzrokowej, zgodnie z którym informacje przekazywane z wyższych warstw kory mózgowej obejmują przewidywania dotyczące aktywności neuronów w jej niższych warstwach.

Na wstępie przedstawiłem już w dużym skrócie ideę przetwarzania predykcyjnego, zgodnie z którą system poznawczy to urządzenie predykcyjne wspierające percepcję i działanie. Rozwinę to w oparciu o pracę Piekarskiego. System poznawczy wytwarza wielopoziomowy, hierarchiczny model generatywny swojego otoczenia. W ramach tego modelu przetwarza informacje w sposób kaskadowy: odgórnie wysyła przewidywania, aby zminimalizować błąd predykcyjny związany z sygnałami oddolnymi. Błąd predykcyjny dotyczy dysproporcji między oczekiwaniami wytworzonymi w modelu generatywnym a zmienną informacją ze zmysłów. Błędy predykcyjne mogą być minimalizowane na dwa sposoby. Jeden sposób opiera się na wnioskowaniu biernym (percepcyjnym), czyli prowadzi do rewizji przekonań we własnym modelu generatywnym. Drugi sposób polega na wnioskowaniu aktywnym, a tym samym na działaniu w środowisku w celu wprowadzenia w nim korzystnych zmian ze względu na wytworzone przewidywania; w tym przypadku przewidywania służą jako instrukcje motoryczne, a działanie minimalizuje niepewność.

Zarysowane tutaj mechanizmy przetwarzania predykcyjnego zasadniczo są wspólne najbardziej znanym odmianom tej koncepcji – a trzeba wiedzieć, że rozwinęły się one w zróżnicowanych kierunkach. Wydaje się, że najbardziej popularną (być może też z różnych względów przydatną) odmianą jest kierunek rozwijany przez Karla Fristona i współpracowników (np. 2009). Wiąże on przetwarzanie predykcyjne z tak zwaną zasadą swobodnej energii. Jak przyjmuje Friston, ogólną zasadą naszego życia jest minimalizowanie tej energii, czyli utrzymywanie porządku, a przez to unikanie zwiększania entropii. Można

jednak różnicować odmiany tej koncepcji z uwagi na stosunek do indywidualizmu metodologicznego, a więc porządkując je według opcji: internalizm–eksternalizm, co u Piekarskiego wiąże się z podziałem tych odmian na konserwatywne i radykalne (rozwinę to niżej).

Jak ma się do takiego ujęcia książka Piekarskiego, jeżeli traktować ją (na razie) jako swoiste kompendium na ten temat? Można powiedzieć, że obejmuje ono – choć nie w pełnym stopniu – trzy pierwsze spośród pięciu rozdziałów tej monografii. Trzeba przy tym wziąć pod uwagę, że spis treści jest dość „gęsty”, więc nie tak łatwo a pierwszy rzut oka zidentyfikować w nim elementy opracowania koncepcji pasującego do schematu, który zarysowałem. Przyczyna tego, jak już pisałem, leży w głównie badawczym, a nie podręcznikowym charakterze tej pracy. Ponieważ jednak doceniam ją również pod kątem przydatności dydaktycznej czy ogólnoakademickiej, nadmienię, że Wprowadzenie i Przedmowa dość przejrzysto i szczegółowo przygotowują czytelników do lektury całości i dostarczają im swoistej mapy.

W Rozdziale I. „Pojęcie przetwarzania predykcyjnego” znajdziemy wyjaśnienie, dlaczego Autor posługuje się terminem „przetwarzanie predykcyjne”, a nie „kodowanie predykcyjne”: z tego drugiego nie chce korzystać, „ponieważ odnosi się ono głównie do takich strategii kodowania, w jakich przewidywane części kodu sygnału wejściowego są odejmowane od rzeczywistego sygnału. [...] W takich ujęciach mniejszą uwagę przywiązuje się do idei hierarchicznego i wielopoziomowego modelu generującego predykcje” (s. 36). Napomknę przy tym, że w książce przewija się skrót „PP” oznaczający tę koncepcję, z którego skorzystam i ja w dalszym toku recenzji. Rozdział ten jest wyjątkowo pojemny i zróżnicowany jak na taki tytuł – ale to również stanowi jego zaletę. Znajdziemy tu bowiem nie tylko najważniejsze kategorie i podstawową strukturę tej koncepcji, lecz także przedstawienie historycznych prac pionierskich w tym względzie (z naciskiem na Kanta, lecz z dużą dozą sceptycyzmu wobec jego obecności w PP), jak również krytykę koncepcji oraz związany z nią podział jej odmian na konserwatywne (gdzie całość przetwarzania predykcyjnego sprostowana jest do pewnych mechanizmów neuronalnych) i radykalne (a więc inaczej niż w pierwszych, stąd też pojawiły się tutaj ujęcia enaktywistyczne czy związane z psychologią ekologiczną, w tym propozycja Nico Orlandi [np. 2017]). Jak można się domyślać, służy to również celom badawczym w książce, a nie tylko wynika z opracowania koncepcji jako takiej. Dzięki takiej zawartości jednak rozdział ten stanowi (w dużej mierze) swoisty przegląd polskojęzycznej na użytek akademicki, choć warto go połączyć – jak wspominałem – ze wstępnymi partiami książki. W rozdziale tym zwraca również uwagę obrona PP przed zarzutem antyrealizmu.

Rozdział II. „Przetwarzanie predykcyjne jako bayesowski model eksplanacyjny” przedstawia niezwykle ważne dla omawianej koncepcji zastosowanie probalistycznej sieci bayesowskiej. Zarówno rozdział, jak i sam temat wymaga od czytelników zwiększonej uwagi. Rzecz w tym, że model generatywny ma

charakter statystyczny, przekłada się na rozkład prawdopodobieństw. Najbliższym mu typem sieci probalistycznej jest właśnie model bayesowski. Samo twierdzenie Bayesa przedstawia, jakie jest prawdopodobieństwo zajścia jakiegoś zdarzenia na podstawie wcześniejszej wiedzy o możliwych warunkach związanych z tym zdarzeniem w świetle dostępnych obserwacji czy danych. Tak zwane wnioskowanie bayesowskie jest zastosowaniem owego twierdzenia i umożliwia aktualizację prawdopodobieństwa zajścia zdarzenia na podstawie pojawiania się nowych danych z nim związanych oraz na podstawie znanych prawdopodobieństw. Piekarski przedstawia tezę, że w systemie poznawczym (zasadniczo mowa tu o mózgu) zaimplementowana jest statystyczna sieć probabilistyczna, która w pewien sposób realizuje wnioskowanie bayesowskie. Dochodzi do tego na drodze jego aproksymacji, czyli z wykorzystaniem innej prostszej funkcji, którą pewni badacze wiążą z wariacyjną energią swobodną (por. Clark, 2013; Friston, 2011) – co bywa przez innych krytykowane. Jednak Piekarski podejmuje się również uzasadnienia, że zawarty w PP model bayesowski ma charakter nietrywialny: należy zakładać tutaj realizm. Innymi słowy, Autor pokazuje, że zjawiska predykcyjne należy wyjaśniać za pomocą twierdzenia Bayesa, gdyż są one bayesowskie z natury. Wyłożenie i uzasadnienie tej tezy nie sprowadza się do tego rozdziału, lecz rozciąga się także na dwa ostatnie.

Natomiast Rozdział III. „Zasada energii swobodnej w przetwarzaniu predykcyjnym” dostarcza szczegółowego rozwinięcia roli tytułowej zasady (wspomianej przeze mnie już wcześniej) w tej koncepcji. Przede wszystkim osadzenie przetwarzania predykcyjnego na zasadzie swobodnej energii (anglojęzyczny skrót: „FEP”) jest jednym z kilku przyjmowanych, możliwych algorytmów dopasowania modelu generatywnego do aktualizowanych danych zmysłowych, aczkolwiek obecnie najbardziej popularnym. Według Fristona, celem samoorganizującego się systemu jest minimalizacja (wariacyjnej) energii swobodnej, tę zaś należy tu rozumieć w sensie statystyczno-informacyjnym, jako rozbieżność między informacją w systemie a faktami. Konsekwencją tego może być szersza teza, że życie zasadniczo polega na minimalizowaniu zaskoczenia. Piekarski podchodzi uważnie i krytycznie do założeń kluczowości FEP dla przetwarzania predykcyjnego, próbując ją traktować raczej jako cenną heurystykę badawczą w PP.

#### **4. Problem**

Autor książki skupia się ostatecznie na problemie badawczym, który łączy jego zainteresowania tematem normatywności z koncepcją przetwarzania predykcyjnego, w oparciu o wyjaśnianie (neo)mechanistyczne. Problematyka normatywności pojawia się w jego publikacjach od połowy ubiegłej dekady (np. Piekarski, 2016), poświęcił jej redagowany przez siebie zbiór tematyczny (zob.

Piekarski, 2018), w międzyczasie też powiązał z zagadnieniem predykcji poznawczych (np. Piekarski, 2017), przeprowadził rozległe analizy, które zaowocowały recenzowaną tu monografią, a w roku bieżącym przedstawił najnowsze studium normatywności mechanizmów predykcyjnych (Piekarski, 2022).

Zarysowany przez Autora na wstępie książki cel to obrona przekonania, że procesy poznawcze obejmują pewien komponent normatywny, zawarty w sposobie, w jaki ludzie, czy ściślej mówiąc, (przynajmniej niektóre) systemy poznawcze „zaopatrzone” w modele generatywne, doświadczają świata, w tym także innych ludzi. Komponent ten przekłada się na pewne mechanizmy poznawcze o charakterze normatywnym, są one bowiem kluczowe dla niektórych ludzkich działań i zachowań. Piekarski podejmuje się analizowania ich jako mechanizmów predykcyjnych, gdyż podstawowa ich funkcja sprowadzać się ma do generowania predykcji. Mechanizm generujący takie predykcje określany jest mianem „mechanizmu normatywnego”. Należy przy tym odnotować naturalistyczne nastawienie Autora do metodologii przedsięwziętych badań, które rozwijane są w kontakcie filozofii z nauką.

W Rozdziale IV. „Funkcje i mechanizmy normatywne w kontekście przetwarzania predykcyjnego” mamy już precyzyjne osadzenie normatywnego PP w filozofii mechanicyzmu. Ta ostatnia, mówiąc pokrótce, związana jest z modelem wyjaśniania w kategoriach mechanizmu, ten zaś to „struktura pełniąca określone funkcje na podstawie części składowych, operacji składowych i ich organizacji. Zorganizowane działanie mechanizmu odpowiada za jedno lub więcej zjawisk” (Bechtel, 2008, s. 13, za: Piekarski, 2020, s. 72). Oczywiście, Piekarski podejmuje się określenia pojęcia i funkcji mechanizmu normatywnego. Argumentuje też za tym, że mechanizmy predykcyjne nie mają charakteru homeostatycznego. Ponadto próbuje wykazać, że pełne wyjaśnienie mechanizmów normatywnych wymaga wyjaśnienia funkcji normatywnych. Te ostatnie są takim typem funkcji biologicznych, który spełniają następujące warunki: (1) pozwalają wyjaśnić dysfunkcje określonego mechanizmu, (2) przyczyniają się do utrzymania stabilności organizmu, a także – w odniesieniu do funkcji reprezentowania i funkcji predykcyjnej – (3) można za ich pomocą wyjaśnić możliwość przypisywania wartości logicznych określonym reprezentacjom bądź predykcjom. Mechanizm realizujący tego rodzaju funkcję został przez Piekarskiego określony jako „mechanizm normatywny”. Mechanizmy predykcyjne, zgodnie z przedstawioną argumentacją, należy więc traktować jako mechanizmy normatywne.

Rozdział V. „Mechanizmy normatywne a działania w przetwarzaniu predykcyjnym” to rozdział kluczowy dla założeń Autora z początku książki. Dla osób nieobeznanych specjalnie z PP istotne jest również, że zawiera on omówienie dyskusowanego w tej koncepcji „problemu ciemnego pokoju”. Piekarski charakteryzuje go słowami Davida Mumforda: „W pewnym sensie jest to stan [znajdowania się w ciemnym pokoju], który kora mózgowa stara się osiągnąć:



doskonale przewidywanie świata, coś jak stan nirwany... kiedy nic cię nie zaskakuje, a nowe bodźce powodują najdrobniejsze falowanie w twojej świadomości” (Mumford, 1992, s. 247, za: Piekarski, 2020, s. 237). Chodzi tutaj o stan, w którym mógłby znaleźć się organizm, gdyby udało mu się zminimalizować sumę wszystkich potencjalnych błędów predykcyjnych. Ten hipotetyczny stan traktowany jest przez wielu krytyków PP jako podstawa do zarzutu, czyli absurdalnej konsekwencji spełnienia się domniemanych dążeń organizmu namiętnie minimalizującemu zaskoczenia. W odpowiedzi Autor, wykorzystując zaproponowaną koncepcję mechanizmów normatywnych i wskazując na istotną rolę ograniczeń środowiskowych, przedstawia alternatywną możliwość wyjaśnienia kompromisu między niepewną eksploracją a bezpieczną eksploatacją. To w efekcie stanowić ma odpowiedź na problem ciemnego pokoju. W dalszej części tego rozdziału pokazuje, że normatywność przetwarzania predykcyjnego opiera się na normatywności samych mechanizmów predykcyjnych, a nie na możliwości istnienia określonych reguł postępowania. Innymi słowy, zgodnie z tezą Autora, przetwarzanie predykcyjne jako bayesowska rama modelowania procesów poznawczych nie dlatego jest normatywne, że formułuje określone normy, reguły bądź zasady postępowania, które system poznawczy „powinien” przestrzegać, ale dlatego jest normatywne, ponieważ normatywne są same mechanizmy predykcyjne, warunkujące wybór takich a nie innych działań przez określone podmioty działania. Argumentacja ta zostaje wzmocniona analizami tak zwanych relacji motywacyjnych, które można opisać w terminach warunkowej zależności predykcji i podejmowanych działań od parametrów modelu, jak i określonych stanów świata. Tak rozumiane relacje motywacyjne mają stanowić podstawę pojawiania się takiej a nie innej motywacji podmiotu<sup>3</sup>. Oznacza to, że wyjaśnienie normatywnych mechanizmów predykcyjnych powinno odwoływać się zarówno do struktury modelu generatywnego i jego parametrów, jak i do ustrukturyzowania środowiska, w którym dany podmiot funkcjonuje. Wreszcie – Piekarski przedstawia hipotezę tak zwanej normatywności rozproszonej, przyznając pewnego rodzaju normatywny charakter strukturze relacyjnej, obejmującej dane elementy z nią związane.

## 5. Podsumowanie

Jak chyba widać było przy omawianiu zwłaszcza ostatnich dwóch rozdziałów, monografia Piekarskiego jest miejscami dość złożona erudycyjnie, co koresponduje z podjętymi celami badawczymi. Ponieważ jednak zdecydowałem się oceniać ją w dużej mierze w kategoriach przydatności dydaktycznej (na poziomie akademickim), wyłożyłem tutaj najpierw swoje uwagi z tym związane.

---

<sup>3</sup> Pełnej analizie tego tematu, znacznie wykraczającej poza ustalenia przedstawione w recenzowanej monografii, poświęcona jest tegoroczna publikacja Autora (Piekarski, 2022).

W tym kontekście można powiedzieć, że wprowadzające i przeglądowe partie książki mogłyby wybrzmieć mniej erudycyjnie. Na przykład część poświęcona Kantowi jest takim stosunkowo wymagającym miejscem, przy założeniu, że sięga po to ktoś po prostu z Kantem niezaznajomiony. Inaczej jest w przypadku przedstawiania konserwatywnych i radykalnych odmian koncepcji przetwarzania predykcyjnego oraz związanych z nimi nurtów w badaniach nad poznaniem: napisane zostały dość przejrzyście i zrozumiale, może dlatego, że tematy te nadal nie są specjalnie obecne w szerszej edukacji akademickiej.

Czy przejścia między fragmentami o charakterze opracowań a rozwijaniem własnych analiz są dostatecznie widoczne? To również pytanie z perspektywy celów, do których książka ta nie powstała przede wszystkim. Wypowiem się tutaj jednak z własnego dydaktycznego doświadczenia: wystarczy odpowiednie wprowadzenie na zajęciach akademickich w połączeniu z sugestiami co do lektury rozdziałów. Przydatna byłaby również odpowiednio wnikliwa lektura Wprowadzenia i Przedmowy z zawartymi tam „mapami” książki – jednak co do uważności czytania (i w ogóle czytania) wstępnych części książek, pozostaje sceptyczny nie tylko w odniesieniu do społeczności studenckiej...

W rezultacie więc największe – względnie największe – moje zastrzeżenie dotyczy terminologii i definiowania. To, co w nich nieco problematyczne, ma zasadniczo rodowód językowy: ich struktury są mocno... „anglojęzyczne”. Oczywiście, nie w dosłownym sensie ani na zasadzie kalek językowych. Podczas gdy w języku angielskim powszechne jest używanie jednego terminu na przykład na określenie zarówno przedmiotu badań, jak i stanowiska badawczego, lub też równie łatwo przechodzi się od pojęcia struktury pełniącej jakąś funkcję do pojęcia tej funkcji, tak w języku polskim rodzić to może różnorakie nieporozumienia, jeżeli nie będziemy pamiętać o tamtej obcojęzycznej tradycji. Bez tego – możemy mieć czasem trudności, gdy „przetwarzanie predykcyjne” czy „PP” odnosi się w zdaniu (czy nawet w śródtytule) po prostu do koncepcji czy jej odmiany, lub gdy mózg – jakby nie było, biologiczny organ – okreśłany jest wprost jako „model generatywny”. Nie czynię z tego wielkiego zarzutu, jako że osoby już wstępnie doświadczone w lekturze kognitywistycznej i z nią związanej powinny sobie z tym poradzić. W przypadku studentów jednak nie zakładałbym tego.

Celownik krytyki przesunę teraz w stronę samej koncepcji i związanej z nią literatury. Wnioski dotyczące zjawisk predykcyjnych pojawiają się niekiedy i koleją rzeczy w badaniach empirycznych nad widzeniem. Innym razem jednak prowadzą do sformułowania bardzo ambitnych stanowisk filozoficznych i szerokiego zakresu tez. Tak jest na pewno w przypadku nurtu w PP związanego z pracami Fristona. Duży rozmach koncepcji w tym przypadku dość chyba naturalnie ściągnął na siebie baczną uwagę i głosy krytyczne – czego przykładów dostarcza przytaczany już wyżej artykuł Piotra Litwina i Marcina Miłkowskiego (2020). A więc koncepcja ta jako ewentualna „teoria wszystkiego”

w kognitywistyce („życie to minimalizowanie zaskoczenia” itd.) nie spełnia wymogów kompletnych, użytecznych teorii naukowych. Odnosi się ją do modeli percepcyjnych, lecz nadal wiele z nich ma ogólnikowy charakter, są szkieletami wnioskowania statystycznego, nadal niezbyt dobrze (bezpośrednio) osadzonymi empirycznie. Natomiast ujęcie bayesowskie (skądinąd i tak krytykowane, zob. np. Bowers i Davis, 2012), ma właściwość dopasowywania się do posiadanego zestawu danych, przez co łatwo o zarzut, że praktycznie wszystko da się oprzeć na mechanizmach przetwarzania predykcyjnego. Zastrzeżenie budzi również osadzenie koncepcji na przyjęciu hierarchicznej budowy mózgu, podczas gdy jest ona częściowo również heterarchiczna. Należy też dodać, że przydatnym „chłopcem do bicia” okazał się tutaj problem ciemnego pokoju, który łatwo w eksperymentowaniu myślowym zdemonizować jako stan faktycznie niepożądany, nieopłaczalny czy wręcz szkodliwy.

W tym kontekście ujawniają się zalety omawianej książki i tym samym warsztatu Autora. Piekarski w żadnym kluczowym przypadku nie poprzestaje na wy-modelowanej argumentacji z krytycznej literatury tematu. Widać to już przy temacie Kanta oraz odpieraniu zarzutu antyrealizmu w polemice z Danem Zahavim. W dalszych partiach książki Autor pokazuje, jak można nietrywialnie ująć bayesowski charakter przetwarzania predykcyjnego. Jest też konsekwentny w zadeklarowanej naturalistycznej analizie tego, co normatywne, w przyjętym zakresie problemowym. I oczywiście nie poprzestaje na tropach krytyków wizji „pojemnika z nirwaną”, czyli eksperymentu myślowego z ciemnym pokojem, lecz buduje odpowiedź mocno osadzoną we własnej koncepcji mechanizmów normatywnych (na ile przekonująco dla czytelników, to już zależy od różnych czynników w naszym procesie ustosunkowywania się do PP).

I tutaj wracam do motywu „Maszyny do ćwierkania” z początku recenzji. Ten rzekomo wpasowujący się w każde niemal zapotrzebowanie interpretacyjne obraz pojawił się między innymi na monachijskiej „Wystawie sztuki zdegenerowanej” w 1933 r., nastawionej na dyskredytowanie czy po prostu ośmieszenie twórczości awangardowej (zob. Jaśkiewicz, 2022). Rzecz jasna, nie chodzi mi tutaj o kreślenie naciąganych analogii między sytuacją PP a historią „Maszyny”. Mam raczej pewne skojarzenia z krytykowaniem i dyskredytowaniem w nauce, które między innymi tradycja analityczna – skądinąd cenna – podniosła do wielkiej rangi w kontekście kognitywistyki. Krytykować zwykle jest łatwiej i efektowniej. Kiedyś fenomenologię, później nurty ucieleśnienia i usytuowania poznania – wciąż było coś, co najprawdopodobniej lądowałoby na „Wystawie nauki zdegenerowanej”, gdyby taką organizowano. Wbrew drugoczącej krytyce jednak, „niedorzeczne” pomysły i perspektywy wracały lub regenerowały się, a nawet jeśli nie zdobywały szczególnych bastionów, to utrzymały się w mainstreamowych dyskusjach i zacnych periodykach. Jest to, oczywiście, temat mocno związany z kontekstem odkrycia w nauce i rozwijanie go tutaj byłoby nadużyciem. A pozycja koncepcji przetwarzania predykcyjnego

w świecie nauki nie jest, póki co, mocniej zagrożona. Można jednak uznać, wyciągając wnioski chociażby z krótkiej historii interdyscyplinarnych badań nad poznaniem, że definitywne oceny tej koncepcji czy jej bayesowskiej podbudowy jako „pasującej do wszystkiego” są po prostu pochopte. Tak pochopte nie dyskredytował PP wspomniany już Zahavi (2018), o czym mieliśmy okazję przekonać się także podczas konferencji fenomenologicznej w Warszawie wiosną 2017.

W tym to kontekście stanowisko Autora książki *Mechanizmy predykcyjne i ich normatywność* odznacza się niezwykle pozytywnie, wyrażając niezgodę na łatwą krytykę i pochopte negowanie. Jeśli więc recenzowana monografia nie przekona kogoś ani do koncepcji przetwarzania predykcyjnego, ani do ujęcia zaproponowanego przez Autora, ani do jego własnego przedsięwzięcia badawczego, powinien ostatecznie docenić stosowaną tam jakość polemizowania w filozofii kognitywistyki.

„Maszyna do ćwierkania” pojawiła się również na początku prowadzonego przeze mnie kursu akademickiego Przetwarzanie Predykcyjne i towarzyszyła studentkom i studentom w czołówce wyświetlanych prezentacji. Podczas pierwszego spotkania zachęciłem do próby interpretacji – choćby i naciąganej – tego obrazu w kontekście nabywanej wiedzy o PP. Jak to jednak bywa, pod koniec semestru, skupieni na zaliczeniach, jakoś nie zdążyliśmy do tego wrócić. Spróbuję następnym razem.

## Bibliografia

- Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. New York: Routledge.
- Bowers, J. S. i Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, 138(3), 389–414. <https://doi.org/10.1037/a0026450>
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181–204. <https://doi.org/10.1017/S0140525X12000477>.
- Feldman-Barrett, L. (2021). *Mózg nie służy do myślenia (7 i ½ wywrotowych lekcji o mózgu)*. Tłum. Z. Lamża. Łódź: Wydawnictwo Feeria Science.
- Foster, D. (2021). *Deep learning i modelowanie generatywne. Jak nauczyć komputer malowania, pisanie, komponowania i grania*. Tłum. R. Meryk. Gliwice: Wydawnictwo Helion.
- Friston, K. J. & Kiebel, S. J. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 1211–1221. <https://doi.org/10.1098/rstb.2008.0300>
- Friston, K. J. (2011). What is optimal about motor control? *Neuron*, 72(3), 488–498. <https://doi.org/10.1016/j.neuron.2011.10.018>

- Gładziejewski, P. (2017). Percepcja jako kontrolowana halucynacja. *Filozofuj!*, 2(14), 13–15.
- Gregory, R. L. (1971). *Oko i mózg. Psychologia widzenia*. Tłum. S. Bogusławski. Warszawa: Warszawa: Państwowe Wydawnictwo Naukowe.
- Grush, R. (1995). *Emulation and Cognition. Dissertation*. University of California, San Diego.
- Jaśkiewicz, J. (2022). Maszyna do świerkania. *iSztuka – Edukacja Kulturalna dla Wszystkich*. Źródło: <http://www.isztuka.edu.pl/i-sztuka/node/699> (dostęp 18.06.-2022).
- Johnson-Laird, P. N. (1994). Mental models and probabilistic thinking. *Cognition*, 50(1–3), 189–209. [https://doi.org/10.1016/0010-0277\(94\)90028-0](https://doi.org/10.1016/0010-0277(94)90028-0)
- Koffka, K. (1922). Perception: an introduction to the Gestalt-Theorie. *Psychological Bulletin*, 19(10), 531–585. <https://doi.org/10.1037/h0072422>
- Litwin, P. i Miłkowski, M. (2020). Unification by Fiat: Arrested Development of Predictive Processing. *Cognitive Science*, 44(7), e12867. <https://doi.org/10.1111/cogs.12867>
- Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biological Cybernetics*, 66(3), 241–251. <https://doi.org/10.1007/BF00198477>
- Orlandi, N. (2017). Predictive perceptual systems. *Synthese*, 195, 2367–2386. <https://doi.org/10.1007/s11229-017-1373-4>.
- Parr, T., Pezzulo, G. i Friston, K. J. (2022). *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/12441.001.0001>
- Piekarski, M. (2016). Od typiki doświadczenia do normatywnej antycypacji. Przyczynek do fenomenologii normatywności. *Filo-Sofija*, 33(2), 71–86.
- Piekarski, M. (2017). Normatywność antycypacji a normatywność predykcji. Dwa podejścia: fenomenologia i teoria przetwarzania predykcyjnego. *Avant*, 8(3), 25–56. <https://doi.org/10.26913/80302017.0112.0001>
- Piekarski, M. (2018). Jedna czy wiele normatywności? *Studia Philosophiae Christianae*, 4(2), 61–81. <https://doi.org/10.21697/2018.54.2.11>
- Piekarski, M. (2020). *Mechanizmy predykcyjne i ich normatywność*. Warszawa: Wydawnictwo Liberi Libri. <https://doi.org/10.47943/lib.9788363487447>
- Piekarski, M. (2022). Motivation, counterfactual predictions and constraints: normativity of predictive mechanisms. *Synthese*, 200, 352. <https://doi.org/10.1007/s11229-022-03837-1>
- Rao, R. P. i Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 1999 2(1), 79–87. <https://doi.org/10.1038/4580>

Wölfflin, H. (2017). *Podstawowe pojęcia historii sztuki. Problemy rozwoju stylu w sztuce nowożytnej*. Tłum. D. Hanulanka. Gdańsk: Wydawnictwo słowo/obraz terytoria.

Zahavi, D. (2018). Brain, mind, world: Predictive coding, neo-Kantianism, and transcendental idealism. *Husserl Studies*, 34, 47–61. <https://doi.org/10.1007/s10743-017-9218-z>

### **Review of the book *Mechanizmy predykcyjne i ich normatywność* (2020)**

**Keywords:** free energy; mechanism; generative model; brain; normativity; predictive processing