

## Dlaczego reprezentacje nie trzymają się modeli dynamicznych?

### Why representations don't stick with dynamic models?

**Abstract:** In this paper I investigate thesis, embraced by proponents of dynamicism in cognitive science, that mind is not representational and explanation of cognition can go without representations. This claim has received serious criticism from cognitive scientists and philosophers of mind, who accuse dynamical explanation of being satisfying only for a narrow class of simple cognitive phenomena. Thus, genuine, representation-free explanation of cognition will always be incomplete. I espouse another strategy and present two arguments saying that the language of pure dynamical systems theory is not rich enough to define any nontrivial notion of representation. If I am right, then at least these phenomena dynamical explanation deals well with are not representational and representation talk can in no way help us understand them.

**Key words:** philosophy of cognitive science, dynamical explanation, representation, cognitive models, Clark, van Gelder

### 1. Na czym polega wyjaśnianie dynamiczne?

Idea wyjaśnienia dynamicznego jest urzekająco prosta: to pewien formalizm (choćby tak prosty, jak równanie algebraiczne z dwiema niewiadomymi), który opisuje, jak manipulacja wartością jednej zmiennej (może być nią np. czas) modyfikuje wartości pozostałych. Choć historia wyjaśnień dynamicznych jest równie długa jak historia nowożytnej nauki i szybko doczekała się spektakularnych sukcesów, takich jak mechanika newtonowska, to jej zastosowanie do badań umysłu jest stosunkowo nowym trendem. Pierwsze próby dynamicznego myślenia o procesach psychicznych podejmowali krótko po drugiej wojnie światowej cybernetycy w Stanach Zjednoczonych i Związku Radzieckim [Wiener 1948], a następnie – już po narodzinach kognitywistyki, jako przedstawiciele jej drugiej fali – koneksjoniści, proponując modele poznawcze o rozproszonej architekturze, takie jak sztuczne sieci neuronowe [Smolensky 1988]. Modele koneksjonistyczne oferowały aparat formalny do badania pewnych na ogół ignorowanych wcześniej aspektów procesów poznawczych (takich jak ich wymiar czasowy czy nieliniowe zależności), zachowując przy tym tradycyjny obraz poznania jako procesu obliczeniowego. Dopiero w latach 90. niektórzy filozofowie

i kognitywiści poszli o krok dalej, zaprzęgając wyjaśnienia dynamiczne do ustanowienia nowego paradygmatu, rozwijanego w opozycji do głównego nurtu kognitywistyki.

Klasycznym przykładem takiego podejścia są badania nad rozwojem poznawczym i motorycznym niemowląt Lindy Smith i Esther Thelen [2003]. W ich ujęciu rozwój polega na wyłanianiu się ze złożonych, nieliniowych współoddziaływań nowych wzorców (*patterns*) zachowania. Przykładowo pojawienie się około 12. miesiąca życia umiejętności postrzegania stałości przedmiotów nie da się wytłumaczyć, postulując pojawiające się *ex machina*, zgodnie z genetycznym programem, nowe kompetencje mentalne; ta umiejętność to wzorzec zachowania wyłaniający się wraz z rosnącą sprawnością motoryczną (chwytywanie przedmiotów) i odpowiednią stymulacją środowiskową. Badaczki proponują psychologicznie wiarygodny i matematycznie elegancki model rozwoju umiejętności percypowania stałości przedmiotów jako pokonywania przez system poznawczy pewnej trajektorii w przestrzeni stanów, która to trajektoria podlega nieustannym modyfikacjom ze względu na lokalne wzorce zachowania (które z kolei współokreśla sama trajektoria).

## 2. Hipoteza dynamicystyczna

Stosowanie wyjaśnień dynamicznych w praktyce badawczej psychologów, robotyków, lingwistów i neuronaukowców to fakt historyczny. Bardziej interesująca z punktu widzenia naszej pracy jest filozoficzna interpretacja nadawana tej praktyce badawczej. Dynamicyzm urósł bowiem w kognitywistyce, jak wspomnieliśmy, do rangi autonomicznego paradygmatu, który postuluje szereg silnych tez empirycznych i teoretycznych (metodologicznych, ontologicznych) i staje w szranki z dominującym paradygmatem obliczeniowym. Moim punktem wyjścia będą manifesty jednego z pierwszych filozoficznych zwolenników dynamicyzmu, Tima van Geldera [1995, 1998; zob. też van Gelder i Port 1995]. Swoje stanowisko nazywa on „hipotezą dynamicystyczną” i formułuje następująco:

(HD) Systemy poznawcze (a) są systemami dynamicznymi i (b) najlepiej opisuje je aparat teoretyczny teorii układów dynamicznych [van Gelder 1998].

Przez system rozumie tu zbiór wzajemnie zależnych od siebie (*coupled*) zmiennych, a przez system *dynamiczny* – system zmieniający swój stan zgodnie z dobrze określoną regułą przejścia, którego przestrzeń stanów jest przestrzenią metryczną (tzn. dla każdej pary stanów jest określona ich odległość od siebie). Bardziej zgodne z duchem dynamicyzmu byłoby może dodanie zastrzeżenia, że chodzi o regułę przejścia opisaną funkcją ciągłą i wielokrotnie różniczkowalną, co presuponuje między innymi, że system może przybierać nieprzeliczalnie nieskończenie wiele stanów, ale van Gelder nie uznaje ciągłości przestrzeni stanów za cechę definicyjną systemu dynamicznego.

HD składa się z (a) komponentu ontologicznego (umysł *jest* systemem dynamicznym) i (b) metodologicznego (umysł *należy opisywać* jako system dynamiczny). Zdaniem van Geldera [1995] HD implikuje co najmniej trzy tezy szczegółowe:

- (1) Poznanie (a) nie jest procesem obliczeniowym i (b) nie należy go opisywać na sposób obliczeniowy.
- (2) Umysł (a) nie ma charakteru reprezentacyjnego i (b) poznanie można rozumieć bez postulowania istnienia wewnętrznych reprezentacji otoczenia<sup>1</sup>.
- (3) Poznanie (a) odbywa się pod presją czasu i (b) w modelach procesów poznawczych należy uwzględnić czas rzeczywisty (ciągły) jako istotną eksplanacyjnie zmienną.

Warto uzupełnić tę listę o jeszcze dwie ważne tezy, nieformułowane *explicitie* przez van Geldera: tezę o aktywnym charakterze i ucieleśnieniu poznania. Postulują je takie nurty kognitywistyki, jak robotyka behawioralna [Brooks 1991], podejście ucieleśnione [Varela, Thompson i Rosch 1991] i enaktywizm [Thompson 2010], które są naturalnymi kontynuacjami gelderowskiego dynamicyzmu.

- (4) System poznawczy (a) nie jest mechanizmem odpowiadającym na egzogenne stymulacje według zaprogramowanego genetycznie algorytmu, ale samorganizującą się i autonomicznie eksplorującą otoczenie całością. (b) Nie należy go rozumieć jako statycznej struktury, ale jako proces związany obustronnymi i obukierunkowymi relacjami przyczynowymi ze sobą i z otoczeniem.
- (5) System poznawczy (a) jest systemem fizycznym i jego materialna konstytucja zarówno ogranicza jego możliwości poznawcze, jak i podsuwa mu gotowe rozwiązania pewnych problemów. (b) Należy uwzględniać niskopoziomowe cechy anatomii lub architektury systemu poznawczego w wyjaśnianiu wysoko-poziomowych procesów poznawczych<sup>2</sup>.

<sup>1</sup> Stosunek van Geldera do tezy (2) (antyreprezentacjonizmu) jest ambiwalentny. Jako jej asercję można interpretować użycie przez niego metafory maszyny Watta jako urządzenia zdolnego do precyzyjnego kontrolowania swojego stanu, a przy tym niewątpliwie niereprezentacyjnego. Chwilę później czytamy jednak: „there is nothing preventing dynamical systems from incorporating some form of representation; indeed, an exciting feature of the dynamical approach is that it offers opportunities for dramatically reconceiving the nature of representation in cognitive systems, even within a broadly noncomputational framework” [van Gelder 1995, s. 376]. W późniejszym artykule [1998, s. 622] van Gelder wyjaśnia, że choć systemy poznawcze nie są ze swej istoty (*inherently*) reprezentacyjne, to jednak postulowanie reprezentacji jest legalnym i czasem pożądanym posunięciem. W szczególności dynamicysta dysponuje szerszą paletą rodzajów reprezentacji, może np. postulować reprezentacje *dynamiczne* na wyższych poziomach organizacji, przykładowo w przestrzeni stanów (por. krytykę tego pomysłu w sekcji 4.1). Mimo ambiwalencji van Geldera co do antyreprezentacjonizmu, na potrzeby tego tekstu uznajemy tezę (2) za istotny składnik HD, ponieważ po pierwsze, bez uznania tej tezy ciężko byłoby dynamicyzmowi określić swoją odrębność od głównego nurtu kognitywistyki oraz, po drugie, akceptuje ją większość innych dynamicystów, takich jak Anthony Chemero, Alva Noë i Evan Thompson (wśród filozofów) czy Linda Smith, Esther Thelen, Kevin O’Regan i Rodney Brooks (wśród badaczy praktyków).

<sup>2</sup> Teza (5) nie jest oryginalna dla dynamicyzmu i wydaje się, że można uczynić jej zadość przy zaprzeczeniu tez (1)–(4), np. w ramach wyjaśnienia obliczeniowego uwzględniającego mechanizm implementacji obliczeń. Nawet jeśli jej oponenci należą już do rzadkości, (5) pozostaje istotną, empirycznie testowalną konsekwencją HD.

Traktując HD jako falsyfikowalną hipotezę, van Gelder kładzie nacisk raczej na jej artykulację niż poparcie argumentami, a te ostatnie opiera bardziej na praktyce badawczej niż analizach pojęciowych. I tak koronnym argumentem za prawdziwością HD są sukcesy eksplanacyjne badań, które przyjęły HD jako hipotezę roboczą (np. badania Smith i Thelen; dla przykładów z innych obszarów kognitywistyki zob. antologię: Port i van Gelder 1995). Argument ten wzmacnia fakt, że wyjaśnienie dynamiczne jest kanonicznym modelem wyjaśnienia naukowego i odnosi sukcesy także (lub przede wszystkim) poza kognitywistyką, np. w naukach fizycznych: HD oferowałyby więc atrakcyjną filozoficznie unifikację kognitywistyki z innymi naukami przyrodniczymi w ramach wspólnego schematu eksplanacyjnego.

Przykładowa linia rozumowania, wychodząca od sukcesów badań Smith i Thelen [2003], może przebiegać następująco. Zakładając, że (1b) kompetencji poznawczych nie powinno się traktować jako zaprogramowanych genetycznie oraz że należy uwzględnić ich (3b) czasowy wymiar i (4b) tendencje do samoorganizacji, badaczki stworzyły model, który pozwala skutecznie przewidywać, jak zachowania eksploracyjne podejmowane przez dziecko wpływają na dynamikę jego dalszego rozwoju. Podobnie Thelen i Smith rozwijają klasyczną teorię rozwoju poznawczego Jeana Piageta, (2b) negując jednocześnie jego tezę, że nabywanie umiejętności percypowania stałości przedmiotu można wyjaśnić przez nabywanie umiejętności operowania na reprezentacjach mentalnych. Okazuje się, że zbudowanie niereprezentacyjnego i spełniającego wcześniejsze warunki modelu dynamicznego pozwala lepiej wytłumaczyć różnice indywidualne rozwoju badanych i gwałtowne zmiany wzorców zachowania (tzw. bifurkacje) przy drobnej rearanżacji sytuacji eksperymentalnej. Wreszcie zakładając, że (5b) rozwój poznawczy jest nierozzerwalnie związany z rozwojem motorycznym, Thelen i Smith tłumaczą, jak rozwój umiejętności chwytania wpływa na rozwój percepcji stałości przedmiotów. Ponieważ przyjęcie postulatów metodologicznych (1b–5b) prowadzi do sukcesów eksplanacyjnych, rozsądne jest wytłumaczenie tych sukcesów przyjęciem tez ontologicznych (1a)–(5a), które są konsekwencjami HD.

Argumentacja sukcesu eksplanacyjnego nie jest, z czego van Gelder zdaje sobie sprawę, rozstrzygająca. Nie przesądza bowiem, że nie istnieje inny zestaw postulatów metodologicznych, który prowadzi do jeszcze lepszych przewidywań, albo że nie ma alternatywnej ontologicznej podbudowy tych postulatów. Rzecz jasna istnieją też niezależne racje empiryczne za tezami (1)–(5) (van Gelder powołuje się np. na istotową czasowość poznania), ale ich omówienie przekracza ramy tego tekstu. Jak zresztą stwierdzono, van Gelder argumentuje nie tyle *explicite* za HD, ile raczej za tym, że nie ma powodów, by *a priori* wykluczyć HD. W szczególności nie jest naszym celem przeprowadzenie systematycznej obrony HD, lecz wykazanie pewnych jej konsekwencji, które skutecznie blokują pewien rodzaj krytyki.

### 3. Pięć argumentów przeciw hipotezie dynamicystycznej

Zacniemy od przytoczenia pięciu argumentów wysuwanych przeciw HD przez bardziej zachowawczo nastawionych badaczy i filozofów kognitywistyki.

(A1) Van Gelder traktuje systemy koneksjonistyczne jako podklasę systemów dynamicznych. Ale systemy koneksjonistyczne mają charakter obliczeniowy i reprezentacyjny, więc teza o dynamicznym charakterze poznania nie pociąga za sobą tezy o jego nieprezentacyjnym i nieobliczeniowym charakterze [Eliasmith 1997].

(A2) Wyjaśnienie dynamiczne nie jest pełnym wyjaśnieniem, bo ignorując strukturę wewnętrzną systemu poznawczego, nie potrafi zdać sprawy z tego, dlaczego niektóre systemy wykonują te same zadania wydajniej niż inne. Formułująca ten argument Melanie Mitchell [1998] powołuje się na swoje badanie nad binarnymi automatami komórkowymi. Ewoluuując za pomocą algorytmu genetycznego różne strategie rozwiązania pewnego zadania (wprowadzenia wszystkich komórek automatu w stan 0 lub 1, który dominował w warunkach początkowych), Mitchell zauważyła, że klasy abstrakcji różnych rozwiązań ze względu na złożoność obliczeniową nie pokrywają się z klasami abstrakcji ze względu na dynamikę. Wyjaśnienia obliczeniowe i dynamiczne są zatem wzajemnie nieredukowalne i komplementarne, więc wyjaśnienie dynamiczne jest niepełne.

(A3) Krytyka reprezentacjonizmu na gruncie zarówno dynamicyzmu [Smith i Thelen 2003], jak i robotyki behawioralnej [Brooks 1991] opiera się na ekwiwokacji pojęcia reprezentacji: ma przesłanki co najwyżej przeciw globalnie dostępnym i manipulowalnym reprezentacjom symbolicznym (*à la* Newell i Simon 1976), a w konkluzji odwołuje się do wszelkich typów reprezentacji. Przykładowo konstrukt taki jak pętla emulacyjna Grusha [2004] stanowi reprezentację w całkiem silnym sensie, a można ją postulować w pełnej zgodzie z tezą o ucieleśnieniu i usytuowaniu poznania, więc jest odporna na argument z biologicznej nieadekwatności reprezentacji i falsyfikuje HD [Clark i Toribio 1994; Clark 1997, 2001].

(A4) Przypadki, na które powołują się dynamicyści i robotycy behawioralni, nie są akurat problemami, które domagają się postulowania reprezentacji, co nie znaczy, że wyjaśnienie wszystkich aspektów poznania może się bez nich obejść. Wrażliwość systemu poznawczego na aktualnie niedostępne percepcyjnie przedmioty lub abstrakcyjne cechy przedmiotów jest na przykład fenomenem, dla którego nie podano dotąd przekonującego wyjaśnienia dynamicznego. HD jest zatem co najmniej przedwczesna [Clark i Toribio 1994; Clark 2001].

(A5) Opis aktywności systemu poznawczego w kategoriach normatywnych zakłada istnienie w nim komponentów o własnościach semantycznych, których przykładem są właśnie reprezentacje. Opis dynamiczny niepostulujący reprezentacji nie potrafi więc odróżnić systemu dysfunkcjonalnego od działającego poprawnie, a zatem jest niepełny [Dennett 1998].

Przed ustosunkowaniem się do tej krytyki spróbujemy wyklarować pozycję teoretyczną, z której będziemy się przed nią bronili. Dennett [1998] słusznie zauważa, że zwa wewnątrznie niespójnych manifestów dynamicyzmu wyłania się właściwie dwóch van Gelderów. Pierwszy z nich, van Gelder bezkompromisowy, jest konsekwentnym



antyrepresentacjonistą, drugi, van Gelder ugodowy, argumentuje tylko na rzecz zastąpienia reprezentacji symbolicznych w stylu GOFAI czymś bardziej biologicznie realistycznym. Bezkompromisowy van Gelder spotyka się z przytoczoną surową krytyką, przed którą nie umie obronić się inaczej, niż wcielając się w swoje ugodowe *alter ego* – które równie dobrze mogłoby uosabiać panujący w kognitywistyce konsens. Wydaje się, że rozterkę van Geldera można rozwiązać, podważając przesłankę, na której się opiera: żeby wyjaśnienie dynamiczne mogło być wartościowe, musi być w stanie wyjaśnić wszystkie aspekty poznania. Nie będziemy tu przedstawiali systematycznej krytyki tak sformułowanego absolutyzmu eksplanacyjnego<sup>3</sup>, pokażemy po prostu, że zakładając roboczo, iż jest fałszywy, można uratować dynamizm bezkompromisowego van Geldera, nie osłabiając go ani trochę.

Spróbujemy więc zinterpretować HD na tyle radykalnie, by była w stanie obronić się przed zarzutami lub z podniesioną głową odpowiedzieć *touché*. To ostatnie wydaje się jedyną rozsądną reakcją na zarzuty (A1), (A2) i (A5) o tyle, o ile van Gelder jak jego bliźniak nie chce rozcieńczać dynamizmu w hybrydowy program badawczy. Zarówno wysokopoziomowe opisy dynamiki sieci koneksjonistycznych, jak i niskopoziomowe opisy ich architektury (postulujące reprezentacje rozproszone) mogą być wartościowe w wyjaśnianiu pewnych aspektów zachowania sieci, a bezradne wobec innych. To, że opis dynamiczny nie może wyjaśnić czegoś, co wyjaśnia opis obliczeniowy, nie jest więc argumentem przeciw dynamizmowi, lecz co najwyżej przeciw absolutyzmowi eksplanacyjnemu. Zabezpieczywszy się tak przed zarzutami typu „bez reprezentacji nie ma wyjaśnienia”, w następnej części tekstu pokażę, że w ramach wyjaśnienia dynamicznego w istocie nie da się mówić o reprezentacjach w żadnym nietrywialnym sensie.

#### 4. Reprezentacje w schemacie wyjaśniania dynamicznego

Zarzuty Clarka (A3 i A4) są w istocie głosami w długim sporze o istnienie reprezentacji. Ponieważ Clark (słusznie) sprzeciwia się konceptualizacji tego sporu jako dychotomii, przytoczymy rozważane przez niego [1997] cztery sensy pojęcia reprezentacji i sformułujemy dwa argumenty, jakie bezkompromisowy van Gelder mógłby przeciwko nim wytoczyć.

(R1) Reprezentacja to stan wewnętrzny systemu poznawczego. (System może zmieniać swój stan pod wpływem przeszłych stymulacji z otoczenia i jego stan może wpływać na reakcje na przyszłe stymulacje)<sup>4</sup>.

(R2) Reprezentacja to stan wewnętrzny systemu poznawczego skorelowany z jego otoczeniem (lub jego wycinkiem).

<sup>3</sup> Celną krytykę absolutyzmu eksplanacyjnego w kognitywistyce (w kontekście wyjaśniania obliczeniowego) przedstawia np. Marcin Miłkowski [2013].

<sup>4</sup> Jak zauważa Clark, reprezentacja w sensie R1 nie ma żadnej wartości eksplanacyjnej. Dalej, mówiąc o nietrywialnych reprezentacjach, będę miał na myśli reprezentacje w sensach R2–R4.

(R3) Reprezentacja to stan wewnętrzny systemu poznawczego, którego funkcją właściwą jest bycie skorelowanym z otoczeniem. (Funkcją właściwą rozumiemy jako określoną przez historię filogenetycznej ewolucji systemu poznawczego i stanowiącą o jego dostosowaniu).

(R4) Reprezentacja to stan wewnętrzny systemu poznawczego skorelowany z otoczeniem, który umożliwia systemowi przewidywanie i analizowanie otoczenia podczas jego sensorycznej niedostępności.

#### 4.1. Argument z braku struktury

Nietrywialne pojęcia reprezentacji zakładają istnienie co najmniej jednego inwariantu: struktury, która jest nośnikiem swoich stanów wewnętrznych i może stanowić statyczny układ odniesienia do szukania korelacji pomiędzy nimi a otoczeniem. Struktura w teorii układów dynamicznych jest własnością emergentną całego systemu (funkcją wielu zmiennych, w tym czasu), a nie czymś gotowym. Żeby mówić o strukturze (w momencie  $t_1$ ), najpierw musimy mieć gotowy model dynamiczny, czyli właściwie gotowe wyjaśnienie – a skoro takie mamy, postulowanie reprezentacji to tylko przyklejanie etykietek, które nie doda do niego niczego, czego już nie wiemy. W szczególności budowanie redukcyjnej teorii reprezentacji na modelu dynamicznym będzie poznawczo jałowe, bo nie jest spełniona przesłanka pragmatyczna postulowania obiektu redukowalnego: że nie mamy efektywnej procedury przewidywania stanów systemu na najniższym poziomie opisu. W modelu dynamicznym taką efektywną procedurą jest manipulacja równaniem stanu układu, w ramach której z definicji można przewidzieć wszystko, co dane jest do przewidzenia (bez potrzeby wchodzenia na wyższy, reprezentacyjny poziom opisu).

Rozważmy strategię ugodowego van Geldera, polegającą na postulowaniu reprezentacji *dynamicznych* [van Gelder 1995]. Taką reprezentacją może być np. wzorzec trajektorii w przestrzeni fazowej czy region w przestrzeni fazowej (np. basen atrakcji) lub abstrahując jeszcze bardziej, region na diagramie bifurkacji. Tego typu reprezentacje są jednak zwykłymi nietrywialnymi reprezentacjami zlokalizowanymi w *statycznej* strukturze modelu dynamicznego: strukturą jest tutaj obiekt matematyczny – przestrzeń stanów danego układu dynamicznego. Samo pojęcie reprezentacji dynamicznej zasadza się więc na fundamentalnej ekwiwokacji modelu i modelowanego zjawiska: postulujemy byt drugiego rzędu na naszym modelu, a nie w modelowanym zjawisku. Przestrzeń fazowa nie jest niczym realnie istniejącym w świecie fizycznym – to geometryczna interpretacja pewnej struktury algebraicznej, która stanowi abstrakt; topologiczne własności przestrzeni fazowej nie muszą się więc przekładać na topologiczne własności modelowanego zjawiska. Równie dobrze moglibyśmy modelować system poznawczy za pomocą innego formalizmu, np. narracji w języku naturalnym, i argumentować, że nasz system poznawczy ma kompetencję językową, bo nasz model ma własności gramatyczne (jest ciągiem zdań złożonych języka polskiego). Reprezentacje drugiego rzędu, zlokalizowane w przestrzeni fazowej, o ile nie umiemy ich zlokalizować w fizycznej instancji modelu, są tylko artefaktem tego modelu.

## 4.2. Argument ze złożoności oddziaływań

Nietrywialne pojęcia reprezentacji zakładają skorelowanie stanów systemu poznawczego z otoczeniem. Załóżmy, że mamy układ dynamiczny złożony z agenta i jego środowiska, a w nim dwie skorelowane zmienne:  $x$  i  $y$ . Jeśli zmienne są skorelowane liniowo, to  $x$  jest funkcją liniową od  $y$  ( $y = ax + b$ , gdzie  $a$  i  $b$  to stałe). Wtedy jedna ze zmiennych jest redundantna: nie osiągamy korzyści eksplanacyjnej, modelując osobno  $x$  i osobno jego izomorficzną reprezentację  $y$ <sup>5</sup>. Jeśli są skorelowane nieliniowo, to  $a$  i  $b$  w poprzednim równaniu też są funkcjami innych zmiennych,  $c$  i  $d$ , dla których przeprowadzamy analogiczne rozumowanie. Wnioskujemy, że jeśli co najmniej dwie zmienne są z sobą skorelowane nieliniowo, to wszystkie zmienne w modelu są z sobą parami skorelowane nieliniowo, więc relacja korelacji (zarówno liniowej, jak i nieliniowej) jest trywialna (wszystko jest ze sobą skorelowane) i nie pozwala zdefiniować nietrywialnej relacji reprezentacji.

Dodatkowe warunki nakładane przez (R3) i (R4) w niczym nie pomagają. Pierwszy (dostosowanie) dlatego, że zakłada adaptacjonizm, który nie cieszy się dobrą sławą wśród dynamicystów. Wolą oni myśleć o ewolucji jako o ko-ewolucji systemu poznawczego i środowiska, w którym dobór naturalny nie jest jedynym czynnikiem [Thompson 2010]. W takim ujęciu złożoność przyczynowa historii filogenetycznej danego systemu poznawczego nie pozwala efektywnie wyróżniać cech adaptacyjnych. Warunki nakładane przez (R4) nie zawężają w żaden sposób denotacji pojęcia reprezentacji, bo (z wymienionych już powodów) nie odwołują się wprost do anatomii bądź architektury systemu poznawczego. Nie mogą więc zapobiec temu, że reprezentacja „rozleje” się na cały system poznawczy.

Okazuje się więc, że fundamentalne dla pojęcia reprezentacji pojęcia struktury i korelacji okazują się nieadekwatne wtedy, kiedy mamy do czynienia z wyjaśnieniem dynamicznym. Argument z braku struktury opiera się na tym, że nie został spełniony warunek pozytywny sensowności definicji struktury (nie istnieje żadna struktura), a argument ze złożoności oddziaływań opiera się na tym, że nie został spełniony negatywny warunek sensowności definicji korelacji (nie istnieje taka para zmiennych, która nie jest skorelowana). Dochodzimy więc do konkluzji, że język teorii układów dynamicznych jest zbyt ubogi, by zdefiniować w nim jakiegokolwiek nietrywialne pojęcie reprezentacji.

Oczywiście do wyjaśnienia dynamicznego oprócz modelu matematycznego potrzebujemy też jego implementacji, o której można myśleć jako o funkcji ze zmiennej w fizyczną własność pewnej instancji modelu. Większość kognitywistów tutaj widziałaby zapewne odpowiedni moment na zapostulowane reprezentacji, ale nie byłoby to posunięcie *fair play*. Przede wszystkim dlatego sam model matematyczny nie określa warunków, kiedy jakiś element jego fizycznej instancji jest reprezentacją, a kiedy nie. Tak jak relacje wyższego rzędu między zmiennymi (domniemane reprezentacje dynamiczne) nie przechodzą na fizyczne instancje modelu, tak struktura fizycznej instancji nie przechodzi na zmienne (o ile nie modelujemy w skali molekularnej).

<sup>5</sup> W pewnym sensie ten krok mojego argumentu to algebraiczne sformułowanie tezy Brooka, że świat jest swoim najlepszym modelem. Jest to sformułowanie życzliwe, zakładające, że bycie czymś różnym od swojego desygnatu nie stanowi cechy definicyjnej modelu.



larnej). Warunki istnienia reprezentacji musiałaby więc zostać określone przez jakąś heterogeniczną teorię – ale wtedy mamy do czynienia z wyjaśnieniem hybrydowym w stylu van Geldera kompromisowego: takie wyjaśnienie postuluje reprezentacje tylko o tyle, o ile uzupełniamy je wyjaśnieniem innego typu (np. koneksjonistycznym czy obliczeniowym). Sam układ dynamiczny jako model matematyczny zobowiązuje nas do uznania reprezentacji co najwyżej w sensie (R1), niezlokalizowanych nigdzie konkretnie w strukturze zjawiska i eksplanacyjnie jałowych.

## 5. Uwagi końcowe

Można zarzucać naszym argumentom, że tak naprawdę nie są argumentami za antyreprezentacjonizmem w sporze teoretycznym przeciw reprezentacjonizmowi, ale argumentami za pluralizmem w sporze *metateoretycznym* z absolutyzmem eksplanacyjnym. Jesteśmy skłonni uznać taką interpretację. Naszym głównym celem było obalenie tezy, że w *czystym* modelu dynamicznym jest miejsce na reprezentacje. Jeśli nie uda się odeprzeć zarzutów (A1), (A2) i (A5), to wynika stąd, że nie jest też prawdą, że model dynamiczny może wyjaśnić wszystkie aspekty poznania, co skłania do uznania pluralizmu eksplanacyjnego albo odrzucenia wyjaśnień dynamicznych. Nie widzimy ani jednego dobrego powodu, który mógłby przemawiać za tym ostatnim krokiem.

## BIBLIOGRAFIA

- Brooks R. (1991). *Intelligence without representation*. „Artificial Intelligence” 47, s. 139–159.
- Clark A., Toribio J. (1994). *Doing without representing?* „Synthese” 101, s. 401–431.
- Clark A. (1997). *The dynamical challenge*. „Cognitive Science” 21, s. 461–481.
- Clark A. (2001). *Mindware: An Introduction to the Philosophy of Cognitive Science*. New York–Oxford: Oxford University Press.
- Dennett D. (1998). *Revolution, no! Reform, si!* „Brain and Behavioral Sciences” 21, s. 636–637.
- Eliasmith C. (1997). *Computation and dynamical models of mind*. „Minds and Machines” 7, s. 531–541.
- Gelder T. van (1995). *What might cognition be, if not computation?* „The Journal of Philosophy” 26, s. 345–381.
- Gelder T. van, Port R. (1995). *It's about time: an overview of the dynamical approach to cognition* [w:] R. Port, T. van Gelder (red.), *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge–London: MIT Press.
- Gelder T. van (1998). *The dynamical hypothesis in cognitive science*. „Behavioral and Brain Sciences” 21, s. 615–665.
- Grush R. (2004). *The emulation theory of representation: motor control, imagery, and perception*. „Behavioral and Brain Sciences” 27, s. 377–396.
- Miłkowski M. (2013). *Explaining the Computational Mind*. Cambridge–London: MIT Press.
- Mitchell M. (1998). *A complex-systems perspective on the “computation vs. dynamics” debate in cognitive science* [w:] M.A. Gernsbacher, S.J. Derry (red.), *Proceedings of the 20<sup>th</sup> Annual Conference of the Cognitive Science Society – CogSci 1998*, s. 710–715.

- Newell A., Simon H. (1976). *Computer science as empirical inquiry: symbols and search*. „*Communications of the ACM*” 19, s. 113–126.
- Port R., Gelder T. van (red.) (1995). *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge–London: MIT Press.
- Smith L., Thelen E. (2003). *Development as a dynamic system*. „*Trends in Cognitive Sciences*” 7, s. 343–348.
- Smolensky P. (1988). *On the proper treatment of connectionism*. „*Behavioral and Brain Sciences*” 11, s. 1–74.
- Thompson E. (2010). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Harvard: Harvard University Press.
- Varela F., Thompson E., Rosch E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge–London: MIT Press.
- Wiener N. (1948). *Cybernetics; or, Control and Communication in the Animal and the Machine*. Paris: Technology Press.

