Studia Philosophiae Christianae
UKSW
62(2026)1

MATEUSZ SABATOWSKI

JOHNA VERVAEKEGO TEORIA REALIZACJI ISTOTNOŚCI

Streszczenie. Artykuł stanowi krytyczną rekonstrukcję teorii realizacji istotności (*relevance realization*, RR) Johna Vervaekego, opracowanej jako odpowiedź na lukę eksplanacyjną w klasycznych, obliczeniowych modelach poznania. Autor analizuje ewolucję teorii od lat 90. XX wieku, wskazując, jak problem ram oraz aporia relacji reprezentacja–relewancja wpływały na rozwój tej teorii w ramach naturalistycznego imperatywu w kognitywistyce. W pracy szczegółowo omówiono mechanizmy regulacyjne RR: zakres, regulowanie i priorytetyzację poznawczą, które za pomocą przetwarzania opozycyjnego mają umożliwić systemom dynamiczne równoważenie kompromisów ekonomicznych. Analizie poddano interdyscyplinarne zastosowania teorii – od przetwarzania predykcyjnego i ogólnej inteligencji, po teorię osobowości i stan flow. W części krytycznej została zidentyfikowana pozostająca luka między deklarowanym przez Vervaekego odrzuceniem rozwiązań obliczeniowych i reprezentacyjnych, a powoływaniem się na zdolność do antycypacji (niosącej konotacje reprezentacyjne) mającej konstytuować realizację istotności. Istniejące trudności ze sformułowaniem teorii realizacji istotności zgodnej z przyjętymi założeniami wskazują na konieczność zrewidowania tych założeń i ponownego zastanowienia się nad tym, czym jest relewancja.

Słowa kluczowe: Vervaeke John; realizacja istotności; problem ramy; problem reprezentacja–relewancja; racjonalność; enaktywizm; przetwarzanie predykcyjne; teleologia; obliczalność; naturalizm

1. Wprowadzenie. 2. Kontekst filozoficzny i naukowy. 2.1. Naturalistyczny imperatyw w kognitywistyce. 2.2. Problem ramy i problem reprezentacja–relewancja. 2.3. Teorie relewancji w literaturze i ich krytyka. 3. Teoria realizacji istotności (RR) – rozwój i główne tezy. 3.1. Relewancja bez reprezentacji. 3.2. Relewancja jako cecha transjektywna i niehomogeniczna. 3.3. Model ekonomiczny teorii RR. 3.4. Przetwarzanie opozycyjne. 3.5. Kompromisy w realizacji istotności. 3.6. Rozwój teorii RR i jej zastosowania. 4. Krytyka i dyskusja. 5. Zakończenie.

1. WPROWADZENIE

Jednym z najpoważniejszych wyzwań współczesnej kognitywistyki i filozofii umysłu jest wyjaśnienie, w jaki sposób systemy poznawcze potrafią rozpoznawać to, co w danym momencie jest istotne,

a co można pominąć. Problem ten, znany jako problem istotności (*relevance problem*) (Fodor, 2008), okazuje się ściśle powiązany z innymi klasycznymi zagadnieniami kognitywistyki i filozofii umysłu: problemem ram w sztucznej inteligencji, kwestią reprezentacji i ich relacji do kontekstu, a także pytaniami o racjonalność i świadomość. John Vervaeke, rozwijając swoje badania od lat dziewięćdziesiątych, twierdzi, że to właśnie zdolność[1] do realizacji istotności (*relevance realization*) (Vervaeke *et al*., 2012) stanowi centralną, choć dotąd niedostatecznie rozpoznaną kompetencję poznawczą, bez której niemożliwe jest zarówno naukowe wyjaśnienie funkcjonowania umysłu, jak i jego poszczególnych funkcji poznawczych.

Punktem wyjścia w badaniach Vervaekego była analiza niepowodzeń klasycznych teorii procesów poznawczych takich jak: kategoryzacja, rozwiązywanie problemów czy pamięć, które w mniej lub bardziej ukryty sposób zakładały zdolność do odróżniania tego, co istotne od tego, co nieistotne. W efekcie powstała luka eksplanacyjna: próba wyjaśnienia tych zdolności opierała się na pojęciach, które same wymagały wyjaśnienia. Vervaeke wskazuje, że to właśnie współzależność między reprezentacją a relewancją odsłania najgłębszą aporię kognitywistyki, a zarazem wyznacza kierunek koniecznej zmiany paradygmatu (Vervaeke, 1997; Vervaeke *et al*., 2012).

Teoria realizacji istotności (RR) stanowi próbę przełamania tej sytuacji. W jej świetle relewancja nie jest ani statyczną cechą reprezentacji, ani jedynie wynikiem stosowania odpowiednich heurystyk, lecz dynamicznym i kontekstowo modulowanym procesem, dzięki któremu możliwe staje się inteligentne działanie i rozumienie. RR rozwija się w pracach Vervaekego od pierwszych analiz problemu ram i racjonalności w latach dziewięćdziesiątych, przez krytyczne

1 W artykule pojęcia takie jak „zdolność", „kompetencja" czy „funkcja" będą używane zamiennie. Nie jest to użycie w pełni rygorystyczne, lecz odpowiada praktyce Johna Vervaeke, który w swoich pracach stosuje na przemian terminy *ability, competence, function, process* czy *mechanism* w odniesieniu do realizacji istotności, bez wyraźnego ich systematycznego rozgraniczania.

polemiki ze Sperberem i Wilson oraz Cherniakiem (Vervaeke, 1997; Vervaeke *et al*., 2012; Chiappe, Vervaeke, 1997), aż po współczesne próby integracji z enaktywizmem i teorią autonomii biologicznej (Jaeger *et al*., 2024), teorią przetwarzania predykcyjnego (Andersen *et al*., 2022), badaniami nad ogólną inteligencją (Vervaeke, Ferraro, 2013a) i enaktywistycznym podejściem do teorii osobowości (Hovhannisyan, Vervaeke, 2021).

Mimo tego, że problem istotności jest wprost omawiany dopiero od niedawna, to intuicje związane z istnieniem wyjątkowej zdolności do selekcji idei pojawiały się już wcześniej. Dobrze niezwykłość tej zdolności i trudności z jej wyjaśnieniem przedstawił Hume w *Traktacie o naturze ludzkiej*: „Nie ma rzeczy bardziej godnej podziwu niż ta gotowość i łatwość, z jaką wyobraźnia poddaje swe idee i stawia je przed umysłem w tej samej chwili, w której stają się one konieczne lub użyteczne. Wyobraźnia przebiega od jednego końca świata do drugiego, zbierając te idee, które należą do danego tematu. Można by pomyśleć, że cały świat intelektualny idei zjawia się naraz przed naszym wzrokiem i że nie potrzebujemy czynić nic więcej jak tylko wybrać te, które są najbardziej odpowiednie dla naszego celu. W istocie rzeczy nie mogą one wszystkie być obecne, wyjąwszy tylko te właśnie, jakie w ten sposób zbiera pewnego rodzaju magiczna władza duszy, która choć jest zawsze najbardziej doskonała u wielkich geniuszy i jest właściwie tym, co nazywamy geniuszem, to przecież nie da się wyjaśnić nawet największymi wysiłkami rozumu ludzkiego." (Hume, 1963)

Celem niniejszego artykułu jest krytyczna rekonstrukcja teorii realizacji istotności poprzez omówienie rozwijających się aspektów tej teorii w latach 1997–2024 wraz z jej zmianami oraz wskazanie na punkty wymagające dalszego wyjaśnienia bądź przeformułowania. Najpierw zostanie omówiona geneza problemu w pracach Vervaekego. Jest to namysł nad naturalizacją poznania i identyfikacją problemu istotności w tym kontekście. Równoległym przejawem tego zagadnienia jest problem ram i krytyka reprezentacji mentalnych oraz

wyjaśnień obliczeniowych. Przytoczone zostaną ówczesne próby rozwiązania problemu istotności wraz z ich krytyką ze strony Vervaekego. Następnym etapem będzie przedstawienie rozwoju teorii oraz jej miejsce w szerszym kontekście kognitywistyki. Na końcu zostanie przedstawiona krytyka i możliwe kierunki dalszego rozwoju teorii. W ostatniej części artykułu zostaje zidentyfikowany rozdźwięk pomiędzy deklarowaną przez Vervaekego krytyką paradygmatu obliczeniowego i proponowaniem rozwiązań enaktywistycznych, a niekonsekwencją terminologiczną. Vervaeke podkreśla, że RR pozwala na nawigację w nieprzewidywalnym „dużym świecie", czego nie są w stanie dokonać sztywne procedury operujące jedynie w sformalizowanych „małych światach". Ewolucja teorii w stronę ujęcia biologicznego ma pozwolić uniknąć błędnego koła, sytuując źródło relewancji w procesach autopojezy i adaptacji, które dopiero na wyższych szczeblach organizacji umożliwiają emergencję reprezentacji. Tak rozumiana istotność jest relacją emergentną, której nie da się w pełni zredukować do procedury obliczeniowej bez utraty jej fundamentalnego sensu. Pozostaje jednak trudność w precyzyjnym wyrażeniu tego podejścia bez odwoływania się do pojęć z bagażem reprezentacjonistycznym . Artykuł kończy się sugestią, że powracające problemy z wyjaśnieniem teorii realizacji istotności mogą wynikać z niewłaściwych założeń dotyczących tego, czym jest sama relewancja.

W niniejszym artykule pojęcia „istotność" i „relewancja" będą używane zamiennie. Polskie słowo „istotność", choć stanowi tłumaczenie angielskiego *relevance*, niesie ze sobą konotacje metafizyczne związane z pojęciem „istoty". Aby uniknąć sugerowania, że chodzi o kategorię ontologiczną w sensie klasycznej metafizyki, a jednocześnie podkreślić niezależność analizowanego pojęcia od tych skojarzeń, oba terminy będą stosowane zamiennie. Takie rozwiązanie pozwala zachować równowagę między wiernością oryginałowi a precyzją filozoficzną w języku polskim.

Wszystkie cytaty z języka angielskiego podaję w tłumaczeniu własnym.

2. KONTEKST FILOZOFICZNY I NAUKOWY

2.1. NATURALISTYCZNY IMPERATYW W KOGNITYWISTYCE

W swoim doktoracie z 1997 roku John Vervaeke zaprezentował ramy metodologiczne, które nazywa naturalistycznym imperatywem kognitywistyki. Zgodnie z nim każde wyjaśnienie procesów poznawczych musi spełniać trzy podstawowe warunki: analizę, formalizację i mechanizację. Po pierwsze, konieczna jest analiza, czyli precyzyjne określenie, jakie zjawisko i jakie jego komponenty mają być uznane za poznawcze i dlaczego. Po drugie, niezbędna jest formalizacja, umożliwiająca przedstawienie wyjaśnienia procesu w kategoriach ścisłych i podatnych na modelowanie. Wreszcie, trzeci wymóg to mechanizacja, która zakłada, że proces kognitywny musi być możliwy do zaimplementowania w rzeczywistym mechanizmie – biologicznym lub sztucznym. Dopiero spełnienie tych trzech kryteriów pozwala mówić o właściwie naturalistycznym wyjaśnieniu zjawiska poznawczego (Vervaeke, 1997).

Tak sformułowany imperatyw ma istotne konsekwencje krytyczne. Vervaeke pokazuje, że wielu tradycyjnych kandydatów na podstawowe procesy poznawcze – takich jak kategoryzacja, rozwiązywanie problemów czy pamięć – nie spełnia tych warunków. Każdy z tych procesów w ostatecznym rozrachunku odwołuje się bowiem do sądów o istotności. Aby zaklasyfikować dane jako należące do określonej kategorii, wybrać odpowiednią heurystykę przy rozwiązywaniu problemu, czy przywołać właściwe wspomnienie, system poznawczy musi najpierw ustalić, co jest relewantne w danej sytuacji. Z tego powodu Vervaeke uznaje, że rozpoznanie istotności jest fundamentalne dla działania wszystkich podstawowych funkcji poznawczych. Dlatego też, jeżeli zdolność do rozróżniania relewancji jest już założona, a nie

wyjaśniona, to analiza, czyli wyznaczenie kluczowych komponentów procesu, a więc i formalizacja i mechanizacja tych procesów, okazują się niepełne (Vervaeke, 1997).

Vervaeke powołuje się na konieczność wyznaczenia *criterion of the cognitive* (Green, 1996) czyli kryterium tego, co uznajemy za zjawisko poznawcze. Pozwoli ono: 1. zidentyfikować, czy dane zjawisko, proces lub stan należy do domeny poznawczej (czy jest „poznawcze", a nie tylko biologiczne, emocjonalne, motoryczne itp.), 2. oddzielić poznanie od innych funkcji psychicznych lub biologicznych (np. od homeostazy, instynktu lub reakcji automatycznej), 3. wyznaczyć ramy opisu i wyjaśniania – tzn. określić, jakie poziomy analizy i jakie typy mechanizmów są właściwe do badania poznania. To kryterium nie może odwoływać się do innych „poznawczych" pojęć bez ich wcześniejszego zdefiniowania. Powinno umożliwiać tworzenie praw i uogólnień – np. pozwolić na odkrywanie reguł rządzących poznaniem. Nie powinno z góry zakładać konkretnej realizacji fizycznej (np. że poznanie musi być „biologiczne" albo „neuronalne"). Musi być zoperacjonalizowane, czyli możliwe do zastosowania w praktyce badawczej (np. eksperymentalnej, neuroobrazowej, symulacyjnej) (Vervaeke, 1997).

W ten sposób naturalistyczny imperatyw staje się nie tylko programem metodologicznym kognitywistyki, ale także punktem wyjścia do wyodrębnienia teorii realizacji istotności jako jedynego kandydata na proces, który mógłby pełnić rolę właściwego kryterium poznawczości.

2.2. PROBLEM RAM I PROBLEM REPREZENTACJA–RELEWANCJA

Jednym z najczęściej dyskutowanych wyzwań, które ujawniły się wraz z rozwojem sztucznej inteligencji, jest tak zwany problem ram. Został on sformułowany w paradygmacie logicznym *Good Old Fashioned Artificial* Intelligence (GOFAI) (McCarthy, Hayes, 1969). Problem wynikał bezpośrednio z wyzwań związanych z logicznym

programowaniem maszyn. Kiedy w świecie fizycznym wykonywana jest jakakolwiek akcja (np. robot przesuwa klocek), zmienia ona tylko bardzo niewielki wycinek rzeczywistości, podczas gdy większość otoczenia pozostaje nietknięta. McCarthy i Hayes zauważyli, że system wymagał nie tylko równań opisujących to, co się zmienia, ale także tzw. aksjomatów ramy (*frame axioms*), które określałyby, co nie ulega zmianie w wyniku danego działania. W najogólniejszym ujęciu polega on na tym, że system poznawczy – czy to sztuczny, czy biologiczny – działając w dynamicznym i złożonym środowisku, musi odróżniać informacje istotne od przeważającej ilości danych nieistotnych. Jeśli system miałby w każdym momencie rozważać wszystkie możliwe konsekwencje swoich działań oraz wszystkie potencjalne zmiany w otoczeniu, uległby eksplozji kombinatorycznej (zbiór elementów, które należałoby sprawdzić przed podjęciem działania byłby zbyt duży, aby dany system mógł go uwzględnić w efektywnie skończonym czasie), która czyniłaby działanie praktycznie niemożliwym. Klasyczne modele sztucznej inteligencji – takie jak General Problem Solver (Newell, Simon, 1972) – właśnie z tego powodu okazały się niewydolne: ich architektury nie zawierały mechanizmu selekcji relewancji, a więc nie były zdolne do efektywnego rozwiązywania problemów w realistycznych warunkach (Vervaeke *et al*., 2012).

Shanahan zaprezentował rozwiązanie problemu ram w ramach rachunku zdarzeń (*event calculus*), wykorzystując w tym celu cyrkumskrypcję (*circumscription*), która pozwala na minimalizację rozszerzeń predykatów odpowiadających za efekty akcji i występowanie zdarzeń. Autor wykazuje, że dzięki zastosowaniu niemonotonicznego domniemania o tzw. prawie bezwładności zdrowego rozsądku, system automatycznie zakłada brak nieoczekiwanych zmian w świecie, co w większości praktycznych zastosowań sprowadza się do techniki dopełniania predykatów. Dzięki temu uznaje się, że problem został rozwiązany w ramach GOFAI (Shanahan, 2016).

Problem ram szybko ujawnił jednak swoje znaczenie wykraczające poza kontekst techniczny. W filozofii i kognitywistyce zaczęto

dostrzegać, że to nie jest jedynie przeszkoda w konstrukcji systemów logicznych mających modelować realne procesy poznawcze, lecz przejaw głębszego ograniczenia teorii poznania (Dennett, 1987; Fodor, 2008). W istocie każda teoria posługująca się pojęciem reprezentacji zakłada już uprzednie rozpoznanie tego, co jest relewantne (Goodman, 1972; Vervaeke, 1997). Każda reprezentacja jest uchwyceniem wybranego aspektu rzeczywistości. Aby określona reprezentacja mogła być skuteczna, system poznawczy musi wiedzieć, które aspekty rzeczywistości należy w niej uwzględnić, a które pominąć. Tymczasem właśnie to rozpoznanie jest najbardziej zagadkowe i wymaga osobnego wyjaśnienia.

Powstaje więc błędne koło: reprezentacja wymaga relewancji, a relewancja wymaga reprezentacji. Jeżeli próbujemy wyjaśnić poznanie przez odwołanie się do reprezentacji, musimy już wcześniej założyć zdolność do dokonywania sądów o istotności, czyli zdolność, której nie wyjaśniamy, a jedynie przesuwamy do wcześniejszego etapu. Problem ramy i problem reprezentacja–relewancja odsłaniają w ten sposób wspólną strukturę aporii: brak teorii, która byłaby w stanie pokazać, jak system poznawczy zyskuje dostęp do tego, co w danej sytuacji ważne, bez uciekania się do założeń o charakterze pozapoznawczym czy arbitralnym. Jak pisze Vervaeke: „W skrócie, problem ram jest problemem reprezentacji i relewancji, co utrudnia rozwiązanie go, ponieważ opiera się mechanizacji. Opiera się mechanizacji, ponieważ opiera się analizie i formalizacji”[2] (Vervaeke, 1997).

To właśnie w obliczu tej aporii Vervaeke wprowadza swoją teorię realizacji istotności. Jej ambicją jest wyjście poza schemat błędnego koła i wskazanie mechanizmu, dzięki któremu relewancja nie jest zakładana, lecz faktycznie wyjaśniana w procesach poznawczych.

---

2 “In short, the frame problem is the problem of representation and relevance and it resists a solution because relevance resists mechanisation. It resists mechanisation because it resists analysis and formalization.”

2.3. TEORIE RELEWANCJI W LITERATURZE I ICH KRYTYKA

Pierwszą poważną próbą ujęcia problemu istotności w kategoriach kognitywnych była koncepcja minimalnej racjonalności Cherniaka (1986). Zgodnie z nią ludzkie zdolności poznawcze są radykalnie ograniczone przez skończoną pamięć i moc obliczeniową. W tym ujęciu selekcja informacji dokonywana jest nie tyle dzięki istnieniu pozytywnego mechanizmu rozpoznającego istotność, ile poprzez eliminację spowodowaną ograniczeniami strukturalnymi pamięci: jednostka nie jest w stanie uwzględnić wszystkiego, więc jej system poznawczy niejako „z konieczności" filtruje dane. Rozwiązanie to proces poznawczy traktuje jako proces obliczeniowy, przetwarzający informację, ale jednocześnie uwzględnia jego fundamentalne ograniczenia, dzięki czemu bierze pod uwagę problem eksplozji kombinatorycznej. Nie pozostaje jednak bez trudności. Po pierwsze, przesuwa ono problem w inne miejsce – z decyzji poznawczej na strukturalne ograniczenia systemu – nie wyjaśniając jednak, jaki proces odpowiada za powstanie takiej struktury ani w jaki sposób organizm rozpoznaje, które informacje należy zachować i użyć w nowych kontekstach. Po drugie, zakłada, że relewancja jest wyłącznie skutkiem braku zasobów (ograniczeniem selektywnym), a nie aktywnym procesem poznawczym (ograniczeniem umożliwiającym), co redukuje jej rolę do pasywnego efektu ubocznego (Vervaeke, 1997).

Znacznie większy wpływ na badania nad istotnością wywarła teoria relewancji Sperbera i Wilson (1995). Ich propozycja, sformułowana na gruncie pragmatyki językowej, zakłada, że interpretacja wypowiedzi polega na dążeniu do optymalnej równowagi między efektem poznawczym a kosztem przetwarzania. Odbiorca, konfrontując wypowiedź z własnym kontekstem wiedzy, ocenia jej potencjalny wkład informacyjny w relacji do wysiłku, jaki musi włożyć w jej interpretację. Dzięki temu teoria trafnie wyjaśnia, jak możliwe jest szybkie i elastyczne rozumienie komunikacji w warunkach

niejednoznaczności i niepełnych danych. Jej sukces polega na dostarczeniu precyzyjnych narzędzi opisu procesów pragmatycznych.

Relewancja jest też jedną z maksym konwersacyjnych Grice'a (1975). Grice uważał, że maksyma relewancji jest tylko jedną spośród czterech zasad komunikacji (ilości, jakości, relewancji i sposobu). Sperber i Wilson argumentują, że pozostałe maksymy można sprowadzić do maksymy relewancji, czyniąc z niej najważniejszą zasadę komunikacji. Relewancja komunikacyjna, która jest centralna dla rozumienia języka i działania systemów symbolicznych, polega na tym, że jednostka przetwarza i przekazuje te informacje, które przynoszą największy efekt poznawczy przy możliwie najmniejszym wysiłku.

Jednak Vervaeke pokazuje, że także ta teoria nie rozwiązuje kluczowego problemu naturalizacji relewancji. Po pierwsze, choć pojęcia „efekt poznawczy" i „nakład przetwarzania" są intuicyjnie zrozumiałe, nie zostały formalnie zdefiniowane w sposób umożliwiający ich implementację w mechanizmach poznawczych. Nie wiadomo, jak system miałby sam rozpoznać, który efekt jest „większy" ani jak obliczyć koszt przetwarzania informacji w dynamicznie zmieniającym się kontekście. Po drugie, teoria Sperbera i Wilson operuje na poziomie funkcjonalnym, ale nie dostarcza realizacyjnych mechanizmów selekcji relewancji – tj. nie mówi, jak mózg lub system poznawczy dochodzi do oceny relewancji danych informacji przed ich przetworzeniem. Vervaeke podkreśla, że brakuje tu połączenia między teorią a implementacją: mówimy o efektach, które obserwujemy, ale nie rozumiemy procesów, które do nich prowadzą. Po trzecie, tak jak u Cherniaka, teoria relewancji zakłada już istnienie struktur reprezentacyjnych – zdolności inferencyjnych, systemów pamięci i zasobów semantycznych – które pozwalają na ocenę relewancji, a więc nie rozwiązuje błędnego koła między reprezentacją a relewancją, lecz jedynie je przesuwa na wyższy poziom (Vervaeke, 1997).

Można wysunąć jeszcze jeden zarzut przeciwko takiemu rozumieniu relewancji. W tej koncepcji relewancja jest sprowadzona do generalizacji. Przykładowo, streszczenie książki to skompresowana

informacja, czyli więcej informacji w stosunku do mniejszego nakładu poznawczego (przeczytanie streszczenia „kosztuje" mniej niż całej książki). To by oznaczało, że tylko streszczenia, uogólnienia są istotne. Jasne jest jednak, że często szczegóły bywają ważne i użyteczne, w zależności od celów i kontekstu w jakim operujemy. Z tego powodu teoria relewancji pomija niezwykle ważny element charakterystyki istotności jakim jest specyfikacja informacji.

3. TEORIA REALIZACJI ISTOTNOŚCI (RR) – ROZWÓJ I GŁÓWNE TEZY

Teoria realizacji istotności Vervaekego powstała jako próba rozwiązania problemu reprezentacja–relewancja i problemu ramy, wskazanych wcześniej jako kluczowe trudności kognitywistyki. Jej celem jest pokazanie, w jaki sposób system poznawczy faktycznie konstytuuje to, co jest istotne w danej sytuacji, nie uciekając się do arbitralnych założeń ani nie popadając w błędne koło. RR traktuje istotność nie jako prostą właściwość reprezentacji ani jako pasywny skutek ograniczeń systemu, ale jako dynamiczny i kontekstualny proces, dzięki któremu poznanie staje się możliwe. Vervaeke uznaje, że poznanie polega na przetwarzaniu informacji. Nie definiuje jednak tego terminu. Informacja jest czymś, co podlega selekcji ze względu na istotność dla agenta poznawczego, posiada strukturę oraz podlega kategoryzacji (Vervaeke *et al*., 2012).

3.1. RELEWANCJA BEZ REPREZENTACJI

Jednym z najbardziej radykalnych punktów teorii RR jest teza, że relewancja nie musi być rozumiana jako cecha reprezentacji. W klasycznym paradygmacie poznawczym opartym na pojęciu podobieństwa jako obiektywnej cechy odpowiedzialnej za istnienie kategorii (Goodman, 1972) zakładano, że reprezentacja zawiera pewne właściwości – mniej lub bardziej trafnie odwzorowujące rzeczywistość – a system poznawczy musi wybrać spośród nich te, które są istotne dla

danego zadania. W ujęciu Vervaekego sytuacja wygląda odwrotnie: to proces realizacji istotności umożliwia powstawanie reprezentacji, nadając im sens i ukierunkowanie. To istotność pewnych elementów pozwala mówić o ich podobieństwie, a nie na odwrót. Reprezentacja jest wtórna wobec relewancji.

Takie odwrócenie perspektywy oznacza, że istotność nie jest dodatkiem do reprezentacji, lecz warunkiem jej sensowności. Dzięki RR możliwe staje się wyjście z błędnego koła reprezentacja–relewancja: system poznawczy nie potrzebuje gotowych reprezentacji, aby określić, co jest ważne, lecz poprzez realizację istotności wypracowuje reprezentacje, które okazują się relewantne.

3.2. RELEWANCJA JAKO CECHA TRANSJEKTYWNA I NIEHOMOGENICZNA

Vervaeke podkreśla, że relewancja ma charakter transjektywny, to znaczy powstaje w wyniku sprzężenia między organizmem a środowiskiem. Nie jest ona ani wyłącznie „w świecie" jako obiektywna cecha rzeczy, ani tylko „w umyśle" jako subiektywne nadanie znaczenia. Relewancja wyłania się w procesie interakcji, który nie daje się sprowadzić do prostego podziału na podmiot i przedmiot (Vervaeke *et al*., 2012).

Jednocześnie istotność jest niehomogeniczna – nie da się jej sprowadzić do jednego kryterium czy wymiaru. To, co jest relewantne w danej sytuacji, może wynikać z bardzo różnych czynników: biologicznych potrzeb, kontekstu społecznego, wcześniejszych doświadczeń czy aktualnych celów. Proces realizacji istotności polega więc na tym, że system poznawczy musi nieustannie integrować wiele heterogenicznych wymiarów i wydobywać z nich to, co w danym momencie umożliwia skuteczne działanie (Vervaeke *et al*., 2012).

Vervaeke na takim rozumieniu realizacji istotności opiera krytykę wyjaśniania poznania jedynie w kategoriach obliczeniowych. Każdy proces obliczeniowy opiera się na zbiorze reguł podstawowych rządzących zmianą fizycznej struktury, która swoim działaniem realizuje

stany przewidziane przez tę logikę (Horsman *et al*., 2014). Według niego relewancja nie jest określona „z góry". To oznacza, że nie można sformułować reguł na podstawie których można wywnioskować lub obliczyć co jest aktualnie istotne. Poza tym, każda reguła jest już reprezentacją. Reprezentacje są już wybrane ze względu na istotność. Zatem RR nie może polegać na procesie obliczeniowym. Ta sama krytyka dotyczy również rozwiązań algorytmicznych. Algorytm rozumiany jest tutaj jako zestaw instrukcji, które są zdefiniowane dla skończonego zbioru stanów i mają określony typ wyniku (Jaeger *et al*., 2024). Bardzo podobną krytykę rozwiązania problemu istotności opartego na zestawie reguł w kontekście paradygmatu aktywnej inferencji (*active inference*) zaproponował Parvizi-Wayne (2025).

Istotność w ujęciu Vervaekego ma strukturalne podobieństwo do procesu adaptacji w ewolucji biologicznej. Tak jak organizmy nieustannie dostosowują swoje cechy do zmiennych warunków środowiskowych, tak system poznawczy reguluje swoje mechanizmy w celu adekwatnego wyróżniania tego, co znaczące w danej sytuacji. W obu przypadkach nie chodzi o realizację jednej, z góry określonej funkcji, lecz o dynamiczne balansowanie między presją selekcyjną a potrzebą zachowania różnorodności. Relewancja, podobnie jak adaptacja, nie jest więc rezultatem pojedynczego aktu decyzyjnego, lecz emergentnym efektem ciągłego procesu regulacji i przystosowania (Vervaeke *et al*., 2012).

Konsekwencje tej podwójnej charakterystyki są dalekosiężne. Po pierwsze, relewancja nie może być zdefiniowana ani jako czysto subiektywna ocena, ani jako obiektywna właściwość świata – musi być rozumiana jako relacyjny proces konstytucji znaczenia. Po drugie, RR nie może być pojmowana jako jeden mechanizm czy algorytm, lecz jako dynamika wielu współdziałających procesów, które wspólnie umożliwiają powstawanie spójnych, choć zawsze kontekstowo ograniczonych sądów o istotności (Vervaeke *et al*., 2012; Vervaeke, Ferraro, 2013a).

3.3. MODEL EKONOMICZNY TEORII RR

Centralnym elementem RR jest tzw. model ekonomiczny. Vervaeke proponuje rozumienie realizacji istotności w kategoriach optymalizacji zasobów poznawczych: system musi nieustannie podejmować decyzje, które równoważą różne koszty i korzyści. Proces ten można rozumieć analogicznie do procesów rynkowych – różne opcje konkurują o uwagę i zasoby, a istotne okazują się te, które zapewniają największy zwrot przy najniższym koszcie w danym kontekście (Jaeger *et al*., 2024; Vervaeke *et al*., 2012; Vervaeke, Ferraro, 2013a).

W przeciwieństwie do prostych heurystyk, model ekonomiczny nie zakłada istnienia stałego kryterium selekcji. Zamiast tego zakłada dynamiczną równowagę, w której relewancja jest stale renegocjowana w zależności od sytuacji. To tłumaczy, dlaczego system poznawczy może działać efektywnie w warunkach niepewności i zmienności, unikając zarówno sztywności, jak i chaosu.

3.4. PRZETWARZANIE OPOZYCYJNE

Kolejnym kluczowym aspektem RR jest teza, że realizacja istotności zachodzi poprzez przetwarzanie opozycyjne (*oppositional processing*). Oznacza to, że system poznawczy nie działa tylko w kierunku maksymalizacji jednej wartości (np. dokładności), lecz balansuje między przeciwstawnymi wymaganiami. Opozycje te pełnią funkcję regulacyjną: ograniczają dominację jednego wymiaru kosztem innych i pozwalają utrzymać elastyczność (Jaeger *et al*., 2024; Vervaeke *et al*., 2012; Vervaeke, Ferraro, 2013a).

Przykładem jest równowaga między szczegółowością a ogólnością reprezentacji: zbyt szczegółowe ujęcie prowadzi do sztywności i braku możliwości generalizacji, zbyt ogólne – do utraty przydatności w konkretnych sytuacjach. Dzięki przetwarzaniu opozycyjnemu system potrafi dynamicznie przesuwać się na kontinuum między tymi skrajnościami, znajdując punkt adekwatny do aktualnych warunków.

3.5. KOMPROMISY W REALIZACJI ISTOTNOŚCI

W ujęciu Vervaekego realizacja istotności nie jest algorytmem przetwarzania informacji, lecz zbiorem wewnętrznych ograniczeń ekonomicznych, które w interakcji wyłaniają istotność jako fenomen wyższego rzędu. Relewancja nie jest więc pojedynczą funkcją, ale rezultatem dynamicznego współdziałania mechanizmów regulacyjnych, które określają, jak dobrze podmiot poznający współgra ze swoim środowiskiem (Vervaeke *et al*., 2012; Vervaeke, Ferraro, 2013a; Andersen *et al*., 2022; Hovhannisyan, Vervaeke, 2022).

Pierwszym z tych mechanizmów jest zakres poznawczy (*cognitive scope*) (CS), czyli zdolność do balansowania między generalizacją a szczegółowością strategii poznawczych. To napięcie – widoczne choćby w problemie przeuczenia (*overfitting)* i technice karania dużych wag (*weight decay)* w sieciach neuronowych – ilustruje, że relewancja powstaje jako efekt równoważenia perspektywy „mikroskopu" i „teleskopu" (Vervaeke *et al*., 2012).

Drugim jest regulowanie poznawcze (*cognitive tempering*) (CT), czyli równowaga między eksploracją a eksploatacją. Mechanizmy analogiczne do wzmacniania i hamowania powrotu (*reinforcement vs. inhibition of return*) pozwalają systemowi nie tylko korzystać ze sprawdzonych rozwiązań, lecz także otwierać się na nowe możliwości. CT umożliwia regulację napięcia między teraźniejszością a potencjalną przyszłością (*projectability*) (Vervaeke *et al.,* 2012).

Trzecim jest priorytetyzacja poznawcza (*cognitive prioritization*) (CP), odpowiadająca za hierarchizację zadań i zasobów w warunkach niepewności. Poznanie zawsze wiąże się z założeniem inwestycji poznawczych; CP pozwala dynamicznie przesuwać priorytety – np. między zaspokajaniem głodu a unikaniem drapieżnika – dzięki procesowi analogicznemu do elastycznej modulacji funkcji kosztu (Vervaeke *et al*., 2012).

Te trzy mechanizmy nie działają niezależnie, lecz tworzą sprzężoną architekturę, w której relewancja wyłania się z ich współdziałania.

Nad nimi działa metapoziomowe napięcie między efektywnością (selekcja, uproszczenie, optymalizacja) a rezyliencją (różnorodność, nadmiarowość, adaptacja). Ich równoważenie przypomina dynamikę ewolucji biologicznej, gdzie selekcja i mutacja wspólnie napędzają rozwój (Vervaeke *et al*., 2012).

W efekcie RR jawi się jako proces poznawczej „ewolucji w działaniu": nie jest redukowalne ani do reprezentacji, ani do jednej heurystyki, lecz jest emergentnym rezultatem interakcji ograniczeń poznawczych. To właśnie ta wielowymiarowa ekonomia poznania umożliwia podmiotowi nieustanne dostosowywanie się do zmiennego i niepewnego świata (Vervaeke *et al*., 2012).

3.6. ROZWÓJ TEORII RR I JEJ ZASTOSOWANIA

Teoria realizacji istotności, początkowo rozwijana jako odpowiedź na problem ram i trudności w naturalizacji pojęcia relewancji, w ostatnich latach znalazła zaskakująco szerokie zastosowanie w różnych dziedzinach kognitywistyki i filozofii umysłu. Jej siła wyjaśniająca polega na tym, że nie proponuje pojedynczego algorytmu czy modułu poznawczego, lecz opisuje ogólną architekturę regulacyjną, w której relewancja wyłania się jako efekt dynamicznego równoważenia opozycyjnych ograniczeń i tendencji. Dzięki temu realizacja istotności okazuje się teorią elastyczną i integracyjną, zdolną do łączenia problematyki tak odległej, jak predykcyjne modele percepcji (*predictive perception models*), psychologia optymalnego doświadczenia (*flow*), enaktywistyczne ujęcia osobowości (*Enactivist Big Five Theory,* EB5T), badania nad ogólną inteligencją, teorie autonomii biologicznej, a także klasyczne zagadnienia filozoficzne dotyczące racjonalności i mądrości. W każdym z tych obszarów RR nie tylko dostarcza nowych narzędzi analitycznych, lecz także wskazuje na wspólne struktury i mechanizmy regulacyjne, pozwalające zrozumieć, w jaki sposób systemy poznawcze skutecznie orientują się w złożoności świata. Ta zdolność integracji i transdyscyplinarnego

zastosowania czyni teorię RR jedną z najbardziej obiecujących propozycji w najnowszej kognitywistyce.

Jednym z najważniejszych obszarów integracji realizacji istotności stało się przetwarzanie predykcyjne (Andersen *et al*., 2022). Obie teorie podzielają przekonanie, że poznanie jest procesem dynamicznej regulacji w warunkach niepewności, ale koncentrują się na różnych jego wymiarach. Przetwarzanie predykcyjne opisuje umysł jako system hierarchicznych modeli generatywnych, których celem jest minimalizacja błędu predykcji. Kluczowy mechanizm ważenia precyzyjnego (*precision-weighting)* decyduje, które błędy traktować jako sygnały, a które jako szum. RR ma odpowiadać właśnie za to, jak system rozpoznaje, które błędy i informacje są w ogóle warte uwagi. Teoria PP zakłada działanie mechanizmu, którego zadaniem ma być de facto realizacja istotności. W tym momencie RR proponuje rozwinięcie wyjaśnienia działania tego mechanizmu za pomocą przetwarzania opozycyjnego. Relewancja staje się tu efektem emergentnej ekonomii kompromisów, a nie prostym parametrem wagowym. W tym sensie RR można traktować jako teorię komplementarną z PP: przetwarzanie predykcyjne wyjaśnia, jak przewidywania są generowane i korygowane, a realizacja istotności – kiedy i dlaczego określone przewidywania mają znaczenie adaptacyjne.

Integracja realizacji istotności z przetwarzaniem predykcyjnym znalazła zastosowanie także w analizie spektrum autyzmu i schizotypii (Andersen *et al*., 2022). W ramach PP zakłada się, że zaburzenia te wynikają z odmiennych sposobów nadawania wagi błędom predykcji: w autyzmie system poznawczy zbyt silnie obciąża błędy sensoryczne, co prowadzi do trudności w abstrahowaniu i nadmiernego przywiązania do szczegółów, natomiast w schizotypii przypisuje się zbyt dużą wagę przewidywaniom wyższego rzędu, co skutkuje halucynacjami i iluzorycznymi korelacjami. Ujęcie RR pozwala doprecyzować ten model, wskazując, że u podstaw tych różnic leży odmienna konfiguracja mechanizmów realizacji istotności. W autyzmie zaburzony jest przede wszystkim balans między kompresją a partykularyzacją

(*cognitive scope*), natomiast w schizotypii – równowaga między eksploracją a eksploatacją (*cognitive tempering*). Z tej perspektywy, spektrum autyzmu i schizotypię można rozumieć jako kontinuum różnych sposobów zakłócenia architektury ekonomii poznawczej, w której emergentna relewancja nie wyłania się w sposób adaptacyjny (Andersen *et al*., 2022).

Zastosowania teorii RR wykraczają jednak daleko poza paradygmat predykcyjny. Interesującym polem integracji okazała się teoria stanu *flow*, rozwinięta przez Csikszentmihalyiego (1996). Stan *flow* można interpretować jako proces poznawczy, a w szczególności jako koherencję mechanizmów RR: zakres poznawczy dopasowany do wyzwania, dynamiczne równoważenie eksploracji i eksploatacji oraz dominacja jednej, wysoko priorytetowej funkcji kosztu. To właśnie takie zestrojenie mechanizmów pozwala na subiektywne doświadczenie płynności, sprawczości i niskiej ceny wysiłku. *Flow* jawi się zatem jako stan optymalnej synchronizacji CS, CT i CP (Vervaeke *et al*., 2018).

Kolejnym zastosowaniem teorii RR w rozumieniu procesów poznawczych stała się teoria osobowości. Integracja teorii realizacji istotności z *Enactive Big Five Theory* otwiera nową perspektywę na rozumienie osobowości (Hovhannisyan, Vervaeke, 2022). EB5T interpretuje wielką piątkę nie jako zestaw statycznych cech, lecz jako dynamiczne wzorce regulacji, które konstytuują się w relacji podmiot–środowisko. W tym ujęciu osobowość nie jest strukturą wewnętrzną zamkniętą w jednostce, lecz formą ucieleśnionej i usytuowanej dynamiki, która kształtuje się w sprzężeniu między potrzebami organizmu, kontekstem społecznym i ekologicznym a afektywnymi modulacjami zachowania.

Teoria realizacji istotności dostarcza mechanizmu pozwalającego uchwycić, jak te wzorce regulacji są możliwe. Istotność, wyłaniająca się z interakcji między poznawczym zakresem, regulowaniem i priorytetyzacją, zapewnia ramy, w których osobowość można rozumieć jako stabilne, choć elastyczne profile w sposobie równoważenia

kompromisów poznawczych i afektywnych. Na przykład wysoka neurotyczność może być rozumiana jako preferencyjne ustawienie priorytetyzacji poznawczej na monitorowanie zagrożeń, a ekstrawersja – jako przesunięcie równowagi w kierunku eksploracji i poszukiwania nagród społecznych (Hovhannisyan, Vervaeke, 2022).

W tym sensie RR pełni rolę architektury metaregulacyjnej osobowości: nie definiuje treści charakterologicznych, lecz opisuje zasady, według których jednostka nadaje znaczenie bodźcom, utrzymuje uwagę na określonych celach i reorganizuje swoje strategie działania w odpowiedzi na zmieniające się warunki. Dzięki temu możliwe jest ujęcie osobowości jako emergentnego wzorca istotności – względnie trwałej konfiguracji sposobów realizowania relewancji w świecie społecznym i biologicznym. Integracja RR z EB5T pozwala zatem przesunąć rozumienie osobowości z poziomu deskryptywnej taksonomii cech na poziom dynamicznej, enaktywnej i transjektywnej teorii regulacji (Hovhannisyan, Vervaeke, 2022).

Analiza związku teorii realizacji istotności z ogólną inteligencją wskazuje również, że RR może stanowić mechanizm wyjaśniający, w jaki sposób inteligencja przejawia się jako zdolność do adaptacji w szerokim zakresie sytuacji (Vervaeke, Ferraro, 2013a). Ogólna inteligencja, rozumiana nie jako pojedyncza zdolność obliczeniowa, lecz jako elastyczność w rozwiązywaniu problemów, wymaga systemu zdolnego do dynamicznego równoważenia sprzecznych wymogów poznawczych. RR dostarcza tu kluczowej ramy: inteligencja polega na umiejętności przełączania się między strategiami poznawczymi poprzez zarządzanie kompromisami – między generalizacją a szczegółowością, eksploracją a eksploatacją czy efektywnością a rezyliencją.

Vervaeke przywołuje badania neurokognitywne na temat korelacji ogólnej inteligencji z zaobserwowanymi wzorcami aktywacji i organizacji sieci neuronowych w mózgu. Jednym z nich jest samoorganizująca się krytyczność (*self-organized criticality*, SOC) - pojęcie z zakresu teorii systemów dynamicznych (Bak, Tang, Wiesenfeld, 1988) i sieć małego świata (*small-world network,* SWN), czyli pojęcie z dziedziny

teorii grafów. SOC i SWN stanowią teoretyczną interpretację działania mózgu związaną ze wskaźnikami ogólnej inteligencji. Vervaeke wykorzystuje te badania, aby pokazać, że tę bazę teoretyczną można zintegrować z zaproponowanym przez niego procesem realizacji istotności (Vervaeke, Ferraro, 2013a).

Zjawisko samoorganizującej się krytyczności można zobrazować prostym przykładem piasku w klepsydrze. Ziarenka spadają jedno po drugim i gromadzą się na dnie, tworząc stopniowo rosnącą górkę. Przez długi czas stos pozostaje stabilny, aż nagle kolejne ziarenko wywołuje lawinę – czasem niewielką, czasem obejmującą całą strukturę. Piasek rozsypuje się (traci wysokość), ale zyskuje zakres (większą powierzchnię), który umożliwia powstanie wyższego stosu. W ten sposób stos piasku może się cały czas powiększać poprzez tymczasowe utraty wysokości na rzecz szerokości. Kluczowe jest to, że system samoczynnie dąży do stanu krytycznego, w którym równowaga między stabilnością a możliwością nagłej reorganizacji utrzymuje się spontanicznie.

Systemy samoorganizującej się krytyczności balansujące między porządkiem a chaosem, wykazują zarówno stabilność, jak i zdolność do nagłych reorganizacji w odpowiedzi na zakłócenia. RR można interpretować jako mechanizm poznawczy SOC: system umysłowy reguluje swoje ograniczenia w taki sposób, aby pozostawać w stanie bliskim krytyczności, co umożliwia mu zarówno efektywne funkcjonowanie w znanych warunkach, jak i szybkie przystosowanie do sytuacji nowych. Właśnie ta zdolność do oscylowania między elastycznością a stabilnością *(efficiency-resiliency tradeoff*) stanowi fundament ogólnej inteligencji (Vervaeke, Ferraro, 2013a; Vervaeke, Ferraro, 2013b).

Drugim ważnym połączeniem między RR i ogólną inteligencją jest struktura sieci małego świata (Bullmore, Sporns, 2009), często przywoływana w badaniach nad organizacją mózgu. Sieci tego typu łączą wysoką lokalną klastrowość z krótkimi ścieżkami globalnymi, co pozwala na szybkie i efektywne rozchodzenie się informacji przy

zachowaniu modularności (Langer *et al*., 2012). W kontekście RR, SWN oferują architekturę neuronalną umożliwiającą realizację istotności: lokalne przetwarzanie sprzyja dopasowaniu do szczegółowych kontekstów, a globalne powiązania pozwalają na uogólnienie i integrację informacji. Inteligencja jawi się tu jako zdolność systemu do reorganizacji przepływów w sieci w taki sposób, by dynamicznie tworzyć i rozwiązywać wzorce relewancji (Vervaeke, Ferraro, 2013a).

Z tej perspektywy ogólna inteligencja nie jest ani prostą sumą zdolności cząstkowych, ani pojedynczym modułem obliczeniowym, lecz emergentnym skutkiem architektury RR działającej w warunkach SOC i SWN. To właśnie RR wyjaśnia, dlaczego system poznawczy potrafi radzić sobie z problemami o bardzo różnej naturze – ponieważ stale reguluje swoje wewnętrzne ograniczenia w taki sposób, aby utrzymywać równowagę pomiędzy stabilnością i adaptacyjnością, lokalnością i globalnością, skupieniem na szczegółach i uchwyceniem ogólności. W ten sposób RR można traktować jako procesowy fundament inteligencji ogólnej.

Kolejnym ważnym zastosowaniem teorii RR stała się teoria racjonalności i mądrości (Vervaeke, Ferraro, 2013b). W klasycznej tradycji filozoficznej racjonalność była utożsamiana przede wszystkim ze zdolnością do poprawnego rozumowania i wyciągania wniosków zgodnych z zasadami logiki. Tak rozumiana racjonalność opiera się na formalnych regułach inferencji, które pozwalają na zachowanie spójności przekonań i adekwatne uzasadnianie twierdzeń. Współczesna kognitywistyka poszerzyła to rozumienie, wskazując na potrzebę rozróżnienia między racjonalnością teoretyczną a praktyczną: pierwsza dotyczy poprawności przekonań, druga – skuteczności działania w świecie. Jednak w obu przypadkach pojawia się pytanie, w jaki sposób podmiot poznawczy rozpoznaje, które informacje są istotne i które przesłanki powinny zostać uwzględnione w danym akcie wnioskowania lub działania.

Warto w tym miejscu zestawić RR z teoriami racjonalności ograniczonej (*bounded rationality,* BR). Simon (1957) podkreślał, że

systemy poznawcze dysponują ograniczonymi zasobami obliczeniowymi i informacyjnymi, dlatego racjonalność w praktyce polega nie na maksymalizacji, lecz na satysfakcjonowaniu (*satisficing*), czyli wyborze wystarczająco dobrych rozwiązań. Gigerenzer (2021) rozwinął ten nurt, wskazując na rolę heurystyk jako „narzędzi adaptacyjnych", które w określonych warunkach środowiskowych działają niezwykle skutecznie. Oba te podejścia zakładają jednak, że system poznawczy musi dysponować mechanizmem rozpoznawania, kiedy dana heurystyka bądź reguła satysfakcjonowania jest adekwatna. Innymi słowy, BR sama zakłada istnienie procesu wyróżniania istotności (Vervaeke, Ferraro, 2013b).

Teoria realizacji istotności wprowadza tu kluczowe doprecyzowanie. Racjonalność – zarówno w wymiarze teoretycznym, jak i praktycznym – zakłada zdolność do dynamicznego rozpoznawania relewancji (Vervaeke, Chiappe, 1997). System poznawczy nie może bowiem traktować wszystkich dostępnych informacji jako jednakowo ważnych – musiałby mierzyć się z nieskończoną złożonością świata, co prowadziłoby do poznawczego paraliżu. Racjonalne działanie polega zatem na umiejętności wyłaniania tych elementów, które w danej sytuacji mają znaczenie, przy równoczesnym pomijaniu pozostałych. RR dostarcza wyjaśnienia tego procesu: racjonalność nie jest jedynie stosowaniem reguł logiki, ale przede wszystkim zdolnością do regulowania własnej ekonomii poznawczej, tak aby istotność wyłaniała się w sposób adaptacyjny.

W tym kontekście odróżnienie racjonalności od mądrości nabiera nowego znaczenia. O ile racjonalność wiąże się z lokalnym i kontekstowym rozpoznawaniem tego, co istotne dla rozwiązywania określonych problemów, o tyle mądrość wymaga zdolności do koordynowania tego procesu w długim horyzoncie czasowym oraz w złożonych, wielowymiarowych sytuacjach egzystencjalnych. Mądrość nie sprowadza się do sprawnego wnioskowania ani skutecznej decyzji w pojedynczym przypadku – obejmuje raczej metapoziom regulacji realizacji istotności. Oznacza to, że „mądry" podmiot nie

tylko dostrzega, co jest istotne tu i teraz, lecz potrafi także rozpoznać, jakie sposoby wyróżniania istotności są adekwatne w różnych kontekstach i w perspektywie całego życia.

Teoria realizacji istotności wypełnia tę lukę, wskazując na wewnętrzną architekturę mechanizmów ekonomicznych (poznawczy zakres, regulowanie i priorytetyzacja), które dynamicznie wyłaniają relewancję w zależności od kontekstu i celu działania. W przeciwieństwie do teorii heurystyk, które opisują już gotowe strategie postępowania, RR proponuje metamodel: wyjaśniający, w jaki sposób w ogóle powstaje i reguluje się przestrzeń zastosowania takich heurystyk. W tym sensie RR można potraktować jako fundament, na którym opierają się wszystkie ujęcia racjonalności ograniczonej – dostarczając brakującego mechanizmu, który tłumaczy, jak podmiot „wie", które ograniczenia są aktualnie relewantne i jaką strategię należy uruchomić (Vervaeke, Ferraro, 2013b).

W najnowszym ujęciu z 2024 roku teoria realizacji istotności została powiązana z koncepcją autonomii biologicznej (Moreno, Mossio, 2015), która wskazuje na współdziałanie trzech podstawowych wymiarów organizacji systemów poznawczych: autopojezy, antycypacji i adaptacji. Współdziałanie tych trzech wymiarów (nazwane trialektyką) ma stanowić uniwersalny fundament dla procesu realizacji istotności (Jaeger *et al.,* 2024). Autopojeza odnosi się do samopodtrzymywania funkcjonowania organizmu, antycypacja do zdolności przewidywania przyszłych stanów, a adaptacja do elastycznego dopasowywania się do zmiennych warunków środowiskowych. W kontekście RR te trzy wymiary jawią się jako fundamentalne ograniczenia regulacyjne, które system musi nieustannie równoważyć, aby mogła wyłaniać się istotność. Ujęcie to współgra z koncepcją sąsiednich możliwości (*adjacent possible*) Kauffmanna, według której systemy żywe rozwijają się poprzez eksplorację przestrzeni możliwych stanów dostępnych dzięki wcześniejszym osiągnięciom organizacyjnym. RR można zatem rozumieć jako mechanizm umożliwiający nawigację po sąsiednich możliwościach – dynamiczne filtrowanie

i wybieranie z nieskończonej liczby możliwości tych, które mają znaczenie w danym kontekście i które wspierają dalsze przetrwanie oraz rozwój systemu (Jaeger *et al*., 2024).

Podobne intuicje rozwijał Rosen (2012) w swoich modelach systemów antycypacyjnych, a także Hofmeyr (Hofmeyr, 2007), wskazując, że życie i poznanie można wyjaśniać poprzez sieci ograniczeń, które są jednocześnie samopodtrzymujące i projektujące się w przyszłość. Włączenie RR do tej trialektycznej perspektywy wpisuje się w szerszy paradygmat emergentnej agencyjności, w którym podmiotowość nie jest traktowana jako dana raz na zawsze własność, ale jako efekt złożonych, dynamicznych procesów organizacyjnych. Agencyjność w tym sensie wyłania się z równoważenia sprzecznych tendencji – samoodniesienia i otwartości, stabilności i zmiany, konserwacji i innowacji – w taki sposób, by system zachowywał zdolność działania i orientacji w świecie (Jaeger *et al*., 2024).

Tak rozumianą realizację istotności można traktować jako architekturę operacyjną emergentnej agencyjności: to właśnie ono sprawia, że organizm nie tylko utrzymuje siebie w czasie (autopojeza), ani nie odpowiada jedynie reaktywnie na zmiany (adaptacja), lecz potrafi także projektować i selekcjonować przyszłe możliwości (antycypacja). RR stanowi zatem mechanizm spajający trialektykę w spójną całość, wyjaśniając, jak istotność wyłania się jako zjawisko wysokiego rzędu w obrębie życia i poznania (Jaeger *et al*., 2024).

4. KRYTYKA I DYSKUSJA

Zaproponowane przez Vervaekego warunki naturalistycznego wyjaśnienia wykazują wyraźne podobieństwo do trójpoziomowej koncepcji wyjaśniania obliczeniowego przedstawionej przez Marra w 1982 roku (Marr, 2010)[3]. Marr wyróżnia poziom obliczeniowy, algorytmiczny oraz implementacyjny. Na poziomie obliczeniowym

3 Dziękuję anonimowemu recenzentowi za zwrócenie mojej uwagi na to podobieństwo.

określa się, jaki problem jest rozwiązywany i jakie funkcje pełnią jego podstawowe komponenty. Poziom algorytmiczny dotyczy konstrukcji modelu opisującego strukturę tych komponentów oraz relacje między nimi. Z kolei poziom implementacyjny lokalizuje ten model w konkretnym nośniku fizycznym, którego działanie realizuje przewidywania modelu. Analogiczna struktura wyjaśniania jest obecna u Vervaekego: identyfikacja funkcjonalnych składników procesu poznawczego, modelowanie ich wzajemnych zależności oraz osadzenie tych procesów w dynamice systemów biologicznych i neuronalnych odpowiada trzem poziomom Marra. Sugeruje to, że naturalizacja w ujęciu Vervaekego może być interpretowana jako szczególna forma wyjaśnienia obliczeniowego, nawet jeśli nie jest ona explicite tak nazywana. Powstaje jednak napięcie między tym strukturalnym podobieństwem a deklarowanym przez Vervaekego dystansem wobec podejść obliczeniowych. Mimo początkowego sceptycyzmu co do możliwości naturalistycznego wyjaśnienia relewancji, podejmuje on próbę jego realizacji, często odwołując się do pojęć takich jak selekcja, optymalizacja czy organizacja funkcjonalna. W konsekwencji jego stanowisko bywa przedstawiane jako antyobliczeniowe, choć w praktyce zachowuje ono istotne cechy schematu obliczeniowego. Można więc argumentować, że Vervaeke nie tyle całkowicie odrzuca język obliczeniowy, ile rekonfiguruje go w ramach bardziej dynamicznego i ucieleśnionego ujęcia poznania.

Autorzy artykułu z 2024 roku, którego współautorem jest Vervaeke (Jaeger *et al*., 2024) twierdzą, że RR nie może być procesem obliczeniowym. Podobne twierdzenia były postulowane już w 2012 roku (Vervaeke *et al*., 2012). Vervaeke przyjmuje za Fodorem, że obliczenia działają na symbolach, które są reprezentacjami (Fodor, 2008), a skoro RR poprzedza reprezentacje to nie może też działać w sposób obliczeniowy. Jeżeli realizacja istotności miałaby zostać w pełni sformalizowana jako proces obliczeniowy np. jako algorytm, to musiałaby działać za pomocą skończonego i określonego zbioru reguł. Tymczasem próba formalizacji prowadzi do trudności

związanych z przejściem od „dużego świata" (*big world*) – rzeczywistości nieograniczonej, otwartej i nieprzewidywalnej – do „małego świata" (*small world*), w którym reguły i stany są jasno określone i możliwe do obliczeniowego przetwarzania. Każdy algorytm działa bowiem jedynie w małym świecie, podczas gdy RR dotyczy zdolności orientowania się w „świecie dużym". Stanowisko to wskazuje, że RR nie może być rozumiana jako algorytm w ścisłym sensie, lecz raczej jako metaarchitektura emergentnych ograniczeń, która umożliwia systemowi balansowanie pomiędzy różnymi strategiami poznawczymi w sposób nie w pełni uchwytny w kategoriach obliczeniowych.

W artykule „Natural & Normative Dynamical Coupling" Hovhannisyan i Dewey (2017) omawiają teorię RR i używają jej do ugruntowania własnego wyjaśnienia naturalistycznej normatywności. Autorzy informują w części *Acknowledgements*, że był on konsultowany z Vervaekem. Można więc uznać, że przynajmniej fragmenty dotyczące teorii RR są zgodne z przekonaniami jej autora. Zawiera on następujący fragment: „Vervaeke i in. (2012) oferują najlepsze obecnie dostępne wyjaśnienie, ponieważ ich teoria realizacji relewancji rozwiązuje (przynajmniej teoretycznie) wszystkie problemy omówione w tej sekcji. Aby to osiągnąć, teoria ta obejmuje podejście heurystyczne, zapewniając, że jeden z procesów przeciwstawnych w każdej parze jest procesem heurystycznym (np. uczenie przez generalizację czy eksploracja). Pozwala jej to rozwiązać problem ram. Ponadto przypisuje procesy heurystyczne względnie ustalonym strategiom, unikając tym samym wspomnianego wcześniej problemu podejścia heurystycznego – że adaptuje się ono zbyt wolno, by system poznawczy mógł przetrwać. Oczywiście same względnie stałe strategie również skazywałyby system poznawczy na błędy w obliczu szybko zmieniającego się środowiska. Problem ten zostaje jednak rozwiązany przez zapewnienie, że jednym z procesów przeciwstawnych w każdej parze jest proces algorytmiczny, który dostosowuje się wystarczająco szybko, aby system poznawczy mógł przetrwać." (Hovhannisyan, Dewey, 2017, 8).

Powyższy cytat wskazuje na to, że RR jednak jest procesem przetwarzania informacji oraz wykorzystuje algorytmy jako elementy przetwarzania opozycyjnego. W kontekście argumentu o poprzedzaniu reprezentacji przez relewancję oraz o fundamentalnej roli RR w formalizacji, wydaje się, że to sformułowanie jest błędne. W końcu samo powstanie algorytmu wymaga realizacji istotności, więc skąd by się miał wziąć algorytm jako część tego procesu? Nie jest również powiedziane, czym jest informacja; pojęcie niesie ze sobą często intuicje reprezentacjonistyczne (Krzanowski, 2020). Rozumienie RR jako procesu przetwarzania informacji, która nie jest reprezentacją wymaga odniesienia do takiej koncepcji informacji, która jest spójna z tym wymogiem. Ale nawet zaproponowanie adekwatnej teorii informacji pozostawia jeszcze problem algorytmów w przetwarzaniu opozycyjnym. Algorytmy wymagają procesu formalizacji, za który ma odpowiadać RR (Jaeger *et al*., 2024; Vervaeke, 1997). Jeśli zatem twierdzimy, że RR sam wykorzystuje algorytmy, to wracamy do błędnego koła, którego od samego początku próbujemy uniknąć. Vervaeke z jednej strony (całkiem przekonująco) argumentuje przeciwko obliczeniowym wyjaśnieniom procesu realizacji istotności, z drugiej jednak strony, posługuje się w swoich propozycjach pojęciami takimi jak informacja czy strategie uczenia. Nie są one precyzyjnie zdefiniowane, a zatem powinniśmy zakładać ich domyślne znaczenie, a niesie ono ze sobą zobowiązania reprezentacjonistyczne i komputacyjne.

Wydaje się, że autor RR zdaje sobie sprawę z tych trudności. W artykule jego współautorstwa (Jaeger *et al*., 2024) rozważane są dwa podejścia – ekonomiczne (oparte na dynamicznej alokacji zasobów do konkurujących celów) (Vervaeke *et al*., 2012) oraz predykcyjne (oparte na ważeniu błędów predykcji w przetwarzaniu sensorycznym) (Andersen *et al*., 2022) – jako potencjalne modele RR. Oba jednak, mimo wartości heurystycznej, zakładają istnienie wewnętrznych celów agenta i pozostają w ramach obliczeniowych, co utrudnia ich generalizację do systemów pozbawionych jawnej intencjonalności, np. prostych organizmów. Autorzy w ten sposób formułują swoje

wątpliwości: „…podejście ekonomiczne i sformalizowana wersja przetwarzania predykcyjnego pozostają całe zanurzone w paradygmacie komputacyjnym, który widzi przypisywanie relewancji (a zatem zasobów lub precyzji wag) jako prosty, powtarzalny i rekurencyjny proces algorytmiczny. To prowadzi do pytania skąd się biorą cele organizmu i jak przewidywanie wyników może wpływać na strategie i podejmowane działania." (Jaeger *et al*., 2024, 7).

Vervaeke przyznaje, że te propozycje rozwiązania problemu istotności również pozostają obliczeniowe. Przypisywanie satysfakcji poszczególnym stanom kompromisów ekonomicznych między strategiami uczenia się zależy od celów organizmu. Brak wyjaśnienia tego, czym jest cel i skąd się bierze stanowi poważną lukę w prezentowanych dotychczas rozwiązaniach problemu istotności. Zapełnienie tej luki, poprzez wyjaśnienie tego czym jest wewnętrzna celowość może, ale nie musi stanowić pełnego wyjaśnienia procesu RR. Aby uniknąć problemu homonkularnego, przed którym sam Vervaeke tak często przestrzega, należy wyjaśnić także, w jaki sposób cele, bez zakładania istnienia świadomego podmiotu oraz bez polegania na jakimś mechanizmie reprezentacji, mogą pełnić rolę weryfikatorów spełnienia warunków „dopasowania" i „adaptacji". Poleganie na reprezentacjach wymagałoby formalizacji, którą jak już było powiedziane wcześniej, próbujemy ostatecznie wyjaśnić.

Ta teoria została rozszerzona na proces życiowy, który nie tylko wspiera funkcjonowanie systemów poznawczych, ale także wyłania i organizuje je na wyższych szczeblach biologicznej organizacji. Wyższy poziom organizacji biologicznej to struktura, w której zbiór autonomicznych jednostek podstawowych (komórek) zostaje poddany nowym regułom i ograniczeniom rozwojowym. Zmuszają one te jednostki do różnicowania się i współdziałania w taki sposób, aby ich aktywność podtrzymywała nowopowstałą całość, nawet jeśli ta nowa, zbiorowa tożsamość wciąż wymaga do istnienia zachowania autonomii przez swoje pojedyncze komórki. (Moreno, Mossio, 2015) W tym sensie RR rozumiany jest jako fundamentalny

mechanizm adaptacyjny, dzięki któremu organizmy utrzymują się przy życiu poprzez dynamiczne dostrajanie się do środowiska. RR nie jest więc pierwotnie procesem obliczeniowym, lecz biologicznym i egzystencjalnym.

Na wyższych poziomach organizacji, zwłaszcza w przypadku ludzi, RR może jednak wykorzystywać reprezentacje i działać w sposób obliczeniowy. Vervaeke i Ferraro (2013a) wskazują, że ludzka zdolność do kategoryzacji, pamięci czy rozumowania zakłada już istnienie struktur poznawczych umożliwiających reprezentację i manipulowanie symbolami. RR na tym poziomie korzysta z modeli predykcyjnych oraz algorytmicznych procedur, które wspierają procesy rozwiązywania problemów i podejmowania decyzji.

Pozostaje kluczowe pytanie: w jaki sposób proces życiowy i adaptacyjny tworzy, organizuje i wykorzystuje reprezentacje mentalne na wyższych poziomach organizacji biologicznej? Odpowiedź na to pytanie jest istotna dla pełnego zrozumienia relacji między życiem, poznaniem i obliczalnością. Wyzwaniem pozostaje więc uchwycenie tego przejścia – od procesów życiowych do reprezentacyjnych i obliczeniowych – w sposób, który nie popadnie w błędne koło, a jednocześnie wyjaśni emergencję zdolności poznawczych u organizmów złożonych.

W kontekście rozszerzenia znaczenia teorii realizacji istotności do procesu organizującego życie w trialektyce autopojezy, adaptacji i antycypacji (Jaeger *et al.*, 2024), problem reprezentacji i relewancji ma zostać rozwiązany nie na płaszczyźnie logicznej, lecz czasowej i adaptacyjnej. To nie algorytmy i reprezentacje mają stanowić pierwotne źródło relewancji, lecz proces adaptacyjnego dostrajania się organizmu do środowiska. Dopiero w toku ewolucji i rozwoju systemów poznawczych, RR wytwarza warunki do powstania reprezentacji mentalnych, które umożliwiają korzystanie z obliczeń w systemach bardziej złożonych, takich jak ludzki mózg.

To jednak nie omija problemu z reprezentacją. Odwołanie do antycypacji jako jednego z koniecznych komponentów trialektyki, która

ma realizować istotność nawet u podstawowych form życia, prowadzi do konieczności uznania istnienia jakiegoś modelu, na podstawie, którego antycypacja mogłaby zachodzić. Model jest reprezentacją pewnego istotnego dla organizmu aspektu rzeczywistości. W ten sposób próba wyjaśnienia realizacji istotność znowu wpada w schemat, którego od początku próbuje uniknąć (por. Parvizi-Wayne, 2025).

Uznanie realizacji istotności za fundamentalną zdolność poznawczą, taką od której zależą wszystkie inne, a która sama już od innych zdolności nie zależy, może prowadzić do zamieszania pojęciowego przy próbie jej wyjaśnienia. Takie zamieszanie identyfikuje Dunn w artykule „Can Rationality Be Naturalistically Explained?" Analizuje on i krytykuje argument Vervaekego i Chiappe z pracy „Fodor, Cherniak and the Naturalization of Rationality" (1997). Vervaeke i Chiappe twierdzą, że naturalizacja racjonalności jest niemożliwa, ponieważ każde wyjaśnienie naukowe zakłada uprzednio zdolność do oceny relewancji, a właściwa ocena relewancji jest główną składową racjonalności, co prowadzi do błędnego koła. Dunn zgadza się, że to ważny problem filozoficzny i uznaje ich argumentację za niespójną.

Jego krytyka opiera się na rozróżnieniu aktu wyjaśniania od treści wyjaśnienia: fakt, że badacz musi racjonalnie stosować teorię, nie oznacza jeszcze, że teoria musi posiadać odniesienie do istnienia racjonalności w swojej treści. Dunn rozdziela też wyjaśnienia identyfikujące (czym coś jest) od przyczynowych (jak coś powstaje), wskazując, że wyjaśnienie mechanizmu powstawania relewancji nie musi odwoływać się do samego aktu jej rozpoznania. Vervaeke i Chiappe utożsamiają zdolność do oceny relewancji – cechę podmiotu poznającego, z relewancją jako cechą rzeczy w kontekście. Wyjaśnienie racjonalności wymaga wyjaśnienia własności rozróżniania istotności, a nie koniecznie wyjaśnienia relewancji jako cechy sytuacji. Można założyć, że istotność już jest pewną cechą sytuacji lub rzeczy i zdolność do jej rozpoznania jest już wtórna.

W drugiej części Dunn odpowiada na tezę, że wyjaśnienia odwołujące się do praw naukowych wymagają tła relewancji. Przyznaje, że

każde wyjaśnienie działa w „istotnych okolicznościach", ale zauważa, że to dotyczy wszystkich teorii, nie tylko racjonalności. Wiele własności – jak smak czy kolor – pozostaje częściowo niezdefiniowanych, a mimo to ich wyjaśnianie nie jest niemożliwe. W konsekwencji racjonalność nie stanowi wyjątku, jak chcą Vervaeke i Chiappe.

Dunn konkluduje, że choć problem relewancji jest realny i trudny, nie ma podstaw, by uznać jego naturalizację za niemożliwą. Argument Vervaekego i Chiappe według Dunna jest zbyt ogólny, miesza poziomy analizy i rozszerza pojęcie presupozycji racjonalności.

Na argument o niewyjaśnialności pojęcia racjonalności można odpowiedzieć krytykując rozróżnienie Dunna na relewancję jako zdolność i jako cechę. Zgodnie z twierdzeniem o transjektywności istotności (Vervaeke *et al.*, 2012), istotność nie jest cechą żadnego przedmiotu ani sytuacji, ale własnością interakcji umysłu ze światem. W takim ujęciu nie możemy rozdzielić istotności jako cechy od zdolności do rozróżniania istotności, ponieważ wszystko, co jest istotne jest takie w czyimś umyśle. Spór o wyjaśnialność takich terminów jak racjonalność i relewancja wydaje się zależeć od decyzji o statusie ontologicznym tych pojęć.

Warto zauważyć, że struktura argumentu, którą Vervaeke i Chiappe zastosowali wobec racjonalności, pojawia się także w jego ujęciu obliczalności. Podobnie jak w przypadku racjonalności, Vervaeke twierdzi, że próba pełnej formalizacji RR w postaci algorytmu prowadzi do błędnego koła. Tak jak wyjaśnienie racjonalności wymaga uprzedniej zdolności do oceny relewancji, tak samo stworzenie algorytmu wyłaniającego relewancję wymaga wcześniejszej wiedzy o tym, co jest relewantne. W obu przypadkach chodzi o niemożność redukcji zjawiska do procedury, która musiałaby już zakładać to, co próbuje wyjaśnić: racjonalność w jednym przypadku, a relewancję w drugim. Vervaeke stosuje zatem tę samą strukturę argumentacyjną: pokazuje, że zarówno racjonalność, jak i obliczalność relewancji nie mogą być ujęte w modelu, który w swojej treści niepostrzeżenie już zakłada to, co ma dopiero wyjaśnić. Dunn pokazuje, że naturalistyczne

wyjaśnienie racjonalności nie musi być niemożliwe – o ile rozróżnimy akt wyjaśniania od treści wyjaśnienia i uznamy, że można badać mechanizmy powstawania racjonalnych zdolności bez zakładania ich ówczesnego istnienia – tak samo można argumentować, że obliczeniowe ujęcie istotności nie jest niemożliwe, jeśli uznamy istotność za obiektywną cechę informacji, przynajmniej częściowo niezależną od podmiotu poznającego.

Dominujące w filozofii podejście polega na precyzyjnym definiowaniu pojęć, a następnie na konsekwentnym wyprowadzaniu wniosków zgodnych z przyjętymi definicjami. Strategia ta zakłada względną stabilność i jednoznaczność analizowanych kategorii. W przypadku pojęcia relewancji założenie to okazuje się jednak problematyczne. Próby jej zdefiniowania jako niehomogenicznej, transjektywnej cechy poznania, czyli takiej, która nie jest ani w pełni subiektywna, ani obiektywna, prowadzą do powtarzających się napięć i sprzeczności.

Sugeruje to, że trudności nie wynikają wyłącznie z niedoskonałości poszczególnych definicji, lecz mają charakter bardziej fundamentalny. Być może właściwym wnioskiem płynącym z teorii realizacji istotności jest konieczność rewizji samego projektu definiowania relewancji w tradycyjnym sensie. Zamiast traktować ją jako własność podlegającą klasycznej analizie pojęciowej, należałoby podjąć próbę jej ontologicznego uchwycenia.

Kluczowy problem polega na tym, że wprowadzenie „trzeciego stanu" pomiędzy subiektywnością a obiektywnością pozostaje niewystarczająco określone. Jeśli relewancja ma być transjektywna, to wymaga bardziej precyzyjnej klasyfikacji niż negatywne odróżnienie od dwóch standardowych kategorii. W przeciwnym razie pozostaje ona pojęciem o niejasnym statusie ontologicznym, które trudno włączyć do spójnego obrazu świata.

Dodatkowa trudność pojawia się, gdy zauważymy, że ani obiektywna struktura rzeczywistości, ani jednostkowe preferencje podmiotu nie determinują w pełni tego, co jest istotne. W takiej sytuacji

wyjaśnienie pochodzenia relewancji staje się problematyczne: brak jednoznacznego źródła prowadzi do wrażenia arbitralności lub ukrytej kolistości. Jeśli bowiem odwołujemy się do mechanizmów poznawczych, które same operują na relewancji, ryzykujemy wyjaśnianie jednego przez drugie.

Najpoważniejsza konsekwencja dotyczy jednak kwestii jedności tego pojęcia. Jeżeli wszystkie rzeczy istotne nie posiadają żadnej wspólnej cechy, to relewancja przestaje być własnością w sensie klasycznym. W takim przypadku nie istnieje kryterium, które pozwalałoby określić, co będzie istotne w nowych sytuacjach. Pojęcie relewancji ulega wówczas rozproszeniu na zbiór lokalnych, kontekstowych rozstrzygnięć, bez możliwości ich uogólnienia.

Z tego punktu widzenia teoria realizacji istotności może być odczytana nie jako próba wyjaśnienia w końcu jak możliwe jest rozróżnienie tego co istotne, od tego co nieistotne, lecz jako wskazanie na konieczność zastanowienia się jeszcze raz nad tym co to znaczy, że coś jest lub nie jest istotne.

5. ZAKOŃCZENIE

W niniejszym artykule podjęto próbę krytycznej rekonstrukcji i analizy teorii realizacji istotności Johna Vervaekego, której celem jest dostarczenie spójnego rozwiązania problemu istotności. Punktem wyjścia było wskazanie, że tradycyjne modele poznawcze, oparte na obliczeniowym przetwarzaniu reprezentacji, nie były w stanie adekwatnie uchwycić dynamiki i kontekstowości ludzkiego poznania. Vervaeke, odwołując się do inspiracji płynących z psychologii poznawczej, neuronauki oraz filozofii umysłu, sformułował koncepcję, w której istotność nie jest nadawana w sposób arbitralny ani algorytmiczny, lecz wynika z procesu dynamicznego dostrajania organizmu do środowiska.

W kolejnych krokach przedstawiono, jak teoria ewoluowała – od wstępnych prób opisu realizacji istotności jako procesu heurystycznego

i informacyjnego, przez jej ujęcie w kategoriach mechanizmów predykcyjnych, aż po rozszerzenie na wymiar enaktywny, autonomii biologicznej i emergentnej agencyjności. Kluczowym elementem tej drogi było odejście od redukcjonistycznych wyjaśnień obliczeniowych na rzecz ujęcia, które łączy mechanizmy neuronalne, procesy kognitywne oraz emergentne właściwości systemów złożonych.

Następnie zwrócono uwagę na związki realizacji istotności z racjonalnością, wskazując, że teoria ta nie tylko redefiniuje sposób rozumienia poznania, lecz również dostarcza nowego spojrzenia na samą naturę racjonalności jako zdolności do adaptacyjnego dostrajania się do sytuacji problemowych. W tym sensie, realizacja istotności okazuje się nie tyle szczegółową teorią procesów poznawczych, co metateorią opisującą warunki możliwości sensownego działania i myślenia.

Co istotne, teoria Vervaekego zyskała oddźwięk także poza wąskim kręgiem dyscypliny i zaczęła być podejmowana szerzej w literaturze. Została omówiona w „The Blind Spot" (Frank *et al*., 2024), gdzie stanowi element refleksji nad granicami obliczeniowych modeli umysłu; przywoływana w kontekście pamięci jako dynamicznego, agencyjnego procesu w pracy Levina (2024); a także uwzględniona w świetle przetwarzania predykcyjnego w artykule Darlinga, Corcorana i Hohwy'ego (2025), gdzie realizacja istotności wskazywana jest jako element umożliwiający rozwiązanie problemu istotności w ujęciu predykcyjnym. Te przykłady pokazują, że teoria Vervaekego inspiruje badania w różnych subdyscyplinach kognitywistyki i filozofii.

Ostatecznie, przeprowadzona krytyka teorii realizacji istotności ukazuje obszary, które wymagają dalszego rozwinięcia i pogłębionej refleksji filozoficznej. W szczególności otwarta pozostaje kwestia spójnej teorii informacji, która nie redukowałaby jej do ujęć reprezentacjonistycznych, a zarazem pozwalała uchwycić informację jako element konstytuujący procesy realizacji istotności. Podobnie, niezbędne jest opracowanie adekwatnego rozumienia celowości, które unikałoby homunkularnych strategii wyjaśniania, a jednocześnie

tłumaczyło, w jaki sposób organizmy i systemy poznawcze generują i weryfikują własne cele. Z tym problemem wiąże się również zagadnienie obliczalności – RR konsekwentnie akcentuje, że algorytmiczne formalizacje nie wystarczają do uchwycenia „dużego świata" otwartej rzeczywistości, jednak sama teoria nadal wymaga wyjaśnienia relacji między obliczeniowością a emergentnymi procesami predykcyjnymi. Jednak nawet to podejście nie unika konotacji reprezentacyjnych z powodu odwołania do zdolności antycypacji.

Ujęcie relewancji jako niehomogenicznej, transjektywnej cechy poznania – ani w pełni subiektywnej, ani obiektywnej – prowadzi do napięć i niejasności. Możliwe, że problem tkwi w założeniu, że relewancję da się uchwycić w ten sposób. Wprowadzenie „trzeciego stanu" między subiektywnością a obiektywnością pozostaje zbyt mało precyzyjne, a brak jednoznacznego źródła relewancji utrudnia jej wyjaśnienie i sprzyja wpadaniu w błąd wyjaśnień kolistych. Jeśli rzeczy istotne nie mają wspólnej cechy, relewancja przestaje być klasyczną własnością i nie daje się wyjaśnić żadną procedurą. Pozostaje tajemnicza, nieuchwytna i prowadząca do sprzeczności.

Traktowanie realizacji istotności jako centralnej zdolności poznawczej prowadzi do głębokiego przesunięcia akcentu filozoficznego: poznanie nie jawi się już przede wszystkim jako proces weredyczny, którego celem jest prawda rozumiana korespondencyjnie, lecz jako proces dynamicznego rozróżniania istotności, od którego zależy zarówno możliwość skutecznego działania, jak i samo powstanie sensownych reprezentacji. W tym kontekście konieczne jest rozwinięcie refleksji nad związkiem między prawdą a istotnością: prawda, choć pozostaje centralnym pojęciem w teorii poznania, nie może być rozumiana w oderwaniu od procesów konstytuujących relewancję. Konieczna jest zatem ponowna fundamentalna refleksja nad znaczeniem pojęcia relewancji i jej związku z podstawowymi pojęciami współczesnej teorii poznania i kognitywistyki: prawdą, celowością, funkcją i reprezentacją. Przeformułowanie podstawowych intuicji na temat tego, czym jest istotność, może nas przybliżyć do takiego

rozwiązania problemu istotności, które nie będzie notorycznie podważało własnych założeń.

OŚWIADCZENIA

Finansowanie badań: Niniejsza publikacja nie otrzymała żadnego zewnętrznego wsparcia finansowego.

Oświadczenie właściwej komisji etycznej: Nie dotyczy.

Oświadczenie o dostępności danych: Nie dotyczy.

Podziękowania: Autor pragnie podziękować Michałowi Piekarskiemu za wsparcie merytoryczne oraz Agnieszce Jaworskiej za redakcję językową.

Deklaracja narzędzi AI: Podczas przygotowywania niniejszego manuskryptu, na podstawie którego powstał niniejszy artykuł, autor korzystał z GPT-4o, Gemini 2.0 Flash w następujących celach: edycja językowa, generowanie tekstu. Autor przejrzał i zredagował wyniki pracy i ponosi pełną odpowiedzialność za treść niniejszej publikacji.

Konflikt interesów: Autor deklaruje brak konfliktu interesów.

BIBLIOGRAFIA

Andersen, B.P., Miller, M., i Vervaeke, J. (2022). Predictive processing and relevance realization: exploring convergent solutions to the frame problem. *Phenomenology and the Cognitive Sciences*, *24*(2), 359–380. https://doi.org/10.1007/s11097-022-09850-6.

Bak, P., Tang, C., i Wiesenfeld, K. (1988). Self-organized criticality. *Physical Review A*, *38*(1), 364–374. https://doi.org/10.1103/PhysRevA.38.364.

Bullmore, E., Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*(3), 186–198. https://doi.org/10.1038/nrn2575.

Cherniak, C. (1990). *Minimal rationality*. MIT Press.

Chiappe, D.L., Vervaeke, J. (1997). Fodor, Cherniak and the Naturalization of Rationality. *Theory & Psychology*, *7*(6), 799–821. https://doi.org/10.1177/0959354397076004.

Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. Harper Perennial.

Darling, T., Corcoran, A.W., i Hohwy, J. (2025). Solving the relevance problem with predictive processing. *Philosophical Psychology*, *39*(4), *1472–1497*. https://doi.org/10.1080/09515089.2025.2460502.

Dennett, D.C. (1987). *Cognitive wheels: The frame problem of AI*. W Z.W. Pylyshyn (red.), *The robot's dilemma: The frame problem in artificial intelligence* (41–64). Ablex Publishing Corporation.

Dunn, J. (b.d.). Can rationality be naturalistically explained? ResearchGate. https://www.researchgate.net/publication/265154524_Can_Rationality_Be_Naturalistically_Explained [data dostępu 21/08/2024].

Fodor, J.A. (2008*). Lot 2: The Language of Thought Revisited*. Oxford University Press.

Frank, A., Gleiser, M., i Thompson, E. (2024). *The blind spot: Why Science Cannot Ignore Human Experience*. MIT Press.

Gigerenzer, G. (2021). Embodied heuristics. *Frontiers in Psychology*, *12*, 711289. https://doi.org/10.3389/fpsyg.2021.711289.

Goodman, N. (1972). *Seven strictures on similarity*. W N. Goodman (red.), *Problems and projects* (437–446). Bobbs-Merrill.

Green, C.D. (1996). Fodor, functions, physics, and fantasyland: is AI a Mickey Mouse discipline? *Journal of Experimental & Theoretical Artificial Intelligence*, *8*(1), 95–106. https://doi.org/10.1080/09528139650042565.

Hofmeyr, J.-H. S. (2007). *The biochemical factory that autonomously fabricates itself: A systems biological view of the living cell*. W F. Boogerd, F.J. Bruggeman, J.-H. S. Hofmeyr, H.V. Westerhoff (red.), *Systems Biology: Philosophical Foundations* (217–242). Elsevier. https://doi.org/10.1016/B978-044452085-2/50012-7.

Horsman, D., Stepney, S., Wagner, R.C., i Kendon, V. (2014). When does a physical system compute? *Proceedings of the Royal Society a Mathematical Physical and Engineering Sciences*, *470*(2169). https://doi.org/20140182.10.1098/rspa.2014.0182.

Hovhannisyan, G., Dewey, C. (2017). Natural & normative dynamical coupling. *Cognitive Systems Research, 43,* 128–139.

Hovhannisyan, G., Vervaeke, J. (2021). Enactivist Big five theory. *Phenomenology and the Cognitive Sciences*, *21*(2), 341–375. https://doi.org/10.1007/s11097-021-09768-5.

Hume, D. (1963). *Traktat o naturze ludzkiej* (tłum. C. Znamierowski). Państwowe Wydawnictwo Naukowe.

Jaeger, J., Riedl, A., Djedovic, A., Vervaeke, J., i Walsh, D. (2024). Naturalizing relevance realization: Why agency and cognition are fundamentally not computational. *Frontiers in Psychology*, *15*, 1362658. https://doi.org/10.3389/fpsyg.2024.1362658.

Krzanowski, R. (2020). Why can information not be defined as being purely epistemic? *Philosophical Problems in Science, 68,* 37–62. https://doi.org/10.59203/zfn.68.494.

Levin, M. (2024). Self-Improvising Memory: A perspective on memories as agential, dynamically reinterpreting cognitive glue. *Entropy*, *26*(6), 481. https://doi.org/10.3390/e26060481.

Marr, D. (2010). Vision. MIT Press. https://doi.org/10.7551/mitpress/9780262514620.001.0001.

McCarthy, J., Hayes, P.J. (1981). *Some philosophical problems from the standpoint of artificial intelligence*. W B.L. Webber, N.J. Nilsson (red.), *Readings in Artificial Intelligence* (431–450). Morgan Kaufmann. https://doi.org/10.1016/b978-0-934613-03-3.50033-7.

Moreno, A., Mossio, M. (2015). *Biological autonomy: A philosophical and theoretical enquiry*. Springer. https://doi.org/10.1007/978-94-017-9837-2.

Newell, A., Simon, H.A. (1972). *Human problem solving*. Prentice-Hall.

Parvizi-Wayne, D. (2025). What active inference still can't do: The (frame) problem that just won't go away. Philosophy and the Mind Sciences, *6*, 12118. https://doi.org/10.33735/phimisci.2025.12118.

Rosen, R. (2012). *Anticipatory systems: Philosophical, mathematical, and methodological foundations* (wyd. 2). Springer. https://doi.org/10.1007/978-1-4614-1269-4.

Shanahan, M. (2016). *The frame problem*. W E.N. Zalta (red.), *The Stanford encyclopedia of philosophy* (Spring 2016 Edition). Stanford University. https://plato.stanford.edu/archives/spr2016/entries/frame-problem/ [data dostępu 20/08/2026].

Simon, H.A. (1957). *Models of man: Social and rational. Mathematical essays on rational human behavior in a social setting*. Wiley.

Sperber, D., Wilson, D. (1995). *Relevance: Communication and cognition* (wyd. 2). Blackwell.

Vervaeke, J.A. (1997). *The naturalistic imperative in cognitive science* [Rozprawa doktorska, University of Toronto]. TSpace Repository. https://tspace.library.utoronto.ca/handle/1807/10696 [data dostępu 20/08/2026].

Vervaeke, J., Ferraro, L. (2013a). *Relevance realization and the neurodynamics and neuroconnectivity of general intelligence*. W I. Harvey, A. Cavoukian, G. Tomko, D. Borrett, H. Kwan, i D. Hatzinakos (red.), *SmartData: Privacy Meets Evolutionary Robotics* (57–68). Springer. https://doi.org/10.1007/978-1-4614-6409-9_6.

Vervaeke, J., Ferraro, L. (2013b). *Relevance, meaning and the cognitive science of wisdom*. W M. Ferrari, N.M. Weststrate (red.), *The Scientific Study of Personal Wisdom* (21–43). Springer. https://doi.org/10.1007/978-94-007-7987-7_2.

Vervaeke, J., Ferraro, L., i Herrera-Bennett, A. (2018). *Flow as spontaneous thought: Insight and implicit learning*. W K. Christoff, K.C.R. Fox (red.), *The Oxford Handbook of Spontaneous Thought: Mind-Wandering, Creativity, and Dreaming* (149–160). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190464745.013.8.

Vervaeke, J., Lillicrap, T.P., i Richards, B.A. (2012). Relevance realization and the emerging framework in cognitive science. *Journal of Logic and Computation*, *22*(1), 79–99. https://doi.org/10.1093/logcom/exp067.

JOHN VERVAEKE'S RELEVANCE REALIZATION THEORY

Abstract. This article carries out a critical reconstruction of John Vervaeke's theory of relevance realization (RR), developed as a response to the explanatory gap in classical computational models of cognition. It traces the evolution of the theory from the 1990s onward, showing how the frame problem and the aporia of the representation–relevance relation have shaped its development within the naturalistic imperative in cognitive science. This paper offers a detailed account of the regulatory mechanisms of RR—cognitive scope, tempering, and prioritization—which, through opponent processing, are intended to enable systems to dynamically balance economic trade-offs. It also examines the theory's interdisciplinary applications, ranging from predictive processing and general intelligence to personality theory and the concept of flow. In the critical section, a persistent tension is identified between Vervaeke's rejection of computational and representational approaches and his appeal to anticipatory capacities (carrying representational connotations) as constitutive of relevance realization. The remaining difficulties in formulating a version of RR consistent with its own assumptions indicate the need to revise these assumptions and to reconsider the nature of relevance itself.

Keywords: Vervaeke John; relevance realization; frame problem; representation–relevance problem; rationality; enactivism; predictive processing; teleology; computability; naturalism

MATEUSZ SABATOWSKI
Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie
(Cardinal Stefan Wyszynski University in Warsaw)
https://orcid.org/0009-0002-6739-625X
ma.sabatowski@gmail.com

https://doi.org/10.21697/spch.2026.62.A.06

Tekst jest udostępniany na zasadach licencji Creative Commons (CC BY-ND 4.0 Międzynarodowe).

Zgłoszono: 29/08/2025. Zrecenzowano: 11/05/2026. Zaakceptowano do publikacji: 08/06/2026.