

BARBARA TOMCZYK

ROZPROSZONY SYSTEM CZŁOWIEK-ARTEFAKT JAKO AUTONOMICZNY SPRAWCA. RACJE I WĄTPLIWOŚCI

Streszczenie. Zwolennicy poznania rozproszonego deklarują odchodzenie od indywidualizmu metodologicznego i poznawczego w stronę ujęcia systemowego, na gruncie którego poznanie traktowane jest jako własność całego szerokiego systemu, a nie pojedynczego człowieka. W artykule skupiam się na problemie sprawczości systemu człowiek-artefakt i wykazuję, że przyjęcie perspektywy indywidualistycznej jest przy jego rozpatrywaniu niezbędne, a ujmowanie go jako systemu rozproszonego nie stanowi przewagi wyjaśniającej. Analizy reprezentacji zewnętrznych, afordancji i dizajnu, przeprowadzone przez Witolda Wachowskiego w książce *Poznanie rozproszone. Od heurystyk do mechanizmów* (2022), nie odchodzą od podejścia skoncentrowanego na indywidualnym podmiocie, a zastosowana heurystyka „od kultury do mózgu” jest zgodna z podejściem podmiotowo-centrycznym. W artykule argumentuję, że wątpliwości związane z odrzuceniem indywidualizmu na rzecz ujęcia systemowego, przy analizie działalności poznawczej systemu człowiek-artefakt, przeważają nad racjami uzasadniającymi takie rozwiązanie.

Słowa kluczowe: autonomiczne działanie; umysł rozszerzony; szeroki system poznawczy; artefakt poznawczy; epistemiczna i moralna odpowiedzialność; doświadczenie sprawczości.

1. Wstęp. 2. Podmiot autonomicznego działania. 3. System człowiek-artefakt jako podmiot autonomicznego działania – racje. 4. System człowiek-artefakt jako podmiot autonomicznego działania – wątpliwości. 4.1. Relacyjna koncepcja artefaktów. 4.2. Doświadczenie podmiotowości jako warunek sprawczości. 4.3. Odpowiedzialność szerokiego systemu człowiek-artefakt. 5. Uwagi końcowe.

1. WSTĘP

Koncepcja poznania rozproszonego jest niezwykle inspirującym stanowiskiem, mającym znaczący wpływ na kształt współczesnej dyskusji filozoficznej i naukowej. Dla indywidualistycznej filozofii, czyniącej podmiotem swych analiz pojedynczego człowieka, jest to pogląd dość kontrowersyjny, gdyż wskazuje na istnienie innych, poza człowiekiem, systemów, przeprowadzających procesy

poznawcze, w wyniku których systemy te zdobywają wiedzę propozycjonalną oraz działają na jej podstawie. Zwolennicy tej idei najczęściej podają jako przykłady takich szerokich systemów grupy ludzi, rozwiązujących pewne zadanie poznawcze przy użyciu artefaktów. Ich zdaniem jednak systemem rozproszonym można nazwać również system złożony z jednego człowieka i wykorzystywanego przez niego artefaktu, a nawet jednego człowieka lub wyłącznie jego mózg, w zależności od tego, jak określona zostanie granica systemu. Poznanie rozproszone, jak wskazuje jego przedstawiciel Edwin Hutchins (2014), stanowi perspektywę, z której należy postrzegać wszelkie procesy poznawcze, niezależnie od tego, gdzie zachodzą. Wyłączają się one bowiem ze złożonych interakcji pomiędzy elementami określonego systemu. Zarówno granice systemu, jak i jego centrum nie są ustalone z góry, ale zależą od tego, jakie zadanie poznawcze jest rozpatrywane. Niektóre systemy mają ściśle określone centra, a niektóre nie posiadają ich wcale lub mają ich wiele.

Analizy, którym poświęcam ten artykuł, dotyczą tych systemów poznawczych, które mają wyraźne centra, mianowicie systemów złożonych z jednego człowieka i artefaktu. Ponieważ centrum to stanowi człowiek, rozważania te można traktować jako prowadzone z perspektywy indywidualizmu poznawczego, rozumianego tak, jak przedstawia to w swej książce *Poznanie rozproszone. Od heurystyk do mechanizmów* Witold Wachowski (2022). Jako indywidualistyczne traktowane są przez niego stanowiska, które są skoncentrowane na podmiocie, a „związane z nim pytania badawcze formułowane są najczęściej w związku z perspektywą lub aktywnością jednostki” (Wachowski, 2022, 19). Moje rozważania, zawarte w tej pracy, skupiają się na systemach, gdzie, jak pisze autor, „ludzki podmiot stanowi centrum, wykorzystujące w toku aktywności poznawczej elementy otoczenia jako elementy swojego osadzenia lub rozszerzenia” (Wachowski, 2022, 19). Poznanie rozumiem, zgodnie z określeniem Wachowskiego, indywidualistycznie „jako własność (...) podmiotu poznawczego, w związku z czym pytania badawcze dotyczą przede

wszystkim podmiotów poznawczych” (Wachowski, 2022, 19). Po lekturze książki omawianego autora wydaje mi się jednak, że nie jestem w tej perspektywie osamotniona nawet wśród zwolenników poznania rozproszonego. Przekonał mnie o tym szczególnie rozdział 2.: „Interakcje w rozproszonych systemach poznawczych”, gdzie autor analizuje systemy podmiot-otoczenie, wskazując wpływ rusztowań poznawczych, zewnętrznych reprezentacji oraz dizajnu na procesy poznawcze, przeprowadzane przez takie systemy. Śledząc kolejne przytaczane przez autora przykłady działań owych systemów, trudno mi oprzeć się wrażeniu, że dotyczą one podmiotu (człowieka), który wykorzystuje (świadomie, bądź nie) zewnętrzne struktury środowiskowe w celu wykonania zadania poznawczego. Teza, że struktury te modyfikują zarówno samo zadanie poznawcze, jak i przeprowadzane w celu jego wykonania procesy, niewiele w tej kwestii zmienia. Procesy poznawcze można uznać za rozproszone na cały system, lecz ich podmiotem jest człowiek, gdyż zarówno afordancje, jak i reprezentacje zewnętrzne są *dla* człowieka, jako sprawcy działania poznawczego. Wątpliwość tę rozwinę jednak w kolejnych częściach artykułu, tutaj chciałam jedynie ją zasygnalizować i jednocześnie podkreślić, że nie podważa ona hipotezy, której Wachowski broni w swej książce, zakładającej „słuszność heurystyki, zgodnie z którą środowisko podmiotu należy uwzględnić na samym początku badania zjawiska poznawczego, a nie dopiero na dalszym jego etapie” (Wachowski, 2022, 26). Można zachować heurystykę „od kultury do mózgu” i jednocześnie skupiać się na badaniu zjawisk poznawczych, których podmiotem jest pojedynczy człowiek, co sam autor *Poznania rozproszonego* w wielu miejscach czyni. W przypadku systemu człowiek-artefakt, heurystyka „od kultury do mózgu” może zostać efektywnie połączona z heurystyką podmiotowo-centriczną.

W prezentowanej pracy przyjmuję rozumienie sprawczości, które zostało rozwinięte w analitycznej filozofii działania. Wypracowane tam warunki autonomicznego działania zostaną przedstawione w kolejnym punkcie, już teraz jednak warto wspomnieć, że spełnić je może

tylko system, który posiada postawy propozycjonalne, czyli stany umysłowe zawierające treść, takie jak przekonania, pragnienia i intencje. Należy jednak zauważyć, że takie reprezentacyjne rozumienie sprawczości nie jest powszechne wśród filozofów, badających zjawisko autonomicznego działania. Inną postawę przyjmują enaktywiści, rozwijający koncepcję autonomii sensomotorycznej, w ramach której posiadanie symbolicznych reprezentacji i stanów umysłowych nie jest konieczne dla sprawczości¹. W tej pracy nie będę się odwoływała do tych koncepcji, bowiem moje rozważania zakładają pojęcie działania, zdefiniowane w terminach stanów umysłowych, posiadających treść².

Zaproponowane w tym artykule analizy skupiają się na zjawisku autonomicznego działania rozproszonego systemu człowiek-artefakt. Mimo że rzecznicy poznania rozproszonego odwołują się w swych rozważaniach głównie do systemów wieloosobowych, pozbawionych centrum, nie wyłączają z grona systemów rozproszonych tych, które są złożone tylko z jednego człowieka. Można zauważyć, że nie są to przypadki marginalne, lecz występujące często w ich analizach. Jak określić podmiot działania, czyli sprawcę, w takim systemie? Przy założeniu, że procesy poznawcze, prowadzące do ukształtowania się postaw propozycjonalnych, kierujących działaniem, są rozproszone na cały taki system, czy same te postawy (intencje, pragnienia, przekonania) również należy uznać za tak rozproszone? Jeśli tak, to czy ich podmiotem, czyli autonomicznym sprawcą, również jest cały system człowiek-artefakt? W przypadku systemu wieloosobowego intuicje zdają się być wyraźniejsze, łatwiej bowiem uzasadnić istnienie systemowych postaw propozycjonalnych i systemowej odpowiedzialności, gdy mamy do czynienia z kilkoma (lub wieloma) ludźmi, którzy

1 Koncepcję taką, budowaną wokół pojęcia autopojetyczności i adaptacyjności, rozwijają między innymi: Varela, 1979; Barandiaran, Di Paolo, i Rohde, 2009; Beer, Di Paolo, 2023. Wśród enaktywistów, krytykujących odwoływanie się do postaw propozycjonalnych i szerzej do pojęcia reprezentacji w wyjaśnianiu procesów poznawczych znajdują się między innymi: Noe, 2004; Chemero, 2011; Hutto, Myin, 2013.

2 Analizie sprawczości systemów rozszerzonych na gruncie koncepcji sensomotorycznych poświęciłam artykuł (Tomczyk, 2024).

wchodzą ze sobą w językową interakcję, niż wtedy, gdy system tworzy jeden człowiek i artefakt³. Ujęcie systemowe poznania, na rzecz którego Witold Wachowski argumentuje w swej książce, odnosi się właśnie do takich „szerszych niż podmiot struktur poznawczych, w których ten podmiot nie odgrywa roli centrum, lecz równorzędnego komponentu struktury” (Wachowski, 2022, 19). Jednak, jak przekonuje Hutchins i z czym autor omawianej książki zdaje się zgadzać (Wachowski, 2022, 131), ujęcie systemowe można, a nawet należy, stosować do wszystkich systemów poznawczych, również jednoosobowych. Wszelako, czy można system człowiek-artefakt traktować jako rozproszony? Co miałyby to oznaczać? Czy tylko interakcję między elementami systemu, co czyni to ujęcie dość trywialnym, czy również szeroką fizyczną realizacją procesów poznawczych i stanów umysłowych, wykraczającą poza organizm człowieka? Jeśli przyjmiemy, że stany umysłowe, które decydują o sprawczości takiego systemu, są częściowo realizowane przez czynniki, zachodzące poza organizmem człowieka, co nas wzbrania przed przyjęciem tezy, że podmiotowość sprawcza takiego systemu należy do niego całego, a nie tylko do jego ludzkiego elementu?

W celu odpowiedzi na powyższe pytania i zajęcia stanowiska wobec sprawczości systemu człowiek-artefakt dziełę rozważania zawarte w tym artykule na dwie główne części. W pierwszej z nich (punkt 3) przedstawiam racje za traktowaniem całego takiego systemu jako autonomicznego sprawcy, w drugiej natomiast (punkt 4) prezentuję wątpliwości wobec takiego rozwiązania. W części podsumowującej uzasadniam swoje przekonanie, że wątpliwości przeważają nad racjami, które nie stanowią wystarczającego argumentu za traktowaniem całego systemu człowiek-artefakt jako sprawcy działania.

³ Argumentację za istnieniem grupowych stanów umysłowych przedstawiają w swoich pracach między innymi Deborah Tollefsen (2015) oraz Georg Theiner i Rob Wilson (2013). Analizy dotyczące podmiotowości poznawczej i epistemicznej systemu grupowego znajdują się w moich artykułach (Trybulec, 2015; 2017; 2018).

Jednak w pierwszej kolejności (punkt 2) przedstawię warunki autonomicznego działania, na które najczęściej zwracają uwagę filozofowie, analizując tę problematykę.

2. PODMIOT AUTONOMICZNEGO DZIAŁANIA

Wśród filozofów badających zjawisko sprawczości panuje względna zgoda co do tego, że autentyczne działanie może być zrealizowane wyłącznie przez autonomiczny podmiot. Zgoda ta nie obejmuje już jednak propozycji wyjaśnienia, co owa autonomia oznacza i jakie są jej warunki. Badacze tej tematyki najczęściej odwołują się do samo-kontroli oraz możliwości inicjowania działania, jako do podstawowych zdolności, którymi powinien wykazać się system poznawczy, aby uzyskać status sprawcy. System taki, innymi słowy, musi być źródłem decyzji o swoim działaniu oraz intencji z nim związanych, nie mogą one pochodzić spoza niego. Niewątpliwie na wszelkie decyzje wpływają czynniki zewnętrzne, nad którymi sprawca nie posiada kontroli, a często nawet ich sobie nie uświadamia. Najważniejszym problemem do rozstrzygnięcia staje się zatem określenie granicy takiego wpływu, po przekroczeniu której działanie systemu nie będzie mogło być dłużej przypisane jego własnej sprawczości. Ów problem kryterium rozróżniającego wpływy środowiskowe, podważające autonomię sprawcy, od tych, które nie stanowią dla niej zagrożenia, jest w czasach szybko rozwijającej się technologii wzmocnień poznawczych jednym z najczęściej podejmowanych⁴. Kryterium owo jest ważne nie tylko dla badaczy autonomicznego działania, lecz również dla epistemologów, psychologów, filozofów moralności, robotyków czy twórców sztucznej inteligencji. Jak zapowiedziałam we *Wstępie*, moje rozważania w tym artykule dotyczą systemu zawierającego jednego człowieka i z tego względu przywołuję poniżej stanowiska,

4 Podaję tę problematykę w artykule *Knower at Risk: Updating Epistemology in the Light of Enhanced Representations* (Tomczyk, 2021).

skupiające się na zjawisku autentycznej sprawczości, związanej z ponoszeniem odpowiedzialności za działanie i tym samym z przypisywaniem podmiotowi działania praw i obowiązków. Interesuje mnie mianowicie odpowiedź na pytanie: co stoi na przeszkodzie, aby objąć tak rozumianą sprawczością cały system człowiek-artefakt, szczególnie przy przyjęciu założenia o jego rozproszeniu?

Debata na temat warunków autonomicznego działania jest bardzo intensywna i uczestniczą w niej przedstawiciele wielu stanowisk. Nie będę ich w tym miejscu przedstawiać, gdyż nie jest to konieczne dla mojej argumentacji⁵. Niezależnie jednak od tego, czy analizy podejmują przedstawiciele internalizmu (Dworkin, 1988; Frankfurt, 1971; Bratman, 1979) czy eksternalizmu (Fischer, Ravizza, 1993; Nelkin, 2007; Christman, 1991), panuje między nimi względna zgoda co do tego, że u podstaw autonomicznego działania leżą: zdolność do rozumowania, posiadanie postaw propozycyjalnych, takich jak przekonania, pragnienia i intencje, które determinują konkretne działanie, oraz znajomość i możliwość oceny ich treści. Takie rozumienie sprawczości obecne jest przede wszystkim w analitycznej filozofii działania, reprezentowanej między innymi przez Donalda Davidsona. Jego zdaniem wszelkie działanie jest uwarunkowane posiadaniem przez sprawcę postaw propozycyjalnych, do treści których ma on zasadniczo dostęp. Takim systemem poznawczym jest wyłącznie człowiek, gdyż tylko on dysponuje językiem pojęciowym, koniecznym dla formowania stanów umysłowych, stanowiących zarówno przyczyny podejmowanych działań, jak i racje w procesie ich uzasadniania (Davidson, 2002). Ze wspomnianych powyżej koncepcji wyłania się obraz sprawcy jako elementu aktywnego, inicjującego działanie na zasadzie wolnego wyboru spośród rozważanych alternatyw. Aby tego dokonywać, sprawca musi być systemem refleksyjnym, a to zakłada świadomość, rozumianą jako posiadanie

5 Przedstawienie najważniejszych stanowisk wobec autonomicznego działania znajduje się w moim artykule (Tomczyk, 2023).

intencjonalności, która sprawia, że procesy poznawcze odsłaniają mu świat. Dzięki świadomości podmiot posiada pierwszoosobową perspektywę, w której prezentują mu się wszelkie poznawane byty, w tym treść jego własnych stanów umysłowych.

Badacze analizujący zjawisko autonomicznego działania poświęcają dużo uwagi problemowi inicjującej działanie mocy sprawczej, której źródłem powinien być sam sprawca jako posiadający kontrolę nad swoimi czynami. Przy założeniu, że stany umysłowe, będące przyczynami działania, są kształtowane przez czynniki środowiskowe, w których podmiot jest zakorzeniony, musi on być ich świadomy i je kontrolować, aby być uznanym za autentycznego sprawcę. Moc sprawcza podmiotu nie może być zredukowana do czynników, których nie jest on świadomy. Zdaniem inkompatybilistów, którzy głoszą takie stanowisko, źródłem autonomicznego działania musi być sam podmiot, w przeciwnym wypadku nie mogłoby być bowiem mowy o wolnym wyborze (Chisholm, 1976; Clarke, Reed, 2015; O'Connor, 2009). Pojawia się jednak w tym miejscu pytanie, w jakim stopniu podmiot musi kontrolować czynniki wpływające na jego decyzje o działaniu, aby nie zagrażały one jego autonomii jako sprawcy.

Na silny, wręcz konstytucyjny wpływ czynników środowiskowych na procesy decyzyjne podmiotu wskazują przedstawiciele relacjonizmu. Ich zdaniem stany umysłowe, kierujące danym działaniem, mogą być zdefiniowane w terminach relacji interpersonalnych i wzajemnych zależności między członkami danej społeczności. Jak wskazuje Mason Cash, „są one ugruntowane w podzielanych, intersubiektywnych normach społecznej i językowej praktyki przypisywania sobie nawzajem stanów intencjonalnych jako racji uzasadniających działanie” (Cash, 2010, 648)⁶. Czynniki decydujące o autonomii sprawczej podmiotu zostają na gruncie relacjonizmu zdecentralizowane, uznaje się je bowiem za konstytuowane przez

6 “[they] are grounded in shared, intersubjective norms of the social and linguistic practice of ascribing intentional states to one another as reasons for actions” (Cash, 2010, 648).

take środowiskowe uwarunkowania, jak etniczność, płeć czy klasa społeczna. Aby zachować autonomię sprawczą, podmiot nie tylko musi posiadać refleksyjny dostęp do swoich własnych stanów umysłowych, lecz również musi mieć świadomość, że wszystkie jego decyzje są kształtowane przez wpływ środowiska społecznego, w którym jest on zakorzeniony. Procesy prowadzące do autonomicznego działania można uznać zatem, na gruncie tego stanowiska, za rozproszone na wiele różnych interakcji międzyludzkich (Christman, 2004).

Stanowisko relacjonizmu wspiera koncepcję poznania rozproszonego, której przedstawiciele skupiają się, między innymi, na sposobie, w jaki różnego typu czynniki środowiskowe determinują stany umysłowe podmiotu decydujące o jego sprawczości. Wpływy owe są tak silne, że skłaniają ku rozpraszaniu sprawczości na cały system, obejmujący zarówno człowieka, jak i owe determinanty, które często traktowane są jako składniki konstytuujące intencje prowadzące bezpośrednio do działania. W kolejnej części przedstawię racje, na które powołują się przedstawiciele rozproszonego poznania, za tym, aby traktować cały system człowiek-artefakt jako sprawcę autonomicznego działania.

3. SYSTEM CZŁOWIEK-ARTEFAKT JAKO PODMIOT AUTONOMICZNEGO DZIAŁANIA – RACJE

Wielu współczesnych badaczy opowiada się za ujmowaniem niektórych działań poznawczych jako wykonywanych przez rozszerzony lub rozproszony zintegrowany system, wykraczający poza indywidualnego człowieka (Clark, 1998; 2008; Rowlands, 1999; Kirsh, 2006; Sutton, 2006; Menary, 2007; Wheeler, 2010). Andy Clark, uznawany za inicjatora dyskusji nad poznaniem rozszerzonym, skupia się na przypadkach wykorzystywania przez indywidualnego człowieka narzędzia poznawczego i wypracowuje warunki, które muszą być spełnione, aby nazwać taką aktywność działaniem systemu rozszerzonego. W jego rozumieniu człowiek w takim systemie stanowi

centrum, a jego procesy poznawcze rozszerzają się w tym sensie, że są realizowane częściowo przez fizyczne zdarzenia zachodzące poza jego ciałem⁷. Edwin Hutchins idzie o krok dalej i argumentuje na rzecz ujmowania wszelkich procesów poznawczych jako rozproszonych, nie tylko tych, które zachodzą w grupie współpracujących ze sobą ludzi, lecz również tych, o których mówi Clark, a nawet tych, które uznawane są za zachodzące wyłącznie w mózgu.

Zdaniem Hutchinsa to, jak określimy granice systemu poznawczego, zależy od przestrzennej i czasowej skali (perspektywy), w której ujmujemy dane zadanie poznawcze. Clark, rozwijając tezę poznania i umysłu rozszerzonego, skupia się tylko na jednym rodzaju poznania, to jest na tym, które wykracza poza ludzki organizm. Posługuje się on przy tym metaforą umysłu jako pojemnika, który może rozszerzać się, przekraczając swoje zwykłe granice⁸. Centrum systemu poznawczego stanowi jednak wciąż organizm (mózg). Poznanie rozproszone nie jest spojrzeniem na jakiś konkretny system poznawczy, lecz jest perspektywą, z której należy rozpatrywać wszelkie przejawy poznania, wszystkie procesy poznawcze są bowiem rozproszone pomiędzy fizyczne zdarzenia, które je realizują, niezależnie od tego, gdzie zachodzą w danym systemie (Hutchins, 2001, 2013). Nawet jeśli definiujemy system poznawczy jako ograniczony do samego mózgu, to procesy poznawcze są rozproszone na realizujące je interakcje, zachodzące pomiędzy neuronami. Granice systemu nie są przy tym sztywno ustalone, lecz mogą się zmieniać w zależności od zadania poznawczego, które aktualnie jest rozpatrywane. Analizę

7 "In rejecting the vision of human cognitive processing as organism bound, we should not feel forced to deny that it is (in most, perhaps all, real-world cases) organism centered... it is the biological human organism that spins, selects, or maintains the webs of cognitive scaffolding that participate in the extended machinery of its own thought and reason" (Clark, 2008, 123).

8 Stosowanie takiej metaforyki zarzucają Clarkowi przedstawiciele tzw. drugiej fali argumentów za tezą umysłu rozszerzonego, do której należą między innymi Richard Menary (2007) oraz Mark Rowlands (2010).

systemów człowiek-artefakt należy również, zdaniem Hutchinsa, prowadzić z perspektywy poznania rozproszonego. Oznacza to, że nie powinno się przyjmować na wstępie założenia o centralnej roli człowieka w takim systemie. To, czy stanowi on takie centrum, jest pytaniem empirycznym i zależy od intensywności przepływu informacji w systemie. Niektóre systemy mają bardzo wyraźne centra, a inne nie mają ich wcale lub mają ich wiele. Hutchins przyznaje jednak, że w przypadku systemu człowiek-artefakt uzasadnione jest, aby traktować mózg jako centrum systemu, gdyż to w nim występuje największe zagęszczenie przepływu informacji (Hutchins, 2013).

Zwolennicy istnienia szerokich systemów poznawczych odwołują się w swych analizach do takich zadań, jak: nawigowanie statkiem przez załogę złożoną z wielu ludzi, wykorzystujących specjalistyczne instrumenty i procedury (Hutchins, 1995), przetwarzanie zamówień w kawiarni za pomocą narzędzi poznawczych, takich jak złożony układ notatek i zapisków (Kirsh, 2006), pisanie artykułu naukowego przy użyciu programów komputerowych (Clark, 2003) czy tworzenie dzieła sztuki przy wykorzystaniu szkicowników (Leeuwen, Verstijnen i Hekkert, 1999). Clark wskazuje, że systemy rozwiązujące tego typu zadania poznawcze konstituowane są przez ciągłą, obustronną relację przyczynową (ang. *the continuous reciprocal causation*) pomiędzy jego elementami (Clark, 1997). Hutchins i inni przedstawiciele poznania rozproszonego zauważają jednak, że zarówno Clark, jak i inni badacze systemów człowiek-artefakt wciąż opierają swoje analizy na kartezyjskiej dychotomii wewnętrzne – zewnętrzne, mimo że ich celem jest jej przezwyciężenie. Za punkt wyjścia przyjmują bowiem wewnętrzny biologiczny rdzeń poznawczy (mózg), który wykorzystuje zewnętrzne narzędzia w celu efektywnego wykonania zadania (Aydin, 2013). Aby tego uniknąć, Ciao Aydin proponuje tezę umysłu artefaktycznego (*The Artifactual Mind Thesis*), zgodnie z którą umysł nie rozszerza się z mózgu na artefakty, lecz rozwija się dzięki nim i jest przez nie kształtowany (Aydin, 2015). To, co uznawane jest za wewnętrzny świat myśli, świadome doświadczenie, wewnętrzne

stany umysłowe jest faktycznie tworzone przez zewnętrzne czynniki środowiskowe, w których zakorzeniony jest podmiot poznający. Skoro tak, to utrzymywanie dychotomii wewnętrzne – zewnętrzne traci sens. Uprzywilejowany pod względem poznawczym wewnętrzny świat umysłu przestaje istnieć i tym samym staje pod znakiem zapytania indywidualistyczne założenie o autonomicznym podmiocie działania, na co zwraca uwagę również Hutchins (Hutchins, 2011).

Przypisanie autentycznej sprawczości indywidualnemu podmiotowi wymaga odwołania się do jego „wewnętrznej przestrzeni”, do jego świadomości siebie jako podmiotu oraz do jego świadomości treści jego własnych stanów umysłowych. To gdzieś w tym „wnętrzu” musi powstawać decyzja o działaniu, która w jakimś sensie powinna być wolna i niezdeteminowana, aby można było mówić o sprawstwie. Teza o konstytutywnym wpływie artefaktów i innych uwarunkowań społecznych na stany umysłowe podmiotu stoi, według zwolenników umysłu rozproszonego, w napięciu wobec warunków nakładanych przez filozofów na autonomiczne działanie. Nasuwa się zatem wniosek, że podmiotowość i sprawczość należą do całego systemu, gdyż to procesy rozproszone po całym systemie realizują odpowiedzialne za nią stany umysłowe. Poczucie jaźni i tożsamości osobowej w czasie rozwija się na skutek interakcji z innymi ludźmi i różnymi uwarunkowaniami społecznymi. Nie istnieje czysto wewnętrzna sfera umysłowa oraz zewnętrzny świat materialnych artefaktów. Dychotomia ta w rozproszonym systemie zostaje rozmyta. W konsekwencji przestają istnieć racje dla przypisywania sprawczości wyłącznie jednemu elementowi (człowiekowi) w rozproszonym systemie człowiek-artefakt.

Jednym ze zwolenników tezy rozszerzonej sprawczości jest Richard Heersmink. Jego zdaniem rozszerzony (w jego terminologii) system poznawczy może jako taki być autonomicznym sprawcą, o ile artefakt jest w pełni transparentny i silnie zintegrowany z procesami poznawczymi jego użytkownika (Heersmink, 2017a, 432). Sprawczości takiej, jak argumentuje omawiany autor, nie posiadają

natomiast same artefakty, nie są one bowiem same z siebie zdolne do poznawania, które wymaga, zgodnie ze standardowym, wypracowanym na gruncie analitycznej filozofii działania stanowiskiem, posiadania postaw propozycyjalnych (przekonań, pragnień, intencji) i dostępu ich treści. Kompetencje te wykraczają daleko poza odbieranie i przetwarzanie informacji zgodnie z zadanymi regułami. Artefakty działają na skutek ludzkich stanów umysłowych i ludzkiej sprawczości, a same są jej wyzbyte.

Heersmink podkreśla, że aby zyskać status sprawcy, system musi wykazywać zdolność do tworzenia i realizowania intencji. Co może być zatem takim systemem? Sam mózg, jego zdaniem, nie jest wystarczającym podłożem dla wyłonienia się takich zdolności, potrzebny jest do tego cały ucieleśniony organizm – człowiek jako jedność umysłu i ciała. Jak jednak uzasadnić tezę, że system człowiek-artefakt stanowi sprawcę, czyli tworzy i realizuje intencje? Heersmink posługuje się w tym celu przykładem niewidomego człowieka z laską i Ottona (człowieka cierpiącego na chorobę Alzheimera) z notatnikiem. Artefakty, które pełnią dla takich ludzi funkcję utraconych zmysłów lub pamięci, nie są przez nich traktowane jako zewnętrzne obiekty, lecz jako część ich samych. Ich strata wiąże się z utratą zdolności działania w świecie lub nawet z utratą części osobowości. Artefakty te są dla nich transparentne, niezauważalne, tak jak części ciała używane w codziennym działaniu (Clark, 2008; Malafouris, 2008; Nakayama, 2013; Belk, 1988; Bernini, 2014; Gallagher, Zahavi, 2015). Zdaniem Heersminka w przypadku tak silnie sprzężonego systemu rozszerzonego powinno się traktować jego całego jako realizującego intencję o działaniu, a nie ograniczać tę zdolność tylko do jednego jego elementu – organizmu człowieka. Cały system człowiek-artefakt jest zatem sprawcą ze względu na transparentność artefaktu i wysoki stopień jego integracji z naturalnymi procesami poznawczymi człowieka (Heersmink, 2017a).

Warto również zauważyć, że artefakty mogą w konstytutywny sposób wpływać na odczuwanie przez podmiot swojej własnej

sprawczości, czyli kształtować poczucie własnej tożsamości i odrębności jako podmiotu. Tożsamość osobową kształtuje w dużym stopniu pamięć, która jest istotna zarówno dla poczucia psychologicznej ciągłości podmiotu w czasie, jak i dla tworzenia przez niego obrazu samego siebie (Locke, 1997; Garrett, 1998; Parfit, 1971). Jeśli w niektórych przypadkach pamięć jest konstytuowana przez elementy zewnętrzne wobec organizmu człowieka, należałoby je uznać za konstytutywne również dla jaźni i osobowości podmiotu. Dotyczy to szczególnie osób cierpiących na różnego rodzaju zaburzenia pamięci, gdzie działanie artefaktów poznawczych może, do pewnego stopnia, zastąpić dysfunkcyjne procesy związane z pamięcią i opóźnić jej dalszą utratę (Clowes, 2015; Sandberg, Bostrom, 2006; Rhodes, 1997). Dobrym przykładem takiej sytuacji jest wspomniany już przypadek osoby cierpiącej na chorobę Alzheimera i jej notatnika (Clark, Chalmers, 1998). Notatnik jest niezbędny Ottonowi w każdym podejmowanym działaniu, zastępuje mu bowiem biologiczną pamięć. Otto nie rozstaje się z nim i aktualizuje na bieżąco informacje w nim zawarte, pełni on dla niego tę samą funkcję, co biologiczna pamięć u zdrowej osoby. Jego strata oznaczałaby brak dostępu do przekonań, pragnień i innych stanów umysłowych, które stanowią o tym, kim Otto jest, jak postrzega siebie samego i jakie decyzje podejmuje. Zatem notatnik konstytuuje jego tożsamość osobową, bez niego Otto straciłby poczucie ciągłości swojej osoby w czasie i nie identyfikowałby działań, które podjął w przeszłości, jako swoich własnych. Artefakt ten, zawierający zewnętrzne reprezentacje umysłowe Ottona, konstytuuje jego osobowość i jaźń, stanowi więc część Ottona jako sprawcy, albo raczej, jak powiedzieliby zwolennicy umysłu rozproszonego, część szerokiego systemu człowiek-artefakt, który jako całość stanowi podmiot działania.

Artefakty odgrywają istotną rolę w kształtowaniu tożsamości osobowej nie tylko ludzi dotkniętych dysfunkcjami poznawczymi, lecz również tych zdrowych. Nasza pamięć jest niemal nieprzerwanie wspierana różnego rodzaju reprezentacjami i strukturami

zewnątrznymi, które zmieniają naturalne zdolności poznawcze i wpływają na wybór podejmowanych działań i sposób ich wykonania. Są nimi na przykład programy komputerowe, mapy, diagramy, modele, kalendarze i wiele innych udogodnień, które pomagają nam zapamiętywać, planować, liczyć, uczyć się – wykonywać większość codziennych zadań poznawczych. Dla badaczy, którzy łączą autonomię sprawczą ze zdolnością do rozumowania, jej wzmacnianie przez artefakty jest codziennym zjawiskiem (Schaefer, Kahane i Savulescu, 2014). Pomagają one sprawcom dokonywać wyboru pomiędzy opcjami, wnioskować o ich konsekwencjach oraz je oceniać, analizować potencjalne cele i metody ich osiągnięcia, rozwiązywać samodzielnie problemy itd. Z tego względu niektórzy badacze są skłonni traktować artefakty jako konstytutywne elementy osobowości podmiotu działania. Heersmink argumentuje, że „tożsamość osobowa nie może zatem zostać zredukowana ani do psychologicznych struktur wytwarzanych przez mózg, ani do biologicznych struktur wytwarzanych przez nasz biologiczny organizm. (...) Powinniśmy rozszerzyć nasze pojęcia jaźni, tak aby zawrzeć w nich społeczne i artefaktyczne struktury, skupić się w empirycznych badaniach nad tożsamością osobową na zewnętrznych systemach pamięci i nie ingerować w ludzkie rozproszone umysły i jaźnie (Heersmink, 2017b, 3149)⁹.

Na wagę czynników społeczno-kulturowych w kształtowaniu osobowości i tożsamości podmiotu zwracają również uwagę badacze pracujący w ramach tak zwanej „drugiej” i „trzeciej fali” argumentów za tezą umysłu rozszerzonego¹⁰. Ich analizy harmonizują z relacjonizmem wobec sprawczości, do którego nawiązałam w poprzednim paragrafie. Filozofowie tacy, jak Richard Menary (2010), Edwin

9 “[p]ersonal identity can thus neither be reduced to psychological structures instantiated by the brain nor to biological structures instantiated by our biological organism... [W]e should broaden our concepts of the self so as to include social and artifactual structures, focus on external memory systems in the (empirical) study of personal identity, and not interfere with people’s distributed minds and selves” (Heersmink, 2017b, 3149).

10 Terminologia ta została zastosowana przez Johna Suttona (2010).

Hutchins (2011) czy Lambros Malafouris (2008) podkreślają, że każde działanie poznawcze, które podmiot podejmuje, jest konstytuowane przez praktyki poznawcze kształtowane przez normy kulturowe i instytucje społeczne. Takie rozumowanie idzie dalej niż teza Clarka, dużo silniej zakorzeniona bowiem podmiotowe procesy poznawcze w środowisku społecznym i kulturowym. W konsekwencji indywidualny podmiot niemal rozmywa się w otaczającym i tworzącym go świecie artefaktów i praktyk poznawczych. Nadal podejmuje decyzje jako indywidualny sprawca i bierze za nie odpowiedzialność, lecz, aby zachować niezbędną do tego autonomię, musi być świadomy oddziałujących na niego sił, przyjmować niektóre z nich i opierać się innym (Meyers, 2005). Jest to dość indywidualistyczny sposób myślenia, gdyż opiera się na założeniu o istnieniu sprawcy, podejmującym autonomiczne decyzje, których źródło jest niejako w nim samym. Od indywidualizmu odchodzi już jednak Mason Cash, dla którego autonomiczny sprawca nie jest obiektywnym faktem, lecz statusem, który członkowie danej społeczności nadają konkretnej osobie, jeśli spełnia ona warunki określone przez podzielane normy i praktyki społeczne. Poczucie sprawczości i doświadczenie bycia sobą rozwijają się właśnie w efekcie tego typu relacji z innymi członkami społeczności (Cash, 2010).

Argumenty przedstawione w tej części artykułu prowadzą do wniosku, że status sprawcy powinien być przyznany całemu systemowi człowiek-artefakt. Tezę tę uzasadnia konstytutywny wpływ artefaktów na procesy poznawcze, wiodące do powstania stanów umysłowych odpowiedzialnych za autonomiczne działanie. Konsekwencje teoretyczne i praktyczne, do których prowadzi ta teza, są jednak tak daleko idące, że w praktyce sprawczość nigdy nie jest ujmowana jako w ten sposób rozszerzona. Nie stanowi to oczywiście argumentu przeciwko tej tezie, gdyby bowiem była ona dostatecznie silnie teoretycznie uzasadniona, należałoby pogodzić się z owymi konsekwencjami. W następnej części wskażę jednak, że istnieją ważne racje za jej odrzuceniem, które podważają wnioskowanie od

konstytutywnego wpływu artefaktów na procesy poznawcze i podstawy propozycyjne szerokiego systemu do przyznania mu statusu sprawcy.

4. SYSTEM CZŁOWIEK-ARTEFAKT JAKO PODMIOT AUTONOMICZNEGO DZIAŁANIA – WĄTLIWOŚCI

Problem podmiotowości i sprawczości szerokich systemów stanowi ważną konsekwencję tezy poznania rozproszonego i, w moim przekonaniu, nie może zostać pominięty. Badacze systemów rozproszonych rzadko go jednak podejmują, skupiając się na opisie przebiegu procesów poznawczych w takich systemach oraz działań, które są ich efektem. W analizach tych brakuje jednak często wyjaśnień dotyczących rozumienia wiedzy, stanów umysłowych (lub stanów, które je zastępują) i mocy sprawczej, inicjującej działanie w takich systemach. O ile dyskusja na ten temat toczy się dość żwawo na gruncie analiz dotyczących systemów grupowych¹¹, o tyle jest ona niemal nieobecna w odniesieniu do systemów człowiek-artefakt. Główne pytanie dotyczy tego, kto (co) jest podmiotem poznania, wiedzy i działania w takich systemach, czy mianowicie należy go ograniczyć do człowieka, czy też objąć nim cały system. W poprzedniej części przedstawiłam racje za szerokim rozumieniem sprawczości systemów człowiek-artefakt, które wynikają z tezy poznania rozproszonego. Poniżej zaprezentuję wątpliwości wobec takiego wnioskowania.

4.1. RELACYJNA KONCEPCJA ARTEFAKTÓW

Skutkiem rozproszonych procesów poznawczych są stany umysłowe, które, zdaniem zwolenników rozszerzonej sprawczości, należy uznać

11 W dyskusji nad systemami grupowymi i ich umysłowością udział biorą między innymi Deborah Tollefsen (2015), Georg Theiner (2013), Thomas Szanto (2014), Orestis S. Palermos (2016).

za realizowane przez cały szeroki system. Cały system człowiek-artefakt jest mianowicie ich podmiotem oraz sprawcą działania będącego ich efektem. Rozumowanie takie, zrównujące status poznawczy człowieka i artefaktu w szerokim systemie, oparte jest na standardowej definicji artefaktu, odwołującej się do dwóch koniecznych warunków. Pierwszy z nich wskazuje na konieczność strukturalnej modyfikacji pierwotnego, naturalnego stanu obiektu, który zyskuje status artefaktu. Drugi warunek wskazuje, że modyfikacja taka musi być intencjonalna, świadomie dokonana przez użytkownika obiektu w pewnym celu. Gdy warunki te zostaną spełnione, obiekt zyskuje status artefaktu i pozostaje nim niezależnie od tego, czy jest używany czy też nie (Hilpinen, 1993; Norman, 1991). Artefakt ponadto oddziałuje na swojego użytkownika w specyficzny i silny sposób, zmieniając naturę samego zadania poznawczego. Donald Norman podkreśla, że informacyjna struktura człowieka i artefaktu łączą się ze sobą oraz z konkretnym zadaniem poznawczym w jeden szeroki system, wykazujący nowe, wzmocnione zdolności poznawcze (Norman, 1991). To nie indywidualne, naturalne zdolności człowieka są wzmocniane przez artefakt, lecz w wyniku jego działania powstaje nowy system poznawczy, o nowych wzmocnionych zdolnościach, rozwiązujący nowe zadanie poznawcze. Taka argumentacja skłania wprost do wniosku, że podmiotem tych nowych zdolności i sprawcą działań poznawczych jest cały szeroki system. Należy jednak zauważyć, że teza ta jest sformułowana z zewnętrznego punktu widzenia, przez obserwatora działania szerokiego systemu. Inna perspektywa otwiera się jednak z punktu widzenia osoby wchodzącej w skład tego systemu. Będąc sprawcą działania, to ona czuje się odpowiedzialna za wykonanie zadania poznawczego, które, w wyniku zastosowania przez nią artefaktu, przyjmuje nową postać i stawia nowe wyzwania, wymagające zdobycia nowej wiedzy lub użycia odmiennych zdolności poznawczych niż wymagało tego pierwotne zadanie (Norman, 1991). Z punktu widzenia użytkownika artefaktu to on jest wykonawcą zadania, a nie cały system. To poczucie sprawczości jest uznawane

przez niektórych badaczy za konieczne dla przyznania systemowi poznawczemu statusu sprawcy (O'Connor, 2009; Bayne, 2008; Nida-Rümelin, 2018). Należy jednak mieć na uwadze, że artefakt może je zaburzyć, co prowadzi do problemów z określeniem podmiotu danego działania. Do tematu tego powrócę w kolejnej części pracy.

Podmiotową (osobową) perspektywę w analizowaniu relacji człowiek-artefakt wykorzystują zwolennicy koncepcji alternatywnej wobec standardowej koncepcji artefaktów poznawczych, nazywanej przez nich realizmem aktywności (ang. *activity realism*), w odróżnieniu od realizmu bytowego (ang. *entity realism*), zakładanego przez standardową koncepcję. Jeden z jej przedstawicieli, Sadjad Soltanzadeh, argumentuje, że metafizyczny podział na obiekty, które są artefaktami i które nimi nie są, ze względu na obiektywne własności, jakie posiadają, jest bezużyteczny w codziennej praktyce stosowania wzmocnień poznawczych. Artefakty nie powinny być ujmowane jako obiektywne byty istniejące niezależnie od innych obiektów w świecie (Soltanzadeh, 2019). Według standardowej koncepcji artefakt zostaje powołany do istnienia poprzez intencjonalną modyfikację, dokonaną przez jego użytkownika, po której pozostaje artefaktem niezależnie od tego, co się z nim dalej dzieje. Soltanzadeh zauważa jednak, że użyteczność pewnych obiektów, decydująca o ich stosowaniu jako pomocy poznawczych, nie zależy od tego, czy posiadają one obiektywne cechy artefaktów, lecz od tego, jaki cel podmiot zamierza za ich pomocą uzyskać w konkretnej sytuacji. Użyteczność artefaktów jest zatem własnością relacyjną, określaną w relacji do konkretnych potrzeb danego użytkownika. Warunki modyfikacji i intencji, nałożone na artefakty przez standardową definicję, nie mówią nic o ich użyteczności, która może być określona tylko w odniesieniu do refleksyjnej istoty, decydującej o użyciu danego obiektu jako pomocy poznawczej. Przy takim rozumieniu artefakty nie są obiektywnie istniejącymi obiektami, lecz są zależne od umysłu człowieka, nie tylko w momencie ich powstawania, lecz przez cały czas ich trwania. Ich istotą jest bowiem użyteczność, która może być określona tylko

w relacji do konkretnych celów danego człowieka w konkretnej sytuacji. Zgodnie z omawianą koncepcją artefakty istnieją tylko w relacji do ludzkiego umysłu. Punktem wyjścia do analizy systemu człowiek-artefakt jest zatem człowiek jako refleksyjna istota. To on decyduje o powstaniu systemu i działaniach przez niego podejmowanych, a także o zastosowaniu danego obiektu jako artefaktu umożliwiającego wykonanie zamierzonego działania.

Z powyższego opisu można wywnioskować, że realizm aktywności jest stanowiskiem skupionym na podmiocie ludzkim, podkreśla bowiem jego moc sprawczą w tworzeniu i utrzymaniu systemu człowiek-artefakt. To od celów, jakie człowiek obiera w swym działaniu, zależy, jak dany obiekt w świecie będzie postrzegany i używany. Papierowa torba może na przykład stać się workiem na śmieci, papierem do pakowania lub notatnikiem, w zależności od aktywności podejmowanej przez podmiot (Nyholm, 2018). Nawet wtedy, gdy mamy do czynienia z zaawansowaną technologią informatyczną, która podejmuje relatywnie autonomiczne decyzje poza bezpośrednią kontrolą człowieka, to człowiek jest elementem, który inicjuje tę aktywność i czerpie z niej korzyści, elementem, dla którego ten artefakt działa. Sprawstwo zatem przysługuje w systemach człowiek-artefakt człowiekowi, gdyż tylko ten ich element posiada niezbędną do tego autonomię, polegającą na świadomej decyzji wyboru działania i użycia obiektu jako artefaktu umożliwiającego jego wykonanie.

Rozumienie artefaktu przedstawione powyżej przywołuje na myśl pojęcie afordancji wprowadzone przez Jamesa Gibsona i rozwinięte między innymi przez Normana i innych badaczy. Afordancje odnoszą się najogólniej do tych własności obiektów, które istnieją tylko w relacji do postrzegającego je podmiotu. Obiekty te niejako „oferują” podmiotowi sposoby ich wykorzystania, w zależności od tego, jakie cele chce on osiągnąć w danym środowisku (Gibson, 1977). Wachowski, analizując szczegółowo to pojęcie w swej książce (Wachowski, 2022), wskazuje, że „interakcje poznawcze oparte na afordancjach wydają się najlepiej charakteryzować to, co

systemowe, a nie skoncentrowane na podmiocie w systemach rozproszonych” (Wachowski, 2022, 98). Trudno mi jednak oprzeć się wrażeniu, że analizy dotyczące reprezentacji zewnętrznych, afordancji i dizajnu, podejmowane przez tego autora, dotyczą właśnie podmiotu i wpływu, jaki elementy środowiskowe wywierają na jego działania poznawcze. Reprezentacje zewnętrzne i afordancje obiektów istnieją „dla” podmiotu, to podmiot podejmuje działanie, wykorzystując ich właściwości do wyznaczonych przez siebie celów. Nawet jeśli działanie to nie jest skutkiem świadomego wyboru, przynależy komuś, kto jest w stanie afordancję dostrzec i wykorzystać. Cała dyskusja wokół tych środowiskowych uwarunkowań zdaje się być skupiona na podmiocie, na tym, jak zmieniają się jego zadania poznawcze i sposoby ich wykonywania w jego interakcji z elementami jego środowiska społeczno-kulturowego. Nie przeczę, że elementy te niekiedy tak silnie konstytuują procesy poznawcze i przez to stany umysłowe, które są ich efektem, że stanowią razem z człowiekiem jeden szeroki system. To wciąż jednak człowiek jest sprawcą działania poznawczego, które wykonuje po to, aby utrzymać swoje istnienie i realizować swoje wartości.

4.2. DOŚWIADCZENIE PODMIOTOWOŚCI JAKO WARUNEK SPRAWCZOŚCI

Ograniczanie podmiotowości do ludzkiego elementu systemu człowiek-artefakt wspierają stanowiska wobec sprawczości, które wiążą ją istotnie ze świadomym odczuciem tego, że jest się sprawcą działania. Jak argumentuje Shaun Gallagher, „(..) jeśli ktoś lub coś sprawia, że coś innego zachodzi, ta osoba lub rzecz nie jest sprawcą (nawet jeżeli może być tego przyczyną) jeśli nie wie, w pewien sposób, że spowodował zajście tego zdarzenia” (Gallagher, 2007, 347)¹². Nie musi

12 “That is, if someone or something causes something else to happen, that person or thing is not an agent (even if they might be a cause) if they do not know in some way that they have caused it to happen” (Gallagher, 2007, 347).

być to wiedza propozycjonalna, zakładająca znajomość treści swoich stanów umysłowych i ich racjonalności. Wystarczy odczucie tego, że jest się przyczyną (autorem) danego działania i że sprawuje się nad nim kontrolę. Innymi słowy podmiot działania postrzega siebie jako odrębnego sprawcę zdarzeń w swoim otoczeniu, zmierzających do realizacji celów, które uznaje za wartościowe. Przy takim rozumieniu sprawczości powstaje pytanie, czy w szerokim systemie człowiek-artefakt nie ogranicza się ona do człowieka, gdyż tylko ten element systemu jest zdolny do posiadania takiego odczucia.

Odczucie sprawczości jest jednym z najbardziej podstawowych doświadczeń w naszym życiu. To, co się na nie składa, jest przedmiotem zaawansowanej dyskusji wśród filozofów. Najczęściej wymienia się doświadczenie celowości działania, posiadania intencji do jego podjęcia czy wysiłku związanego z osiągnięciem wyznaczonego celu. Aby podjąć zamierzone działanie, sprawca musi zainwestować energię i siłę woli, a to wymaga odczucia siebie jako źródła mocy koniecznego do realizacji planów. Filozofowie, którzy opowiadają się za tezą przyczynowości mentalnej (ang. *mental causation thesis*), głoszą, że na doświadczenie siebie jako autora danego ruchu składa się odczucie intencji jako jego przyczyny (Wegner, 2002; Hohwy, 2004). Z kolei zwolennicy tezy przyczynowości podmiotowej (ang. *agent causation thesis*) argumentują, że istotą doświadczenia sprawczości jest odczucie siebie jako źródła działania. Niektóre intencje mogą być bowiem odczuwane przez ludzi jako znajdujące się poza ich kontrolą (Chisholm, 1976; Taylor, 1966; Horgan, Tienson i Graham, 2003). Sprawca, innymi słowy, ma poczucie „mojności” (ang. *mineness*) działania, bycia jej autorem, inicjatorem. To stanowisko zdaje się być niezgodne z teorią mentalnej przyczynowości, gdyż zakłada, że to sam podmiot jest przyczyną swych działań, a nie jego stany umysłowe (Bayne, 2008; Nida-Rümelin, 2018)¹³.

13 Interesujące rozwiązanie tego problemu proponuje John Bishop, wprowadzając ideę podmiotu jako przyczyny na grunt przyczynowej teorii działania (Bishop, 1997).

Poczucie autorstwa działania wspiera intuicję, że decyzje podmiotu nie są przyczynowo zdeterminowane, lecz są wolne, co przez wielu filozofów było wskazywane jako warunek moralnej odpowiedzialności (Kant, 1996; Taylor, 1966; Chisholm, 1995). Intuicja ta leży u podstaw stanowisk inkompatybilistycznych i libertariańskich, zgodnie z którymi wolne działanie nie może być skutkiem żadnych zdarzeń mentalnych lub fizycznych¹⁴. To, czy możliwe jest doświadczenie siebie samego jako niezdeterminowanego przez nic sprawcę działania, jest problemem wymagającym odrębnej dokładnej analizy, nie obejmującej jednak tematyki tego artykułu (Bayne, Levy, 2006). W tym miejscu wystarczy założyć, że odczucie autorstwa działania, bycia jego przyczyną, jest istotne dla doświadczenia sprawstwa¹⁵. Ważne w tym względzie jest również odczucie wysiłku. Mianowicie, aby podjąć działanie, podmiot musi zainwestować energię, czyli musi być świadomy siebie jako źródła mocy sprawczej, która skutkuje działaniem (Bayne, 2008)¹⁶.

Czy takie doświadczenie sprawczości, o którym tutaj mowa, nie może zostać zrealizowane częściowo przez elementy leżące poza organizmem człowieka? Jeśli działanie artefaktu konstytuuje, jak to ma miejsce w przypadku Ottona, stany umysłowe, odpowiedzialne za tożsamość osobową, czy nie należy traktować go jako część jednego szerokiego podmiotu, posiadającego świadomość siebie jako przyczyny podejmowanych działań? Przy założeniu, że Otto bez notatnika nie jest w stanie wykonać żadnego działania, czyli nie stanowi

14 Propozycję uzgodnienia libertarianizmu i kompatybilizmu proponuje Martin Nida-Rümelin. Argumentuje on, że w aspekcie ludzkich wolnych działań istnieje specyficzny rodzaj przyczynowania, które zachodzi pomiędzy osobami i zdarzeniami (Nida-Rümelin, 2018).

15 Odczucie to jest trafnie opisane przez Terry'ego Horgana poprzez przykład z zaciskaniem pięści: "You experience your arm, hand, and fingers as being moved by *you yourself* – rather than as experiencing their motion either as fortuitously moving just as you want them to move, or passively experiencing them as being caused by your own mental states. You experience the bodily motion as generated by *yourself*" (Horgan, 2007, 187).

16 Więcej na temat różnych stanowisk stanowiących problematykę odczucia sprawczości znajduje się w moim artykule (Tomczyk, 2023).

sprawcy, czy nie należy traktować całego systemu Otto + notatnik jako autonomicznego, samodzielnego, świadomego podmiotu?

Przeciwko takiemu rozwiązaniu opowiadają się sami inicjatorzy tezy aktywnego eksternalizmu – Clark i Chalmers. Co prawda przyznają oni, że w przypadku Ottona można mówić o rozszerzonej jaźni, rozumieją ją jednak jako przekraczającą granice świadomości (Clark, Chalmers, 1998). Jaźń, jako tożsamość osobowa w czasie, konstituowana jest przez pamięć, która w przypadku Ottona wykracza poza jego organizm w postaci jego przekonań dyspozycyjnych, zapisanych w notatniku. Otto nie jest ich jednak aktualnie świadomy. Świadomość tymczasem, zdaniem Chalmersa, koreluje z fizycznymi procesami, które czynią daną informację bezpośrednio dostępną dla globalnej kontroli (ang. *direct available for global control*). Owe procesy zachodzą wyłącznie w mózgu, gdyż te, które rozszerzają się poza organizm poprzez percepcję lub działanie, zapewniają globalnej kontroli jedynie pośredni dostęp do danej informacji. Informacja taka musi bowiem przebyć wtedy drogę złożoną z trzech etapów: od obiektu do oka, od oka do kory wzrokowej, a stamtąd do obszarów mózgowych odpowiedzialnych za realizację globalnej kontroli. Tymczasem wewnętrzne, neuronalne korelaty świadomości muszą przebyć tylko pewną część ostatniego etapu. Zakładając, że świadomość wymaga bezpośredniej dostępności globalnej kontroli do informacji, nie może być ona rozszerzona poza mózg (Chalmers, 2019). Przekonanie Ottona o adresie Muzeum Sztuki Współczesnej jest dyspozycyjne, nie wymaga zatem dostępu globalnej kontroli do jego treści i dlatego ten stan umysłowy może być rozszerzony. Zatem mimo, że percepcja i działanie nie stanowią granic dla procesów poznawczych, stanowią granice dla świadomości. Przy założeniu, że świadomość jest konieczna dla sprawczości, podmiot działania również pozostaje w granicach organizmu, a właściwie nawet w granicach mózgu.

4.3. ODPOWIEDZIALNOŚĆ SZEROKIEGO SYSTEMU CZŁOWIEK-ARTEFAKT

Traktowanie szerokiego systemu człowiek-artefakt jako autonomicznej sprawy prowadzi do daleko idących konsekwencji, dotyczących odpowiedzialności zarówno epistemicznej, jak i moralnej za podejmowane przez system działania. Pokazuje to wyraźnie Jessica Swallow, odwołując się do eksperymentu myślowego z Ottonem w roli głównej (Swallow, 2013). Wyobraźmy sobie mianowicie, że Otto popełnia zbrodnię. Zgodnie z oryginalnym scenariuszem pamięta on ten fakt tylko dzięki dostępowi do swojego notatnika, w którym zapisuje swoje przekonania. Bez notatnika Otto nie posiada poczucia ciągłości swojej osobowości w czasie, nie postrzega siebie jako sprawcy swych przeszłych działań, gdyż ich nie pamięta. Dopiero wraz z notatnikiem, do którego ma dostęp, staje się on autonomicznym sprawcą, który może ponieść odpowiedzialność za swoje działanie. To prowadzi jednak do praktycznie nieakceptowalnej konsekwencji, że osobą popełniającą zbrodnię jest Otto + notatnik, notatnik bowiem jest konieczny dla Ottona dla podjęcia decyzji o zbrodni oraz dla zachowania pamięci o tym zdarzeniu. Sam Otto stanowi inną osobę, która nie jest, zgodnie z założeniem eksperymentu, zdolna do żadnych przemyślanych działań. Winny jest zatem Otto + notatnik i to ten szeroki system powinien być obciążony odpowiedzialnością za zbrodnię, jakkolwiek miałyby to wyglądać. Tak wygląda konsekwencja przyjęcia tezy umysłu rozszerzonego: jeśli pamięć jest realizowana (częściowo) przez elementy spoza organizmu człowieka (przekonania zapisane w notatniku) i jeśli podmiot działania jest konstytuowany (częściowo) przez pamięć, to on również wykracza poza człowieka. Nie sposób jednak przyjąć takiego wniosku. Szerokiemu podmiotowi nie mogą przysługiwać prawa i obowiązki przypisywane podmiotom ludzkim. Obciążanie kogokolwiek winą ma bowiem sens tylko wtedy, gdy osoba ta jest świadoma zarówno podejmowanego czynu, jak i jego konsekwencji. Nawet jeśli Otto potrzebuje dostępu do notatnika, aby być obciążonym winą, to on, a nie on + notatnik jest podmiotem

odpowiedzialnym za zbrodnię. Sprawczość przynależy wyłącznie Ottonowi, gdyż to on podejmuje decyzję o dokonaniu zbrodni i doświadcza jej popełniania, on również decyduje o używaniu notatnika jako nośnika swoich przekonań.

Problem ze sprawczością systemu złożonego z człowieka i artefaktu nie jest jednak tylko domeną filozofów lubujących się w eksperymentach myślowych, lecz stanowi przedmiot analiz badaczy opisujących faktyczne przypadki stosowania nowoczesnej technologii. Dotyczy to między innymi interfejsów mózg-komputer (BCI), które pozwalają człowiekowi na używanie urządzenia, na przykład wózka inwalidzkiego, z pominięciem obwodowego układu nerwowego lub systemów motorycznych. Badacze zauważają mianowicie, że używanie takich artefaktów może prowadzić do zaburzeń autonomii sprawczej podmiotu. Dzieje się tak wtedy, gdy kontrola nad działaniem niejawnie „przełącza się” z człowieka na urządzenie. Użytkownik w takich przypadkach przestaje czuć pewność, że działanie zostało zainicjowane przez niego samego, że jest on jego autorem (Heselager, 2013). Takie systemy podzielonej kontroli (ang. *shared control systems*) były przedmiotem badań naukowców pracujących w projekcie MAIA (Mental Augmentation through determination of Intended Action) (Vanacker *et al.*, 2007; Milla'n *et al.*, 2010; Tonin *et al.*, 2010). Badacze ci argumentują, że tworzone przez nich systemy wykazują podzielną autonomię (ang. *shared autonomy*), gdyż trudno odpowiedzieć na pytanie kto – człowiek, maszyna czy oba te elementy – kontrolują działanie systemu w danym momencie. Sytuacja ta wpływa silnie na poczucie sprawczości człowieka, posiadającego się artefaktem, co powoduje pojawienie się problemu z odpowiedzialnością za działanie. W czasie interakcji mózg – komputer – urządzenie może dojść bowiem do różnego rodzaju zaburzeń, w wyniku których urządzenie w sposób nieuzasadniony przejmuje kontrolę nad działaniem lub zachowuje się w sposób błędny, niezgodny z intencją użytkownika. Tracąc poczucie kontroli nad działaniem, człowiek przestaje postrzegać siebie jako jego przyczynę i tym samym jego

autonomia sprawcza zostaje podważona. Pod znakiem zapytania staje wtedy kwestia odpowiedzialności zarówno epistemicznej, jak i moralnej za działanie takiego systemu. Skoro sprawstwo nie może zostać przypisane tylko i wyłącznie człowiekowi, nie może mu być również w pełni przypisana zasługa lub wina za podjęte działanie. Kto jest jednak wtedy za nie odpowiedzialny?

Problem odpowiedzialności staje się szczególnie ważny w dobie artefaktów, które podejmują coraz bardziej autonomiczne działania przy coraz słabszej ludzkiej interwencji. Ich przykładami mogą być tzw. autonomiczne samochody lub systemy broni. Obarczanie jakiegokolwiek człowieka lub grupy ludzi odpowiedzialnością za krzywdy spowodowane przez takie zaawansowane technologie zdaje się być niesprawiedliwe, gdyż nikt nie posiada kontroli nad konkretnymi działaniami podejmowanymi przez te maszyny. Z drugiej strony nie sposób przypisać odpowiedzialności samym artefaktom. Ta tzw. luka w odpowiedzialności (ang. *responsibility gap*) (Matthias, 2004) stanowi poważne wyzwanie nie tylko dla filozofów teoretyków, lecz przede wszystkim dla twórców prawa. Niektórzy argumentują za przypisywaniem w takich sytuacjach ludziom i sztucznym systemom wspólnej lub rozproszonej sprawczości, zaznaczając jednak, że kontrolną funkcję pełnią w takiej współpracy konkretni ludzie (Nyholm, 2018). Czy pojęcie rozproszenia jest tu jednak w jakikolwiek sposób pomocne? Czy stanowi przewagę nad pojęciem indywidualnego sprawcy? Kwestia sprowadza się ostatecznie do pytania, czym różni się ludzka sprawczość od sprawczości autonomicznych systemów i na ile człowiek jest odpowiedzialny za ich działanie.

Tym, co odróżnia człowieka od autonomicznych maszyn, jest możliwość refleksji nad racjami kierującymi jego decyzjami. Maszyna nie wyznacza sobie sama celów, są one zadane przez człowieka, jeśli nie szczegółowo, to w ogólnym zarysie. Na przykład autonomiczny samochód nie tworzy sam zasad bezpieczeństwa i ruchu drogowego, nie zadaje też sam sobie celu podróży. Działa on pod kontrolą człowieka lub grupy ludzi na skutek ludzkich intencji, nie stanowi

zatem autonomicznego sprawcy (Nyholm, 2018; Soltanzadeh, 2002; Heersmink, 2017a; Formosa, 2021). Można by w takim przypadku mówić o rozproszonej sprawczości, nie rozwiąże to jednak problemu z przypisaniem jakiemuś podmiotowi odpowiedzialności za skutki działania maszyny. Odpowiedzialność taką można przypisać tylko podmiotowi moralnemu, czyli takiemu, który jest świadomy możliwości wyboru pomiędzy alternatywnymi działaniami oraz ich konsekwencji dla siebie i otoczenia w kontekście przyjętej w danej społeczności hierarchii wartości. Co więcej, podmiot taki może świadomie zdecydować o działaniu wbrew zastanym zasadom, a nawet wbrew zasadom, które sam wyznaje. Tylko taki system może być obarczony odpowiedzialnością zarówno epistemiczną, jak i moralną za podjęte działanie i stanowi go obecnie wyłącznie człowiek.

5. UWAGI KOŃCOWE

W artykule starałam się pokazać, że wątpliwości związane z odchodzeniem od indywidualizmu metodologicznego i poznawczego na rzecz ujęcia systemowego przy analizie działalności poznawczej systemu człowiek-artefakt, przeważają nad racjami uzasadniającymi takie rozwiązanie. Należy oczywiście pokreślić, że nie jest to rodzaj systemów standardowo analizowanych przez zwolenników poznania rozproszonego, takimi bowiem są te złożone z wielu ludzi i artefaktów. Z drugiej jednak strony nie można zgodzić się z tezą, że jest to system poznawczy przywoływany rzadko w tego typu badaniach. Świadczy o tym chociażby uwaga, jaką w swej książce Witold Wachowski poświęca roli reprezentacji zewnętrznych, afordancji i dizajnu w działalności poznawczej człowieka, szczególnie w rozdziale 2. „Interakcje w rozproszonych systemach poznawczych”. Rzecz jasna nie każda opisywana w tym kontekście relacja pomiędzy człowiekiem a elementem jego środowiska stanowi przykład sprzężonego systemu pretendującego do miana szerokiego sprawcy. Takie systemy stanowią zdecydowaną mniejszość. Ważna jest jednak ogólna

heurystyka stosowana do opisu i analizy zachodzącego w takiej sytuacji działania poznawczego. Jest to mianowicie heurystyka skupiona na indywidualnym podmiocie, mimo że realizowana jest w kierunku „od kultury do mózgu”. Decydująca w tej kwestii jest odpowiedź na pytanie, kto jest sprawcą działania, o którym mowa: człowiek, czy jakaś większa całość? Kto mianowicie tworzy i wykorzystuje reprezentacje zewnętrzne, aranżuje przestrzeń, zarządza czasem i pamięcią zewnętrzną? Samo mówienie o „zewnętrznych” reprezentacjach zakłada istnienie indywidualnego podmiotu, stanowiącego centrum systemu poznawczego. Pojęcia afordancji i dizajnu również są osadzone w heurystyce podmiotowo-centrycznej. Afordancje rozumiane są jako własności relacyjne danego systemu podmiot-otoczenie, które oferują podmiotowi określone możliwości działania. Kim (czym) jest jednak ów podmiot, który stanowi centrum całej tej dyskusji?

Nie mam wątpliwości co do tego, że czynniki kulturowe wpływają na treść stanów umysłowych, decydujących o podejmowanych przez człowieka działaniach poznawczych. Wpływ ten jest nieraz tak znaczny, że można mówić o zachodzeniu relacji współkonstytuowania (wraz z elementami wewnętrznymi wobec organizmu). Nawet jednak przy przyjęciu założenia o zaistnieniu szerokiego systemu poznawczego, w którym przebiegają rozproszone procesy poznawcze, a stany umysłowe są realizowane częściowo przez elementy zewnętrzne wobec organizmu człowieka, to człowiek stanowi sprawcę działania poznawczego, on jest bowiem jedynym kandydatem na podmiot intencji urzeczywistniającej wybór obarczony odpowiedzialnością. Ujęcie systemowe nie stanowi zatem w analizie problematyki sprawczości systemu człowiek-artefakt przewagi wyjaśniającej nad ujęciem indywidualistycznym (skupionym na podmiocie).

O ograniczeniu sprawczości do człowieka w szerokich systemach człowiek-artefakt decydują ostatecznie dwa argumenty. Pierwszy dotyczy roli świadomości w autentycznym działaniu. Jednym z często podkreślanych warunków sprawczości, na co zwróciłam uwagę w poprzednim punkcie, jest odczucie bycia przyczyną swoich własnych

działań, sprawowania nad nimi kontroli. Niezbędny ku temu jest dostęp do treści własnych postaw intencjonalnych, które kierują działaniem. Doświadczenie sprawczości jako stan świadomy zdaje się być jednak realizowany przez procesy mózgowy, na co wskazuje między innymi jeden z inicjatorów aktywnego eksternalizmu – David Chalmers¹⁷. Drugi argument przeciwko rozszerzonej sprawczości dotyczy różnicy pomiędzy procesami i stanami umysłowymi, a procesami fizycznymi, które je realizują. Nawet przy założeniu, że stany umysłowe, odpowiedzialne za sprawczość, są realizowane częściowo przez elementy wykraczające poza organizm człowieka, nie jest konieczne przyjęcie tezy, że one same również są tak rozszerzone. Własności wyższego rzędu na gruncie nieredukcyjnego fizykalizmu mogą zostać potraktowane jako нефизyczne, nie posiadające przestrzennej lokalizacji¹⁸. Jako systemowe powinny być, zgodnie z przyjętym rozumieniem, uznane za cechy całego systemu, jednak w przypadku takiej własności, jak sprawczość systemu człowiek-artefakt, jej przypisanie do konkretnego obiektu należy podeprzeć innymi racjami niż lokalizacja fizycznego realizatora. Podmiot autentycznego działania powinien stanowić również podmiot odpowiedzialności epistemicznej i moralnej, a takim systemem może być tylko system świadomy, ograniczony w przypadku szerokiego systemu człowiek-artefakt do człowieka.

W literaturze omawianego problemu można odnaleźć wiele eksperymentów myślowych, wskazujących na absurdalne scenariusze, do których prowadzi przypisywanie sprawczości, a przez to odpowiedzialności, zasługi i winy całemu systemowi człowiek-artefakt (Olson, 2011). Należy jednak podkreślić, że te problematyczne, praktyczne

17 W literaturze dotyczącej tej tematyki nie brakuje pozycji, argumentujących za rozszerzaniem tożsamości osobowej i jaźni poza organizm człowieka. Jedną z nich jest praca Russela W. Belka *Extended Self in the Digital World* (2013). Dyskusja na ten temat zawarta jest w mojej książce *Podmiotowość rozszerzonych systemów poznawczych* (Tomczyk, 2022).

18 Ten sposób argumentacji przeciwko tezie rozszerzonej jaźni rozwija Lynne Rudder Baker (2009).

konsekwencje nie mogą same stanowić argumentu przeciwko tezie rozszerzonej sprawczości, mogą jedynie motywować taką argumentację. Jeśli istniałyby silne racje przeciwko ograniczaniu sprawczości do ludzi w szerokich systemach, praktyczne konsekwencje tej tezy musiałyby zostać przyjęte i należałoby stawić im czoła. W artykule tym miałam na celu pokazanie, że nie ma znaczących powodów, aby podejmować taki wysiłek, przeciwnie, istnieją racje, aby tego nie robić. Sprawczość może być zatem przypisana wyłącznie systemowi, który jej świadomie doświadcza, a doświadczenie to nie jest silnie zaburzone przez dysfunkcje neurologiczne, farmakologię i inne osiągnięcia współczesnej technologii. Podmiot autonomicznego działania odczuwa to, że jest jego źródłem oraz że posiada nad nim kontrolę. Ta pierwszoosobowa perspektywa jest wskazywana przez większość filozofów działania jako niezbędna dla przypisania odpowiedzialności epistemicznej i moralnej. Refleksja nad procesem powstawania intencji działania, samo-kontrola, świadomość zmian treści swoich stanów umysłowych i ich źródeł, odczucie wysiłku, satysfakcji, dumy lub porażki, wstydu i lęku przed karą – żaden z tych stanów umysłowych nie może być rozproszony na artefakty i przypisany całemu systemowi człowiek-artefakt. Z tego powodu nawet przy założeniu, że niektóre intencje i decyzje mogą być częściowo realizowane przez procesy wykraczające poza organizm człowieka, nie stanowią one własności całego szerokiego systemu, lecz przynależą człowiekowi. To Otto przywołuje treść swoich stanów umysłowych do swojej świadomości i na ich podstawie decyduje o podejmowanym działaniu, nawet przy założeniu, że jego pamięć znajduje się poza nim.

BIBLIOGRAFIA

- Aydin, C. (2015). The artifactual mind: overcoming the 'inside-outside' dualism in the extended mind thesis and recognizing the technological dimension of cognition. *Phenomenology and the Cognitive Sciences*, 14(1), 73-95.
- Baker, L.R. (2009). Persons and the Extended Mind Thesis. *Zygon*, 44(3), 642-658.

- Barandiaran, X.E., Di Paolo, E., i Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior*, 17(5), 367-386.
- Bayne, T. (2008). The Phenomenology of Agency. *Philosophy Compass*, 3(1), 182-202.
- Bayne, T., Levy, N. (2006). *The Feeling of Doing: Deconstructing the Phenomenology of Agency*. W N. Sebanz, W. Prinz (red.), *Disorders of Volition* (49-68). The MIT Press.
- Beer, R.D., Di Paolo, E.A. (2023). The theoretical foundations of enaction: Precariousness. *Biosystems*, 223, 104823.
- Belk, R.W. (1988). Possessions and the Extended Self. *Journal of Consumer Research*, 15, 139-168.
- Belk, R.W. (2013). Extended Self in the Digital World. *Journal of Consumer Research*, 40(3), 477-500.
- Bernini, M. (2014). Supersizing Narrative Theory: In Intention, Material Agency, and Extended Mind-Workers. *Style*, 48(3), 349-449.
- Bishop, J. (1997). *Naturalising Mental Action*. W G. Holmström-Hintikka, R. Tuomela (red.), *Contemporary Action Theory*, Volume 1: *Individual Action* (251-266). Springer.
- Chalmers, D. (2019). *Extended cognition and Extended Consciousness*. W M. Colombo, E. Invine, M. Stapleton (red.), *Andy Clark and His Critics* (9-20). Oxford University Press.
- Chemero, A. (2001). Dynamical Explanation and Mental Representations. *Trends in Cognitive Sciences*, 5(4), 141-142.
- Chisholm, R. (1976), *The Agent as Cause*. W M. Brand, D. Walton (red.), *Action Theory* (199-211). D. Reidel Publishing Co.
- Chisholm, R. (1995). *Agents, Causes, and Events: The Problem of Free Will*. W T. O'Connor (red.), *Agents, Causes and Events: Essays on Indeterminism and Free Will* (95-100). Oxford University Press.
- Clark, A. (2003). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford University Press.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford University Press.
- Clark, A., Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7-19.
- Clowes, R. (2015). Thinking in the Cloud: The Cognitive Incorporation of Cloud-Based Technology. *Philosophy and Technology*, 28, 261-296.
- Davidson, D. (2002). *Casual Relations*. W D. Davidson, *Essays on Actions and Events* (149-162). Clarendon Press.

- Formosa, P. (2021). Robot Autonomy vs. Human Autonomy: Social Robots, Artificial Intelligence (AI), and the Nature of Autonomy. *Minds and Machines*, 31, 595-616.
- Gallagher, S. (2007). The Natural Philosophy of Agency. *Philosophy Compass*, 2(2), 347-357.
- Gallager, S., Zahavi, D. (2015). *Fenomenologiczny umysł*. PWN.
- Garrett, B. (1998). *Personal Identity and Self-Consciousness*. Routledge.
- Gibson, J.J. (2014). *The Ecological Approach to Visual Perception*. Psychology Press.
- Haselager, P. (2013). Did I Do That? Brain-Computer Interfacing and the Sense of Agency. *Mind and Machines*, 23(3), 405-418.
- Heersmink, R. (2012). *Mind and Artifact: A Multidimensional Matrix for Exploring Cognition-Artifact Relations*. W M. Bishop (red.), *Proceedings of the 5th AISB Symposium on Computing and Philosophy* (54-61). Yasemin Erden.
- Heersmink, R. (2017a). Distributed Cognition and Distributed Morality: Agency, Artifacts and Systems. *Science and Engineering Ethics*, 23(2), 431-448.
- Heersmink, R. (2017b). Distributed selves: personal identity and extended memory systems. *Synthese*, 194, 3135-3151. <https://doi.org/10.1007/s11229-016-1102-4>.
- Hilpinen, R. (1993). Authors and Artifacts. *Proceedings of the Aristotelian Society*, 93, 158-178.
- Hohwy, J. (2004). The Experience of Mental Causation. *Behavior and Philosophy*, 32, 377-400.
- Horgan, T. (2007). Mental Causation and the Agent-Exclusion Problem. *Erkenntnis*, 67, 183-200.
- Horgan, T., Tienson, J., i Graham, G. (2003). *The Phenomenology of First-Person Agency*. W S. Walter, H.-D. Heckmann (red.), *Physicalism and Mental Causation: The Metaphysics of Mind and Action* (323-340). Imprint Academic.
- Hutchins, E. (1995). *Cognition in the wild*. The MIT Press.
- Hutchins, E. (2001). *Cognition, Distributed*. W N. Smelser, P. Baltes (red.), *The International Encyclopedia of the Social and Behavioral Sciences* (2068-2072). Elsevier.
- Hutchins, E. (2011). Enculturating the supersized mind. *Philosophical Studies*, 152(3), 437-446.
- Hutchins, E. (2013). The cultural ecosystem of human cognition. *Philosophical Psychology*, 27(1), 34-49.
- Hutto, D., Myin, E. (2013). *Radicalizing Enactivism: Basic Minds without Content*. The MIT Press.
- Kant, I. (1996). *The Groundwork of the Metaphysics of Morals*. Cambridge University Press.

- Kirsh, D. (2006). Distributed cognition: A methodological note. *Pragmatics and Cognition*, 14(2), 249-262.
- Leeuwen, van C., Verstijnen, I., i Hekkert, P. (1999). *Common unconscious dynamics underlie uncommon conscious effects: a case study in the interaction of perception and creation*. W J. Jordan (red.), *Modeling Consciousness Across the Disciplines* (179-218). University Press of America.
- Locke, J. (1975). *An Essay Concerning Human Understanding*. Clarendon Press.
- Malafouris, L. (2008). *At the Potter's Wheel: An Argument for Material Agency*. W C. Knappet, L. Malafouris (red.), *Material Agency. Towards a Non-Anthropocentric Approach* (19-36). Springer.
- Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, 6(3), 175-183.
- Menary, R. (2007). *Cognitive integration: Mind and cognition unbounded*. Palgrave McMillan.
- Meyers, D.T. (2005). *Decentralizing autonomy: Five faces of selfhood*. W J. Christman, J. Anderson (red.), *Autonomy and the challenges to liberalism* (27-55). Cambridge University Press.
- Nakayama, J. (2013). The extended mind and the extended agent. *Procedia – Social and Behavioral Sciences*, 97, 503-510.
- Nida-Rümelin, M. (2018). Freedom and the Phenomenology of Agency. *Erkenntnis*, 83(1), 61-87.
- Noë, A. (2004). *Action in Perception*. The MIT Press.
- Norman, D.A. (1991). *Cognitive artifacts*. W J.M. Carroll (red.), *Designing interaction: Psychology at the human-computer interface* (17-38). Cambridge University Press.
- Nyholm, S. (2018). Attributing Agency to Automated Systems: Reflections on Human-Robot Collaborations and Responsibility. *Science and Engineering Ethics*, 24, 1201-1219.
- O'Connor, T. (2009). *Agent-Causal Power*. W T. Handfield (red.), *Dispositions and Causes* (184-214). Oxford University Press.
- Olson, E.T. (2011). The Extended Self. *Mind and Machines*, 21(4), 481-495.
- Palermos, S.O., Pritchard, D. (2016). *The Distribution of Epistemic Agency*. W P.J. Reider, (red.), *Social Epistemology and Epistemic Agency: De-Centralizing Epistemic Agency* (109-126). Rowman and Littlefield International.
- Parfit, D. (1971). Personal Identity. *Philosophical Review*, 80, 3-27.
- Rhodes, B. (1997). The wearable remembrance agent: A system for augmented memory. *Personal Technologies*, 1, 218-224.
- Rowlands, M. (1999). *The body in mind: Understanding cognitive processes*. Cambridge University Press.

- Rowlands, M. (2010). *The New Science of the Mind. From Extended Mind to Embodied Phenomenology*. A Bradford Book – The MIT Press.
- Sandberg, A., Bostrom, N. (2006). Converging cognitive enhancements. W W.S. Bainbridge, M.C. Roco (red.), *Progress in convergence: Technologies for human wellbeing* (201-227). Blackwell Publishing.
- Schaefer, O., Kahane G., i Savulescu, J. (2014). Autonomy and Enhancement. *Neuroethics*, 7, 123-136.
- Soltanzadeh, S. (2019). A Practically Useful Metaphysics of Technology. *Techné: Research in Philosophy and Technology*, 23(2), 232-250.
- Soltanzadeh, S. (2022). Strictly Human: Limitations of Autonomous Systems. *Minds and Machines*, 32, 269-288.
- Sutton, J. (2006). Distributed cognition: Domains and dimensions. *Pragmatics and Cognition*, 14(2), 235-247.
- Swallow, J. (2013). *Sharing the blame: Implications of the hypothesis of extended cognition for personal identity and ethics*. (BA Philosophy Dissertation). University of Exeter.
- Szanto, T. (2014). How to share a mind: Reconsidering the group mind thesis. *Phenomenology and the Cognitive Sciences*, 13, 99-120.
- Taylor, R. (1966). *Action and Purpose*. Prentice-Hall.
- Theiner, G., Wilson, R. (2013). *Group Mind*. W B. Kaldis (red.), *Encyclopedia of Philosophy and the Social Sciences* (401-404). Sage Publications.
- Tollefsen, D.P. (2015). *Groups as Agents*. Polity Press.
- Tomczyk, B. (2021). Knower at Risk: Updating Epistemology in the Light of Enhanced Representations. *Studia Semiotyczne*, 35(1), 35-54.
- Tomczyk, B. (2023). The Locus of Agency in Extended Cognitive Systems. *Journal for General Philosophy of Science*. <https://doi.org/10.1007/s10838-023-09666-1>.
- Tomczyk, B. (2024). I Act therefore I Live? Autopoiesis, Sensorimotor Autonomy and Extended Agency. *Diametros*. DOI 10.33392/diam.1847.
- Trybulec, B. (2015). Grupa jako podmiot stanów intencjonalnych – indywidualizm a obiektywizm. *Zagadnienia Naukoznawstwa*, 2, 147-156.
- Trybulec, B. (2017). Group as a distributed subject of knowledge. Between radicalism and triviality. *Avant*, 7, 183-207.
- Trybulec, B. (2018). It takes effort to be (Collectively Rational): Group as a Reasoning Agent. *Poznań Studies in the Philosophy of the Sciences and the Humanities*, 111, 54-70.
- Vanacker, G., Millan, del R., Lew, E., Ferrez, P.W., Moles, F.G., Philips, J., van Brussel H., i Nuttin, M. (2007). Contextbased filtering for assisted brain-actuated wheelchair driving. *Computational Intelligence and Neuroscience*. <https://downloads.hindawi.com/journals/cin/2007/025130.pdf>.

- Varela, F. (1979). *Principles of Biological Autonomy*. North Holland.
- Wachowski, W. (2022). *Poznanie rozproszone. Od heurystyk do mechanizmów*. Wydawnictwo UMCS.
- Wegner, D. (2002), *The Illusion of Conscious Will*. The MIT Press.
- Wheeler, M. (2010). *In defense of extended functionalism*. W R. Menary (red.), *The Extended Mind* (245-270). The MIT Press.

REASONS FOR AND AGAINST THE DISTRIBUTED HUMAN-ARTEFACT SYSTEM AS AN AUTONOMOUS AGENT

Abstract. Supporters of distributed cognition declare a departure from methodological and cognitive individualism toward a dynamic systems approach, on the basis of which cognition is treated as a property of the entire wide system, and not of a single person. In this article, I focus on the problem of the agency of the human-artifact system and show that the distributed approach does not provide an explanatory advantage over the individualistic perspective on this issue. The analysis of external representations, affordances and design carried out by Witold Wachowski in his book *Distributed Cognition. From Heuristics to Mechanisms [Poznanie rozproszone. Od heurystyk do mechanizmów]* (2022) does not stray far from an individual-centered approach, despite his use of the “from culture to brain” heuristic. This heuristic does not contradict the subject-centric approach. In this article, I argue that the reasons for rejecting individualism outweigh the reasons that justify a distributed approach.

Keywords: autonomous agency; extended mind; wide cognitive system; cognitive artifact; epistemic and moral responsibility; experience of agency

BARBARA TOMCZYK

Uniwersytet Marii Curie-Skłodowskiej w Lublinie
(Maria Curie-Skłodowska University in Lublin, Poland)
ORCID <https://orcid.org/0000-0001-8145-7755>
barbara.tomczyk@mail.umcs.pl

DOI 10.21697/spch.2024.60.A.02



Tekst jest udostępniany na zasadach licencji Creative Commons (CC BY-ND 4.0 Międzynarodowe).
Zgłoszono: 22/03/2023. Zrecenzowano: 10/01/2024. Zaakceptowano do publikacji: 24/04/2024.