



Szczeliny determinacji

Filozoficzna polemika z Robertem M. Sapolskim

Maciej URBANEK*

ABSTRACT

Fissures in determination. A philosophical polemic with Robert M. Sapolsky: The following article engages in a philosophical polemic with selected theses from the book *Determined: The Science of Life Without Free Will* by the eminent neurobiologist Robert M. Sapolsky. The aim of this text, though, is not to refute the hard determinism advocated by Sapolsky, but rather to problematise the cognitive horizon in which is the backdrop for his arguments for the nonexistence of free will. Therefore, the thesis I will formulate at the end of this discussion — as a form of opposition to Sapolsky's claim that science has reliably and unequivocally proven the impossibility of the existence of free will — is not so much the statement that “free will exists” but rather the assertion that “the belief that science has already resolved the free will debate — thereby rendering it moot — is false”. In subsequent sections of the text, I analyse certain philosophical assumptions and conclusions that are central to the coherence of the argument presented in *Determined*. I begin by discussing Sapolsky's understanding of the concept of *free will* and demonstrating that the meaning given to this term by the neuroscientist is *de facto* paradoxical. Next, I critically examine the concept of determination. I am primarily interested in whether its universal and widespread character truly opens the way to explaining all phenomena. Finally, in the last sections of the text, I will attempt to question the validity of the mechanical-reductive approach to explaining human activity.

KEYWORDS

determinism, free will, neurobiology, philosophy

* Dr, adiunkt, Instytut Socjologii, Kognitywistyki i Filozofii, Uniwersytet Komisji Edukacji Narodowej w Krakowie. E-mail: maciej.urbanek@uken.krakow.pl.

I. UWAGI WSTĘPNE

W 2023 roku ukazała się książka autorstwa wybitnego amerykańskiego badacza, profesora biologii i neurologii na Uniwersytecie Stanforda Roberta M. Sapolsky'ego zatytułowana *Determined. A Science of life without free will*¹, która wywołała sporo zamieszania w środowiskach filozoficznych na całym świecie. Powód zainteresowania tekstem biologa wydaje się przy tym dosyć oczywisty. W swojej pracy, utrzymanej w tonie syntetycznej prezentacji i komentarza do wyników najnowszych badań naukowych (przede wszystkim z dziedziny neurobiologii i fizyki²), stwierdza on mianowicie, iż współczesna nauka nie pozostawia absolutnie żadnych wątpliwości co do faktu, że wszelka aktywność w świecie i każdy jego poszczególny stan są nieuchronnie zdeterminowane czynnikami biologicznymi oraz fizycznymi. Powszechność i konieczność determinacji oznacza zaś, że wolna wola nie istnieje. Rezygnacja z wiary w ten filozoficzny koncept (choć w praktyce być może trudna do realizacji, co autor książki wielokrotnie podkreśla) jest przy tym nie tylko uzasadniona i teoretycznie konieczna, lecz ma również pożądane konsekwencje praktyczne. Prowadzi bowiem do „innego [oczywiście pełniejszego niż dotychczas — M.U.] spojrzenia na kwestię odpowiedzialności moralnej, winy i nagrody, oraz na kwestię bycia wolnym sprawcą. [...] Lecz przede wszystkim do zmiany fundamentalnych aspektów naszego zachowania” (Sapolsky, 2024: 312). Prowokująca do dyskusji jest przy tym nie tyle sama treść powyższych stwierdzeń — nie wykraczają one bowiem poza ramy wielokrotnie już dyskutowanego filozoficznego inkompatybilizmu³

¹ Książka została przetłumaczona na język polski w 2024 roku i ukazała się pod tytułem: *Zdeterminowany. Jak nauka tłumaczy brak wolnej woli* (przeł. M. Popławska. Poznań: Media Rodzina). Tekst polskiego tłumaczenia będzie podstawą cytatów w niniejszej polemice. Ponieważ jednak odbiega ono w co najmniej kilku miejscach od oryginału i nieco zniekształca niektóre fragmenty książki (widać to już w samym tłumaczeniu podtytułu *Determined*), to w przypadku rozbieżności między przywoływanym przeze mnie fragmentem i tłumaczeniem polskim, które zmieniałyby sens bądź wydzwikał cytowanych fragmentów, będę podawał tłumaczenie własne lub przywołał tekst oryginału w przypisie — wszystkie takie miejsca zostaną odpowiednio zaznaczone.

² Jak czytamy: „[...] prawie każdy fakt w tej książce został odkryty w ciągu ostatnich pięćdziesięciu lat, a prawdopodobnie połowa z nich w ciągu ostatniego roku” (Sapolsky, 2024: 283). Cytat przytaczam wiernie za polskim tłumaczeniem, choć wdał się w niego błąd. W tekście oryginalnym czytamy bowiem: „[...] nearly every fact in this book was discovered in the last fifty years, probably half in the last five” (Sapolsky, 2023: 224).

³ Sapolsky podpisuje się pod tym właśnie stanowiskiem. Jak podkreśla: „Świat jest deterministyczny i wolna wola nie istnieje. W tym ujęciu, jeśli pierwsze założenie jest prawdziwe, drugie również musi takie być; determinizm i wolna wola nie są kompatybilne. Jestem zwolennikiem tej perspektywy zwanej twardym inkompatybilizmem” (Sapolsky, 2024: 19). W przypisie do tego fragmentu Sapolsky dodaje, że „twardy inkompatybilizm” utożsamia z „twardym

— ile raczej kategoryczność i konsekwencja, z jaką autor *Determined* wykazuje ich zasadność oraz to, że argumentuje swoje stanowisko niemal wyłącznie na podstawie materiału dostarczanego przez nauki empiryczne. Warto w tym miejscu zaznaczyć, że na siłę oddziaływania interesującego nas tekstu wpływa dodatkowo fakt, iż Sapolsky prowadzi swoją narrację w sposób polemiczny, tj. antycypując i rozbrajając wiele z potencjalnych zarzutów, które można podnosić w kontrze do prezentowanego przez niego stanowiska⁴. Sposób, w jaki prowadzi swój wywód, jest przy tym na tyle przemyślany i systematyczny (pomimo popularno-naukowego charakteru całości), a materiał wykorzystany do krytyki kompatybilizmu (bo z nim głównie dyskutuje⁵) na tyle obszerny i sugestywny, że czytelnik może odnieść wrażenie, iż uznanie nieistnienia wolnej woli faktycznie jest jedynym możliwym rozwiązaniem, które byłoby jednocześnie zgodne z efektami wielu wieków postępu naukowego i ze zdrowym rozsądkiem (*sic!*).

Pozwolę sobie w tym miejscu na pewien osobisto-egzystencjalny wtęret. Otóż trzeba oddać autorowi *Determined*, że konstruowana przez niego mozaika argumentów i idei jest na tyle dobrze przemyślana, że również jako autor niniejszego tekstu (względnie wykształcony filozoficznie i śledzący toczące się współcześnie debaty nad pojęciem wolnej woli) w wielu momentach czułem, jak wartość i siła uznawanych przeze mnie dotychczas za obowiązujące rozstrzygnięć zaczyna słabnąć, aby niejednokrotnie zupełnie upaść. Stąd też niniejsza polemika z tekstem *Determined* ma do pewnego stopnia wymiar osobisty. Z jednej strony zrodziła się ona bowiem z potrzeby przeanalizowania centralnych punktów argumentacji Sapolsky'ego — zrozumienia stanowiska, które od wielu lat wydawało mi się cokolwiek ezoteryczne, a z którego odrzuceniem czy podaniem w wątpliwość po lekturze *Zdeterminowanego* zacząłem mieć nie-mały problem. Z drugiej natomiast z potrzeby usystematyzowania i klaryfikacji powodów, dla których, pomimo kilkukrotnej lektury książki i przemyślenia tez formułowanych na jej kartach, ostatecznie w dalszym ciągu trwam w przekonaniu, że problem istnienia (bądź nieistnienia) wolnej woli nie został

determinizmem”. W kilku miejscach pracy deklaruje ponadto, że filozoficznie najbliższe jest mu stanowisko Sama Harrisa, który w jednym z tekstów poświęconych problemowi wolnej woli zauważa m.in.: „Problemem jest nie tylko to, że wolna wola nie ma sensu obiektywnie [...]; nie ma ona również sensu subiektywnie” (Harris, 2012: 43).

⁴ W poszczególnych partiach tekstu stara się pokazać m.in., że za istnieniem wolności woli nie przemawiają nieprzewidywalność, zmienność czy przypadkowość zdarzeń w świecie. Przekonuje również, że fenomenowi tego nie znajdziemy na poziomie kwantowego indeterminizmu, emergentnych struktur czy teorii chaosu.

⁵ Libertarianie podejścia do kwestii istnienia/nieistnienia wolnej woli nie są przez niego w zasadzie brane pod uwagę (Sapolsky, 2024: 19–20).

ostatecznie rozstrzygnięty⁶. Nie oznacza to oczywiście, że poniższe rozważania mają charakter autobiograficzny czy że zdają sprawę jedynie z subiektywnych, emocjonalno-intelektualnych odczuć jednego czytelnika. Wręcz przeciwnie — właściwe zrozumienie węzłowych punktów argumentacji autora *Behave*⁷, otwierające jednocześnie możliwość ich analizy, to proces, który rozpoczął się — w moim przypadku — dopiero w momencie przezwyciężenia mentalnego oporu, jaki w pierwszej chwili wzbudziła perspektywa zapomnienia o sobie i uznania, że żadna z chwil przeżytego przeze mnie dotychczas życia nie zależała tak naprawdę ode mnie (jako od przeżywającego ją podmiotu), a jedynie wydarzyła się w świecie. Być może oczywiście nie każdemu taka *epoché* będzie potrzebna, jednak w moim przypadku umożliwiła ona niejako poddanie się prowadzonej w książce narracji i „wzięcie w nawias” (na czas lektury) własnego poglądu w przedmiotowej kwestii. Kolejnym etapem procesu narodzin argumentacji przedstawionej w niniejszym artykule był oczywiście powrót do „własnego” sposobu myślenia i ponowna, krytyczna tym razem, lektura *Determined*. Uzupełniona o konfrontację tekstu Sapolsky’ego z innymi podejmującymi podobną tematykę, pozwoliła nie tyle odrzucić tezy formułowane w interesującej nas pracy, ile zidentyfikować, gdzie w postulowanym na jej kartach obrazie człowieka jako istoty w sposób nieuchronny zdeterminowanej we wszelkich swych działaniach znajdują się potencjalne „szczeliny”, w które wolność decydowania i wybierania mogłaby się wsunąć bądź w których być może już się kryje.

Co do zasady zatem celem niniejszej polemiki nie jest dyskusja z inkompatybilizmem jako takim (a przynajmniej nie bezpośrednio), a raczej problematyzacja horyzontu poznawczego, na gruncie którego neurobiolog Robert M. Sapolsky stara się argumentować na rzecz nieistnienia wolnej woli. O ile bowiem nie czuję się kompetentny, aby kwestionować wnikliwość, z jaką prezentuje on i interpretuje wyniki badań neurobiologicznych czy fizycznych (nie uważam też, aby taka krytyka była rolą filozofa), o tyle wydaje się, że szereg założeń i wniosków natury filozoficznej formułowanych w toku prowadzonych przez niego analiz można (a wręcz powinno się) opatrzyć znakiem zapytania.

Dlatego też stanowiskiem, jakie przeciwstawię Sapolsky’emu, będzie ostatecznie nie teza „wolna wola istnieje”, a raczej stwierdzenie: „teza jakoby współczesna nauka rozstrzygnęła spór o wolną wolę i tym samym uczyniła go bezprzedmiotowym jest nieprawdziwa”.

⁶ Choć faktycznie przekonanie to jest obarczone obecnie znacznie większą dozą niepewności.

⁷ Książka napisana przez Sapolsky’ego w 2017 roku. Przetłumaczona na polski i opublikowana w 2021 roku pt. *Zachowuj się. Jak biologia wydobywa z nas to, co najgorsze, i to, co najlepsze* (przeł. P. Szymczak. Poznań: Media Rodzina).

II. PORZĄDEK WYWODU

Jak pisałem wyżej, podstawowym celem książki Sapolsky'ego jest pokazanie, że nauka potwierdza prawdziwość „twardego determinizmu”⁸, którego istota daje się sprowadzić do treści następujących twierdzeń: „ponieważ [1] świat jest deterministyczny, [2] wolna wola nie może istnieć, a więc [3] pociąganie ludzi do moralnej odpowiedzialności za ich działania jest błędem” (Sapolsky, 2024: 22; numeracja tez — M.U.). Rzeczywistość człowieka jest przez badacza widziana jako jednoznacznie zdeterminowana i zależna od czynników znajdujących się całkowicie poza jego kontrolą:

[...] ewolucja wytwarza geny podlegające procesom epigenetycznym związanym z uwarunkowaniami wczesnego środowiska, które produkują białka, a te z kolei, wspomagane przez hormony w określonym kontekście, działają w mózgu, aby uformować twoją osobę. Oto płynne kontinuum nie pozostawiające żadnych szczelin, w które mogłaby się wślizgnąć wolna wola (Sapolsky, 2024: 280).

Tym, co może dziwić, jest natomiast fakt, że ten uznany naukowiec na przestrzeni kilkusetstronicowego tekstu nie formułuje ani nie przytacza w zasadzie żadnej właściwej definicji wskazanych wyżej terminów. Posługuje się nimi na mocy jedynie lakonicznych omówień i szeregu anegdot, które, choć są na tyle „obrazowe”, aby czytelnik uchwycił intencję autora, nie są jednocześnie wystarczająco precyzyjne, aby się nimi właściwie posłużyć w przypadku bardziej szczegółowej analizy. Dokładniejsza lektura treści pierwszych kilku rozdziałów książki odśladania zresztą bardzo wyraźnie słabe strony takiego sposobu wprowadzania pojęć⁹.

Ponieważ jednak prawomocność tez stawianych w *Determined* jest w znacznym stopniu uzależniona właśnie od sposobu rozumienia terminów takich jak wolna wola i determinacja, wydaje się konieczne, aby w ramach niniejszej

⁸ Problem zakresu determinacji ludzkich działań w świecie stawiany był już w starożytności (m.in. przez Demokryta i stoików, choć nie posługiwali się oni jeszcze pojęciem wolnej woli). Narodziny współczesnej postaci tego problemu należy jednak wiązać z ugruntowaniem się w XVII wieku newtonowskich zasad fizyki (por. Pink, 2004: 14).

⁹ Z jednej strony brak precyzji pojęciowej i pomijanie pewnych stron zagadnienia są zupełnie zrozumiałe w przypadku popularno-naukowego opracowania. Z drugiej jednak, a widać to w szeregu fragmentów pracy, Sapolsky wydaje się nadmiernie bagatelizować, a niekiedy wręcz wyśmiewa drobiazgowość, z jaką filozofowie analizują różne strony pojęć. Sam zachowuje przy tym „klucz” pozytywistyczny, tj. wyklucza ze swoich rozważań jako bezprzedmiotowe i niezwiązane z omawianym zagadnieniem wszystko to, co nie znajduje przełożenia na „obserwowalne” bądź wyprowadzone z empirii mechanizmy i procesy.

polemiki z tekstem biologa sprowadzić właśnie sposób ich rozumienia oraz mechanizm, który pozwala Sapolskiemu je ze sobą wiązać. Stąd też poniższy wywód przeprowadzony zostanie w trzech częściach. Najpierw postaram się pokazać, że wolna wola, której nieistnienia Sapolski dowodzi na kartach *Determined*, jest *de facto* czymś paradoksalnym, tj. nie może istnieć, nawet jeżeli autor ten nie ma racji. Później skoncentruję się na kwestii zakresu postulowanej determinacji — w szczególności na wyeksplikowaniu jej nieograniczonego charakteru. I wreszcie przedstawię kilka problemów wynikających z mechanistyczno-redukcyjnego sposobu tłumaczenia zachowań człowieka.

III. PARADOKSY NIEOBECNEJ WOLNOŚCI

I tak istota wolnej woli w *Determined* dookreślana jest na przykładzie tego, co dzieje się z człowiekiem naciskającym spust pistoletu¹⁰:

Oto wyzwanie dla zwolennika wolnej woli: znajdź neuron [w mózgu pociągającego za spust — M.U.], [...] w którym potencjał czynnościowy pojawił się bez powodu, pomimo iż żaden inny neuron nie „odezwał się” do niego wcześniej. Następnie udowodnij, że na działanie tego neuronu nie miały wpływu czynniki takie, jak zmęczenie, głód, stres czy ból odczuwane przez mężczyznę. Udowodnij, że funkcje tego neuronu nie uległy zmianie za sprawą tego, co mężczyzna widział, słyszał i czuł w minutach poprzedzających pociągnięcie za spust, że nie miały na niego wpływu poziomy hormonów utrzymujące się w jego mózgu w poprzednich godzinach lub dniach, ani to, czy w ostatnich miesiącach lub latach doświadczył wydarzenia zmieniającego życie. Udowodnij, że na rzekomo swobodne funkcjonowanie tego neuronu nie miały wpływu geny mężczyzny lub trwające przez całe życie zmiany w ekspresji tych genów spowodowane doświadczeniami z dzieciństwa. Ani też poziomy hormonów, na które był narażony jako płód, kiedy jego mózg znajdował się w fazie tworzenia. Ani historia i ekologia, które ukształtowały kulturę, w której się wychował. Wskaż neuron będący pierwotną przyczyną w obliczu tych wszystkich czynników. Wybitny filozof Alfred Mele z Uniwersytetu Stanowego Florydy, zwolennik kompatybilizmu, stanowczo stwierdził, że wymaganie czegoś takiego w odniesieniu do wolnej woli jest ustawianiem poprzeczki „absurdalnie wysoko”. Jednak ta poprzeczka nie jest ani absurdalna, ani ustawiona zbyt wysoko. Jeśli wskażesz neuron (lub mózg), który generuje zachowanie niezależne od sumy czynników związanych z jego biologiczną przeszłością, dla celów tej książki uznamy, że wykazałeś istnienie wolnej woli (Sapolski, 2024: 24–25).

¹⁰ Zacytuję nieco dłuższy fragment, gdyż jego treść wydaje się kluczowa dla zrozumienia wyводу Sapolskiego, a ponadto poszczególne jego momenty będą stanowić punkt odniesienia dla kilku dalszych partii niniejszego tekstu.

Tylko — retorycznie zapytam — czy wysuwanie takiego żądania faktycznie nie jest absurdalne?

Otóż zauważmy, że powyższe przedstawienie wpisuje interesujący nas problem w ramy dualizmu, który Sapolsky jednocześnie z góry wyklucza. Postrzega on bowiem jako wolne jedynie te działania, które całkowicie wymykają się warunkowaniu biologicznemu (czy fizycznemu), czyli *de facto* stanowią element zewnętrzny wobec rządzonego przyrodniczymi prawami świata. Z tej perspektywy za czysto retoryczne można uznać wezwanie: jeżeli chcesz wykazać istnienie wolnej woli, to pokaż, w jaki sposób w świecie złożonym jedynie z „gier” neuronalnych potencjałów czynnościowych, genetyki i czynników środowiskowych uruchamiających i dostrajających działanie tych genów możliwe jest coś zupełnie niezależnego i niedającego się sprowadzić do sumy tych uwarunkowań. W praktyce oznacza to bowiem, że przed zwolennikiem istnienia wolnej woli staje zadanie wskazania klasy elementów (bądź działań) istniejących — z perspektywy reguł rządzących światem, w którym formułowane jest to wymaganie — *ex nihilo*. Co więcej, autor *Determined* domaga się, aby owym pozaświatowym czynnikiem był „neuron (lub mózg)”, czyli element stanowiący integralną część biologicznego systemu, a to — jak się wydaje — faktycznie kwalifikuje się jako „absurdalnie wysoko podniesiona poprzeczka”. Oczywiście wydaje się bowiem, że aby element biologicznego systemu, jakim jest układ nerwowy, miał pojawić się/działać całkowicie niezależnie od reszty komponentów biologicznych, to musiałby nie tylko przestać być częścią tego systemu, lecz również „odmienić” swoją naturę, tj. przestać być elementem biologicznym, co czyni wyzwanie Sapolsky’ego z góry niemożliwym do realizacji. Podkreślmy przy tym, że niemożliwość ta wynika nie z faktu nieobecności takiego elementu — w takim przypadku bowiem wymagania neurobiolga należałoby uznać za uzasadnione, a jedynie niespełnione (bądź trudne do spełnienia). Problemem jest tutaj raczej to, że żądanie autora *Determined* mógłby spełnić jedynie przedmiot paradoksalny — tj. będący biologicznym komponentem poddanym warunkowaniu przyczynowo-skutkowemu, a jednocześnie funkcjonujący jako nieuwarunkowany i poza-biologiczny wyjątek od tych reguł. Wolna wola, której w ten sposób domaga się Sapolsky, byłaby zatem faktycznie czymś albo niemożliwym, albo paradoksalnym, albo magicznym.

Warto zwrócić w tym miejscu uwagę, że taki sposób przedstawienia problemu można potraktować poniekąd jako przykład „walki z wiatrakami”. Jak trafnie zauważa Albert Mele, przywoływany zresztą w cytowanym wyżej fragmencie, o ile wśród naukowców (w tym neurobiologów, jak Sapolsky) względnie częste jest stanowisko czyniące z wolności woli element „magiczny” czy „nienaturalny”, np. jako produkt wyłączonej z porządku świata zjawisk duszy (Mele, 2014: 84–88),

o tyle większość badanych laików (Mele, 2014: 87), a także „przeważająca większość profesorów filozofii zajmujących się obecnie problemem wolnej woli [...] odrzuciłaby takie opisy jako dziwaczne” (Mele, 2014: 85). Co ciekawe, sam Sapolsky zwraca na to pośrednio uwagę w pierwszym rozdziale swojej książki, gdy przedstawiając specyfikę podejść kompatybilistycznych, twierdzi:

Nie ma magii ani wróżkowego pyłu, nie ma dualizmu substancji, według którego mózg i umysł są odrębnymi bytami. [...] Pogląd ten podziela w przybliżeniu pięćdziesiąt procent filozofów i prawników (Sapolsky, 2024: 20).

Nasuwa się tym samym (nieco złośliwe) pytanie: czy Sapolsky chce nam powiedzieć, że na rzecz inkompatybilizmu przemawia koniec końców również to, iż pokazuje on nieistnienie wolnej woli w wersji magicznej czy będącej wyrazem jakiejś formy dualizmu, których i tak niemal nikt nie traktuje poważnie (nawet kompatybiliści)?

Kończąc ten wątek, zwrócę jeszcze uwagę na fakt, że ze sformułowanego w *Determined* opisu specyfiki ludzkiego działania wynika, iż wolna wola — gdyby istniała — musiałaby być czymś w konkretny sposób zlokalizowanym. Obraz „wolnego neuronu” (jakkolwiek jest oczywiście tylko metaforą) sugeruje mianowicie jakąś formę substancjalnej podstawy czy biologicznego (fizycznego?) mechanizmu fundującego możliwość wyłączenia spod powszechnych praw determinizmu przyrodniczego. I faktycznie, jak wskazywałem wyżej, trudno nie zgodzić się z Sapolskim, że takiej podstawy znaleźć się póki co nie udało. Wydaje się jednak, że autor książki zbyt pobieżnie i za szybko odrzuca przy tej okazji alternatywy dla swojego sposobu rozumienia wolności woli. Fakt nieistnienia zjawiskowego komponentu czy nośnika wolności nie musi oznaczać wszakże, że koncept ten jest z gruntu niekompatybilny z deterministycznie pomyślanym światem¹¹. Ciekawej alternatywy dla takiego oglądu dostarcza choćby Michael S. Gazzaniga. W książce *Who's in charge? Free will and the science of the brain* (2011)¹² proponuje mianowicie, aby neuronauki zamiast na poszukiwaniach konkretnego rejonu mózgu umożliwiającego wolne działanie jednostek skupiły się raczej na interakcjach wielkiej liczby systemów i podsystemów konstytuujących mózg oraz na wpływach, jakie na te interakcje

¹¹ Przykładowo Daniel Dennett w książce *Freedom evolves* (2003) stara się pokazać wolną wolę jako produkt ewolucji biologicznej pojawiający się dopiero na pewnym poziomie skomplikowania żyjących organizmów, co czyni z niej w zasadzie bezpośredni wynik działania biologicznie pojętej determinacji.

¹² Polskie tłumaczenie: Gazzaniga, M. (2013). *Kto tu rządzi — ja czy mój mózg? Neuronauka a istnienie wolnej woli*. (Przeł. A. Nowak-Młynikowska). Sopot: Smak Słowa.

mają kulturowo-społeczne otoczenie człowieka i obecność innych ludzi. Jak bowiem zauważa, to właśnie na tym polu należałoby szukać źródeł umysłu, który niejako na zasadzie sprzężenia zwrotnego oddziaływałby następnie na swoją neurologiczną (mózgową) podstawę. Jak czytamy:

Moja argumentacja sprowadza się do tego, że wszystkie doświadczenia życiowe — osobiste i społeczne — wpływają na nasz emergentny system psychiczny. Owe doświadczenia są potężnymi siłami, które kształtują nasz umysł. [...] neuronauka będzie jednak musiała przyrzeć się temu, w jaki sposób reguły i algorytmy, które kierują działaniem poszczególnych, rozproszonych modułów, wspólnie wytwarzają to, co nazywamy ludzką kondycją (Gazzaniga, 2013).

Wolna wola i wynikająca z niej możliwość ponoszenia odpowiedzialności za swoje działania mogą być zatem efektem nie tyle jednego, konkretnego elementu czy systemu elementów, ile emergentnym skutkiem współdziałania niezliczonych części składowych. Niewidoczne na poziomie mikro (np. komórkowym czy neuronalnym), odsłaniałyby się na poziomie makro — niczym krajobraz wyłaniający się z tysięcy kolorowych plam na obrazach pointylistów.

Sapolsky krytycznie, co zrozumiałe, odnosi się do tego rodzaju rozwiązań. Komentując pracę Gazzanigi, zauważy, że:

Takie podejście generuje dwa poważne problemy: po pierwsze, wolnej woli i odpowiedzialności nie należy rozpatrywać jedynie na poziomie społecznym tylko dlatego, że każdy tak twierdzi [...]. Po drugie, społeczność, interakcje społeczne, organizmy wchodzące ze sobą w społeczne relacje są końcowym produktem biologii oddziałującej ze środowiskiem w takim samym stopniu jak kształt twojego nosa (Sapolsky, 2024: 109).

Szczególnie drugi z punktów domagałby się rozwinięcia, ale Sapolsky nie pogłębia swojej krytyki. Powodów jego sprzeciwu wobec stanowiska wyrażonego w *Kto tu rządzi?* można jednak szukać w partiach *Determined* poświęconych krytyce idei emergencji jako „mechanizmu” wyłaniania się wolnej woli.

O ile bowiem autor oczywiście zgadza się, że emergentne struktury występują w przyrodzie, a całości, które tworzą, nie są prostymi sumami części składowych, o tyle, co wielokrotnie i na różne sposoby podkreśla: „Systemy emergentne nie mogą sprawić, by cegły, które je zbudowały, przestały być ceglaste” (Sapolsky, 2024: 237). Wydaje się przy tym, że taki argument jest tylko częściowo zasadny. Gazzaniga nie postuluje przecież w swojej pracy zmiany natury czy charakteru interakcji na poziomie części składowych. Postuluje

jednakże, że efektem jednoznacznie zdeterminowanych oddziaływań na poziomie mikro może być uformowanie na wyższym poziomie organizacyjnym struktury, która następnie sama stanie się czynnikiem determinującym sposób organizacji innych poziomów tego systemu.

Obraz taki jest dla Sapolsky'ego trudny do zaakceptowania choćby w tym punkcie, w którym zakłada możliwość determinacji odgórnej. Pogłębienie tych wątków domaga się jednak wyjaśnienia sposobu rozumienia przez biologa pojęcia determinacji, któremu to wątkowi poświęcona jest kolejna część niniejszej polemiki.

IV. NIEOGRANICZONY POCHÓD DETERMINANT

Podobnie jak w przypadku pojęcia wolnej woli, kluczowy dla przedstawienia sposobu, w jaki autor *Determined* posługuje się pojęciem „determinacja”, jest eksperyment myślowy przedstawiony na początku książki. Sapolsky opisuje w nim ceremonię ukończenia studiów i dwóch jej uczestników — absolwenta oraz osobę sprzątającą po uroczystości:

Wybierz losowo jednego z absolwentów. Użyj trochę magii, aby człowiek z obsługi rozpoczął życie z genami absolwenta. Zapewnij mu również podobne warunki płodowe, w których spędził dziewięć miesięcy, oraz związane z nimi konsekwencje epigenetyczne mające wpływ na całe jego życie. [...] Wprowadźmy zmiany nie tylko w życiu śmieciarza, ale podarujmy jego przeszłość absolwentowi. Zamień każdy czynnik, nad którym nie mieli kontroli, a okaże się, że kto inny będzie stał w todze absolwenta, a kto inny będzie zbierał śmieci. To właśnie mam na myśli, mówiąc o determinizmie (Sapolsky, 2024: 27).

Prezentowany w ten sposób obraz nie wydaje się przy tym w najmniejszym stopniu kontrowersyjny. Biolog nie podkreśla tu bowiem nic innego jak to, że określony układ zdarzeń w świecie bezpośrednio wynika z jego wcześniejszych stanów i okoliczności, które do nich doprowadziły. Absolwent odbiera dyplom, ponieważ odpowiednie wyposażenie genetyczne oraz warunki życia pozwoliły mu zdobyć wykształcenie i ukończyć szkołę; osoba obsługująca uroczystość miała mniej szczęścia i w loterii genetyczno-środowiskowej wyciągnęła gorszy los, co jest przyczyną jej słabszej pozycji startowej¹³. Obraz ten jest przy

¹³ Wchodząc w dyskusję z Dennettem, który w książce *Elbow room* stara się wykazać, że „szczęście — jak głosi sportowa maksyma — zwykle się wyrównuje” (Dennett, 1984: 94–95), Sapolsky słusznie zauważa, że zdanie takie jest niezgodne z doświadczeniem. Jest raczej tak,

tym sugestywny, gdyż w formie uogólnionej może być potraktowany jako opis codziennego doświadczenia, które uczy nas, że cokolwiek się wydarza ma swoją przyczynę i samo prowadzi do pojawienia się kolejnych skutków.

Oczywistość towarzysząca takiemu opisowi znika jednak w chwili, gdy wchodzimy głębiej w sens tego, co faktycznie próbuje nam powiedzieć biolog. Postaram się to pokazać w ramach krótkiego komentarza do przywołanego powyżej eksperymentu.

Zacznijmy od tego, że Sapolsky ma oczywiście rację, twierdząc, iż zmiana przeszłości i warunków życia osoby wpływa na to, kim ona ostatecznie się stanie. Jednak implikowane w przytoczonym eksperymencie stwierdzenie, że zamiana warunków życia Absolwenta i Sprzątacza mogłaby spowodować, iż zamienią się oni miejscami już tak jednoznaczna nie jest. Rozpatrzmy dwa przypadki.

PRZYPADEK A:

Zakres zmian, którym poddana byłaby każda z osób, jest całkowity, a determinacja, której podlegają, zupełna (tj. każdy element życia Sprzątacza zostaje wymieniony na element życia Absolwenta) — w takim przypadku nic się *de facto* nie zmieni ani w obserwowanej przez nas sytuacji, ani z perspektywy uczestniczących w niej osób. Sprzątacze i Absolwenci będą dokładnie tam, gdzie są obecnie i (co wydaje się zrozumiałe) żaden z nich nie będzie miał poczucia, że powinien być gdzie indziej albo że jest kimś innym. Wiedza taka mogłaby być dostępna jedynie zewnętrznemu obserwatorowi, który świadom zamiany mógłby czysto konceptualnie uchwycić fakt, że aktualny Sprzątacze to potencjalny Absolwent i *vice versa*. Samym zainteresowanym takiej wiedzy przypisać by się jednak nie dało¹⁴. Już zresztą samo twierdzenie, że doszło tutaj do jakiejś zamiany byłoby nie do końca słuszne.

że: „Mniejsza liczba wspomnianych korzyści uczyni twoje dorosłe życie bardziej stresującym, a to z kolei sprawi, że twój mózg będzie sobie gorzej radził z procesami takimi, jak wytrzymałość, kontrola emocjonalna, refleksja i poznanie... Pech nie zostanie zrekompensowany przez dobro. Zwykle przybiera na sile aż do momentu, gdy jesteś już poza polem gry, które wymaga wyrównania” (Sapolsky, 2024: 101).

¹⁴ Można oczywiście szukać w tej hipotetycznej sytuacji dylematu etycznego, analogicznego do tego, który analizuje Derek Parfit (por. Parfit, 2012: 243–244), omawiając przypadek teleportacji (tj. zniszczenia jednego zgrupowania cząstek i odtworzenia go w innym miejscu w dokładnie takiej samej konfiguracji), jednak analizy tego rodzaju zostawiam tutaj bez kontynuacji.

PRZYPADEK B:

Zakres zmian, którym poddana byłaby każda z osób, nie jest całkowity, np. możemy przyjąć, że determinacja zdarzeń nie oznacza, iż ten sam układ początkowy daje każdorazowo ten sam układ końcowy¹⁵; albo zmienimy jedynie te okoliczności życia Sprzątacza i Absolwenta, na które mamy wpływ; albo jedynie te, co do których wiemy, że są/mogą być istotne dla kształtu finalnej sytuacji, itp. W takim przypadku jednak nasze przewidywania odnośnie do zakresu i charakteru zmian sytuacji obu osób zaczynają być niepewne i poddane jedynie statystycznym prawidłowościom. Nie wiemy, ani czy zamienią się oni miejscami, ani jakie dokładnie zmiany w ich aktualnej sytuacji będziemy obserwować. Możemy oczywiście zakładać, że jakieś zmiany się dokonają, ale nie potrafimy przewidzieć jakie.

Dodajmy przy tym, że o ile przypadek A zakłada w zasadzie quasi-boski charakter wiedzy i kontroli nad okolicznościami, przez co pozostaje zupełnie abstrakcyjny i daleki od naszej rzeczywistości, o tyle przypadek B jest czymś, co można do pewnego stopnia uznać za faktycznie możliwe¹⁶ — ten zresztą wariant byłby właściwy samemu Sapolski'emu. Cóż jednak takie przedstawienie wnosi do naszej analizy pojęcia determinacji?

Otóż nieprzewidywalność efektów zmian, jakim poddane byłyby osoby w przypadku B, wynika z faktu, że w praktyce liczba czynników, które mogą wpływać na konkretne ludzkie działanie czy stan danego systemu, może być nieograniczona, a w efekcie również poznawczo nieuchwytna. Samo w sobie nie jest to oczywiście dla autora *Determined* ani zaskakujące, ani problematyczne. Jak słusznie stwierdza: „Nieprzewidywalność systemu nie oznacza, że jest zaczarowany” (Sapolsky, 2024: 179), a złożoność czy niepoznawalność nie podważa jego deterministycznego charakteru. Jednak w praktyce przyjęcie takiej optyki mogłoby oznaczać, że żadnego wydarzenia nie da się w sposób dostateczny i prawidłowy wyjaśnić, a przynajmniej nie jedynie opierając się na dostępnych poznaniu związkach przyczynowo-skutkowych. Jeżeli przy tym dodamy — za Elizabeth Anscombe — że wyjaśnienia zjawisk można szukać

¹⁵ Stanowisko takie uznaje zresztą sam Sapolsky, pisząc: „Współczesne poglądy na temat determinizmu muszą brać pod uwagę fakt, że pewne rodzaje przewidywalności okazują się niemożliwe (rozdziały 5 i 6), a niektóre aspekty wszechświata w rzeczywistości nie podlegają determinizmowi (rozdziały 9 i 10)” (Sapolsky, 2024: 24).

¹⁶ Nie możemy oczywiście nadpisać przeszłości, ale mając coraz większą wiedzę na temat prawideł rządzących światem, w coraz większym stopniu możemy projektować skutki naszych działań (zarówno indywidualnych, jak i na poziomie społecznym). Przykłady tego rodzaju inżynierii na poziomie genetycznym czy farmakologicznym rozważał np. Francis Fukuyama w *Końcu człowieka* (por. Fukuyama, 2004).

na co najmniej dwóch poziomach: a) badając łańcuchy determinant w porządku czasowym — od wydarzeń przeszłych do chwili obecnej, oraz b) na drodze posuwania się w głąb struktury zdarzeń — tj. od powierzchni aż po konfigurację jej najbardziej elementarnych części składowych (Anscombe, 2005: 84–85), to proponowany przez Sapolsky'ego obraz determinacji mocno się komplikuje. Przyjmując bowiem, że

[...] wydarzenia mające miejsce od sekundy do milionów lat wcześniej decydują o tym, czy twoje życie i miłość rozwijają się w pobliżu szemrzących strumieni czy może maszyn dławiących cię dymem sadzy. Czy podczas ceremonii ukończenia szkoły masz na sobie togę i biret, czy może wynosisz śmieci? Czy według innych zasługujesz na długie spełnione życie czy też długi wyrok więzienia? (Sapolsky, 2024: 460)

– przyjmujemy, że za każdym z niezliczonych aspektów składających się na „kształt” konkretnego wydarzenia stoi faktycznie niepoznawalna liczba ciągów przyczynowo-skutkowych, które sięgają zarówno w najodleglejszą przeszłość, jak i do najgłębszych struktur rzeczywistości. Kształt tych powiązań zależy przy tym od praw, które znamy jedynie częściowo i jedynie częściowo rozumiemy. Jeżeli zatem fizycy jako nierozwiązywalny uznają obecnie problem zachowania trzech wpływających na siebie ciał, to propozycja Sapolsky'ego jest w zasadzie propozycją spotęgowania tego problemu do niewyobrażalnych wręcz porcji.

Czy oznacza to, że taki obraz świata nie jest deterministyczny? Bynajmniej. Jak pisałem wyżej, Sapolsky ma rację, twierdząc, że nic nie stoi na przeszkodzie, aby oddziaływania tego typu (nawet jeżeli ich liczba byłaby faktycznie nieograniczona) postrzegać jako rządzone niezmiennymi prawami, które stopniowo będziemy odkrywać. Co więcej, przekonująco wskazuje (zob. rozdz. 6), że nawet chwilowy indeterminizm elementów systemu nie oznacza jeszcze wolnej woli, a jedynie chaotyczność/nieprzewidywalność zachowania tych elementów. Czy można jednak faktycznie wykluczyć, że w granicach tak złożonego systemu nie działają również prawa powodujące (choćby okazjonalne) wyłonienie się struktur emergentnych, w ramach których wyłączenie określonej grupy zjawisk spod działania powszechnego determinizmu dawałoby tym elementom jakąś formę swobody w kształtowaniu własnego zachowania?¹⁷

¹⁷ Sapolsky prawdopodobnie pominąłby milczeniem takie pytanie. Jak bowiem prześmiewczo pisze o filozofach podnoszących tego typu wątpliwości, ich argumentację można sprowadzić do trzech zdań: „a. Wow, doszło do wielu niesamowitych odkryć w dziedzinie neuronauki i wszystkie potwierdzają pogląd, że nasz świat jest deterministyczny. b. Niektóre z tych odkryć

V. MECHANIKA MYŚLENIA O DETERMINACJI

Ostatnim wątkiem, który chciałbym poruszyć w ramach tej dyskusji z Sapolskim, jest kwestia możliwości wyjaśniania działań człowieka w ramach propozycji formułowanych na kartach *Zdeterminowanego*.

I tak wydaje się, że przyjmowany przez biologa obraz rzeczywistości można określić jako mechanistyczno-redukcyjnistyczny. Sprowadza on bowiem poszczególne byty do odpowiedniej dla nich konfiguracji elementów składowych oraz mechaniki ich powiązań (por. Sapolsky, 2024: 153), a schemat ten (z drobnymi wyjątkami; por. rozdz. 5–6) stosuje do wyjaśniania każdego elementu świata. Na przykładzie omówionego wcześniej eksperymentu widać przy tym, że nie wymyka się mu również człowiek. Zarówno Absolwent, jak i Sprzątacze są koniec końców jedynie konfiguracjami bezosobowych sił i czynników współgrających ze sobą od początków wszechświata i żaden z nich nie ma w tej kwestii nic do powiedzenia. Ich „los” jest czymś, co się po prostu dzieje, a nie czymś, co zostało przez kogokolwiek spowodowane. Nasuwa się tutaj jednak zasadnicze pytanie: czy zachowanie człowieka faktycznie daje się sprowadzić do takiego mechanicznego powiązania przyczyn i skutków? Czy wyjaśnianie ono wszelkie jego istotne aspekty?¹⁸

Charles Taylor w tekście *How is mechanism conceivable* stwierdza, że „podstawą potocznych wyjaśnień życia jest obraz podmiotu działającego ze względu na cele, pragnienia, skłonności oraz różne rodzaje wrażliwości” (Taylor, 1985: 165). Jak przy tym zauważa, motywy tego typu wymykają się obiektywizacji i mogą zdecydowanie się różnić między sobą w zależności od kultury, w której wychowywał się człowiek, od sensu, jaki przypisuje sobie i innym, a nawet od tego, jakich słów używa do jego wyrażenia. W odróżnieniu od takich opisów wyjaśnienia mechanistyczne (a za takie Taylor uznaje również wyjaśnienia neurobiologiczne) „wymijają” niejako — zdaniem filozofa — tę niejednoznaczność. Chcą pokazać, że wielość postaci, jakie przybiera życie wewnętrzne człowieka, jest jedynie epifenomenem bądź efektem niezmiennych praw warunkowania, któremu jednostka podlega ze strony środowiska czy na mocy innych czynników. Stąd też konstatacja filozofa, że „zwyczajowy opis zachowania przedstawia

w dziedzinie neuronauki tak mocno podważają nasze poglądy dotyczące sprawczości, odpowiedzialności moralnej oraz kwestii zasługiwania na karę bądź nagrodę (ang. deservedness), że należałoby dojść do wniosku, iż wolna wola nie istnieje. c. E tam, wolna wola istnieje i już” (Sapolsky, 2024: 22).

¹⁸ Kwestii tej nie będę omawiał szczegółowo, a tym bardziej nie jestem w stanie jej tutaj rozstrzygać. Chciałbym jednak dać w niniejszym tekście wyraz jednemu z ciekawszych argumentów, które — moim zdaniem — można przywołać w kontekście tego problemu, a który wydaje się istotny z perspektywy prowadzonych w tej polemice rozważań.

je jako działanie, a mechanicyzm jest nim zainteresowany jako ruchem” (Taylor, 1985: 167).

Różnica między tymi rodzajami opisu ogniskuje się w perspektywie, z jakiej widziana jest aktywność człowieka. Wyjaśniając działania, opisujemy *de facto* postępowanie intencjonalnego podmiotu, a wyjaśniając ruch, wskazujemy na sekwencję następujących po sobie zdarzeń. Nie musi to oczywiście oznaczać niezgodności tych opisów, wręcz przeciwnie: „każdy akceptowalny opis [naukowy — M.U.] musi być skoordynowany z naszym opisem potocznym, aby «ocalić fenomen»” (Taylor, 1985: 174) będący przedmiotem tego opisu. Jednak w tym miejscu pojawiają się istotne dla naszych rozważań pytania.

Mianowicie: jeżeli wyjaśnienia mechaniczne mają być zgodne z wyjaśnieniami w kategoriach sensu, to w którym momencie opis fizycznego ruchu komponentów zaczyna przekładać się na opis warstwy przeżyciowej jednostki? I dalej: jeżeli — przyjmując optykę Sapolsky’ego — uważamy, że opis mechaniczny (tj. formułowany w języku oddziaływania między biologicznymi bądź fizycznymi elementami) jest właściwym i podstawowym opisem zdarzeń, to w jaki sposób tłumaczy on również stany mentalne towarzyszące działaniom?

Sapolsky może oczywiście odpowiedzieć na takie wątpliwości, wskazując na wyniki badań epigenetycznych. Podkreśla zresztą:

[...] całe twoje dzieciństwo — poczynając od tego, w jaki sposób zaopiekowano się tobą w kilka minut po urodzeniu — było uwarunkowane kulturowo, oraz przez czynniki ekologiczne, które przez wiele wieków wpływały na aspekty kulturowe rozwijane przez twoich przodków (Sapolsky, 2024: 12).

Uznaje tym samym społeczno-kulturowe otoczenie jednostki (np. charakter opieki sprawowanej nad dzieckiem) za czynnik wpływający na ekspozycję określonych genów, a tym samym znaczący w perspektywie możliwości wyjaśniania zachowania ludzkiego¹⁹. Odpowiedź taka nie dotyka jednak istoty problemu stawianego przez Taylora.

Kanadyjskiego filozofa interesuje bowiem nie tyle podstawa korelacji między czynnikami społeczno-kulturowymi a częstością występowania określonych zachowań czy reakcji organizmu (por. Sapolsky, 2024: 93–94), ile raczej faktyczna możliwość sprowadzenia opisu całego bogactwa stanów mentalnych jednostki do szeregu pobudzeń nerwowych czy układu elektronów w mózgu. Podkreśli wręcz, że mechanistyczną teorię zachowania wtedy dopiero będzie

¹⁹ Problemu tego dotyczy rozdział 3, pt. *Skąd się bierze zamiar?* (w org. *Where does intent come from?*).

można uznać za wiarygodną, gdy uwzględni ona sposoby generowania „kulturowo specyficznych słowników, w ramach których rozumiemy nasze działania i uczucia” (Taylor, 1985: 179).

Tak sformułowanego wymagania mechanistyczno-redukcyjny obraz człowieka postulowany przez Sapolsky’ego wydaje się (jak na razie?) nie spełniać.

IV. UWAGI KOŃCOWE

W ramach podsumowania tych rozważań chciałbym przypomnieć, że celem niniejszej polemiki nie było *de facto* obalenie inkompatybilizmu — ani w ogólności, ani w wersji proponowanej przez Roberta M. Sapolsky’ego. Było nim jedynie wzruszenie przekonania, które może rodzić się po lekturze *Determined*, że „jeśli zbierzemy razem wszystkie wyniki naukowe ze wszystkich odpowiednich dyscyplin naukowych, okaże się, że nie ma w nich miejsca na wolną wolę” (Sapolsky, 2024: 17). Innymi słowy, analizując sposoby rozumienia przez Sapolsky’ego terminów „wolna wola” czy „determinacja” oraz starając się pokazać problem mechanicznego tłumaczenia zachowań ludzkich, chciałem jedynie pokazać, że w nieprzerwanym i nieograniczonym pochodzie wzajemnie determinujących się czynników może pojawić się „szczelina”, która nie podważając sensu determinacji, której sama jest przecież produktem, pozwoli jednocześnie na chwilowe uchylenie nieuchronności przyczynowo-skutkowych związków.

Czy jednak udało mi się osiągnąć założony cel? Nie jestem pewny, nie mogę oprzeć się bowiem wrażeniu, że tekst ten — pomimo wszelkich moich starań — został przez Sapolsky’ego zdeterminowany.

BIBLIOGRAFIA

- Anscombe, G.E.M. (2005). *Human life, action and ethics: Essays by G.E.M. Anscombe*. Imprint Academic.
- Dennett, D.C. (1984). *Elbow room: The varieties of free will worth wanting*. Clarendon Press.
- Dennett, D.C. (2003). *Freedom evolves*. Viking.
- Fukuyama, F. (2004). *Koniec człowieka: Konsekwencje rewolucji biotechnologicznej*. (Przeł. B. Pietrzyk). Kraków: Znak.
- Gazzaniga, M.S. (2013). *Kto tu rządzi — ja czy mój mózg? Neuronauka a istnienie wolnej woli*. (Przeł. A. Nowak-Młynikowska). Sopot: Smak Słowa.
- Harris, S. (2012). *Free will*. New York: Free Press.
- Mele, A.R. (2014). *Free: Why science hasn't disproved free will*. Oxford: Oxford University Press.

- Parfit, D. (2012). *Racje i osoby*. (Przeł. W.M. Hensel i M. Warchala). Warszawa: Wydawnictwo Naukowe PWN.
- Pink, T. (2004). *Free will: A very short introduction*. Oxford: Oxford University Press.
- Sapolsky, R.M. (2023). *Determined: A science of life without free will*. New York: Penguin Books.
- Sapolsky, R.M. (2024). *Zdeterminowany. Jak nauka tłumaczy brak wolnej woli*. (Przeł. M. Popławska). Poznań: Media Rodzina.
- Taylor, C. (1985). *Human agency and language*. Cambridge: Cambridge University Press.