

TOMASZ GRZEGOREK *

Nie ma badań jakościowych

Do rozróżnienia na badania ilościowe i jakościowe przywiązuje się w psychologii bardzo dużą wagę. Są oddzielne podręczniki do metodologii, oddzielne kursy w programach studiów, wyodrębniło się środowisko badaczy uprawiających „jakościówkę”. Niektórzy mówią wręcz o dwóch osobnych psychologiach, które mają się wzajemnie dopełniać. Prezentowany tekst nie jest kolejną krytyką tak zwanych badań jakościowych z pozycji metodologicznego rygorysty. Broni tezy bardziej podstawowej, że samo dzielenie badań empirycznych na jakościowe i ilościowe nie ma sensu. Po pierwsze, nie istnieją ku temu wystarczające racje merytoryczne. Po drugie, lansowanie tego niejasnego podziału i przypisywanie mu istotnego znaczenia psuje psychologię.

1. Pojęcie badania jakościowego

Wiadomo, że każda dyscyplina naukowa ma własne pojęcia i metody, lecz wszystkie są zbudowane na fundamencie wspólnych reguł racjonalności. To ze względu na przestrzeganie tych podstawowych reguł bardzo różne przedsięwzięcia nazywa się naukowymi i odróżnia od artystycznych, politycznych albo religijnych. Prawie cały dwudziesty wiek filozofowie próbowali skodyfikować ogólną metodologię nauki, ale problematyczny okazał się już wybór drogi do celu: czy wywodzić ją z samego rozumu, kierując się aprioryczną epistemologią i logiką, czy rekonstruować na podstawie faktycznego postępowania uczonych, przecież zdumiewająco owocnego. Żadna z tych dróg, ani normatywna, ani deskryptywna, nie doprowadziła choćby do zdefiniowania akceptowanego przez wszystkich kryterium demarkacji pozwalającego odróżniać naukę od tego, co nią nie jest. Pomimo braku konkluzji dyskusja była pożyteczna, ponieważ rozwiązała niektóre złudzenia i wiadomo już przynajmniej, co nie jest istotą nauki. Różnica zdań nie zamażała ogólnego poglądu, że nauka jest przedłużeniem zdrowego rozsądku, a badanie naukowe jest wyrafinowanym sposobem zdobywania wiedzy, które w codziennym życiu polega na tym samym: na szukaniu dowodów. Już starożytni uważali, że przekonania przyjętych bez uzasadnienia, choćby towarzyszyło im silne poczucie pewności, nie należy nazywać wiedzą. Wiedza naukowa to nic innego jak przekonania uzasadnione najlepiej. Nauka stawia najwyższe wymagania nie tylko potencjalnym dowodom, lecz również

* Dr Tomasz Grzegorek (tomasz.grzegorek@ignatianum.edu.pl), Uniwersytet Ignatianum w Krakowie

dowodzonym przekonaniom. Jeśli teza jest mglista i niezrozumiała, to nawet najbardziej wyrazisty i bezsporny wynik obserwacji nie będzie dowodem. Terminy naukowe są pedantycznie definiowane, żeby twierdzenia z nich budowane były maksymalnie jednoznaczne, choć naturalnie ich znaczenie może być nieuchwytnie dla laików. Ogólną zasadę racjonalności naukowej da się więc ująć krótko: zawsze wypowiadać się jasno i nigdy bezpodstawnie. O szczegóły można się spierać, kiedy wypowiedź jest sensowna, kiedy uzasadnienie uznać za wystarczające, czy wolno uzasadnioną tezę nazwać górnolotnie prawdziwą etc.

Użycie nauki do poznawania człowieka jest problematyczne i nie przypadkiem historycznie pierwszym tematem dociekań naukowych był odległy nieboskłon. Nie sprawia kłopotów ludzkie ciało, które podlega tym samym prawom, co reszta przyrody, lecz jakiś aspekt człowieczeństwa, który nie mieści się w granicach skóry. Dyscypliny, które się tym aspektem zajmują, nazywane są społecznymi i należy do nich też psychologia. Nieadekwatność metod chemii czy biologii kompensują wymyślaniem własnych oryginalnych, jednak z ambicją pozostania w obrębie nauki. Wymaga to przestrzegania zasad racjonalności naukowej mimo odmiennego sposobu prowadzenia badań. Część innowacji jest wprowadzana pod nazwą „badania jakościowe”, „metody jakościowe”, „podejście jakościowe” lub „paradygmat jakościowy”. Użycie tego czy innego ogólnego szyldu dla różnych praktyk badawczych sugeruje, że coś je łączy. Niestety definicje wykraczające poza zdawkowe stwierdzenie, że wszystkie one mają charakter nieilościowy – co jest zwykłym *ignotum per ignotum* – są zawsze mętne¹. Juliet Corbin i Anselm Strauss piszą w podręczniku: „Badania jakościowe to forma badań, w której badacz lub wyznaczony współbadacz zbiera i interpretuje dane, co sprawia, że badacz staje się integralną częścią procesu badawczego, tak samo jak uczestnicy i dostarczane przez nich dane. Badania jakościowe wykorzystują otwartą i elastyczną strukturę, co stoi w sprzeczności z koncepcją rygoru, tak istotną w badaniach ilościowych” (2015, s. 29). Na podstawie tak sformułowanego kryterium trudno sklasyfikować nawet badanie Milgrama. Jak każdy nurt, również intelektualny zwany paradygmatem jakościowym, ma dwa brzegi: prawy postpozytywistyczny oraz lewy postmodernistyczny. Epigonom pozytywizmu zależy na usprawiedliwieniu własnych nieortodoksyjnych metod jako mimo wszystko naukowych. Postmoderniści zachwyceni ideą konstruktywizmu społecznego otwarcie kontestują nawet ideał bezinteresownego dążenia do prawdy. Na przykład Norman Denzin i Yvon-

¹ Czasem sobie przeczą: można przeczytać, że synonimem „jakościowe” jest „interpretatywne” (Juszczyk 2013, s. 10), a gdzie indziej „badacz jakościowy nie musi być interpretatywistą” (Konecki 2000, s. 23). Żeby nie brnąć w wieloznaczne przezwiska, jako kryteria wskazywane są cechy konkretne, lecz drugorzędne, np. sposób doboru próby i jej wielkość (Maison 2022). Prawie tak, jakby próba złożona ze studentów filologii kazała zaliczyć badanie do psychologii humanistycznej. Tu i ówdzie pojawia się w końcu ton rezygnacji, że „nie wolno nam dokonywać zbyt ostrego rozróżnienia badań ilościowych i jakościowych” (Silverman 2013, s. 37).

na Lincoln piętnują słowo „badanie” jako „nierozzerwalnie związane z europejskim kolonializmem i imperializmem” (2009, s. 19) i ogłaszają, że w metodach jakościowych chodzi o zaangażowanie „w obronę sprawiedliwości społecznej, równości, braku przemocy, pokoju i uniwersalnych praw człowieka” (s. 38). Tak radykalni bywają prawie wyłącznie pracownicy wydziałów socjologicznych, których absmak dla „imperialistycznej” nauki o dziwo nie skłania do rezygnacji z posad akademickich. Natomiast psychologowie przedstawiający się jako zwolennicy podejścia jakościowego przeważnie nie negują wartości eksperymentów laboratoryjnych i nie potępiają statystyki jako narzędzia przemocy intelektualnej. Twierdzą tylko, że istnieje alternatywny i równoprawny sposób uprawiania psychologii naukowej. Tak naprawdę jednak nie chodzi wyłącznie o sposób, lecz także o tematykę badań. W historii psychologii już nie raz otwierano nową szkołę, której specyfika sprowadzała się do oryginalnych zainteresowań, choćby psychologię transpersonalną przedstawianą wtedy jako „czwarta siła”². Tematem dominującym w czasopiśmie *Qualitative Psychology*, wydawanym pod patronatem Amerykańskiego Towarzystwa Psychologicznego, są doraźne problemy społeczne związane z imigracją, dyskryminacją i rolami płciowymi³. O wiele trudniej znaleźć raporty z zastosowania metod jakościowych w badaniu pamięci, inteligencji czy percepcji. Stosunkowo dużo tekstów poświęconych jest filozoficznym założeniom paradygmatu jakościowego, co wyróżnia to środowisko na tle innych badaczy zainteresowanych bardziej otaczającym światem niż własnym podejściem do świata.

Nie ma powszechnie obowiązującej ścisłej definicji metod jakościowych. Większość ogólnych charakterystyk zawiera określenia „idiograficzne”, „rozumiejące” i „fenomenologiczne”. Chyba takie skojarzenia ma przeciętny psycholog, który nie interesuje się sprawą bliżej, nie posługuje się skrótami QUAL i QUAN, nie wspomina z upodobaniem teorii ugruntowanej i nie dyskutuje o zastąpieniu rzetelności i trafności pojęciem wiarygodności (*truthworthiness*). Niestety sens trzech wyżej wymienionych określeń jest notorycznie zniekształcany.

1.1. Jakościowe, czyli idiograficzne?

Wprowadzony przez neokantystę Wilhelma Windelbanda podział dyscyplin akademickich na idiograficzne i nomotetyczne dotyczy celu prowadzenia badań. Celem każde-

² Podejście jakościowe doczekało się godności „piątej siły” w psychologii (Ponterotto 2002) i paru innych wstrząsających przydomków, np. „zmiana tektoniczna” (O'Neill 2002).

³ Wydawany od 2013 roku periodyk jest organem Towarzystwa Badań Jakościowych w Psychologii należącym do oddziału 5. APA. „*Qualitative Psychology* publikuje badania reprezentujące szeroką gamę podejść metodologicznych, w tym narrację, analizę dyskursu, historię życia, fenomenologię, etnografię, badania w działaniu i studium przypadku. Czasopismo zajmuje się ponadto dyskusjami na temat nauczania badań jakościowych i szkolenia badaczy jakościowych”. dostęp: <https://www.apa.org/pubs/journals/qua>

go przedsięwzięcia naukowego jest poznanie samej natury rzeczy, czyli wiedza ogólna o rodzajach nazywana po łacinie *scientia* (jako przekład greckiego *episteme*). Naukowiec geolog chce poznać naturę kwarcu jako rodzaju minerału, a nie tylko opisać właściwości bryłki leżącej przed nim na stole. Tak samo naukowiec psycholog chce odkrywać prawdy o ludzkim umyśle, a nie tylko o psychice tego czy innego osobnika. Zaobserwowany akt zachowania traktuje zawsze jako przypadek typu aktywności, a konkretną osobę badaną jako przedstawiciela populacji, np. psychotyków w remisji albo ojców wychowujących samotnie dziecko. Nie ma czegoś takiego jak psychologia idiograficzna, podobnie jak nie ma biologii idiograficznej, mimo że każdy pies, a nawet każda mrówka jest inna. Jako typowa metoda idiograficzna psychologii wymieniane jest studium przypadku, a jako sztandarowy przykład zastosowania tej metody wskazywane są prace Zygmunta Freuda. Jest to przykład niefortunny, ponieważ Freud miał jednocześnie nastawienie idiograficzne i nomotetyczne. Analizował życie Idy Bauer po to, żeby zrozumieć tę jedyną i niepowtarzalną osobę, ponieważ chciał ją wyleczyć. Jednocześnie analizował jej życie, żeby poznać mechanizm powstawania nerwicy, nieważne czyjej. Po prostu robił naraz dwie rzeczy, które na ogół nie są łączone. Każda była nawet opłacana osobno: za pierwszą zapłacił ojciec Idy, za drugą wydawca publikacji „Bruchstück einer Hysterie – Analyse”. Jako psychoterapeuta Freud diagnozował konkretną osobę, a jako uczony formułował hipotezy na temat działania umysłu w ogóle, czego przeciętny psychoterapeuta nie robi. Esej biograficzny bez prawdziwych nazwisk nie ma żadnego sensu, diagnoza kliniczna też musi być zaadresowana do konkretnej osoby. Natomiast Freud w swoich studiach przypadku bez szkody dla ich naukowej wartości zastępował nazwiska wymyślonymi pseudonimami, właśnie po to, żeby nadać im walor uniwersalności. Analiza przypadku „Dory” jest badaniem nomotetycznym, chociaż na bardzo małej próbie $n = 1$. Uogólnianie z małych prób jest problematyczne, ale nie zlekceważono z tego powodu ani odkrycia Broki po autopsji Louisa Leborgne'a, ani generalizacji Piageta z obserwacji trójki własnych dzieci. Badanie naukowe wychodzące od pytania ogólnego ma zawsze charakter nomotetyczny, niezależnie od procedury i wielkości próby. Anonsowanie w psychologii odrębnej grupy metod idiograficznych jest więc nieporozumieniem. Wywiad albo analiza dzienników mają charakter idiograficzny tylko wtedy, gdy celem jest zrozumienie unikalnej osoby jako *indywiduum* a nie jako *exemplum*. Taki jest horyzont zainteresowań historyka, dlatego pisze teksty z nazwami własnymi, zwłaszcza z nazwiskami ludzi. Owocem badań historycznych jednakże nie jest *scientia*, a nazywa się je naukowymi tylko w luźniejszym sensie, gdyż są metodyczne i dostarczają nietrivialnych informacji. Przyczyną całego nieporozumienia, jak to często bywa, jest utarty sposób mówienia, mianowicie stosowanie tego samego określenia „badanie psychologiczne” w odniesieniu do dwóch zupełnie różnych zajęć: dociekania naukowego oraz stawiania indywidualnej diagnozy. Mówi się o przełomowym „badaniu psycholo-

gicznym” na temat przyczyn nadpobudliwości, ale także mówi się, że nadpobudliwy uczeń potrzebuje „badania psychologicznego”. To drugie badanie ma charakter idiograficzny, ale nie ma nic wspólnego z działalnością naukową.

1.2. Jakościowe, czyli rozumiejące?

Wskazywanie na interpretację i szukanie sensu jako cechę wyróżniającą badania jakościowe jest niewiele bardziej uzasadnione niż łączenie ich z nastawieniem idiograficznym. Rozróżnienie na wyjaśnianie i rozumienie odzwierciedla metafizyczny podział na kulturę i naturę. Jak pisał Dilthey, naturę się wyjaśnia (*erklären*), a kulturę rozumie (*verstehen*). Wyjaśnianiem zajmuje się nauka, a rozumieniem humanistyka. Naturą albo przyrodą jest wszystko to, co człowiek zastał na tym świecie, a kulturą to, co wytworzył razem z innymi ludźmi, a więc język, technologie, obyczaje itd. Biorąc się za empiryczne badanie umysłu, psychologowie od początku mieli rozterkę, czy traktować go bardziej jak wytwór przyrody czy kultury. W drugim przypadku pytanie o wpływ kultury na myślenie i emocje staje się niedorzeczne, jak pytanie o wpływ temperatury na ocieplenie klimatu. Stosunek pomiędzy obiema rzeczywistościami jest trudny do pojęcia, ponieważ z jednej strony fakty kultury muszą być ucieleśnione, z drugiej strony zdają się być niezależne od rzeczywistości materialnej. Na przykład prawdziwość zdania „Ze ściany spadło godło” zależy od fizycznego biegu zdarzeń: czy pewien średniej wielkości obiekt uderzył w podłogę czy nie. Z drugiej strony jakby kompletnie od tego nie zależy, ponieważ nawet po zajściu tego zdarzenia fizycznego zdanie będzie fałszywe, jeśli społeczność nie traktuje wyjątkowo wizerunku ptaka z koroną na głowie. Można powiedzieć, że godło nie podlega prawu grawitacji i ogólnie fakty kultury nie wchodzą w zależności przyczynowe typowe dla świata przyrody. Podany wyżej opis tego, co się stało, zakłada rozpoznawalne obiektywne znaczenie pewnego obiektu. Dla psychologii ważne jest to, że coś znaczyć mogą nie tylko obiekty, ale również ludzkie zachowania. Poruszenie kończyn albo wyemitowanie dźwięku ma taki sam związek z aktem zachowania, jak zadrukowany papier z godłem państwowym: konwencjonalny. Ten sam gest może w jednej kulturze być zachowaniem agresywnym, a w innej przyjaznym, a nawet może wcale nie być odbierany jako działanie tylko przypadkowy ruch. Społeczeństwo może cieszyć się niskim poziomem agresji dzięki przyjętym regułom interpretacji zachowań, bez ingerencji przyczynowej w rodzaju powszechnej lobotomii. I na odwrót, po Freudzie, Foucaulcie i wymyśleniu pojęcia biernej agresji wiele dotąd niewinnych zachowań da się interpretować jako „przemocowe”, dlatego teraz jest na świecie obiektywnie więcej przemocy, chociaż niekoniecznie ludzie stali się gorsi.

Różnica między wyjaśnianiem a rozumieniem ludzkich zachowań nie polega na odseparowaniu beznamietnej analizy od wczuwania się w położenie bliźniego, nie chodzi o przeciwstawienie serca rozumowi. Problem dotyczy tego, czy opisywać i tłumaczyć

aktywność człowieka, uwzględniając jej sens czy nie. Jako rozwiązanie na samym początku Wilhelm Wundt zaproponował psychologię dwudzielną: jedna miała zajmować się tym, co przyrodnicze i mechaniczne (bez-sensowne), druga tym, co kulturowe. Granica była niełatwa do ustalenia, ponieważ nie przebiegała pomiędzy tematami, lecz je przecinała, np. pamięć ma aspekty dające się tłumaczyć mechanizmem, co pokazał Ebbinghaus oraz aspekty niemożliwe do wyjaśnienia bez uwzględnienia semantyki języka i znaczenia pamiętanych wydarzeń. Współcześnie problem jest formułowany jako pytanie o granice stosowalności metafory komputerowej, według której umysł jest rodzajem oprogramowania, czyli systemem czysto syntaktycznym. W takim czy innym kostiumie zagadnienie wyjaśniania i rozumienia od początku jest dla psychologii osiowe i sięga głębiej niż typologia danych. Możliwe jest badanie na danych jakościowych zmierzające do odkrycia przyczyny albo badanie z liczbowymi wynikami pomiarów zmierzające do zrozumienia jakiegoś zjawiska poprzez jego interpretację w ramach obiektywnej struktury znaczeń.

1.3. Jakościowe, czyli fenomenologiczne?

Słowo „fenomenologia” jest kojarzone najbardziej z Edmundem Husserlem, chociaż wymyślił je w połowie osiemnastego wieku Johann Heinrich Lambert przy okazji badań z zakresu optyki, żeby oddzielić rozważania o tym, jak rzeczy wyglądają, od opisu ich prawdziwych własności. Na przykład prosty kij zanurzony do połowy w wodzie jawi się jako krzywy. Takie zjawisko, nazywane z greckiego fenomenem (*φαινόμενον*), jest powtarzalne i podlega prawidłowościom, więc możliwa jest swego rodzaju nauka o fenomenach, czyli właśnie fenomenologia. Zjawisko nie jest tym samym co złudzenie, w końcu przymiotnik „prosty” też opisuje jedynie to, jak kij się jawi, a nie jaki jest: jawi się prosty w kontakcie z dłonią oraz oczom po wyciągnięciu z wody. Tylko niektóre zjawiska są uznawane za złudzenia i zawsze w wyniku porównania z innymi zjawiskami, nigdy z rzeczą samą w sobie. Ale trudno żyć z przeświadczeniem, że rzeczywistość kończy się na zjawiskach. Platon wierzył, że można wyjść z jaskini cieni i cieszyć się poznaniem prawdziwej rzeczywistości dzięki filozoficznej kontemplacji. Mniej więcej to samo powtórzył Husserl, którego fenomenologia jest osobliwą filozoficzną metodą poznawania samej istoty rzeczy (*εἶδος*). Przebieg badania fenomenologicznego tylko w pierwszej fazie przypomina psychologiczną introspekcję. Zadanie polega wtedy na opisanu tego, co jest w świadomości bezpośrednio dane, z pominięciem całej reszty. Fenomenologiczne *epoché*, to znaczy wzięcie w nawias wszystkiego, co wiadomo o świecie łącznie z przekonaniem o jego istnieniu, jest zbieżne z napiętnowaniem przez dawnych psychologów eksperymentalnych tak zwanego błędu bodźca. Popołnia go introspektor, kiedy charakteryzuje bodziec zamiast swojego doświadczenia wywołanego przez ten bodziec, np. opisuje jabłko zamiast swojego wrażenia wzrokowego wywołanego przez pokazane

jabłko. Błąd bodźca narusza zasadę metodologicznego solipsyzmu, żeby każdy stan umysłu określać definitywnie przez związki z innymi stanami organizmu, bez wspominania o obiektach w otoczeniu fizycznym czy społecznym. Jednak Husserla metoda analizowania fenomenów nie zatrzymuje się na pominięciu przedmiotu, lecz konsekwentnie pomija również empiryczny podmiot, żeby została sama treść doświadczenia. Psychologia rzecz jasna nie może tego zrobić, skoro ten właśnie podmiot ma badać i o nim zdobywać wiedzę. Świadomy podmiot doświadczeń w takiej mierze, w jakiej stanowi przedmiot wiedzy naukowej lub naiwnej samowiedzy, jest jedną z wielu rzeczy na tym świecie, dlatego zawieszenie sądu o istnieniu świata zawiesza też sąd o istnieniu świadomości. Końcowy wynik analizy fenomenologicznej jest prawdą o samym bycie, a nie o ludzkim umyśle, podobnie jak udowodnione twierdzenie w matematyce jest prawdą o liczbach, a nie o intelekcie matematyka. Według Husserla poznany byt istnieje wprawdzie tylko jako korelat świadomości, ale świadomości czystej, a nie tej empirycznej produkowanej przez mózgi niektórych kręgowców. Fizyka, biologia, psychologia zaczynają swoją działalność dopiero wtedy, kiedy odpowiednie dziedziny przedmiotowe są gotowe i tylko czekają na zbadanie. Fenomenologia jako najbardziej bezzakoładowe przedsięwzięcie poznawcze dostarcza podstaw każdej dyscyplinie naukowej, ponieważ ujawnia, jak jej dziedzina przedmiotowa jest konstruowana. Jednak nie można mówić z tej okazji o wykorzystaniu wyników analizy fenomenologicznej w nauce – czyli również w psychologii – ponieważ znajomość głębokich podstaw praktyki nie wpływa na samą praktykę. Dzieci też nie zmieniają sposobu oddychania, gdy dowiedzą się w szkole, jak w szczegółach przebiega ten proces i co od urodzenia tak naprawdę robią. Fenomenologia jako zuchwała misja dotarcia do fundamentu bytu jest bezużyteczna dla nauki. Psychologowie deklarujący afiliację z fenomenologią nie mają powodu powoływać się na Husserla, już prędzej na jego mentora z młodości Carla Stumpfa.

Swoim zainteresowaniem fenomenami psychologia wcale nie wyróżnia się na tle innych dyscyplin naukowych. Wszystkie zaczynają od tego, co się jawi w doświadczeniu – dlatego są nazywane empirycznymi – a różnią się sposobem opisywania. Fizyka czy biologia opisują doświadczenie w taki sposób, jakby nikt go nie posiadał i w rezultacie dostają świat niezawisły od podmiotu poznającego, podczas gdy psychologia konceptualizuje doświadczenie na odwrót – jakby nie było żadnego świata, tylko podmiot doświadczający. Na przykład, zdanie „Jabłko leży na stole” wyraża fakt fizyczny, natomiast zdanie „Widzę jabłko leżące na stole” fakt psychologiczny. Warunkiem prawdziwości pierwszego zdania jest realna obecność gdzieś jabłka i stołu, natomiast prawdziwość drugiego nie wymaga istnienia niczego poza podmiotem. Dlatego żaden aspekt treści indywidualnego doświadczenia nie stanowi logicznego kryterium odróżniania halucynacji od percepcji, można w treści znaleźć jedynie wskazówki. Opis samego doznania wzrokowego odbywa się bez charakterystyki fal elektromagnetycznych docierających do siatkówki oka,

opis panicznego strachu przed pajakiem nie zawiera prawdziwych danych na temat pająka. Abstrahowanie od źródła doświadczenia konstytuuje umysł jako swego rodzaju osobną rzeczywistość. Stumpf uważał, że dopóki psychologia tylko opisuje tę rzeczywistość, jest *de facto* fenomenologią, która według niego polega na relacjonowaniu fenomenów i na niczym więcej. W podobny sposób posługują się wyrażeniem „opis fenomenologiczny” współcześni psychologowie. Tego typu opis ma zdawać sprawę z indywidualnego doświadczenia po to tylko, żeby ukazać – jak to się mówi – świat osobistych przeżyć. Z definicji nie da się mieć przeżycia bez zdawania sobie z tego sprawy, a nie można zdawać sobie sprawy, że się coś przeżywa, nijak nie nazywając tego, co się przeżywa. Jeśli opis fenomenologiczny ma być wierną relacją z żywego doświadczenia, to jego forma językowa nie może daleko odbiegać od pierwszego spontanicznego opisu dla siebie. Frazeologia i stylistyka niosą informację o cechach doświadczenia i koekstensja słów nie wystarcza do zachowania prawdziwości opisu, skoro ktoś przestraszył się „obleśnego pająka”, a nie „kątnika domowego”. Z tego względu zbieranie i analizowanie rozbudowanych opisów fenomenologicznych przeradza się w tak zwaną psychologię narracyjną.

Stumpf uważał, że fenomenologia odgrywa wobec psychologii rolę propedeutyczną, jak zresztą wobec wszystkich nauk empirycznych, które równo zaczynają od danych doświadczenia. Ale nieuprzedzony opis fenomenów jest zaledwie wstępnym etapem procesu zdobywania wiedzy autentycznie naukowej. W toku rozwoju każda dyscyplina odchodzi od języka potocznego i wytwarza własną sztuczną terminologię do wyrażania coraz bardziej abstrakcyjnych pojęć zakotwiczonych w skomplikowanych konstrukcjach teoretycznych, a już nie bezpośrednio w empirii. Opis fenomenologiczny w żadnej nauce nie jest celem, na przykład celem psychologii jest zdaniem Stumpha wiedza o funkcjach psychicznych takich jak chcenie czy rozumowanie. Tymczasem niektóre teksty z obszaru psychologii narracyjnej wyglądają bardziej jak literacko słabe reportaże, ponieważ kończą na zreferowaniu, co ktoś myśli albo czuje. Badacza umysłu interesuje inteligencja osoby badanej, a nie jej mniemanie na temat własnej inteligencji (chyba że ono jakoś świadczy o inteligencji). Psychologia nie jest od tego, żeby prezentować subiektywność, lecz żeby ją obiektywizować.

Fenomenologia rozumiana skromnie jako opis tego, co się zjawia w świadomości, odgrywa ważną rolę nie tylko w tak zwanych badaniach jakościowych. Jest na przykład częścią zmatematyzowanej psychofizyki, gdzie pomiarom fizycznej własności bodźca towarzyszą raporty werbalne obserwatora interpretowane jako dane o intensywności subiektywnych wrażeń zmysłowych. Zresztą również w naukach przyrodniczych mierzenie czegoś jest odmianą standaryzowanej obserwacji, której przedmiotem jest zawsze zjawisko, a nie rzecz sama w sobie. W gruncie rzeczy zjawiskiem jest też bezpośredni wytwór procedury pomiarowej, czyli wskazanie miernika. Nic nie stoi na przeszkodzie,

żeby opis fenomenów zawierał liczebniki i dostarczał danych ilościowych. Z takiej perspektywy fenomenologia okazuje się wpisana w każde badanie empiryczne.

2. Pomiar naukowy i dwa rodzaje danych

Empiryczne badanie naukowe rozciąga się na trzy etapy: zbieranie danych, analizowanie ich i wyciąganie wniosku. Wyrażenie „wynik badania” oznacza wniosek sformułowany na podstawie rezultatów analizy danych. Między pierwszym i trzecim etapem wiele się dzieje, niektóre informacje zawarte w danych są wyciągane na pierwszy plan, inne spychane do kontekstu, a jeszcze inne zupełnie pomijane, za to pojawiają się informacje nieobecne w danych, pochodzące z dedukcji oraz uogólnień. Jeśli dystynkcja na badania jakościowe i ilościowe ma być fundamentalna, to jej kryterium musi odwoływać się do fundamentu, czyli do tego, co jest na samym początku. Wbrew bezmyślnie powtarzanym sformułowaniom na pewno nie chodzi tutaj o dwa rodzaje metod badawczych. Metody to narzędzia, więc mówienie o metodach jakościowych i ilościowych jest jak mówienie o garnkach warzywnych i mięsnych. Takie przymiotniki odnoszą się do składników, czyli w przypadku działalności naukowej do danych. Tylko dane mogą być jakościowe lub ilościowe. Na przykład wywiad nie jest żadną metodą jakościową, lecz po prostu metodą zbierania danych. Przeważnie danych jakościowych, ale czasem ilościowych, gdy indagowany podaje swój wiek, zarobki albo liczbę rodzeństwa.

Kryterium podziału danych na ilościowe i jakościowe również można sformułować prosto. Uczynienie rzeczywistości – najpierw będącej jedynie środowiskiem życia – przedmiotem namysłu wymaga jej wyartykułowania. Uprawianie nauki jest działaniem dyskursywnym, zaczyna się od zdań opisujących pojedyncze obserwacje i kończy na zdaniach teorii. Materiałem do budowania teorii naukowej nigdy nie są surowe fakty, lecz fakty już jakoś opisane, najlepiej precyzyjnie i bezstronnie. Skoro dane to opisy faktów, to różne opisy jednego faktu stanowią różne dane⁴. Należy przyjąć, że deskrypcje słowne to dane jakościowe, a deskrypcje liczbowe to dane ilościowe. Jednak nie samo użycie cyfr i numerów robi różnicę, w końcu każdy numer można zapisać słowami. Dana ma charakter ilościowy tylko wtedy, gdy za znakiem „5” albo „pięć” stoi prawdziwa liczba, inaczej pozostaje daną jakościową. Nauka wprowadza liczby do opisu faktów przy pomocy dwóch procedur: zliczania oraz mierzenia. Przedmiotem zliczania są obiekty, natomiast przedmiotem pomiaru jest zawsze wybrana cecha obiektu, która może mieć różną wielkość. Zliczanie jest banalne i na ogół nie wymaga specjalnych narzędzi, natomiast mierzenie jest czynnością wyrafinowaną i znakiem firmowym nauki.

⁴ Niefortunnie się składa, że słowo „dane” można czytać jak imiesłów albo jak rzeczownik, co prowadzi do nieporozumień. Tym, co dane (po angielsku *given*) są fakty, natomiast dane (*data*) to opisy faktów.

2.1. Klasyczna definicja pomiaru

Według klasycznej definicji, pomiar polega na odkrywaniu stosunków różnych wielkości tej samej cechy. Na przykład pomiar cechy zwanej długością polega na ustaleniu stosunku wielkości długości jednego przedmiotu do wielkości długości drugiego przedmiotu. Normalnie z wygody nie mówi się „wielkość długości przedmiotu”, tylko po prostu „długość przedmiotu”, ale trzeba pamiętać, że słowo „długość” oznacza rodzaj cechy, którą różne obiekty mogą posiadać w różnej ilości: obiekt długi ma jej więcej, a krótki mniej. Ustalony dzięki porównaniu stosunek jest liczbą rzeczywistą, stąd wynik mierzenia jest zawsze liczbą. Przy założeniu, że rzeczy w świecie i ich cechy istnieją niezależnie od człowieka, również relacje między różnymi wielkościami tej czy innej cechy – czyli liczby – istnieją niezależnie od człowieka. Mierzenie w takim razie jest odkrywaniem obiektywnie istniejących liczb. Żeby wyniki kolejnych pomiarów były *nomen omen* współmierne, wszystkie mierzone obiekty muszą być porównywane do tego samego jednego obiektu, pełniącego rolę wzorca albo standardu. Obiekt taki jest wybrany arbitralnie i ze względu na jedną cechę, np. pręt z Sèvres ma pewną masę i zapach, ale za istotną uznaje się tylko jego długość. Ta przypadkowa wielkość wyróżnionej cechy stanowi jednostkę pomiarową, dla której wymyślana jest nazwa, np. długość francuskiego pręta przezwano „metrem”. Mierzenie polega więc na porównaniu pod względem tej samej cechy mierzonego obiektu ze standardem, a wynik pomiaru jest liczbą jednostek, np. „2 metry” oznacza, że stosunek długości wybranego obiektu do długości standardu ma się jak 2 : 1⁵. Standardy się zmieniają, z powodu nietrwałości duże materialne artefakty są zastępowane procesami czy zdarzeniami w mikroświecie, które w naszych czasach da się obserwować w laboratorium. Obecnie pomiar długości polega na porównaniu tego rodzaju cechy wybranego przedmiotu z drogą przebytą w ułamku sekundy przez światło w próżni, więc pręt z Sèvres prawdopodobnie nie ma już dokładnie jednego metra. Oczywiście w codziennym życiu takie porównanie jest zapośredniczone przez kolejne mierniki. Zaufanie do plastikowej linijki przykładanej do mierzonego obiektu opiera się na zaufaniu do precyzyjnego skalibrowania maszyny, która tę linijkę wyprodukowała i nadrukowała podziałkę, z kolei wiarygodność tego skalibrowania opiera się na zaufaniu do maszyny produkującej godne zaufania maszyny do produkcji linijek i tak dalej. Pierwszym ogniwem w tym łańcuchu dziedziczenia autorytetu jest wynik laboratoryjnego doświadczenia ze światłem. Tak czy inaczej wynik pomiaru jest zawsze liczbą wyrażającą stosunek dwóch wielkości.

⁵ Międzynarodowe Biuro Miar i Wag definiuje pojęcie jednostki w systemie SI następująco „The value of a quantity is generally expressed as the product of a number and a unit. The unit is a particular example of the quantity concerned which is used as a reference, and the number is the ratio of the value of the quantity to the unit.” (BIPM 2019, s. 127).

W kontekście klasycznej definicji pomiaru podział danych naukowych na dwie klasy jest konsekwencją struktury rzeczywistości, która zwyczajnie ma aspekty ilościowe i jakościowe, np. masa jest atrybutem ilościowym jabłka, a smak atrybutem jakościowym. Natura cechy jest niezależna od obserwatora, który może ją tylko rozpoznać lub nie. Rozpoczęty przez Galileusza projekt matematyzacji przyrody polega między innymi na sukcesywnym odkrywaniu, że atrybuty na pierwszy rzut oka jakościowe są w istocie ilościowe, jak to się okazało w pewnym momencie z prędkością czy lepkością. Cecha jest ilościowa, kiedy różne jej wielkości wchodzą ze sobą w ścisłe stosunki. Trzeba to sprawdzać testem na konkatenację, która jest jakby empiryczną interpretacją operacji dodawania⁶. Z arytmetycznego punktu widzenia stosunek jest ilorazem, który jako odwrotność iloczynu można rozpisać w sumę. Zlewając do jednej miski dwie szklanki ciepłej wody, można udowodnić, że objętość jest cechą ilościową, a ciepło cechą jakościową: objętości się zsumują, temperatury nie. Z tego powodu mówi się, że cechy ilościowe są ekstensywne, a jakościowe nie są ekstensywne⁷. Temperaturę i niektóre inne atrybuty jakościowe można mierzyć wtórnie (*derived measurement*) dzięki wykrytym w eksperymentach prawom łączącym je z cechami ilościowymi. Termometr rtęciowy skonstruowano po obserwacji, że rtęć pod wpływem ciepła zwiększa regularnie swoją objętość. Kreski na szklanej rurce termometru wyznaczają różne poziomy rtęci, więc nadrukowana podziałka mierzy długość (tj. cechę słupka rtęci), z której dopiero można wywnioskować – biorąc za przesłankę odkryte prawa przyrody – wielkość temperatury (cechę powietrza). Żeby ułatwić zadanie użytkownikowi, gotowe wnioski są zapisane na szkle jako stopnie Celsjusza. Możliwość przeprowadzania tego rodzaju pośrednich pomiarów jeszcze nie pozwala sądzić, że wszystko da się zmierzyć. Dopóki nie ma empirycznego dowodu na ilościowy charakter danej cechy i nie są znane prawa łączące ją z inną cechą o udowodnionym charakterze ilościowym, dopóty uznaje się daną cechę za niemierzalną.

2.2. Reprezentacyjna definicja pomiaru

Kryterium podziału danych na ilościowe i jakościowe sięga głęboko do filozoficznego pytania o to, czym jest liczba. Za klasyczną definicją pomiaru stoi odpowiedź, że liczba to – jak pisał Newton – „nie tyle mnogość jednostek, co wyabstrahowany stosunek jednej wielkości do innej wielkości tego samego rodzaju” (Newton 1728 za: Michell 1993, s. 63). Liczba jest relacją i chociaż nie zajmuje określonego miejsca w przestrzeni, to jest faktem poznawalnym zmysłowo. Podobnie widać gołym okiem, że jeden dom

⁶ W sensie ścisłym dodawać można tylko liczby, ewentualnie symbole liczbowe, cukier do herbaty można co najwyżej dosypać.

⁷ Po angielsku mówi się wtedy *intensive*, ale przymiotnik „intensywny” ma już w polszczyźnie utarte inne znaczenie.

jest wyższy od drugiego, aczkolwiek samego faktu bycia wyższym od nie da się zlokalizować. Tak rozumiane liczby istnieją w świecie samodzielnie bez człowieka, zgodnie z wiarą pitagorejczyków przenikają rzeczywistość. Ale możliwa jest jeszcze inna odpowiedź, że liczby to abstrakcyjne idee nieobecne w doświadczeniu zmysłowym. Takie liczby wprowadza do świata człowiek, który wcześniej albo je wymyślił (psychologizm), albo uchwycił rozumem bytujące gdzieś w zaświatach (platonizm). Ma to oczywiste implikacje dla definicji pomiaru i podziału danych na ilościowe i jakościowe. Żadne napotymane w doświadczeniu cechy obiektów same w sobie nie są ilościowe, taki charakter nadaje im dopiero opis obserwatora. Bez człowieka wszystko jest jakościowe, różnica długości dwóch prętów ma taki sam charakter jak różnica smaku dwóch owoców. Pomiar nie może wtedy polegać na odkrywaniu liczb w obiektach, lecz na przypinaniu ich do obiektów. Po jednej stronie jest jakiś sformalizowany system symboliczny, a po drugiej stronie jakościowa rzeczywistość empiryczna. Akt pomiaru polega na przypisaniu konkretnego symbolu zaobserwowanej wielkości atrybutu. Wystarczy, że symbol nie jest wybrany z zestawu przypadkowo, lecz zgodnie z regułą. Reguły mogą być różne, najlepsze umożliwiają homomorficzne odwzorowanie całej struktury empirycznej w systemie symbolicznym. System pełni rolę skali pomiarowej, symbole są wartościami (stopniami) skali. Dyktując dozwolone operacje matematyczne na symbolach, Stanley Stevens pogrupował skale pomiarowe w cztery typy: nominalny, porządkowy, przedziałowy i stosunkowy. Symbole skali nominalnej są tylko nazwami (etykietami). Symbole skali porządkowej można przekształcać monotonicznie, dlatego są już nazywane stopniami. Na symbolach dwóch ostatnich skal wolno wykonywać operacje arytmetyczne, więc są uznawane za liczby. Warunkiem użycia skali stosunkowej jest możliwość zaobserwowania braku mierzonego atrybutu, który wtedy otrzymuje symbol zera bezwzględnego. Wiadomo, na czym polega brak długości, dlatego w pomiarach korzysta się ze skal stosunkowych, np. metrycznej. Natomiast trudno powiedzieć, kiedy nie ma wcale temperatury, dlatego w mierzeniu tego atrybutu korzysta się ze skal przedziałowych z zerem umownym, np. skali Celsjusza⁸. Umowność stopnia zerowego nie pozwala uznać żadnego dnia za dwa razy cieplejszy od innego dnia, tak jak można uznać pręt za dwa razy dłuższy od innego pręta. Warto zauważyć, że potocznie mówi się „mierzenie” tylko o czynnościach kończących się ogłoszeniem liczby, dlatego zamoczenie dłoni w wodzie jest zaledwie „sprawdzaniem” temperatury. Uznanie czterech systemów symbolicznych za typy skal pomiarowych (*scales of measurement*) zrywa ze zwyczajem językowym, ponie-

⁸ Zdefiniowanie temperatury przez ruch cząstek pozwala mówić sensownie o braku temperatury, mianowicie jako bezruchu. Ponieważ ruch jest względny, absolutny brak ruchu można stwierdzić tylko w próżni. W ten sposób wyznaczono zero skali Kelvina. Żeby podkreślić niearbitralność skal z zerem bezwzględnym, tradycyjnie ich jednostki noszą nazwy własne, często utworzone od nazwiska wynalazcy, np. wolty, paskale, kelwiny (ale nie celsjusze).

waż łączyć nazywanie miarą również przypisywanie etykiet nieliczbowych, np. diagnozowanie kogoś jako flegmatyka to inaczej pomiar temperamentu na skali nominalnej.

Reprezentacyjna definicja pomiaru zaciera różnicę między liczbą i znakiem do oznaczania liczby, czyli cyfrą lub wielocyfrowym numerem. Będąc entuzjastą operacjonizmu, Stevens używał dwuznacznego słowa „number” bez zajmowania stanowiska w kwestii filozoficznej, czy matematyka opisuje relacje między liczbami (realizm) czy jest grą symbolami (formalizm). Przyjęło się skale nominalne i porządkowe obdarzać epitetem „jakościowe”, a przedziałowe i stosunkowe nazywać „ilościowe”. Taki podział skal stanowi zarazem kryterium podziału danych na jakościowe i ilościowe, gdyż dane to wyniki pomiarów. Wykluczenie skal porządkowych z kategorii ilościowych może wydawać się dziwne z dwóch powodów. Pierwszy jest błahy i związany tylko ze sposobem mówienia. Na przykład nazwy stanowisk w administracji państwowej można potraktować jako stopnie władzy w skali o charakterze porządkowym, ponieważ nie ma pewności, że kolejne miejsca w hierarchii reprezentują zawsze dokładnie ten sam ubytek władzy. Niemniej wolno mówić po polsku o ilości władzy: pytać „ile” władzy ma urzędnik i czy ma jej „więcej” niż ktoś inny. Nazywanie różnicy pod względem ilości różnicą jakościową brzmi paradoksalnie⁹. Uznanie skali porządkowej za jakościową może się wydawać dziwne również dlatego, że jej stopnie bywają oznaczane numerami. Trzeba podkreślić, że zastąpienie słów oznaczających poziom wykształcenia „podstawowe”, „średnie”, „wyższe” arabskimi znakami „1”, „2”, „3” nie zmienia magicznie natury skali, jest to zabieg kosmetyczny i nie powinien wpływać na sposób analizowania danych. Dokładnie tak samo jest z kafeterią odpowiedzi typu „nie lubię 1-2-3-4-5 uwielbiam” i średnia arytmetyczna z kilku takich odpowiedzi w ankiecie ma tyle samo sensu co wyliczona w grupie średnia ze stopni wykształcenia. Jednak w niektórych sytuacjach arytmetyczne przekształcenia ponumerowanych stopni są użyteczne. Nauczyciele pomagają sobie w wystawianiu uczniowi końcowej oceny średnią arytmetyczną wyliczoną z poszczególnych ocen, które są w istocie jakościowe. Czytając wypracowanie polonista nie ma czego liczyć, przede wszystkim ocenia jakość stylu i treści, a potem wyraża tę ocenę za pomocą przymiotników „mierna” albo „dobra” stanowiących oficjalne stopnie skali. Cyfry są podstawiane za przymiotniki głównie po to, żeby zapisy ocen mieściły się w rozmaitych rubrykach. W szkolnictwie anglosaskim w tej samej roli występują litery alfabetu, ale nauczyciel też może zastąpić je na własny użytek cyframi i wyliczyć średnią, żeby jakoś zobiektywizować sobie obraz ucznia i postawić w końcu sprawiedliwą literę. W pewnych okolicznościach średnia arytmetyczna wyliczona nawet z numerów na plecach piłkarzy może okazać się przydatna. W nauce dopuszczalność takich działań jest dyskusyjna, ponieważ

⁹ Zasady poprawnej polszczyzny ograniczają użycie słowa „ilość” do rzeczowników niepoliczalnych. Najwyraźniej jest bliżej angielskiego „amount” niż „quantity”, dlatego w odniesieniu do skali pomiarowej lepszym przekładem „quantitative” jest przymiotnik „liczbowa” niż „ilościowa”.

celem nauki jest uzyskanie wiedzy o rzeczywistości. Skoro numery na koszulkach piłkarzy są tylko identyfikatorami i nie oznaczają ilości czegokolwiek, to wynik matematycznych operacji na takich numerach nie będzie mówił niczego o rzeczywistości, nie będzie stał za nim żaden fakt. Stevens uważał, że gdzie jak gdzie, ale w nauce nie należy wykonywać operacji na pseudoliczbach skali nominalnej lub porządkowej, na przykład dodawać do siebie, wyliczać w zbiorze średniej albo kalkulować dla dwóch zbiorów współczynnika korelacji.

Wyprowadzanie prawdziwych liczb z jakościowych relacji pomiędzy stopniami skali porządkowej jest możliwe, ale skomplikowane. Taką możliwość stwarza teoria pomiaru łącznego (*conjoint measurement*), która dostarcza formalnego uzasadnienia dla przeprowadzania pomiarów pierwotnych bez konkatenacji, a więc umożliwia mierzenie atrybutów nieekstensywnych (ale tylko na skali przedziałowej). Na przykład wielkości temperatury i ciśnienia gazu da się precyzyjnie ustalić poprzez obserwację ich łącznego wpływu na objętość gazu. Dzięki obserwacji pewnych jakościowych relacji łączących objętość ze zmianami temperatury i ciśnienia można skonstruować addytywne reprezentacje temperatury i ciśnienia bez korzystania z jakiejś wcześniejszej metody pomiaru objętości. Taką procedurę dedukowania liczby z samych relacji „mniejsze-niż”, „większe-niż” i „równe” można przeprowadzić na dowolnym tercie atrybutów.

3. Liczby w psychologii

Trudno sobie wyobrazić operację konkatenacji na czymś tak ulotnym jak stan psychiczny. Jako dowodzenie charakteru ilościowego musi to być czynność powtarzalna i intersubiektywnie kontrolowalna, tymczasem stany umysłowe są prywatne, np. doznanie smakowe jednego człowieka jest niedostępne dla innych. Nawet gdyby siłę takiego doznania dało się jakimś sposobem empirycznie sumować, to obserwowałaby to za każdym razem tylko jedna osoba, więc przeprowadzone doświadczenie nie byłoby żadnym dowodem. Ponadto psychologia nie dysponuje ścisłymi prawami psychofizycznymi, które mogłyby stanowić podstawę mierzenia atrybutów psychicznych pośrednio. Więc nawet stosowane w psychofizyce techniki ustalania wielkości subiektywnych wrażeń zmysłowych nie są pomiarami z prawdziwego zdarzenia, co autorytatywnie stwierdziła w 1940 roku tak zwana Komisja Fergusona. Można co najwyżej ustalać, które wrażenie jest silniejsze i w ten sposób uporządkować od najsłabszego do najsilniejszego, otrzymując na końcu np. skalę subiektywnej głośności podobną do skali twardości minerałów Mohsa, która jest zbudowana na prostej zasadzie, że twardszy kamień zostawia rysę na miększym. Krótko mówiąc, nie trzeba być romantykiem, aby powątpiewać w możliwość zmierzenia siły miłości, głębi mądrości, bólu tęsknoty i innych stanów ludzkiej duszy.

Żeby chwalić się umiejętnością mierzenia czegokolwiek, psychologia musiała przyjąć inną niż nauki przyrodnicze definicję pomiaru. Pomiar łączny został opracowany

późno i nie upowszechnił się z powodu wyzwań matematycznych towarzyszących jego przeprowadzaniu (teoria perspektywy Tversky'ego i Kahnemana jest jednym z bardzo niewielu przykładów zastosowania tego rodzaju pomiaru w naukach społecznych). Od połowy XX wieku podręczniki do metodologii badań psychologicznych oficjalnie podają, że pomiar to po prostu przypisywanie liczb według reguły. Taką regułą jest na przykład klucz zliczania punktów w teście psychologicznym. Wyniki ogólne testu to liczby nie-przypadkowo przypisywane ludziom, a ściślej różnym wielkościom tej samej cechy posiadanej przez różnych ludzi. Na przykład podany przez Davida Wechslera klucz punktowania odpowiedzi każe jednemu człowiekowi przypisać poziom inteligencji 96, a innemu 97. W interpretacji uzyskanych liczb jest zawsze doza swobody, ale same liczby są produktem sztywnej reguły. Dlatego test psychologiczny nazywany jest instrumentem pomiarowym, a wynik testu jest traktowany w psychologii jak wskazanie linijki czy woltomierza.

Reprezentacyjna teoria pomiaru pozwala twierdzić, że nie ma na świecie niczego, co byłoby zasadniczo niemierzalne, są tylko właściwości, których na razie nie opisujemy ilościowo, ponieważ nie potrafimy skonstruować miernika lub w opracowaniu takiego opisu nie mamy żadnego interesu. W pewnym momencie społeczeństwo kapitalistyczne zobaczyło interes akurat w tym, żeby opisywać liczbami różne zdolności istot ludzkich. Nie przypadkiem taki pomysł narodził się w ogarniętej rewolucją przemysłową Anglii i szybko został rozwinięty na wielką skalę w Stanach Zjednoczonych. Wprawdzie projektodawca psychologii naukowej Wilhelm Wundt stał na stanowisku, że większość problemów interesujących psychologów nie nadaje się na przedmiot badań eksperymentalnych i mierzył wyłącznie atrybuty fizyczne, najchętniej wpływ czasu (między bodźcem i reakcją). Był jednak w tej godnej szacunku powściągliwości osamotniony, jego uczeń McKeen Cattell już nie miał skrupułów. Zaproponował baterię testów obejmujących dziesięć zadań, spośród których dziewięć sensomotorycznych miało prowadzić do ustalenia jakiejś wielkości fizycznej: upływu czasu w sekundach (potrzebnego na nazwanie koloru lub na reakcję na dźwięk), siły w niutonometrach (zaciskania dłoni) albo długości w milimetrach (odcinków otrzymanych w bisekcji lub dystansu między uciskanymi punktami na skórze). Ostatnie zadanie w zestawie polegało na głośnym powtórzeniu wcześniej usłyszanego ciągu liter i w porównaniu do poprzednich miało stanowić „*more purely mental measurement*” (Cattell 1890, s. 374). Umieszczając na jednej liście zadanie pamięciowe z zadaniami podpadającymi pod twarde pomiar fizyczny, a w tytule artykułu wyrażenie „testy umysłowe” tuż obok słowa „pomiar”, Cattell wykreował skojarzenie, które w tamtym czasie nie było jeszcze oczywiste. Kilkanaście lat później Alfred Binet i Théodore Simon skonstruowali pierwszy test zdolności intelektualnych składający się z zadań podobnych do tego dziesiątego, ale nie przedstawili go jako narzędzia pomiarowego. Miał być wyłącznie pomocą w diagnozowaniu i klasyfikowaniu

dzieci pod kątem wymagań francuskiego szkolnictwa, początkowo nawet bez możliwości precyzyjnego policzenia wyniku ogólnego. Optymizm w sprawie mierzenia umysłu zaczął uskrzydlać psychologię, po tym jak Henry Goddard przetłumaczył test na angielski, a William Stern zaproponował wyrażanie wyniku w postaci ilorazu inteligencji. Od tej pory sposób użycia testu psychologicznego zaczął przypominać korzystanie z miernika fizycznego, psycholog wkładał badanemu arkusz testowy do ręki niczym termometr pod pachę, a potem stwierdzał dwu lub trzycyfrowy IQ jak liczbę stopni Celsjusza. Simon nazwał pomysł kalkulowania ilorazu inteligencji i przypisywania ludziom liczbowych wyników zdradą ideałów Bineta, który już tego nie doczekał. Pragmatyczni Amerykanie szybko rozszerzyli kwantyfikację na sferę motywacji i osobowości. Dzisiaj porównanie człowieka do chodzącego komputera przestaje być traktowane z przymrużeniem oka i myślenie o ludzkich cechach jak o ilościowych parametrach wydaje się wszystkim naturalne. Dzieje się to między innymi za sprawą przenikania żargonu naukowego do mowy potocznej, ludzie spontanicznie mówią o „poziomie” głupoty, „ładunku” emocji, stosując retorykę ilościową w wyrażaniu opinii („sześć na dziesięć”) itd.

Z perspektywy klasycznej definicji pomiaru wszystkie dane psychologiczne są jakościowe. Wyników testowania i ankietowania nie można traktować tak serio jak liczb odkrywanych w pomiarach przez fizyków. Są one rezultatem co najwyżej uporządkowania wielkości empirycznych, a nie wyrażają stosunków pomiędzy wielkościami. Psychologowie nagminnie lekceważą zakaz Stevensa, żeby nie wykonywać operacji arytmetycznych na takich danych. Argumentują, że w toku przekształceń informacje w nich zawarte nie muszą ulec inflacji i skala pomiarowa nie determinuje wyboru techniki statystycznej. W innym razie musieliby ograniczyć się w analizach prawie wyłącznie do technik nieparametrycznych. Jedyne prawdziwe liczby wykorzystywane w psychologii pochodzą bowiem ze zliczania częstości przypadków oraz z pomiarów fizycznych (fizjologicznych). Trudno uznać te liczby za dane psychologiczne w ścisłym tego słowa znaczeniu. Wyniki mierzenia czasu reakcji albo częstotliwości przebiegu fal mózgowych to dane czysto fizyczne, chociaż wykorzystywane w psychologii. Liczba milisekund albo herców staje się częścią wiedzy o umyśle dopiero w komplecie z odpowiednimi jakościowymi danymi o wykonaniu jakiegoś zadania albo o samopoczuciu, uzyskanymi poprzez obserwację albo wywiad. Liczba wyrażająca częstość też nie stanowi w ścisłym sensie danej psychologicznej. Zaobserwowany fakt musi zostać najpierw opisany, zanim zostanie policzony, dlatego liczbę w tabeli frekwencji należy uznać za predykat odnoszący się do pierwotnej danej psychologicznej, a nie za daną psychologiczną. Z pewnością obecność liczb wyrażających częstości nie czyni jeszcze badania ilościowym, w *Qualitative Psychology* zdarzają się całe tabelki z liczbami generowanymi przez specjalny software (NVivo, MaxQda), który „koduje” numerycznie materiał słowny oraz zlicza istotne słowa czy wątki w transkrypcji wypowiedzi osób badanych.

Zdecydowana większość artykułów publikowanych dzisiaj w czasopismach psychologicznych zawiera rozbudowaną część statystyczną. Niewątpliwie ma to związek z kompleksem niższości wobec nauk przyrodniczych i typową dla dyscyplin społecznych metodolatrią wyśmianą przez Richarda Feynmana w anegdocie o kulcie cargo¹⁰. Ale bezpośrednio jest spowodowane tym, że psychologowie chętniej obserwują niż eksperymentują, ponieważ to jest łatwiejsze. Nie przypadkiem rozwój nowoczesnej statystyki splótł się historycznie z rozpoczętym przez Francisa Galtona programem masowego gromadzenia faktów za pomocą kwestionariuszy i testów. Im bardziej wyrafinowane stawały się techniki analizy statystycznej, tym mniejszą wagę przywiązywano do wyposażenia laboratoriów, od których kiedyś zaczynało budowanie instytutów psychologii¹¹. Wybitni eksperymentatorzy jak Wertheimer, Tolman, Skinner korzystali co najwyżej z prostej statystyki opisowej, ponieważ uzyskiwane w bardzo przemyślany sposób surowe dane mówiły same za siebie. W kuchni naukowej obowiązuje ta sama zasada co w zwykłej: im gorszej jakości składniki, tym mocniej trzeba je przetwarzać. Chyba większość publikowanych dzisiaj raportów badawczych to opracowania „metodą korelacyjną” odpowiedzi na pytania zadawane najczęściej przez Internet. Ogromna część naukowej wiedzy o ludzkiej psychice zgromadzonej od początku dwudziestego wieku wisi na włosku umowy, żeby przyznawać w pewien sposób punkty w wymyślonych kwestionariuszach i o skonstruowanych w ten sposób liczbach mówić „poziom stabilności emocjonalnej”, „poziom empatii” itd.¹² To ma być twardy grunt faktów, na którym dumnie stoją spekulacje przedstawiane jako prawda o człowieku.

¹⁰ Biali zbudowali w dzikim terenie pasy do lądowania dla samolotów, które dostarczały im jedzenie. Tubylcy postanowili zbudować podobne pasy, żeby skłonić bogów do przysyłania także im żelaznych ptaków z jedzeniem. Oczywiście żaden nie przyfrunął. Biali ludzie to badacze z dyscyplin przyrodniczych, lądowisko to sformalizowana metodologia, jedzenie to prawdziwa wiedza, tubylcy wiadomo.

¹¹ Przykładowo laboratorium psychologiczne Uniwersytetu Cornell w 1900 roku miało na wyposażeniu 373 różne urządzenia, wiele w niejednym egzemplarzu: wahadło z kamertonem do demonstrowania echa („powidoku”) słuchowego, keratometr, dwuosiowy mikser kolorów, perymetr, chronoskop Hippa, tropostereoskop, kampimetr, tachioskop, estezjometr, dyski Meumanna do pomiaru poczucia czasu, algesjometr, episkotister Krilla, sfigmograf, aparat Leuba do wyznaczania progu percepcji figury na tle, olfaktometr, zwykłą wagę z odważnikami, połańdowane lustro i wiele innych (Titchener 1990).

¹² W obiegu akademickim jest lekko ponad tysiąc testów psychologicznych. Psychologowie mierzą znane ludzkie cechy oraz cechy ludzkie zwykłym ludziom nieznane. Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego oferuje około 180 profesjonalnych instrumentów do mierzenia m.in. ilości nadziei podstawowej (BHI-12), stopnia pozytywności stosunku do własnej osoby i innych ludzi (KU), nasilenia poczucia kontroli w pracy (CWP), natężenia spontanicznej gotowości do posługiwania się wymiarem płci w odniesieniu do siebie i świata (IPP), poziomu uogólnionej własnej skuteczności (GSES).

A przecież nie trzeba jak Kant widzieć w matematyzacji istoty naukowości, chociaż bezsprzecznie dzieje fizyki, chemii i biologii to obejmowanie siecią równań coraz szerszego zakresu zjawisk. Używanie liczb jest na pewno najlepszym sposobem na jasne i precyzyjne opisywanie rzeczywistości. Niemniej żaden kompletny wywód naukowy, z wyjątkiem dowodu matematycznego, nie kończy się podaniem liczby albo równania. W historii fizyki bywało tak, że równanie pojawiało się na papierze jakby samo, natomiast uczony, będąc wtedy bardziej świadkiem niż autorem, potrzebował trochę czasu, żeby je zrozumieć. Ostatnim zdaniem w nauce jest interpretacja, ponieważ teoria zawsze wychodzi poza dane. Interpretacja wcale nie jest specjalnością wyłącznie jakiejś metodologii jakościowej.

4. Podsumowanie

Cechą badań jakościowych nie może być nastawienie idiograficzne, jeśli mają być naukowe. Nie wyróżnia ich ani korzystanie z opisów fenomenologicznych, ani dążenie do rozumienia zamiast do wyjaśniania. Nie da się również uzasadnić podziału badań na jakościowe i ilościowe odmiennością metod. Funkcjonujący w dyskursie akademickim ogólny podział metod badawczych jest ufundowany na dychotomicznej klasyfikacji danych. W rzeczywistości jednak żadna metoda nie jest powiązana na wyłączność z danymi jednego rodzaju, np. wywiadem można zbierać dane ilościowe, a statystycznie analizować dane jakościowe. Co więcej, merytorycznie zasadny podział danych nie stosuje się do danych psychologicznych, ponieważ te mają charakter wyłącznie jakościowy. Psychologowie nie potrafią niczego zmierzyć w klasycznym sensie tego słowa, co na szczęście nie wyklucza ich jeszcze z grona naukowców. Decydujące jest przestrzeganie reguł racjonalności naukowej: formułowanie zrozumiałych tez i logiczna argumentacja powołująca się na rezultaty systematycznych badań. Tam gdzie zaczyna się wypowiadanie twierdzeń bez pokrycia, a tym bardziej wieloznaczne bajdurzenie, tam kończy się nauka. To się niestety zdarza w psychologii, a nieproporcjonalnie często w „jakościówce”. Nie wolno zapomnieć, że jedyny adekwatny komplement, jakim można obdarzyć teorię naukową, to nazwanie jej prawdziwą, a nie zdumiewająco oryginalną czy niesłychanie inspirującą. Odpowiedzialność za słowo to jedyne zobowiązanie uczonych wobec społeczeństwa, które finansuje ich działalność i obdarza nadzwyczajnym zaufaniem.

Psychologowie powinni porzucić kategorię „badania jakościowe”, tak jak biolodzy porzucili „robaki” (*vermes*). Obie kategorie są niepoważne, druga przydaje się w dyskusjach już tylko wędkarzom. Rzecz jasna, to nie mówienie jest fatalne, tylko sposób myślenia. Deklarowanie podejścia jakościowego, które może oznaczać bardzo różne rzeczy, zbyt często pełni rolę usprawiedliwienia marnej roboty badawczej i bywa legalnym sposobem wprowadzania ideologii na teren nauki. Zerwanie szyldu nie oznacza re-

zygnacji z zainteresowań, wszystko co teraz jest badane „jakościowo”, można po prostu porządnie badać, bez tej całej szkodliwej gadaniny o alternatywnym paradygmacie. Badania jakościowe nie istnieją.

Literatura

- Bickhard M. (1992). Misconceptions of Science in Contemporary psychology. *Theory and Psychology* 2(3): 321–337.
- Boer de J. (1995). On the History of Quantity Calculus and the International System. *Metrologia* 31: 405–429.
- Boring E. (1950). *A History of Experimental Psychology*. Prentice-Hall.
- Boring E. (1961). The Beginning and Growth of Measurement in Psychology. *Isis* 52(2): 238–257.
- Brown J. (2011). Likert items and scales of measurement? *Testing & Evaluation* 15(1): 10–4
- Bureau International des Poids et Mesures BIPM (2019). *The International System of Units (SI)*. 9th edition.
- Cattell J. McK. (1890). Mental tests and measurements. *Mind* 15: 373–381.
- Corbin J., A. Strauss (2015). *Basics of Qualitative Research*. Sage Publications Ltd.
- Danziger K. (1990). *Constructing the Subject*. Cambridge University Press.
- Denzin N., Y. Lincoln, (red.) (2009). *Metody badań jakościowych*. Tom 1. Warszawa: PWN
- Diez J. (1997). A Hundred Years of Numbers. *Studies in History and Philosophy of Science* 28(1): 167–185.
- Ferguson A., C. Myers, R. Bartlett, H. Banister, F. Bartlett, W. Brown, N. Campbell, K. Craik, J. Drever, J. Guild, R. Houstoun, J. Irwin, G. Kaye, S. Philpott, L. Richardson, J. Shaxby, T. Smith, R. Thouless, W. Tucker (1940). Quantitative estimates of sensory events: final report of the committee appointed to consider and report upon the possibility of quantitative estimates of sensory events. *Advancement of Science* 1: 331–349.
- Fraser C. (1980). Measurement in Psychology. *British Journal of Psychology* 71: 23–34.
- Gaito J. (1980). Measurement Scale and Statistics: Resurgence of an Old Misconception. *Psychological Bulletin* 87(3): 564–567.
- Husserl E. (2024). *Kryzys nauk europejskich a transcendentna fenomenologia*. Warszawa: PWN.
- Humphry S. (2013). Understanding measurement in light of its origins. *Frontiers in Psychology* 4(113): 1–8.
- Jovanović G. (2010). Historizing Epistemology in Psychology. *Integrative Psychological and Behavioral Science* 44(4): 310–328.
- Juszczyk S. (2013). *Badania jakościowe w naukach społecznych: szkice metodologiczne*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Kołakowski L. (1990). *Husserl i poszukiwanie pewności*. Warszawa: Aletheia.
- Konecki K. (2000). *Studia z metodologii badań jakościowych. Teoria ugruntowana*. Warszawa: PWN.
- Krasnodębski Z. (1986). *Rozumienie ludzkiego zachowania. Rozważania o filozoficznych podstawach nauk humanistycznych i społecznych*. Warszawa: PIW.
- Kyngdon A. (2011). Psychological Measurement Needs Units, Ratios, and Real Quantities: A Commentary on Humphry. *Measurement* 9: 55–58.

- Laudan L. (1998). Zgon problemu demarkacji. [w:] *Z badań nad prawdą, nauką i poznaniem*. Z. Muszyński (red.), Lublin: Wydawnictwo UMCS: 63–79.
- Lord F. (1953). On the Statistical Treatment of Football Numbers. *The American Psychologist* 8: 750–781.
- Maison D. (2022). *Jakościowe metody badań społecznych*. Warszawa: PWN.
- Maraun M. (2012). Validity and Measurement. *Measurement* 10: 80–83.
- Michell J. (1993). The Origins of the Representational Theory of Measurement: Helmholtz, Hölder, and Russell. *Studies in History and Philosophy of Science* 24(2): 185–206.
- Michell J. (2005). *Measurement in psychology*. Cambridge University Press.
- Michell J. (2008). Is Psychometrics Pathological Science? *Measurement* 6: 7–24.
- Newton I. (1728/1967). Universal arithmetic: or, a treatise of arithmetical composition and resolution, [w:] *The Mathematical Works of Isaac Newton*, Vol. 2. D. Whiteside (red.). New York: Johnson Reprint Corp.
- Nicolas S., B. Andrieu, J.-C. Croizet, R. Sanitioso, J. Burman (2013). Sick? Or slow? On the origins of intelligence as a psychological object. *Intelligence* 41: 699–711.
- O'Neill P. (2002). Tectonic change: The qualitative paradigm in psychology. *Canadian Psychology* 43: 190–194.
- Ponterotto J.G. (2002). Qualitative research methods: The fifth force in psychology. *The Counseling Psychologist* 30(3): 394–406.
- Putnam H. (1998). Znaczenie wyrazu „znaczenie”. [w:] H. Putnam, *Wiele twarzy realizmu i inne eseje*. Warszawa: PWN.
- Sady W. (2020). *Struktura rewolucji relatywistycznej i kwantowej w fizyce*. Kraków: Universitas.
- Shepard R. (1966). Matric Structures in Ordinal Data. *Journal of Mathematical Psychology* 3: 287–15.
- Silverman D. (2013). *Prowadzenie badań jakościowych*. Warszawa: PWN.
- Tal E. (2015). *Measurement in science*. Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/entries/measurement-science/>
- Titchener E. (1900). The Equipment of a Psychological Laboratory. *The American Journal of Psychology* 11(2): 251–65.
- Trendler G. (2009). Measurement Theory, Psychology and the Revolution That Cannot Happen. *Theory & Psychology* 19(5): 579–99.
- Windelband W. (2011). Normy i prawa przyrody. [w:] *Neokantyzm badeński i marburski. Antologia tekstów*. A. Noras, T. Kubalica (red.). Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Nie ma badań jakościowych

Niepotrzebnie dzieli się badania empiryczne na jakościowe i ilościowe. Nie istnieją ku temu wystarczające racje merytoryczne, a propagowanie tego niejasnego podziału potęguje w psychologii pojęciowy zamęt. Na dodatek powoływanie się na odrębną metodologię jakościową nierzadko staje się po prostu usprawiedliwieniem dla słabych badań. Określenie „jakościowe” należy odnosić tylko do danych, a nie do metod ich zbierania czy analizowania, ani tym bardziej do jakiegoś paradygmatu badawczego konkurencyjnego dla nomotetycznej nauki. Mówienie o badaniach jakościowych w opozycji do ilościowych jest nieporozumieniem również

dlatego, że w psychologii te ostatnie nie są wcale takie „ilościowe”, za jakie uchodzą.

Słowa kluczowe: paradygmat jakościowy, podejście fenomenologiczne, metody idiograficzne, pomiar psychologiczny

There are no qualitative studies

Dividing empirical research into qualitative and quantitative makes no sense. There are no sufficient substantive reasons for such a distinction, and promoting this vague division only deepens the conceptual confusion in psychology. Even worse, proclaiming a separate qualitative methodology is sometimes used to justify poor science. The term "qualitative" should refer only to the data, not to the methods of collecting or analyzing them, and certainly not to some overarching research paradigm that is supposedly an alternative to nomothetic science. Defining qualitative research in psychology by contrasting it with quantitative research also fails because the latter is not as "quantitative" as it is often perceived to be.

Key words: qualitative paradigm, phenomenological approach, idiographic methods, psychological measurement

