

nych. Howard Gardner (1992) przekonująco uzasadnia, że niebawem psychologia jako dziedzina wiedzy może przestać istnieć. Jego zdaniem pojawiające się nowe dyscypliny zajmujące się procesami kognitywnymi i procesami neurologicznymi wchłoną podejście poznawcze i teorie uczenia się. Interdyscyplinarna wiedza o zjawiskach kulturowych wchłonie podejście społeczno-kulturowe. Inne gałęzie psychologii (np. psychologia kliniczna, psychologia przemysłowa i konsumencka) będą naukami stosowanymi; badania psychoanalityczne i badania osobowości zostaną przeniesione do literatury i wiedzy humanistycznej.

Nie chcemy przyłączać się do przewidywań dotyczących przyszłości psychologii jako nauki. Jesteśmy jednak przekonane, że najlepszym sposobem zrozumienia każdego tematu jest ukazanie, jak rozmaite podejścia psychologiczne czy szkoły myślenia rozwiązują konkretne problemy społeczne. W ostatnim rozdziale pokażemy, jak można to uczynić. W dogłębnej analizie dwóch żywotnych problemów połączymy ze sobą wszystkie podejścia, starając się opisać całego słonia.

Jeśli jesteś gotowy dzielić z nami zainteresowanie badaniem problemów psychologicznych; jeśli lubisz tajemnice i chciałbyś wiedzieć nie tylko, kto zabił, ale również dlaczego; jeśli masz ochotę przemyśleć to, co myślisz... możesz przystąpić do czytania naszego podręcznika.

Badania nad zachowaniem człowieka

Wyobraź sobie, że właśnie urodziła ci się córeczka. Uważasz, że jest najpiękniejszym dzieckiem, jakie kiedykolwiek widziałeś, i z optymizmem patrzysz w jej przyszłość. W ciągu najbliższych tygodni twoja radość zaczyna jednak gasnąć. Mała nie zachowuje się normalnie: zamiast gaworzyć i uśmiechać się jak większość niemowląt, twoja córeczka sztywnieje i odwraca głowę, kiedy bierzesz ją na ręce. Z upływem miesięcy, gdy jest coraz starsza, twoje obawy rosną, dziecko w dalszym ciągu nie nawiązuje kontaktu, nie zaczyna mówić. Wykonuje natomiast dziwne i niepokojące czynności. Godzinami kiwa się w tył i w przód, bez celu wpatruje się przez palce w światło słoneczne i rytmicznie uderza głową o ścianę. Szukasz pomocy u lekarzy i psychologów, po czym poznajesz diagnozę: twoja córeczka jest dotknięta autyzmem, jednym z najgłębszych upośledzeń wieku dziecięcego.

W tej bolesnej sytuacji chciałbyś na pewno znać przyczynę choroby dziecka. Dowiedziałbyś się, że autyzm jest uważany za schorzenie neurologiczne i choć jego przyczyny nie są do końca znane, istnieje wiele prawdopodobnych hipotez. Gdyby jednak autystyczne dziecko urodziło ci się jeszcze niedawno temu, lekarze i psychiatrzy pogłębiliby twój smutek, twierdząc, że winę ponoszą rodzice, a zwłaszcza matka. Swego czasu wielu klinicystów ulegało wpływowi pism Brunona Bettelheima, słynnego psychoanalityka, który uważał, że autyzm spowodowany jest lękiem dziecka przed brakiem akceptacji i odrzuceniem przez chłodną matkę. Dowody Bettelheima były kruche. W swej książce *The Empty Fortress* (1967) opisał trzy przypadki dzieci autystycznych, których matki miały w przeszłości problemy psychiczne. Wspomniał również o trzydziestu siedmiu innych przypadkach, ale nie przedstawił faktów dotyczących któregośkolwiek z nich. Sława Bettelheima była jednak tak wielka, że jego opinie przyjmowano bezkrytycznie, mimo słabości argumentów.

Tymczasem niektórzy badacze zaczęli podważać teorię Bettelheima, że przyczyną autyzmu są problemy psychiczne matki. Postanowili dokładnie porównać rodziców dzieci autystycznych z tymi, którzy mają zdrowe dzieci. Zamiast, jak Bettelheim, opierać się na subiektywnych wrażeniach, posługiwali się standaryzowanymi testami psychologicznymi, a uzyskane informacje poddawali analizie statystycznej. Rezultaty były jednoznaczne: czy to w wyniku badań cech osobowości, czy stosunków małżeńskich bądź życia rodzinnego nie stwierdzono, średnio rzecz biorąc, aby rodzice dzieci autystycznych różnili się od rodziców dzieci normalnych (DeMyer 1975; Koegel i in. 1983). Bruno Bettelheim nie miał racji. Wskutek jego

poglądów tysiące rodziców, którym specjaliści kazali uwierzyć, że są odpowiedzialni za chorobę własnego dziecka, niepotrzebnie cierpiało z powodu poczucia winy i wyrzutów sumienia.

Postaramy się wykazać, dlaczego metody badawcze, o których mówimy w tym rozdziale, są tak istotne dla psychologów. Metody te są narzędziami rzemiosła psychologicznego. Dają możliwość uporządkowania sprzecznych poglądów. Pozwalają psychologom odrzucić niedorzeczne pytania (Na czym właściwie polega owo straszne zjawisko chłodnej matki?) i skierować uwagę na nowe problemy (Czy autyzm jest chorobą neurologiczną?). Dementują błędne hipotezy, które nieraz wyrządzają ogromną krzywdę. Jak dowiemy się z niektórych fragmentów książki, nowatorska albo szczególnie pomysłowa metoda może dostarczyć odpowiedzi na podstawowe pytania dotyczące ludzkich zachowań, które niegdyś uważano za niemożliwe do zbadania.

Wiemy, że studenci często się niecierpliwią, kiedy mają uczyć się o metodach badawczych. „Dlaczego nie przejść od razu do wyników?” dziwią się. Wprawdzie tymi, którzy używają metod badawczych, są naukowcy, jednak laicy również powinni je poznać, jeśli chcą krytycznie myśleć i trafnie oceniać odkrycia psychologii (lub innych nauk). Nieustannie jesteśmy zasypywani sprzecznymi opiniami o sprawach, które głęboko wpływają na nasze życie — są to na przykład wskazówki, jak zerwać ze złymi nawykami, jak opanować emocje, stosownie się ubierać, rozwiązywać konflikty, przezwyciężyć nieśmiałość, naprawić swoje życie uczuciowe albo zredukować stres. Nie wszystkie badania dotyczące tych zagadnień są prawidłowe, niektóre porady samozwańczych specjalistów od formowania osobowości opierają się — jak u Bettelheima — na wątpliwych dowodach albo w ogóle są nieuzasadnione.

Alfred Forsyth i współpracownicy wykazali, że studenci, którzy nie mają wystarczającej wiedzy o metodach badawczych, skłonni są akceptować badania w zależności od tego, na ile ich wyniki potwierdzają ich własne oczekiwania. Jeśli rezultaty są zgodne z oczekiwaniami, zazwyczaj przyjmują je z zaufaniem, nawet gdy metoda badań owego zaufania nie uzasadnia. Kiedy jednak potrafią już oceniać poszczególne warunki eksperymentu (np. dobór uczestników), przestają bezkrytycznie akceptować zastosowane procedury przed wyciągnięciem wniosków (Forsyth, Arpey, Stratton-Hess, 1992). Mamy nadzieję, że nasi czytelnicy, słysząc lub czytając coś na temat psychologii, będą również dzięki wskazówkom z niniejszego rozdziału zastanawiać się, w jaki sposób uzyskano informacje i jak interpretuje się wyniki.

CO STANOWI O NAUKOWOŚCI BADAŃ

Gdy mówimy o psychologach jako naukowcach, niekoniecznie mamy na myśli ludzi w białych fartuchach, posługujących się skomplikowanymi przyrządami i aparatami (choć niektórzy to czynią). O podejściu naukowym świadczy raczej stosunek do zagadnienia i sposób działania niż używana aparatura. Filozofowie i naukowcy napisali już wiele opasłych ksiąg na temat cech odróżniających naukę od innych źródeł poznania. Nie uwzględniając wszystkich, podajemy poniżej kilka najistotniejszych cech idealnego uczonego.

1. Precyzja. Naukowcy zazwyczaj rozpoczynają od **hipotezy**, czyli wypowiedzi, która jest próbą opisu lub wyjaśnienia zachowania. Początkowo hipoteza może być sformułowana bardzo ogólnie, na przykład: „W nieszczęściu szukamy towarzystwa”. Jednak przed wykonaniem jakichkolwiek badań hipotezę trzeba uściślić. Można ją zastąpić sformułowaniem: „Ludzie, którzy czują lęk w sytuacji zagrożenia, zwykle szukają innych osób, narażonych na podobne niebezpieczeństwo”.

Niektóre hipotezy wynikają z wcześniejszych odkryć lub obserwacji zjawisk. Inne buduje się na podstawie ogólnej **teorii**, czyli uporządkowanego systemu twierdzeń i zasad, które służą wyjaśnianiu pewnych wydarzeń i zależności pomiędzy nimi. Teoria naukowa nie jest zlepkiem domysłów i opinii (choć mówi się potocznie: „To tylko teoria”). Teorie, które mają zyskać akceptację większej części społeczności uczonych, muszą być zgodne z wieloma różnymi obserwacjami, sprzeczne zaś tylko z niektórymi (Stanovich 1992).

Hipoteza formułuje przypuszczenia, co się stanie w określonej sytuacji. Terminy wieloznaczne, takie jak lęk i sytuacja zagrożenia, precyzuje się za pomocą **definicji operacyjnych**, które określają, w jaki sposób dokonuje się obserwacji i pomiaru badanych zachowań lub sytuacji. Poziom lęku można na przykład określić za pomocą wyniku w kwestionariuszu do badania lęku, a sytuację zagrożenia uściślić jako zagrożenie wstrząsem elektrycznym. Hipoteza może mieć wówczas następującą postać: jeżeli u badanych zwiększamy poziom lęku, mówiąc, że grozi im wstrząs elektryczny, a następnie każemy im wybierać, czy chcą oczekiwać tego zdarzenia sami czy z tymi, którzy są w podobnej sytuacji, decyzja oczekiwania z innymi jest w tym wypadku bardziej prawdopodobna niż u osób, które nie odczuwają lęku. Założenie jest następnie weryfikowane za pomocą dokładnych i systematycznych procedur badawczych.

Natomiast pseudonaukowcy posługują się często nieściśłymi i wieloznacznymi pojęciami, a ich założenia są wręcz bezsensowne. Czego właściwie dowiadujesz się z horoskopu, który mówi, że dziś jest dobry dzień na załatwienie nie dokończonych spraw? Czyżby nie każdy dzień był dobry, aby wykreślić którąś pozycję z listy spraw do załatwienia?

2. Sceptycyzm. Naukowcy nie przyjmują poglądów na wiarę lub z powodu autorytetu. Ich mottem jest: „Udowodnij mi to”. Kilku największych odkryć naukowych dokonali ludzie, którzy odważyli się wątpić w to, co każdy uważał za prawdziwe: że Słońce obraca się wokół Ziemi, że chorobę można wyleczyć, przykładając do skóry pijawki, że szaleństwo jest oznaką opętania przez demony. W codziennym życiu uczonych sceptycyzm oznacza ostrożne podejście do twierdzeń naukowych zarówno najnowszych, jak i dawnych. Ostrożność musi być jednak zrównoważona otwartością na nowe idee i uzasadnienia. W przeciwnym razie badacz może w końcu stać się tak krótkowzroczny jak słynny fizyk lord Kelvin, który pod koniec XIX wieku twierdził podobno z wielkim przekonaniem, że radio nie ma przyszłości, promienie rentgenowskie to oszustwo, a cięższe od powietrza maszyny latające nie są możliwe.

3. Poleganie na dowodach empirycznych. W przeciwieństwie do dramatu i poezji naukowe teorie i hipotezy nie są oceniane pod względem ich wartości estetycznych lub rozrywkowych. Niektóre koncepcje mogą początkowo wydawać się fascynujące, ponieważ są wiarygodne, pomysłowe lub atrakcyjne. Niezależnie od tego, jak bardzo wydają się prawdziwe lub słuszne, i tak muszą zostać zweryfikowane, jeśli

hipoteza

Wypowiedź będąca próbą przewidywania lub wytłumaczenia zespołu zjawisk. Hipotezy naukowe określają zależności pomiędzy zdarzeniami lub zmiennymi i są potwierdzane lub odrzucane w wyniku badań empirycznych.

teoria

Uporządkowany system twierdzeń i zasad, które służą wyjaśnianiu określonego zespołu zjawisk i współzależności pomiędzy nimi.

definicja operacyjna

Ścisła definicja terminu w hipotezie, określająca sposób obserwacji zjawiska i jego pomiar.

chce się traktować je poważnie. Jak pisał noblista Peter Medawar (1979): „Siła przekonania, że hipoteza jest prawdziwa, nie ma wpływu na to, czy jest ona prawdziwa czy nie”. Ponadto, jak zauważyliśmy w rozdziale 1, prawdziwość koncepcji naukowej musi być stwierdzona empirycznie, tzn. musi się opierać na dokładnej i systematycznej obserwacji. Nie jest dowodem zbiór prywatnych poglądów lub anegdot ani odwołanie się do autorytetu.

4. Gotowość formułowania ryzykownych przewidywań. Naukowiec musi formułować koncepcję w taki sposób, aby można było ją odrzucić lub obalić za pomocą przeciwnych dowodów. Ta zasada, zwana **zasadą falsyfikacji**, nie oznacza, że teoria zostanie odrzucona, lecz jedynie, że może być odrzucona, jeśli zostaną odkryte pewne fakty. Inaczej mówiąc, naukowiec musi zakładać nie tylko to, co się zdarzy, ale i to, co się nie zdarzy. Gotowość formułowania „ryzykownych przewidywań” zmusza do poważnego traktowania dowodów negatywnych. Badacz, który wzbrania się iść do końca, aby nie ryzykować braku potwierdzenia, nie jest prawdziwym naukowcem.

Ta ostatnia cecha może wydawać się trochę niezrozumiała, posłużmy się zatem przykładem. Niektórzy ludzie wierzą, że można znaleźć podziemne źródła, spacerując z różdżką w dłoni, która zacznie się wówczas wyginać w kierunku ukrytej żyły wodnej. Jedni różdżkarze posługują się specjalną różdżką stalową, inni wolą tak prozaiczne przedmioty, jak wieszak na ubrania albo rozwidlona gałąź. Różdżkarze wierzą, że są psychicznie nastroszeni na obecność wody, co powoduje reakcję różdżki, czyli przyczynia się do jej mimowolnych ruchów w ich rękach. Iluzjonista i zawodowy demaskator James Randi (1982) przed laty rzucił wyzwanie tym przesładczeniom, przeprowadzając kontrolowane testy z zastosowaniem procedur naukowych, na które różdżkarze wyrazili zgodę. Randi w dalszym ciągu oferuje nagrodę w wysokości kilku tysięcy dolarów dla różdżkarza, który podważy jego przekonanie o pseudonaukowym charakterze różdżkarstwa. Dotąd nie stracił ani centa; różdżkarze niezmiennie osiągają wynik nie przekraczający stopnia prawdopodobieństwa. Mimo niepowodzeń rzadko tracą wiarę we własne zdolności. Winią raczej układ planet, plamy na Słońcu albo złe wibracje obserwatorów. Nie ma żadnego znaczenia, jak zakończy się pokaz sztuki różdżkarskiej, ponieważ sam różdżkarz z góry ma na wszystko wytłumaczenie.

Wielu krytyków psychoanalizy oskarża ją o podobne naruszanie zasady falsyfikacji. Jeżeli pacjent przywołuje pewne konflikty z dzieciństwa, które psychoanalitycy uważają za powszechne, stwierdzają, że potwierdza to ich teorię. Jeśli natomiast pacjent nie przypomina sobie konfliktów, mówią, że zostały one z pewnością wyparte (zepchnięte do podświadomości, tak że nie można ich pamiętać) lub badany nie przyznaje się do nich. Nie ma zatem dowodów, które mogłyby świadczyć przeciwko teorii, nie ma też możliwości obalenia jej, nawet jeśli jest fałszywa. Każda teoria, która rzekomo wyjaśnia wszystko, co tylko może się zdarzyć, jest nienaukowa.

Z pewnością znajdziesz wiele przykładów naruszania zasady falsyfikacji w środkach masowego przekazu i w codziennym życiu. FBI i policja nie potrafią na przykład wykazać domniemanej zbrodniczej działalności wyznawców szatana, co nie przeszkadza niektórym z nich wierzyć w istnienie takiego kultu (Hicks 1991). Oficerowie i inni, którzy wierzą, że kult szatana istnieje, twierdzą, że nie dziwi ich brak dowodów, ponieważ sataniści ukrywają swoje czyny, zjadając ciała, albo w inny sposób pozbywają się ich bez śladów. Brak dowodów uważany jest właśnie za

oznakę powodzenia satanistów. Zastanówmy się jednak nad tym twierdzeniem! Jeżeli brak dowodów ma być argumentem, to co miałyby być kontrargumentem?

5. Otwartość. Naukowcy muszą w każdej chwili umieć wyjaśnić, skąd zaczerpnęli swoje pomysły, jak je sprawdzali i jakie uzyskali rezultaty. Muszą wypowiadać się jasno i precyzyjnie, aby inni badacze mogli powtórzyć, czyli odtworzyć ich badania oraz zweryfikować wyniki.

Odtwarzanie jest ważnym elementem procesu naukowego, ponieważ czasami coś, co wydaje się niewiarygodnym osiągnięciem (np. zimna fuzja), okazuje się jedynie szczęśliwym przypadkiem. Wiele lat temu grupa badaczy zajmująca się bezkręgowcami należącymi do typu płazińców, nauczyła je reagować skurczem na błysk światła. Następnie zwierzęta uśmiercono, starto na papkę, którą karmiono drugie pokolenie. Jak doniesiono, kanibalistyczna dieta przyspieszała utrwalanie reakcji skurczu w następnym pokoleniu (McConnell 1962). Można się domyślać, że odkrycie to wywołało ogromne podniecenie. Jeżeli bezkręgowce mogą się uczyć szybciej, połykając cząstki pamięci swoich towarzyszy, to czy dzieli nas jeszcze wiele od pigulek na pamięć? Studenci wymyślali dowcipy o mądrze z profesorów, profesorowie dociepkowali o przeszczepianiu mózgow studentom. Okazało się jednak, że rezultaty trudno odtworzyć i nadzieje na pigułki poprawiające pamięć ostatecznie wygasły.

Czy psychologowie i inni naukowcy zawsze spełniają stawiane im wysokie wymagania? Oczywiście, że nie. Będąc ludźmi, czasami za bardzo ufają własnym doświadczeniom. Niekiedy sami się oszukują. Pozwalają, aby ambicje brały górę nad szczerością. Czasami niezbyt dokładnie sprawdzają własne teorie. Zawsze łatwiej być sceptycznym wobec cudzych koncepcji niż wobec ulubionych własnych idei. Nawet Albert Einstein ukrywał nieraz dane, które mogłyby podważać jego domysły.

Przywiązanie do jakiejś idei samo w sobie nie jest niczym złym. Pasja jest motorem postępu. Jak zauważyliśmy w rozdziale 1, entuzjazm motywuje do odważnego myślenia, obrony niepopularnych koncepcji i przeprowadzania szczegółowych badań, które są często niezbędne do poparcia nowej teorii. Pasja może jednak również zniekształcać prawdziwy obraz, w niektórych przypadkach prowadzi wręcz do podstępów i oszustwa. Dlatego nauka musi być pod społeczną kontrolą. Oczekuje się, że naukowcy będą dzielić z innymi swoje spostrzeżenia i działania, że udostępnią własne wyniki czasopismom naukowym i prześlą do oceny specjalistom, zanim zaczną je rozpowszechniać. Powinni nadto przekonać sceptycznych kolegów, że ich stanowisko jest dobrze uzasadnione. Środowisko naukowców — w naszym wypadku psychologów — działa jak jury, analizując i podważając dowody, przyjmując pewne punkty widzenia, a inne wyrzucając na śmietnik nauki. Taki publiczny proces nie jest rozwiązaniem doskonałym, ale przynajmniej zapewnia nauce wewnętrzny układ kontroli i równowagi. Ludzie nie zawsze są obiektywni, uczciwi, a nawet racjonalni, dlatego nauka zmusza ich do uzasadniania własnych poglądów.

zasada falsyfikacji

Zasada, według której teoria naukowa musi formułować założenia na tyle szczegółowe, aby możliwe było jej zaprzeczenie, to znaczy musi przewidywać nie tylko, co się stanie, ale i co się nie stanie.

Psychologowie posługują się różnymi metodami badawczymi, w zależności od rodzaju pytań, na które chcą znaleźć odpowiedź. Metody te nie wykluczają się na-

metody opisowe

Metody, które dostarczają opisów zachowań, natomiast nie zawsze wyjaśniają przyczyny.

wzajem. Jak detektyw wykorzystuje w swej pracy szkło powiększające, odciski palców oraz przesłuchiwanie podejrzanych, chcąc wykryć przestępcę, tak psychologowie na różnych etapach swoich poszukiwań sięgają również po rozmaite techniki badawcze.

Wiele metod psychologii ma charakter opisowy. **Metody opisowe** pozwalają badaczom opisywać i przewidywać zachowania, choć niekoniecznie wybrać jedno wyjaśnienie spośród wielu możliwych. Niektóre z tych metod używane są przede wszystkim przez klinicystów do opisywania zachowań jednostek. Inne stosowane są najczęściej przez badaczy do porównywania grup ludzkich i formułowania uogólnień na temat zachowań człowieka. Część metod wykorzystują zarówno jedni, jak i drudzy. Poniżej przedstawimy najbardziej znane metody opisowe. Czytając, możesz wypisać sobie na kartce zalety i wady każdej z nich. Kiedy skończysz czytać ten podrozdział i dwa następne, porównaj własny wykaz z zamieszczonym w tabeli 2.1.

Studium przypadku

Studium przypadku (inaczej historia przypadku) to szczegółowy opis konkretnej jednostki. Może opierać się na wnikliwej obserwacji lub na formalnych testach psychologicznych. Może zawierać informacje o dzieciństwie badanego, jego snach, marzeniach, doświadczeniach i nadziejach, czyli o wszystkim, co pomaga uzyskać wgląd w ludzkie zachowanie. Jak się przekonamy przy podejściu psychoanalitycznym, badania przypadków są przez klinicystów stosowane najczęściej. Czasami posługują się nimi jednak również teoretycy. Badania te są szczególnie cenne przy podejmowaniu nowych tematów. Wielu badaczy języka zaczynało od prowadzenia szczegółowego dziennika rozwoju mowy własnych dzieci. Studium przypadku może być bogatym źródłem hipotez dla przyszłych prac badawczych.

Badania przypadku ilustrują twierdzenia psychologii w sposób, jaki nie jest dostępny abstrakcyjnym uogólnieniom i chłodnym statystykom. Odtwarzają też dokładniejszy obraz jednostki, niż można to otrzymać, stosując inne metody. Często jednak badania przypadku zależą od ludzkiej pamięci przeszłych wydarzeń, a taka pamięć może być i selektywna, i niedokładna. Ponadto badania przypadku skoncentrowane są na jednostkach, dlatego ich użyteczność dla psychologów zmierzających do dokonywanych uogólnień ludzkiego zachowania jest ograniczona. Osoba stanowiąca obiekt badania nie musi przypominać większości innych osób, co do których badacz chce wyciągnąć wnioski (przypomnijmy, iż jednym z błędów Bettelheima było twierdzenie, że matki, którymi się zajmował, były reprezentatywne dla wszystkich rodziców dzieci autystycznych). Czasami nie wiadomo, na jakiej zasadzie wybrać jedną interpretację przypadku spośród wielu innych.

Historie przypadków mogą być przydatne, jeżeli względy praktyczne lub etyczne nie pozwalają na zbieranie informacji w inny sposób lub jeżeli nietypowe okoliczności rzucają światło na zagadnienie generalne. Biopsychologowie badają na przykład przypadki pacjentów z uszkodzeniem mózgu, aby lepiej poznać jego funkcjonowanie. Stwierdzają, że uszkodzenia mózgu mogą mieć bardzo swoiste skutki w zależności od miejsca uszkodzenia. Niektórzy pacjenci rozpoznają sztuczne wytwory, takie jak fotografie, narzędzia i książki, ale nie potrafią rozpoznać przedmiotów naturalnych. Inni rozpoznają większość przedmiotów naturalnych, ale nie rozróżniają zwierząt, owoców albo roślin (Damasio 1990). W jednym z badań (Cu-

belli 1991) dwóch pacjentów nie potrafiło prawidłowo napisać samogłosek w wyrazach, nie mieli natomiast kłopotów ze spółgłoskami!

Paradoksalnie, nietypowe przypadki mogą zatem przyczyniać się czasami do wyjaśnienia ogólnych zagadnień dotyczących czynności człowieka. Większość badań przypadków stanowi jednak raczej źródło hipotez niż ich weryfikacji. Należy być bardzo ostrożnym wobec poradników psychologicznych, które jako „dowody” przedstawiają jedynie świadectwa i barwne opisy przypadków.

Obserwacja

W badaniach obserwacyjnych badacz dokładnie i systematycznie obserwuje i rejestruje zachowanie, niczego nie narzucając badanym osobom. W przeciwieństwie do badań przypadków obserwacje obejmują zazwyczaj wiele różnych osób. Często obserwacja jest pierwszym etapem programu badawczego, umożliwia prawidłowy opis zachowania przed podjęciem prób jego wyjaśnienia.

Głównym celem obserwacji jest opis zachowania zachodzącego w naturalnym środowisku. Etnologowie, Jane Goodall oraz w późniejszych pracach Dian Fossey, posługiwały się tą metodą w badaniach nad małpami i innymi zwierzętami żyjącymi w lesie. Psychologowie przeprowadzają obserwacje we wszystkich miejscach, gdzie przebywają ludzie — w domu, na boisku, w szkołach i w biurach.

W jednym z badań wykorzystującym obserwacje psycholog społeczny wraz ze swymi studentami odwiedzał bary. Studenci chcieli się dowiedzieć, czy pijący w barach spożywają więcej alkoholu, kiedy są w grupie czy wtedy, kiedy są sami. Odwiedzili wszystkie trzydzieści dwa puby w średniej wielkości mieście, zamawiali piwo i zapisywali na serwetkach i kawałkach gazety, ile piją obserwowani goście. Stwierdzili, że pijący w towarzystwie spożywają więcej alkoholu niż samotni. Pijący w grupach nie pili szybciej, po prostu dłużej przesiadywali w barze (Sommer 1977).

Zauważmy, że studenci w swoich badaniach nie ograniczali się do wrażeń lub wspomnień, ile pili obserwowani ludzie. W trakcie obserwacji istotne jest obliczanie, szacowanie lub pomiar. Dzięki tym działaniom można ograniczyć skłonność większości obserwatorów do dostrzegania tylko tego, czego oczekują lub co chcą widzieć. Szczegółowa rejestracja zapewnia dokładność i pozwala na porównywanie obserwacji. Porównywanie jest konieczne, aby się upewnić, że obserwacja jest prawdziwa i zgodna ze spostrzeżeniami różnych osób. Zauważmy również, że badacze starali się nie zdradzić, czym się zajmują. Gdyby wkroczyli z kamerami wideo i ogłosili, że są studentami psychologii, obserwowane osoby mogłyby nie zachowywać się jak zazwyczaj. W innych badaniach obserwatorzy w ogóle pozostawali w ukryciu. Przy zachowaniu ostrożności obserwacja daje prawdziwy obraz badanych w ich normalnym otoczeniu społecznym. Nie mówi jednak, co powoduje dane zachowania. Wyniki uzyskane w barze niekoniecznie świadczą na przykład o tym, że będąc w grupie, ludzie zaczynają więcej pić. Możliwe, że przyłączają się do grupy, gdy mają ochotę się napić i uważają, że przyjemniej jest spędzać czas przy barze z innymi.

Czasami pożądane lub konieczne jest przeprowadzanie obserwacji w laboratorium zamiast w sytuacji naturalnej. W obserwacjach laboratoryjnych psycholog ma większą kontrolę. Może się posłużyć skomplikowanym wyposażeniem, jak również określić, ile osób będzie obserwowanych jednocześnie, może zapewnić dobre warunki obserwacji itd. Załóżmy na przykład, że chcemy się dowiedzieć,

jak zachowują się niemowlęta w różnym wieku, kiedy pozostają w obecności obcych. Można obserwować dzieci w żłobku, lecz większość z nich będzie już prawdopodobnie w wieku, w którym rozpoznaje swoich opiekunów. Można odwiedzać prywatne domy, ale to wymaga dużo czasu i nie jest wygodne. Rozwiązaniem jest zatem zaproszenie rodziców z dziećmi do laboratorium, obserwowanie ich razem przez jednostronne lustro, wprowadzenie obcej osoby, a po paru minutach wyproszenie rodziców. Można wówczas rejestrować oznaki niepokoju, interakcje z obcym i inne zachowania niemowląt. Stosując tę technikę, zauważymy, że podczas nieobecności rodziców najmłodsze niemowlęta pogodnie zajmują się czymkolwiek, dopiero dzieci mniej więcej od ośmiu miesięcy wybuchają płaczem lub wykazują inne oznaki tak zwanego w języku psychologii niepokoju rozłąki (Ainsworth 1979).

Jednym z ograniczeń obserwacji laboratoryjnej jest fakt, iż obecność badaczy i specjalistycznego sprzętu powoduje, że badani zachowują się inaczej, niż zachowywaliby się w swoim normalnym otoczeniu. Ponadto obserwacje laboratoryjne, podobnie jak przeprowadzane w sytuacjach naturalnych, są bardziej przydatne do opisu zachowań niż do ich wyjaśniania. Gdy obserwujemy dzieci reagujące sprzeciwem na wyjście rodziców z pomieszczenia, nie możemy stwierdzić z pewnością, dlaczego tak się zachowują. Czy dlatego, że przywiązały się do rodziców i są uzależnione od ich bliskości, czy też po prostu nauczyły się z doświadczenia, że płacz zwraca uwagę i wzbudza współczucie? Trudno odpowiedzieć na te pytania wyłącznie na podstawie obserwacji.

Testy

Testy psychologiczne stosowane są przez psychologów zajmujących się diagnozowaniem psychologicznym, czyli pomiarem i oceną cech osobowości, stanów emocjonalnych, uzdolnień, zainteresowań czy wartości. Typowe testy wymagają odpowiedzi na wiele pytań, pisemnych lub ustnych. Odpowiedzi mogą być sumowane, tworząc pojedynczy wynik liczbowy lub zestaw wyników, które mówią coś na temat badanej osoby. Niektóre testy badają przekonania, uczucia lub zachowania, których badany jest świadomy. Inne przeznaczone są do wykrywania nieświadomych uczuć lub motywów.

Zapewne prawie każdy poddawał się kiedyś testom psychologicznym, na przykład testowi na inteligencję, testowi osiągnięć lub testom przydatności zawodowej. Być może byłś poddawany testom, starając się o pracę, wstępując do wojska albo rozpoczynając psychoterapię. Setki testów psychologicznych lub „instrumentów diagnostycznych” opracowano dla pracowników przemysłu, szkolnictwa, nauki i służb socjalnych. Niektóre dotyczą jednostek, inne dużych grup. Pomiary pozwalają wyjaśnić różnice pomiędzy jednostkami, a także różnice w reakcjach niektórych jednostek w różnych okolicznościach lub w różnych okresach życia. Przyczyniają się do lepszego zrozumienia samego siebie, są pomocne w programach rozwoju osobowości, w badaniach naukowych zaś pozwalają sformułować ogólne prawidłowości ludzkich zachowań.

Dobrze skonstruowane testy psychologiczne mają znaczną przewagę nad zwykłą samooceną, ponieważ wiele osób ma zaburzony pogląd co do własnych zdolności i cech. Większość ludzi na przykład skłonna jest uważać siebie za „ponadprzeciętnych”, nawet jeżeli tacy nie są (Myers 1980). Testy mają także przewagę nad nieformalnymi opiniami, jakie wyrażamy o innych. W. Grant Dahlstrom (1993) za-

uważa, że w ocenie innych popełniamy często pewne podstawowe błędy. Jeśli na przykład dostrzega się u kogoś jakąś cechę negatywną, to na ogół zakłada się, że jest ich więcej — tendencja ta nazywana jest efektem negatywnego halo. Dahlstrom wspomina, że kiedy był terapeutą na oddziale psychiatrycznym dużego szpitala, została do niego skierowana przez lekarza domowego piętnastoletnia dziewczyna, nazywana Mary. Mary była biednie ubrana, nieatrakcyjna, zaniedbana i niekomunikatywna. Psychiatra, który przyjmował dziewczynę, stwierdził ponadto, że cierpi ona na schizofrenię i jest opóźniona w rozwoju, złośliwie określając ją jako przykład „nieszczęśnej protoplazmy”. Kiedy jednak Dahlstrom poddał Mary serii testów psychologicznych, przekonał się, że wcale nie jest schizofreniczką ani upośledzoną umysłowo; jej zahamowania wynikały stąd, że ojciec znęcał się nad nią emocjonalnie. Z powodu efektu negatywnego halo psychiatra błędnie ocenił stan umysłu Mary. Stopniowo pod wpływem terapii i wsparcia stan psychiczny Mary poprawił się, a kiedy Dahlstrom słyszał o niej ostatni raz, była wicedyrektorem miejscowego instytutu sztuki.

Kryterium sprawdzającym rzetelność testu jest pytanie, czy jest on **standaryzowany**, to znaczy, czy obowiązują jednolite zasady przy jego przeprowadzaniu i interpretacji. Nie byłoby przecież w porządku, gdybyśmy niektórym uczestnikom udzielili szczegółowych instrukcji i dali mnóstwo czasu, innym zaś jedynie mętne wskazówki i czas ograniczony. Przeprowadzający test muszą dokładnie wiedzieć, jak wyjaśnić zawarte w nim pytania, ile czasu trzeba pozostawić na odpowiedź i jakich użyć materiałów. Ocena wyników dokonywana jest zazwyczaj w odniesieniu do **norm**, czyli ustalonych standardów wykonania. Normy określa się na ogół, przeprowadzając test na dużej grupie osób mających podobne cechy jak osoby, dla których test jest przeznaczony. Normy mówią użytkownikom testu, które wartości ocenić wysoko, nisko lub średnio.

Skonstruowanie, przeprowadzenie i interpretacja testu wymagają specjalistycznego przygotowania. Po pierwsze, test musi być **rzetelny**, to znaczy musi zapewniać spójne wyniki badań przeprowadzanych w różnych miejscach i różnym czasie. Test przydatności zawodowej nie jest wiarygodny, jeżeli Tom dowiadyuje się, że mógłby być świetnym inżynierem, ale słabym dziennikarzem, natomiast po tygodniu rozwiązuje test ponownie i uzyskuje odwrotne wyniki. Ponadto test musi być **trafny**, to znaczy musi sprawdzać to, co zgodnie z założeniami przewidziane jest do sprawdzenia. Test twórczego myślenia nie jest trafny, jeżeli sprawdza na przykład płynność wypowiedzi.

Psychologowie kontrolują rzetelność testu różnymi sposobami. Badają na przykład tzw. stabilność bezwzględną (stałość) metody: rest-retest, testując dwukrotnie tę samą grupę, a następnie porównując statystycznie obydwa zbiory wyników. Jeśli test jest wiarygodny, indywidualne wyniki powinny być podobne w obydwu sesjach. Metoda ta ma jednak pewną wadę. Wykonując test po raz drugi, badani osiągają na ogół lepsze wyniki, ponieważ są już zaznajomieni z wymaganymi zadaniami i pytaniami testu. Innym sposobem jest określenie rzetelności form alternatywnych za pomocą przedstawienia różnych wersji tego samego testu tej samej grupie w dwóch sesjach. Zadania w obu wersjach są podobnie sformułowane, ale nie są identyczne pod względem treści. Wyniki uzyskane z użyciem tej metody nie poprawiają się wskutek zaznajomienia z pytaniami, choć i tak przeważnie za drugim razem są nieco lepsze, ponieważ badani znają już ogólne strategie i techniki, których się od nich oczekuje.

Trafność testu sprawdza się także na różne sposoby. Jednym z nich jest kon-

standaryzacja
W odniesieniu do konstrukcji testu jest to opracowanie jednolitych procedur przeprowadzania badania i oceny.

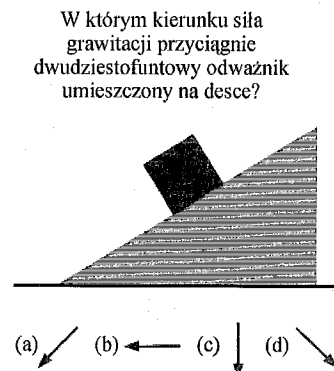
normy
W konstruowaniu testów ustalone standardy ich wykonania.

rzetelność
W odniesieniu do konstrukcji testu jest to spójność wyników uzyskanych w różnych miejscach i w różnym czasie.

trafność
Pewność, że test bada to, co zgodnie z założeniami powinien badać.

trolowanie trafności treściowej (wewnętrznej) — czy pytania są zdecydowanie reprezentatywne dla badanej cechy. Wyobraźmy sobie, że opracowujesz test, który ma służyć do oceny satysfakcji zawodowej pracowników. Jeżeli twój test obejmuje szeroki zakres istotnych przekonań i zachowań („Czy wydaje ci się, że osiągnąłeś w pracy granice możliwości?”, „Czy jesteś znudzony swoimi zadaniami?”) dotyczących rozmaitych aspektów satysfakcji zawodowej, oznacza to wyższą trafność treści. Jeśli natomiast zawiera jedynie pytania o satysfakcję z osiągniętej płacy, może nie być trafny i okazać się mało użyteczny. Pracownicy otrzymujący wysokie pensje nie zawsze są zadowoleni z pracy, natomiast osoby mało zarabiające nie zawsze odczuwają brak satysfakcji.

Większość testów sprawdza się również pod względem trafności kryterialnej, czyli możliwości potwierdzenia wyniku za pomocą innych niezależnych metod lub kryteriów w odniesieniu do analizowanej cechy. Test może być trafny kryterialnie, a tym samym użyteczny, nawet gdy niektóre zadania nie wydają się testowanym osobom trafne ani znaczące. Kryterium zdolności ucznia mogą być oceny szkolne, kryterium nieśmiałości — zachowania w sytuacjach społecznych. Chcąc sprawdzić, czy test satysfakcji zawodowej jest trafny pod względem kryteriów, można powtórzyć go po roku, aby przekonać się, czy właściwie przewidywał nieobecności w pracy, zwolnienia lub prośby o zmianę pracy. W ostatnich latach pracowano nad sformułowaniem bardziej adekwatnych testów przydatności zawodowej, wprowadzając pytania mające większe znaczenie dla predyspozycji wymaganych w danym zawodzie. Sprawdź na przykład, czy potrafisz odpowiedzieć na pytanie z testu dla strażaków (ryc. 2.1).



Ryc. 2.1

W procesie szacowania przydatności tej pozycji oceniający stwierdził, że choć studenci fizyki prawidłowo wybraliby odpowiedź (c), odpowiedź (a) może być bardziej użyteczna przy gaszeniu pożaru!

Niestety, nauczyciele, rodzice i pracodawcy nie zawsze są przekonani o trafności testów, zwłaszcza gdy ich wyniki podsumowuje się jedną liczbą, gdy idzie na przykład o iloraz inteligencji lub ocenę przydatności zawodowej. Zmyleni wynikiem sądzą po prostu, że test bada to, co im się wydaje. Robert Sternberg (1988) zauważa, że jest to domniemanie szczególnie powszechne w odniesieniu do testów sprawności umysłowej, choć, jak twierdzi, testy tego rodzaju sprawdzają jedynie ograniczony zakres predyspozycji mających znaczenie dla inteligentnego zachowania. „To magia ściśle brzmiących liczb — uważa Sternberg. — Iloraz inteligencji 119, współczynnik SAT 580, wynik sprawności umysłowej 74 procent — wszystko tak bardzo precyzyjne... Precyzja nie zastępuje jednak trafności”.

Kontrola trafności i rzetelności wykazuje, że niektóre nawet bardzo rozpowszechnione testy nie odpowiadają tym kryteriom. Potwierdza to historia wzlotu i upadku nowoczesnego urządzenia o nazwie poligraf, inaczej wykrywacz kłamstwa. Wynalazek poligrafu opierał się na przekonaniu, że osoba, która jest winna i odczuwa lęk, wykaże swoją winę przez objawy fizjologiczne — zmianę rytmu serca, przyspieszony oddech oraz zwiększone przewodnictwo elektryczne skóry — kiedy będzie podawać nieprawdziwe odpowiedzi na pytania. Testy z zastosowaniem wykrywacza kłamstw cieszyły się dużą popularnością, ponieważ organy rządowe, pracodawcy, współmałżonkowie i szpiegdy chcieli mieć niezawodny sposób poznawania prawdy. Do niedawna w wielu przedsiębiorstwach stosowano testy z wykrywaczem kłamstw jako jeden z elementów rekrutacji kandydatów do pracy. Niektórzy pracodawcy używają tego urządzenia rutynowo wobec pracowników podejrzanych o narkomanię lub złodziejstwo. W Stanach Zjednoczonych Senacka Komisja Pracy stwierdziła, że do roku 1988 przeprowadzano za pomocą poligrafu około dwóch milionów testów rocznie.

Kiedy jednak psychologowie poddali to urządzenie ocenie naukowej, okazało się, że wyniki testu nie spełniają kryterium trafności. Żadna z reakcji fizjologicznych nie jest bowiem specyficzna dla kłamstwa. Maszyny nie odpowiedzą nam na pytanie, czy czujemy się winni, rozdrażnieni, zdenerwowani, przestraszeni lub podnieceni po pełnym wrażeń dniu. Ktoś całkiem niewinny może silnie zareagować na słowo „bank” nie dlatego, że go obrabował, ale dlatego, że ma ostatnio debet na swoim koncie. Maszyna nigdy nie zarejestruje „kłamstwa”. Często popełnia natomiast przeciwny błąd. Niektórzy wytrawni kłamcy potrafią oszukiwać bez mrugnięcia okiem, inni starają się „przechytryć” maszynę, napinają mięśnie albo myślą o podniecających wydarzeniach, odpowiadając na obojętne pytania (Lykken 1981).

Test z zastosowaniem poligrafu nie jest też dostatecznie rzetelny; osoby, które przeprowadzają testy, często błędnie interpretują wyniki. Nie zawsze zgadzają się z osądami kogoś innego w odniesieniu do tych samych wyników i — co gorsza — są raczej skłonne niewinnego oskarżyć o kłamstwo, niż przyznać mu rację (Gale 1988; Klienmuntz, Szucko 1984). Ze względu na niedostateczną trafność i rzetelność badań za pomocą poligrafu Amerykańskie Towarzystwo Psychologiczne (APA) sprzeciwiło się stosowaniu wykrywaczy kłamstw i domaga się całkowitego zakazu ich używania. Kongres Stanów Zjednoczonych nie zastosował się do tego zalecenia, jednak w roku 1988 zabronił rutynowego wykorzystywania poligrafów przy testowaniu kandydatów do pracy lub do losowej kontroli pracowników. Jak stwierdził z oburzeniem jeden z senatorów: „Każdego roku około 320 000 uczciwych Amerykanów pomawianych jest o kłamstwo”.

Krytycyzm i weryfikacja testów psychologicznych zapewniają uczciwość i naukową dokładność charakterystyki psychologicznej. Psychologowie, którzy owe testy układają, oczekują, że rezultat ich pracy zostanie sprawdzony. Musimy natomiast zdawać sobie sprawę, że „testy psychologiczne” zamieszczane w różnych pismach i magazynach na ogół nie są kontrolowane pod względem trafności i rzetelności. Pytania często są poprzedzane chwytliwym nagłówkiem, żeby lepiej się sprzedać, np. „Oblicz swój współczynnik asertywności” albo „Czy jesteś samodestrukcyjna?” lub „Siedem rodzajów kochanka”. Tego rodzaju „testy” to nic innego jak lista pytań, które komuś wydały się atrakcyjne. Jest to, innymi słowy, pseudonauka.

Testy psychologiczne na ogół charakteryzują badanych tylko pośrednio. Kwestionariusze i wywiady, w których informacje uzyskuje się, pytając badanych wprost o ich doświadczenia, postawy i opinie nazywamy sondażami. Większość czytelników zapewne spotkała się z sondażami opinii publicznej, chociażby Instytutu Gallupa. Sondaże przeprowadza się na wiele tematów, poczynając od upodobań konsumenckich do preferencji seksualnych.

Sondaże dostarczają mnóstwa danych, ale nie jest je łatwo przygotować. Największa trudność to dobranie **próby**, czyli grupy badanych, która byłaby **reprezentatywna** dla większej populacji, którą badacz chce opisać. Założmy, że chcemy się czegoś dowiedzieć na temat używania narkotyków wśród studentów drugiego roku. Możemy przepytąć wszystkich studentów drugiego roku w całym kraju, ale to nie jest praktyczne; należy raczej dobrać próbę. Konieczne są przy tym specjalne techniki selekcji, które gwarantują, że próba będzie obejmowała kobiety, mężczyzn, czarnych, białych, biednych i bogatych, katolików, żydów itp., w takim stosunku, w jakim grupy te występują w całej populacji studentów drugiego roku.

Wielkość próby jest mniej istotna niż jej reprezentatywność; małe, lecz reprezentatywne próby mogą dostarczać nadzwyczaj dokładnych wyników. Przeciwnie, ankiety i sondaże opracowywane bez odpowiednich metod doboru próby dają na ogół rezultaty wątpliwe. Trudno mówić o naukowym sondażu, kiedy prezynterzy programu telewizyjnego proszą widzów, aby głosowali telefonicznie, udzielając odpowiedzi „tak” lub „nie” na kontrowersyjne pytanie. Wiadomo, że zadzwonią tylko osoby mające emocjonalny stosunek do poruszanych problemów (oczywiście, jeśli akurat oglądają program), a skoro są emocjonalnie zaangażowane, raczej będą wypowiadać się stronnictwo. Psychologowie i statystycy powiedzieliby w odniesieniu do takiej sytuacji, że próba wykazuje **błąd dobrowolności**. Zgłaszający się ochotniczo prawdopodobnie mają różną opinię od tych, którzy się nie wypowiadają.

Również popularne ankiety w kolorowych czasopismach — „Redbook”, „Cosmopolitan”, „Playboy”, „The Ladies’ Home Journal” — dotyczące zwyczajów i poglądów czytelników, obarczone są błędem dobrowolności. Czytelnicy, którzy decydują się wypełnić ankietę na temat seksu, bywają na ogół bardziej (może również mniej) aktywni seksualnie niż ci, którzy na ankietę nie odpowiadają. Ponadto osoby, które regularnie czytają magazyny, są zazwyczaj młodsze, lepiej wykształcone i zamożniejsze niż populacja Amerykanów jako całość. Cechy te mogą również wpływać na wyniki ankiet. Kiedy czytamy ankietę, zawsze powinniśmy pomyśleć, kto brał w niej udział. Stronnicza, niereprezentatywna próba nie zawsze świadczy, że ankietę jest bezwartościowa lub nieciekawa, oznacza jednak, że jej wyniki nie mogą być przenoszone na inne grupy.

Inny problem wynikający ze stosowania ankiet polega na tym, że ludzie czasami kłamią. Jest to prawdopodobne, zwłaszcza gdy ankietę dotyczy tzw. drażliwych tematów („Co? Ja miałbym postąpić tak obrzydliwie/nieuczciwie/brutalnie? Nigdy!”). Prawdopodobieństwo kłamstwa zmniejsza się, gdy respondenci mają zapewnioną anonimowość. Są też pewne metody kontrolowania kłamstwa — na przykład przez zadawanie tego samego pytania kilka razy na różne sposoby. Nie we wszystkich jednak ankietach wykorzystuje się tę technikę, zresztą nawet gdy ludzie nie kłamią umyślnie, mogą dobrze nie pamiętać wydarzeń z przeszłości albo źle zrozumieć pytania ankiety (Tanur 1992).

próba

Grupa osób wybranych spośród populacji do badań mających na celu określenie pewnych cech populacji.

próba reprezentatywna

Próba odpowiadająca populacji pod względem istotnych cech, takich jak wiek i płeć.

błąd dobrowolności

Nieadekwatność wyników pochodzących z próby obejmującej ochotników zamiast z próby reprezentatywnej.

Kiedy czytamy wyniki ankiety lub sondażu, powinniśmy zwrócić uwagę na sformułowanie pytań. Niejasne słowa i wyrażenia mogą mieć różne znaczenie dla różnych osób, na przykład, jeśli pytanie brzmi: „Jak wyglądało twoje małżeństwo w ciągu minionych lat?”, to co należy rozumieć przez „minione” — trzy lata czy może piętnaście? Co więcej, stronnicze ujęcie pytania może skłaniać badanych do odpowiadania w zasugerowany sposób. Weźmy na przykład pytanie: „Czy uważasz, że można zmniejszyć przestępczość, stosując karę dożywocia dla zbrodniarzy, którzy dokonali kilku rozbojów?” Takie sformułowanie wywoła prawdopodobnie inną odpowiedź, niż uzyskalibyśmy, proponując kilka możliwych rozwiązań problemu przestępczości, i zapytali, którą opcję badany popiera. Pytania dobrane tak, aby wyniki były zgodne z oczekiwaniami, spotykamy często w sondażach politycznych.

Ankiety mogą być bardzo dobrym źródłem informacji, jeśli są starannie przygotowane i właściwie zinterpretowane. Pozostaje jednak pewien często nie dostrzegany problem, na który powinniśmy zwrócić uwagę. Czasami sondaże nie tylko odzwierciedlają nasze postawy i zachowania, ale również wpływają na nie. Często chcemy być „jak wszyscy”; dowiadując się zatem z sondażu, że wobec jakichś zagadnień społecznych lub politycznych reprezentujemy pogląd mniejszości, popadamy w apatię, stwierdzając, że „nie ma sensu cegokolwiek zmieniać”. Czasami też postępujemy wbrew sobie albo zmieniamy poglądy, ponieważ „każdy” (według sondażu) postępuje lub uważa tak a tak, a my nie chcemy się odróżniać. Socjolog Elisabeth Noelle-Neumann (1984) twierdzi, że tendencja do nieświadomego dostosowywania własnych opinii i wartości do panujących trendów sprawia, iż „opinia publiczna” staje się potężnym narzędziem społecznej kontroli. Opisuje wybory w Niemczech, w których dwie partie miały takie same prognozy, dopóki sondaż nie wykazał lekkiej przewagi jednej z nich. „Właśnie wtedy, na samym końcu, ludzie dołączyli do wygranych — pisze. — Jak porwane z prądem, 3 do 4 procent głosów przepłynęło na stronę ogólnie oczekiwanego zwycięzcy”.

Zresztą nie tylko sondaże, ale każdy typ badań może wpływać na sposób, w jaki postrzegamy własne zachowanie. Jak byś zareagował, gdybyś w tej książce znalazł dowody na to, że reprezentujesz opinie i zwyczaje mniejszości?

Badanie korelacji

W badaniach opisowych często psychologów interesuje, czy dwa lub więcej zjawisk wykazuje jakiś związek ze sobą, a jeśli tak, to w jakim stopniu. Taką informację można uzyskać dzięki badaniom korelacyjnym. Jeżeli wśród studentów przeprowadza się ankietę, aby dowiedzieć się, ile godzin tygodniowo oglądają telewizję, nie jest to badanie korelacyjne. Z badaniem korelacyjnym mamy do czynienia wtedy, gdy szukamy zależności, na przykład pomiędzy godzinami spędzonymi przed telewizorem a średnią ocen.

Słowo **korelacja** jest często używane jako synonim „zależności”. Z technicznego punktu widzenia korelacja oznacza jednak miarę liczbową stopnia zależności pomiędzy dwoma czynnikami. Czynnikami są w tym przypadku zjawiska, liczby i wszystko, co można zmierzyć i oszacować. W badaniach psychologicznych takie czynniki nazywa się **zmiennymi**, ponieważ mogą zmieniać się w wymierny sposób. Wzrost, waga, wiek, dochód, poziom inteligencji, liczba słów powtórzonych w teście zapamiętywania, liczba uśmiechów w określonym czasie — wszystko, co da się zmierzyć, oszacować lub obliczyć, może stanowić zmienną.

korelacja

Miara zależności pomiędzy dwiema zmiennymi; wyraża się statystycznie przez współczynnik korelacji, który określa się za pomocą wartości od -1,00 do +1,00.

zmienne

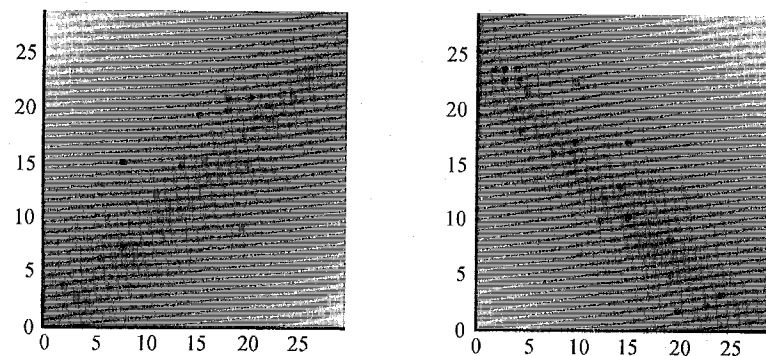
Cechy zachowania lub doświadczenia, które można zmierzyć lub opisać za pomocą skali liczbowej. Zmienne są przedmiotem manipulacji i obliczeń w badaniach naukowych.

Korelacje zachodzą zawsze pomiędzy zbiorami obserwacji. Czasami są to spostrzeżenia dotyczące jednej osoby. Wyobraźcie sobie, że kilka razy dziennie mierzycie badanemu temperaturę oraz koncentrację uwagi. Aby stwierdzić zależność pomiędzy temperaturą a koncentracją, potrzeba kilku pomiarów, czyli wartości, dla każdej zmiennej. Oczywiście wyniki będą odnosić się tylko do badanej osoby. W badaniach psychologicznych zbiory związanych ze sobą spostrzeżeń z reguły pochodzą od wielu osób lub są wykorzystywane do porównań grup badanych. W badaniach nad źródłami inteligencji psycholog szuka na przykład zależności pomiędzy współczynnikiem inteligencji rodziców i dzieci. W tym celu musi zebrać dane od grupy rodziców oraz dzieci każdej pary. Nie można określić korelacji, jeśli znamy tylko współczynniki inteligencji jednych rodziców i ich dziecka. Aby stwierdzić istnienie korelacji, potrzebujemy porównania więcej niż jednej pary wartości.

Korelacja dodatnia oznacza, że wysokie wartości jednej zmiennej są związane z wysokimi wartościami drugiej, natomiast niskie wartości jednej zmiennej są związane z niskimi wartościami drugiej. Korelacja dodatnia zachodzi na przykład pomiędzy wzrostem a wagą lub ilorazem inteligencji a ocenami szkolnymi. Korelacje rzadko kiedy bywają całkowite. Niektórzy wysocy ważą mniej od niektórych niskich, osoby ze średnim ilorazem inteligencji bywają prymusami w klasie, natomiast uczniowie o wysokim I.I. dostają słabe stopnie.

Rycina 2.2 przedstawia wykres ilustrujący potwierdzoną w badaniach dodatnią korelację pomiędzy poziomem wykształcenia a rocznym dochodem. Każda kropka przedstawia jedną osobę. Poziom wykształcenia danej osoby odczytujemy, przeprowadzając linię poziomą od kropki do osi pionowej. Dochody odczytujemy, wykreślając linię pionową od kropki do osi poziomej.

Korelacja ujemna oznacza, że wysoka wartość jednej zmiennej skojarzona jest z niską wartością drugiej. Rycina 2.3 ilustruje korelację ujemną pomiędzy średnim dochodem a występowaniem próchnicy zębów w grupach liczących po sto rodzin. Każda kropka przedstawia jedną grupę. Ogólnie, jak widać, im wyższe dochody, tym mniej kłopotów z zębami (Wright 1976).



Ryc. 2.2
Ryc. 2.3

W handlu samochodami im starszy wóz, tym niższa cena, z wyjątkiem modeli muzealnych i poszukiwanych przez kolekcjonerów. Jeśli chodzi o ludzi, ogólnie biorąc, im starsi dorośli, tym mniej kilometrów mogą przebiec, tym mniej przestępstw skłonni są popełnić i tym mniej włosów mają na głowie. Zastanów się, czy potrafisz znaleźć inne zmienne, pomiędzy którymi zachodzi korelacja ujemna. Pamiętaj jednak, że korelacja ujemna oznacza zawsze istnienie jakiegoś związku.

Jeśli nie ma związku pomiędzy dwiema zmiennymi, mówimy że nie są skorelowane. Rozmiar buta i iloraz inteligencji nie są skorelowane.

Wskaźnik statystyczny stosowany do wyrażenia korelacji nazywany jest współczynnikiem korelacji. Liczba ta wyraża zarówno wielkość korelacji, jak i jej kierunek. Całkowita korelacja dodatnia ma współczynniki $+1,00$. Całkowita korelacja ujemna ma współczynnik $-1,00$. Jeśli słyszymy, że korelacja pomiędzy dwiema zmiennymi wynosi $+0,80$, oznacza to, że bardzo silnie od siebie zależą. Jeśli słyszymy, że korelacja wynosi $-0,80$, zależność jest równie wysoka, lecz ujemna. Jeśli nie ma związku pomiędzy dwiema zmiennymi, to współczynnik wynosi zero lub jest bliski zero. Oto kilka przykładów korelacji dodatnich z codziennego życia (Kenny 1987):

niska (około 0,10)
stres-choroby somatyczne
iloraz inteligencji-oceny w szkole podstawowej

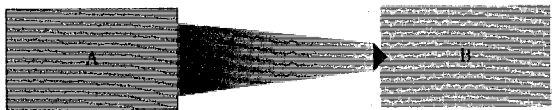
średnia (około 0,30)
poczucie własnej wartości-stopnie w szkole
podobne systemy wartości-atrakcyjność dla drugiej osoby

wysoka (około 0,50)
zadowolenie żony z małżeństwa-zadowolenie męża z małżeństwa
wiara w Boga-chodzenie do kościoła

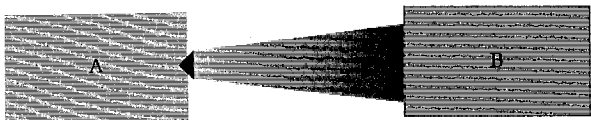
Korelacja pozwala badaczom stosującym metody statystyczne wyciągać ogólne wnioski na temat określonej zmiennej, jeśli wiedzą, jak jest ona związana z inną zmienną. Ponieważ jednak korelacja rzadko jest całkowita, dlatego wnioski w odniesieniu do konkretnej jednostki mogą być niesłuszne. Jeśli wiemy, że ktoś ma dobre wykształcenie, możemy wnioskować, nie mając innych informacji, że osoba ta dobrze zarabia, ponieważ wykształcenie i dochody wykazują korelację dodatnią. Nie możemy stwierdzić dokładnie, ile zarabia, lecz spodziewamy się, że jej dochody są stosunkowo wysokie. Możemy jednak się pomylić, ponieważ korelacja ta jest daleka od doskonałości. Niektórzy ludzie z doktoratami odbierają niskie pensje, bywają natomiast ludzie z wykształceniem podstawowym, którzy zbijają fortuny.

Badania korelacji w naukach społecznych są powszechne i często relacjonowane w mediach. Należy jednak zachować ostrożność, gdyż korelacje mogą być mylące. Przede wszystkim trzeba pamiętać, że korelacja nie wskazuje przyczyny. Skłonni jesteśmy sądzić, że jeśli A pozwala przewidzieć B, to A musi być przyczyną B, ale niekoniecznie tak jest. Liczba bocianów w niektórych wioskach europejskich wykazuje dodatnią korelację z liczbą urodzeń dzieci w tych wsiach. Innymi słowy, wiedząc, kiedy przylatują bociany, można przewidzieć, kiedy nastąpi więcej urodzeń niż zazwyczaj. Oczywiście nie oznacza to, że bociany przynoszą dzieci ani że dzieci przyciągają bociany. Liczba urodzeń wydaje się nieco większa w pewnych porach roku (możecie zastanowić się nad przyczynami), a jej wartość maksymalna przypada właśnie na okres przylotu bocianów.

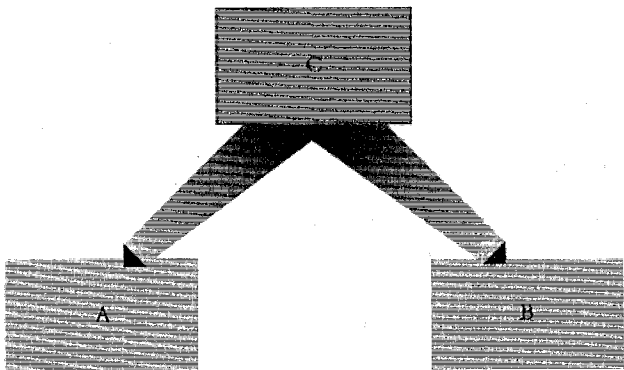
Przypadkowy charakter korelacji pomiędzy przylotem bocianów a liczbą niemowląt wydaje się oczywisty, w innych wypadkach jednak zdarza nam się bezpodstawnie wysnuwać wnioski o przyczynowości. Oglądanie telewizji jest na przykład dodatnio skorelowane z agresywnością u dzieci. Dlatego na ogół sądzi się, że oglądanie telewizji (A) wywołuje agresję (B):



Z drugiej strony, być może wysoki stopień agresji (B) przyczynia się do tego, że dzieci częściej oglądają telewizję (A):



Jest jednak jeszcze inna możliwość. Być może wychowanie w brutalnym środowisku (C) sprawia, że dzieci są agresywne i oglądają telewizję:



Psychologowie ciągle zastanawiają się, który z tych związków przyczynowych jest najsilniejszy; w rzeczywistości wszystkie trzy można udowodnić (APA Comision on Violence and Youth 1993; Eron 1982; Eron, Huesmann 1987; Oskamp 1988). Morał z tej historii jest następujący: jeśli dwie zmienne są ze sobą związane, to jedna zmienna może, choć nie musi być spowodowana przez drugą zmienną.

BADANIA EKSPERYMENTALNE

eksperyment
Kontrolowane sprawdzanie hipotezy, w którym manipuluje się jedną ze zmiennych, aby określić jej wpływ na inną zmienną.

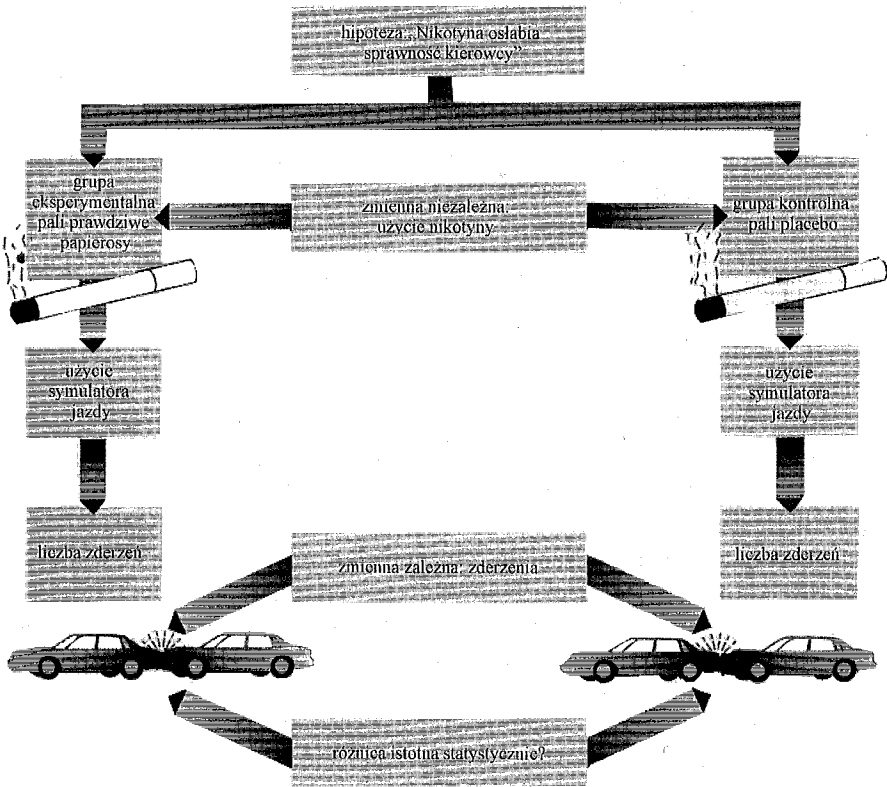
Naukowcy często próbują wyjaśnić zachowania na podstawie badań opisowych, chcąc jednak naprawdę zgłębić ich przyczyny, opierają się w głównej mierze na badaniach eksperymentalnych. **Eksperyment** pozwala badaczom kontrolować sytuację, którą się zajmują. Zamiast biernie rejestrować, co się dzieje, badacze podejmują działania, które, ich zdaniem, mogą mieć wpływ na zachowanie badanej osoby, a następnie obserwują, jakie są skutki tych działań. Techniki te pozwalają eksperymentatorom formułować wnioski na temat przyczyn i skutków.

Wyobraź sobie, że jesteś psychologiem i przeczytałeś właśnie, że palacze powodują więcej wypadków samochodowych niż niepalący, nawet gdy uwzględni się różnice w spożyciu alkoholu, wiek i inne czynniki (DiFranza i in. 1986). Związek ten nie świadczy jednak, że palenie prowadzi do wypadków. Być może palacze są po prostu większymi ryzykantami niż niepalący, czy to narażając się na raka płuc, czy to jadąc przy czerwonym świetle. Być może bardziej niż samo palenie zależność tę tłumaczy rozproszenie uwagi przy strzepywaniu popiołu albo sięganiu po zapalki. Poślaniasz zatem przeprowadzić eksperyment. Próbę wykonujesz w laboratorium, nakazując palaczom prowadzić samochód po krętej drodze z użyciem komputerowego symulatora jazdy wyposażonego w drążek zmiany biegów i pedał gazu. Wyjaśniasz, że celem jest jak najdalsza jazda z jak największą prędkością oraz uniknięcie zderzenia z samochodem jadącym z przodu. Na twoją prośbę niektórzy zapalają papierosa natychmiast po zajęciu miejsca za kierownicą. Drudzy tego nie robią. Interesuje cię porównanie, ile wypadków „zdarzy się” w obydwu grupach. Podstawowy model eksperymentu pokazuje rycina 2.4.

Aspekt sytuacji doświadczalnej, manipulowany lub zmieniany przez badacza, nosi nazwę **zmiennej niezależnej**. Reakcja badanych — zachowanie, które badacz stara się przewidzieć — nazywana jest **zmienną zależną**. W każdym eksperymencie występuje co najmniej jedna zmienna zależna i co najmniej jedna zmienna nieza-

zmienna niezależna
Zmienna poddawana manipulacji przez eksperymentatora.

zmienna zależna
Zmienna, która według przewidywań eksperymentatora ma ulegać wpływom manipulacji zmienną niezależną.



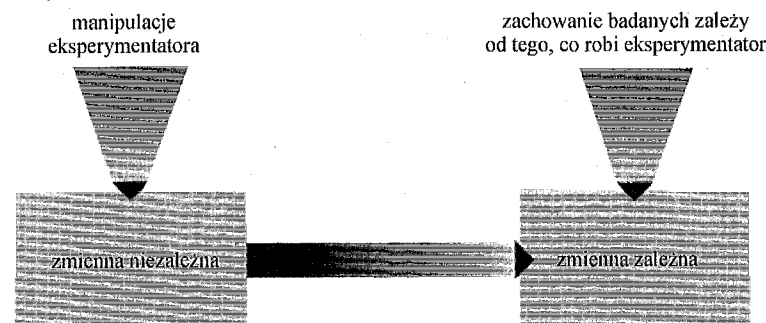
Ryc. 2.4

leżna. W naszym przykładzie zmienną niezależną jest użycie nikotyny: papieros w odniesieniu do braku papierosa. Zmienną zależną jest liczba zderzeń.

W warunkach idealnych wszystko związane z sytuacją eksperymentalną — poza zmienną niezależną — powinno być stałe, to znaczy takie samo dla wszystkich podmiotów. Nie można dopuścić, aby niektórzy posługiwali się dźwignią, inni zaś automatyczną skrzynią biegów, chyba że typ zmiany biegów jest zmienną niezależną. Podobnie, nie można zgodzić się, aby niektórzy odbywali próbę w samotności, a inni przed widownią. Zachowując niezmienną wartość wszystkiego poza zmienną niezależną, gwarantujemy, że cokolwiek się zdarzy, będzie związane wyłącznie z manipulacją badacza. Pozwala to wykluczyć inne interpretacje.

Nie należy się dziwić, że studenci mają czasami kłopoty z rozróżnianiem zmiennych zależnych i niezależnych. Zmienna zależna — wynik badania — zależy od zmiennej niezależnej. Kiedy psycholog przygotowuje eksperyment, myśli: „Jeśli zrobię to i to, badani zrobią to i to”. Pierwsze „to i to” oznacza zmienną niezależną, drugie natomiast zmienną zależną (patrz ryc. 2.5).

Ryc. 2.5



Większość tych samych zmiennych może być w jednej sytuacji zależna, w innej zaś niezależna, ze względu na to, czym manipuluje eksperymentator i co chce udowodnić. Jeżeli chcemy wiedzieć, czy jedzenie czekolady powoduje zdenerwowanie, to ilość zjedzonej czekolady jest zmienną niezależną. Jeżeli zaś interesuje nas, czy zdenerwowanie skłania do jedzenia czekolady, to ilość zjedzonej czekolady jest zmienną zależną.

Warunki kontrolne

warunki kontrolne

W eksperymencie są to warunki porównawcze, w których badani nie podlegają zmianom lub manipulacjom zmiennej niezależnej jak w warunkach eksperymentalnych.

Eksperymenty wymagają zazwyczaj zarówno warunków eksperymentalnych, jak i porównania, czyli inaczej **warunków kontrolnych**. W warunkach kontrolnych badanych traktuje się właściwie tak samo jak w warunkach eksperymentalnych, nie są jednak poddawani takim samym zmianom lub manipulacjom zmiennej niezależnej. Bez grupy kontrolnej nie byłoby pewni, czy zachowanie, które nas interesuje, nie pojawi się w każdym przypadku, nawet bez manipulacji. W niektórych badaniach te same osoby poddawane są zarówno warunkom eksperymentalnym, jak i kontrolnym — mówi się wówczas, że stanowią własną kontrolę. W innych doświadczeniach badanych dzieli się na grupę eksperymentalną i grupę kontrolną.

W doświadczeniu z papierosami osoby, które palą przed jazdą, stanowią grupę eksperymentalną, a osoby wstrzymujące się od palenia tworzą grupę kon-

trolną. Należy się postarać, aby obie grupy miały mniej więcej takie same umiejętności prowadzenia samochodu. Nie można dopuścić, aby w grupie eksperymentalnej byli sami brawurowi rajdowcy, grupa kontrolna zaś składała się wyłącznie z niedzielnych kierowców. Powinniśmy też zadbać, aby obie grupy reprezentowały podobny poziom inteligencji, wykształcenia, podobne nawyki palenia papierosów i inne cechy, wykluczając wpływ którejkolwiek z tych zmiennych na wyniki eksperymentu. W tym celu musimy dokonać **losowego przydziału** osób do obydwu grup. Można na przykład losowo nadać każdej osobie numer, a następnie skierować wszystkich z numerami parzystymi do grupy eksperymentalnej, a wszystkich z numerami nieparzystymi do grupy kontrolnej. Na początku eksperymentu wszystkie osoby mają taką samą szansę przydziału do danej grupy. Jeśli mamy dostateczną liczbę osób, to najprawdopodobniej indywidualne różnice pomiędzy nimi zostaną w obu grupach zrównoważone. Jeśli chodzi o niektóre cechy (np. płeć), to możemy odstąpić od losowego przydziału. Do każdej grupy można wówczas celowo wybrać tę samą liczbę osób z każdej kategorii (np. mężczyzn i kobiet).

Czasami badacze posługują się kilkoma różnymi grupami eksperymentalnymi lub kontrolnymi. W naszym doświadczeniu z papierosami możemy na przykład chcieć sprawdzić oddziaływanie różnej dawki nikotyny, mając w grupie eksperymentalnej osoby, które wypaliły przed jazdą jednego, dwa lub trzy papierosy i porównując je między sobą oraz z osobami z grupy kontrolnej, które nie paliły w ogóle. Na razie jednak poprzestaniemy na eksperymencie z osobami, które wypaliły jednego papierosa.

Mamy zatem dwie grupy. Mamy też pewien problem. Chcąc wypalić papierosa, palacze muszą najpierw go zapalić, a potem zaciągnąć się dymem. Czynności te mogą się wiązać z pewnymi oczekiwaniami — odprężenia, zdenerwowania, opowania itp. Oczekiwania te mogą z kolei wpływać na sprawność kierowcy. Dlatego byłoby lepiej, gdyby grupa kontrolna wykonywała wszystko to co grupa doświadczalna oprócz palenia papierosów. W tym celu trzeba trochę zmienić model eksperymentu. Zamiast poprzestać na niepaleniu w grupie kontrolnej, rozdajemy w niej **placebo**, czyli fałszywe papierosy. Placebo, stosowane często w badaniach leków, może mieć postać tabletek lub zastrzyków nie zawierających składników czynnych. Założmy, że w eksperymencie z papierosami możliwe jest zastosowanie sztucznych papierosów, które smakują i pachną jak prawdziwe, ale nie zawierają aktywnych substancji. Badani z grupy kontrolnej nie wiedzą, że ich papierosy są fałszywe i nie mają możliwości odróżnienia ich od papierosów prawdziwych. Jeśli się okaże, że mają istotnie mniej kolizji niż grupa doświadczalna, to możemy z przekonaniem stwierdzić, że nikotyna zwiększa prawdopodobieństwo wypadku samochodowego (nawiasem mówiąc, placebo czasami wywołuje efekt równie silny lub niewiele słabszy niż prawdziwy lek; fałszywe zastrzyki są, na przykład, niekiedy zaskakująco skuteczne w usuwaniu bólu — efekt placebo jest zagadką oczekującą naukowego wyjaśnienia).

przydział losowy

Procedura przydziału osób do grupy eksperymentalnej i grupy kontrolnej, zgodnie z którą każdy z badanych ma taką samą szansę przydziału do określonej grupy.

placebo

Substancja obojętna lub zabieg pozorny stosowane do kontroli w eksperymencie lub podawane pacjentom.

Wpływ eksperymentatora

Ponieważ na wynik eksperymentu mogą mieć wpływ oczekiwania, badani nie powinni wiedzieć, czy są w grupie eksperymentalnej czy w grupie kontrolnej. W takiej sytuacji (najczęstszej) mamy do czynienia z **pojedynczą ślepa próbą**. Nie tylko osoby badane przynoszą jednak do laboratorium swoje oczekiwania, dotyczy to rów-

pojedyncza ślepa próba

Eksperyment, w którym badani nie wiedzą, czy znajdują się w grupie eksperymentalnej czy w grupie kontrolnej.

Tabela 2.1. Metody badawcze w psychologii: zalety i wady

Metoda	Zalety	Wady
Historia przypadku	Dobre źródło hipotez. Zapewnia pogłębioną wiedzę o jednostce. Nietypowe przypadki mogą wyjaśniać sytuacje i problemy, w których inne sposoby badania są nieetyczne lub niepraktyczne.	Jednostka może nie być reprezentatywna lub typowa. Trudno ocenić, która interpretacja subiektywna jest najlepsza.
Obserwacja w naturalnym środowisku	Pozwala opisać zachowania, jakie spotyka się w środowisku naturalnym. Często użyteczna w pierwszych etapach programu badawczego.	Zapewnia badaczowi co najwyżej ograniczoną kontrolę nad sytuacją. Spostrzeżenia mogą być tendencyjne. Nie pozwala na ścisłe wnioski o przyczynach i skutkach.
Sondaż	Dostarcza dużej ilości informacji o dużej liczbie osób.	Jeśli próba nie jest reprezentatywna lub obciążona błędem, nie można uogólniać wyników badania. Odpowiedzi mogą być niedokładne lub nieprawdziwe.
Badanie korelacji	Wskazuje, czy dwie lub więcej zmiennych jest ze sobą powiązanych. Umożliwia ogólne prognozy.	Nie pozwala na rozpoznanie przyczyn i skutków.
Eksperyment	Pozwala kontrolować sytuację. Umożliwia określenie przyczyn i skutków.	Sytuacja jest sztuczna, nie zawsze można uogólniać wyniki do warunków realnego życia. Czasami trudno uniknąć wpływu eksperymentatora.
Obserwacja laboratoryjna	Umożliwia większą kontrolę niż obserwacja w środowisku naturalnym. Pozwala na stosowanie specjalistycznych urządzeń.	Zapewnia jedynie ograniczoną kontrolę nad sytuacją. Spostrzeżenia mogą być obciążone błędem. Nie pozwala na ścisłe wnioski o przyczynach i skutkach. Zachowanie w laboratorium może różnić się od zachowania w naturalnym środowisku.
Test	Dostarcza informacji na temat cech osobowości, stanów emocjonalnych, postaw, zdolności.	Trudno opracować trafne i rzetelne testy.

niez badaczy. Ich oczekiwania (i nadzieje na pozytywny wynik) mogą mieć mimowolny wpływ na reakcje uczestników badania wywołane przez wyraz twarzy, postawę, ton głosu lub inną wskazówkę.

Wiele lat temu Robert Rosenthal (1966) zademonstrował, jak istotny może być **wpływ eksperymentatora**. Polecił studentom nauczyć szczury poruszania się w labiryncie. Połowę studentów poinformowano, że ich zwierzęta mają wrodzoną zdolność pokonywania labiryntu, natomiast drugiej połowie powiedziano, że otrzymali zwierzęta z wrodzonym upośledzeniem tej sprawności. W rzeczywistości nie było

wpływ eksperymentatora
Niezamierzone zmiany
w zachowaniu badanych
spowodowane mimowolnymi
wskazówkami eksperymentatora.

różnic genetycznych pomiędzy obiema grupami zwierząt. W trakcie doświadczenia szczury uznane za zdolniejsze rzeczywiście szybciej poznawały labirynt, najwidoczniej ze względu na sposób traktowania przez studentów. Jeśli oczekiwania eksperymentatora wpływają na zachowanie gryzoni, stwierdził Rosenthal, to z pewnością mogą też oddziaływać na zachowanie człowieka. Pogląd ten starał się potwierdzić w wielu innych badaniach. W jednym z eksperymentów okazało się, że wskazówki udzielane badanym przez eksperymentatora mogą być subtelne jak uśmiech Mona Lizy — rzeczywiście uśmiech może być wskazówką. Rosenthal zauważył, że mężczyźni przeprowadzający eksperyment częściej uśmiechają się do badanych kobiet niż do mężczyzn. Ponieważ na uśmiech zazwyczaj odpowiada się uśmiechem, zachowanie eksperymentatora może stanowić przeszkodę w badaniach nad uprzejmością lub współpracą.

Jednym z rozwiązań problemu wpływu eksperymentatora jest **podwójna ślepa próba**. W takim badaniu osoba przeprowadzająca eksperyment (mająca bezpośredni kontakt z badanymi) nie wie, które osoby należą do której grupy, dopóki nie zostaną zebrane wyniki. Procedura podwójnej ślepej próby jest standardem w badaniach leków. Różne porcje leku są kodowane w określony sposób, przy czym osoba podająca lek nie jest poinformowana o znaczeniu kodów przed uzyskaniem danych. Chcąc przeprowadzić doświadczenie z papierosami w formie podwójnej ślepej próby, nie możemy informować osoby rozdającej papierosy, które są prawdziwe, a które stanowią placebo. W badaniach psychologicznych podwójne ślepe próby są często trudniejsze do zorganizowania niż pojedyncze ślepe próby. Celem każdego eksperymentu jest jednak zawsze kontrolowanie wszystkiego co możliwe.

Ponieważ eksperymenty pozwalają wnioskować o przyczynach i skutkach, od dawna stanowiły w psychologii metodę z wyboru. Jak wszystkie metody eksperymenty mają jednak pewne ograniczenia. Czasami wpływ na wynik badania mogą mieć procedury zastosowane dla wygody badaczy. Zwróćmy uwagę na trudność, jaka pojawia się w badaniach na zwierzętach. Niedawno stwierdzono, że większość badań na szczurach, myszach i chomikach jest przeprowadzana w ciągu dnia, przy zapalonym świetle. Gryzonie te są jednak zwierzętami nocnymi, to znaczy, normalnie są aktywne w nocy, a śpią w ciągu dnia. Jak zauważyli krytycy: „Psychologii zarzuca się często, że opiera się w przeważającej mierze na badaniach samców białych szczurów [...] Być może należałoby uzupełnić zarzut, dodając: »śpiących samców białych szczurów«!” (Brodie-Scott, Hobbs 1992).

W badaniach nad ludźmi laboratorium przyczynia się do powstawania swego rodzaju zależności od eksperymentatora, polegającej na tym, że badacz decyduje, jakie stawia się pytania i jakie zachowania rejestruje, a badani wykonują tylko jego polecenia. Z chęci współpracy mogą zachowywać się w sposób odbiegający od postępowania na co dzień. Psychologowie eksperymentalni stają przed następującym problemem: im więcej kontroli nad sytuacją, tym mniej przypomina ona realne życie. Dlatego wielu z nich widzi potrzebę rozszerzania badań pozalaboratoryjnych, rzetelnej obserwacji zachowań w kontekstach naturalnych.

podwójna ślepa próba
Eksperyment, w którym zarówno badani, jak i prowadzący doświadczenie nie wiedzą przed uzyskaniem wyników, którzy z badanych są w grupie kontrolnej, a którzy w grupie doświadczalnej.

DLACZEGO PSYCHOLOGIA POSŁUGUJE SIĘ STATYSTYKĄ

Jeśli jako psycholog przeprowadziłeś właśnie obserwację, ankietę lub eksperyment, twoja praca dopiero się rozpoczęła. Mając w ręku wyniki, musisz jeszcze wykonać

trzy zadania: opisać dane, ocenić, na ile są wiarygodne i znaczące oraz zastanowić się nad ich wyjaśnieniem.

Statystyka opisowa: pytanie o fakty

statystyka opisowa

Statystyka, która porządkuje i podsumowuje wyniki badań.

Założmy, że w doświadczeniu z nikotyną trzydzieści osób paliło prawdziwe papierosy, a trzydzieści placebo. Zarejestrowaliśmy liczbę kolizji, którą spowodował każdy z badanych. Mamy więc sześćdziesiąt liczb. Co możemy z nimi zrobić?

Pierwszy krok to podsumowanie danych. Nikogo nie interesuje, ile kolizji spowodowała każda osoba. Ważne jest, co się działo w grupie nikotynowej jako całości w porównaniu z tym, co nastąpiło w grupie kontrolnej. Aby uzyskać tę informację, musimy podsumować nasze dane. Takie podsumowanie nazywamy **statystyką opisową** i często przedstawiamy za pomocą wykresów i tabel.

Jednym ze sposobów podsumowania danych jest obliczenie wartości przeciętnej w grupie. Najczęściej wyraża się ją za pomocą średniej arytmetycznej. Średnią arytmetyczną oblicza się, sumując wszystkie indywidualne wyniki i dzieląc sumę przez liczbę wyników. Średnią dla grupy nikotynowej obliczymy, dodając do siebie trzydzieści wyników oznaczających liczbę kolizji i dzieląc sumę przez trzydzieści. To samo możemy zrobić dla grupy kontrolnej. Trzydzieści liczb sprowadziliśmy do dwóch. Założmy do celów naszego przykładu, że przeciętna liczba kolizji w grupie nikotynowej wynosiła dziesięć, a w grupie kontrolnej tylko siedem.

Musimy być jednak ostrożni z interpretacją wartości przeciętnych. Średnia arytmetyczna nie zawsze oznacza to co typowe; stanowi jedynie podsumowanie zbioru danych. Jest na przykład możliwe, że nikt z grupy nikotynowej nie miał akurat dziesięciu kolizji. Być może połowa osób z tej grupy to piraci drogowi, którzy mieli po piętnaście kolizji, natomiast pozostali jeżdżą ostrożnie i mieli ich tylko po pięć. Możliwe, że prawie wszyscy mieli dziewięć, dziesięć lub jedenaście kolizji. Możliwe, że liczba zderzeń była rozłożona równomiernie od zera do piętnastu.

Średnia nie informuje nas o takich różnicach w zachowaniach badanych. Do tego mamy inne wskaźniki statystyczne. Jednym z nich jest rozstęp, czyli różnica pomiędzy najwyższym a najniższym wynikiem spośród uzyskanych w badaniu. Jeśli najmniejsza liczba kolizji w grupie nikotynowej wynosiła pięć, a największa piętnaście, to obszar zmienności wynosi dziesięć. Innym wskaźnikiem, dającym lepszy obraz sytuacji, jest **wariancja**, która informuje nas, w jakim stopniu indywidualne wyniki są rozproszone lub skupione wokół średniej. Im bardziej są rozproszone, tym mniej „typowa” średnia. Niestety, gdy w doniesieniach prasowych lub w wieczornych wiadomościach mówi się o wynikach badań naukowych, zazwyczaj brana jest pod uwagę wyłącznie średnia.

wariancja

Miara rozproszenia wyników wokół średniej.

Wnioskowanie statystyczne: pytanie o wnioski

W tym punkcie naszych badań nad nikotyną mamy jedną grupę, w której średnia kolizji wynosi dziesięć i grupę drugą ze średnią siedem. Czy mamy już otwierać szampana? Wystąpić w telewizji? Zadzwoić do rodziców?

Na razie poczekajmy. Statystyka opisowa nie mówi nam, czy wynik jest osiągnięciem, którym warto się pochwalić. Może gdyby w jednej grupie przeciętna wynosiła piętnaście, a w drugiej tylko jeden, byłoby czym się cieszyć. Badania psycho-

logiczne rzadko jednak podsuwają nam jednoznaczne rezultaty. Na ogół nie można wykluczyć, że różnica pomiędzy dwiema grupami zaistniała przypadkowo. Może zdarzyło się po prostu, że osoby w grupie nikotynowej były w trakcie doświadczenia nieco bardziej skłonne do popełniania błędów, lecz ich zachowanie nie miało nic wspólnego z paleniem. Dziwne byłoby, gdyby obie grupy miały dokładnie tę samą liczbę kolizji.

Aby stwierdzić, na ile istotne są uzyskane wyniki, psychologowie posługują się **wnioskowaniem statystycznym**. Pozwala ono na wyciąganie wniosków (twierdzeń opartych na dowodach) dotyczących wyników doświadczenia. Mamy do wyboru różne formy wnioskowania statystycznego w zależności od rodzaju badania i celu, który nas interesuje. Podobnie jak statystyka opisowa wnioskowanie statystyczne ujmuje dane w formuły matematyczne.

Wnioskowanie statystyczne nie tylko opisuje lub podsumowuje dane, lecz dostarcza również wskaźników informujących badacza, na ile jest prawdopodobne, że uzyskany rezultat badania jest tylko przypadkowy. Dokładniej, określają one prawdopodobieństwo uzyskania efektu równego lub większego niż obserwowany, jeśli manipulowanie zmienną niezależną nie ma rzeczywistego wpływu na badane zachowanie. Niemożliwe jest całkowite wykluczenie roli przypadku. Jednak gdy prawdopodobieństwo, iż wynik uzyskano przypadkowo, jest bardzo niewielkie, mówimy, że wynik jest **statystycznie istotny**. Oznacza to, że prawdopodobieństwo rzeczywistej różnicy jest istotne. Zauważmy, że nie jest to pewność, lecz jedynie znaczne prawdopodobieństwo. Dla wygody psychologowie uznają wynik za istotny, jeśli można spodziewać się jego uzyskania pięć lub mniej razy w stu powtórzeniach eksperymentu (inaczej można powiedzieć, że wynik jest istotny na poziomie 0,05).

Wnioskowanie statystyczne odkrywa, że fakty na pozór nieprawdopodobne niekiedy wcale nieprawdopodobne nie są. Na przykład na ile wydaje się prawdopodobne, że spośród dwudziestu pięciu osób znajdujących się w jednym pomieszczeniu przynajmniej dwie obchodzą urodziny w tym samym dniu? Większość ludzi przypuszcza, że jest to bardzo mało prawdopodobne, okazuje się jednak, że szansa jest większa niż pół na pół. Nawet gdy w pomieszczeniu jest tylko dziesięć osób, prawdopodobieństwo wynosi 1:9. Spośród prezydentów USA dwaj mieli urodziny w tym samym dniu (Warren Harding i James Polk), a trzech zmarło czwartego lipca (John Adams, drugi prezydent; Thomas Jefferson, trzeci; James Monroe, piąty). Zaskakujące? Nie. Takie zbiegi okoliczności dla statystyki nie są wcale rewelacją.

Za pomocą wnioskowania statystycznego możemy ustalić, czy wynik eksperymentu byłby rzeczywiście rzadkim zdarzeniem, gdyby działał tylko przypadek. W naszej próbie z nikotyną dane wyprowadzone na drodze wnioskowania statystycznego informują nas, na ile jest prawdopodobne, że różnica pomiędzy grupą nikotynową a grupą placebo pojawiła się przypadkowo. Jeśli można się spodziewać, że różnica pojawi się przypadkowo w sześciu próbach na sto, to musimy przyznać, że wyniki nie są potwierdzeniem hipotezy. Jesteśmy zmuszeni stwierdzić, że uzyskana przez nas różnica mogła się zdarzyć po prostu przypadkowo. Możemy oczywiście przeprowadzić dalsze badania, aby się co do tego upewnić. Widzimy więc, jak trudno przekonać psychologów, aby poważnie potraktowali uzyskany niegdyś jednorazowy wynik.

Dane statystycznie istotne pozwalają na ogólne przewidywanie zachowań człowieka. Przewidywania zazwyczaj mają postać prawdopodobieństwa („Spodziewamy się, że przeciętnie 60 procent studentów robi X, Y, Z”). Nie mówią nam jednak z pewnością, co w określonej sytuacji uczyni konkretna jednostka. Prognozowanie

wnioskowanie statystyczne

Obliczenia statystyczne, które pozwalają ocenić, na ile jest prawdopodobne, że określony wynik pojawił się jedynie przypadkowo.

statystycznie istotny

Określenie używane w odniesieniu do wyniku, gdy jego pojawienie się na zasadzie przypadku jest wysoce nieprawdopodobne.

jest typowe dla wszystkich nauk. Badania medyczne wykazują, na przykład, że u osób palących istnieje znaczne prawdopodobieństwo zachorowania na raka płuc. Ponieważ jednak każdy pojedynczy przypadek raka powstaje wskutek współdziałania różnych zmiennych, naukowcy nie powiedzą nam na pewno, czy ciocia Bessie, która wypala dwie paczki dziennie, zostanie dotknięta chorobą.

Bardzo szczegółowe przewidywania dotyczące jednostek są w zasadzie niemożliwe. Nowe prace matematyków nad teorią chaosu wskazują, że nawet drobne, ledwo dostrzegalne zdarzenie może mieć głęboki, lecz nieprzewidywalny wpływ na zachowanie złożonego systemu — a człowiek z pewnością jest złożonym systemem (Barton 1994). Przypuśćmy, że nie możesz dostać się na kurs, w którym naprawdę chciałbyś uczestniczyć, ponieważ lista jest już zamknięta, zapisujesz się więc tam, gdzie są jeszcze wolne miejsca — na wstępny kurs astronomii. W rezultacie dziesięć lat później pracujesz dla NASA, czego nie wyobrażałeś sobie nawet w najbardziej fantastycznych marzeniach. Teoria chaosu mówi nam jednak również, że istnieją pewne ogólne wzorce zjawisk naturalnych oraz zachowań człowieka, które można rozpoznać.

Nawiasem mówiąc, próba z nikotyną, przypominająca nasz hipotetyczny przykład, została naprawdę przeprowadzona, choć z zastosowaniem trochę innych i bardziej skomplikowanych procedur (Spilich, June, Renner 1992). Palacze, którzy zapalili papierosa przed jazdą, dojechali wprawdzie nieco dalej na symulowanym torze, mieli jednak przeciętnie znacząco więcej kolizji (10,7) niż osoby, które chwilowo zrezygnowały z papierosa (5,2) i osoby niepalące (3,1). Dowiedziawszy się o tych badaniach, szef Federal Express zabronił palenia podczas pracy dwunastu tysiącom kierowców zatrudnionym w jego przedsiębiorstwie (George J. Spilich, informacja własna).

Interpretacja rezultatów

Ostatnim etapem każdego badania jest wyjaśnienie, co oznaczają uzyskane rezultaty. Próba zrozumienia zachowań na podstawie nie zinterpretowanych wyników przypomina próbę opanowania języka walijskiego poprzez czytanie słownika. Tak jak w nauce języka konieczna jest gramatyka, aby nauczyć się łączyć słowa, tak w psychologii potrzebne są hipotezy i teorie, aby wyjaśnić, w jaki sposób wiążą się ze sobą fakty stwierdzone w badaniu.

Czasami trudno wybierać spośród konkurencyjnych wyjaśnień. Czy nikotyna przeszkadza w prowadzeniu samochodu, ponieważ powoduje zaburzenia koordynacji? Sprzyja rozproszeniu uwagi? Zakłóca przetwarzanie informacji? Zniekształca ocenę i postrzeganie niebezpieczeństwa? Ogólnie, najlepsze wyjaśnienie powinno tłumaczyć jak największą liczbę wyników i umożliwiać jak najdokładniejsze przewidywanie nowych rezultatów.

Często wyjaśnienie wyników badań wymaga uwzględnienia wielu różnych aspektów. W rozdziale 1 przekonaliśmy się na przykład, że doznanie przemocy w dzieciństwie niekoniecznie jest przyczyną stosowania przemocy przez rodziców. Różne czynniki, m.in. stresy, oglądanie przemocy w telewizji oraz charakter doświadczanej przemocy współdziałają ze sobą w skomplikowany sposób, wpływając na to, jakimi ostatecznie stajemy się rodzicami (Widom 1989). Na szczęście, za pomocą specjalnych procedur statystycznych psychologowie często mają możliwość dokonania

Tabela 2.2. Zestawienie metod badań psychologicznych

Psychologowie mogą stosować różne metody, aby uzyskać odpowiedzi na różne pytania dotyczące jakiegoś zagadnienia. Dla ilustracji poniżej zestawiono niektóre zastosowania metod badania różnych kwestii związanych z agresją. Wymienione metody nie zawsze się wykluczają. Czasami w badaniach nad jakimś zagadnieniem można korzystać z dwóch lub więcej metod. Jak mówiliśmy w tekście, wyniki uzyskane jedną metodą mogą poszerzać, potwierdzać lub wykluczać wyniki uzyskane inną metodą.

Metoda	Cel	Przykład
Historia przypadku	Zrozumienie rozwoju zachowań agresywnych u konkretnej jednostki; sformułowanie hipotez co do pochodzenia agresji.	Historia rozwoju wielokrotnego mordercy.
Obserwacja w naturze	Opis charakteru działań agresywnych we wczesnym dzieciństwie.	Obserwacja, klasyfikacja i opis bójk, popychania się itd. w czasie zabaw w przedszkolu.
Obserwacja laboratoryjna	Stwierdzenie, czy agresja w parach dzieci tej samej oraz odmiennej płci różni się pod względem częstotliwości i intensywności.	Obserwacja przez jednostronne lustro par tej samej oraz odmiennej płci wśród przedszkolaków. Pary muszą rozstrzygać, które z dzieci dostanie atrakcyjną zabawkę obiecaną obojgu.
Test	Porównanie cech osobowości osób agresywnych i nieagresywnych.	Przeprowadzenie testów osobowości wśród więźniów stosujących przemoc i nie stosujących przemocy.
Sondaż	Stwierdzenie, w jakim stopniu przemoc w rodzinie jest powszechna.	Kwestionariusz przeznaczony dla anonimowych respondentów (w próbie reprezentatywnej dla populacji USA) na temat stosowania kar cielesnych w ich rodzinach.
Badanie korelacji	Badanie zależności pomiędzy agresywnością a oglądaniem telewizji.	Przeprowadzenie testu agresywności z wykorzystaniem rysunków oraz wywiad dotyczący liczby godzin w tygodniu spędzanych przed telewizorem. Obliczenie współczynnika korelacji.
Eksperyment	Stwierdzenie, czy wysoka temperatura pobudza agresywne zachowania.	Zaaranżowanie sytuacji, w której uczestnicy karzą „ucznia” (w rzeczywistości pomocnika eksperymentatora), wywołując wstrząs, gdy znajdują się w pomieszczeniu o temperaturze 8 lub 25°C.

analizy, w jakim stopniu każda zmienna przyczynia się do uzyskanego wyniku i jak współdziałają w tym różne czynniki.

Interpretując konkretne badania, nie należy ignorować rzeczywistości. Może się zdarzyć, że kilka wyjaśnień równie dobrze tłumaczy te same fakty, co oznacza konieczność dalszych badań w celu uzyskania najlepszego wyjaśnienia. Rzadko kiedy jedna próba pozwala sprawdzić wszystko. Dlatego powinniśmy być podejrzliwi wobec tytułów, które zwiastują „wielki przełom w nauce!!!” Postęp naukowy zazwyczaj dokonuje się stopniowo, a nie od razu. Czasami najlepsza interpretacja pojawia

się dopiero, gdy hipoteza zostanie sprawdzona na kilka różnych sposobów. Choć każda z opisanych przez nas metod na ogół przeznaczona jest do innych celów (patrz tab. 2.2), różne metody mogą się wzajemnie uzupełniać. Jedna metoda może zatem potwierdzać, zaprzeczać lub uzupełniać rezultaty uzyskane inną metodą. Jeśli wyniki badań przeprowadzanych różnymi metodami pokrywają się, to jest to najlepszy powód, aby im zaufać. Z drugiej zaś strony, jeśli są ze sobą niezgodne, to badacze powinni zmodyfikować hipotezę lub przeprowadzić więcej badań.

Oto jeden z przykładów. Porównując wyniki testów sprawności umysłowej osób młodych i starszych, psychologowie zazwyczaj stwierdzają, że młodzi mają przewagę nad starszymi. Badania, w których różne grupy porównywane są w tym samym czasie, nazywamy **badaniami poprzecznymi**. Niektórzy badacze podejmują **badania podłużne** (longitudalne), których celem jest śledzenie zmian sprawności umysłowej w miarę starzenia się. W badaniach podłużnych te same osoby są obserwowane przez dłuższy okres i poddawane kontroli w pewnych odstępach czasu. W przeciwieństwie do badań poprzecznych badania podłużne wykazują, że w wielu rodzajach testów wyniki wcale nie pogarszają się z wiekiem. Ogólny spadek sprawności umysłowej, jeśli w ogóle następuje, to na ogół nie wcześniej niż w wieku siedemdziesięciu lub osiemdziesięciu lat (Baltes, Dittman-Kohli, Dixon 1984; Schaie 1993). Dlaczego rezultaty tych dwóch typów badań są sprzeczne? Najwidoczniej w badaniach poprzecznych ujawnia się różnica pokoleniowa, młodsze pokolenie przewyższa starszych prawdopodobnie dlatego, że jest lepiej wykształcone lub lepiej zna typ zagadnień będących przedmiotem testów. Bez badań podłużnych moglibyśmy wysnuć niesłuszny wniosek, że sprawność umysłowa z wiekiem maleje.

Czasami psychologowie zgadzają się co do rzetelności i interpretacji wyników, lecz nie są jednomyślni w ocenie ich znaczenia dla teorii i praktyki. Sama istotność statystyczna nie daje odpowiedzi, ponieważ to, co istotne statystycznie, nie zawsze ma znaczenie w realnym życiu. Wynik może być statystycznie znaczący na poziomie 0,05, lecz równocześnie mieć znikomy udział w codziennym życiu (Rosnow, Rosenthal 1989). Psychologowie wykorzystują obecnie inne rodzaje danych statystycznych do określenia, na ile za zmienność wyników w badaniu odpowiada zmienność niezależna. Często okazuje się, że „wielkość efektu” jest znikoma.

Jedną z użytecznych nowych technik, zwaną **metaanalizą**, łączy i analizuje wyniki wielu badań, wykorzystując dane statystyczne, zamiast każde badanie oceniać osobno. Metaanaliza informuje, na ile zmienność wyników we wszystkich badaniach można wyjaśnić za pomocą określonej zmiennej. Często dostarcza ciekawych niespodzianek. Niektórzy badacze stwierdzili na przykład, że uznawane dotąd za niepodważalne różnice pomiędzy mężczyznami a kobietami, jeżeli chodzi o sprawność językową i wyobraźnię przestrzenną, agresywność i uleganie wpływom wprawdzie rzeczywiście istnieją, lecz są niewielkie. W większości przypadków płeć powoduje tylko 1-5 procent zmienności wyników (Eagly, Carli 1981; Feingold 1988, Hyde 1981, 1984; Hyde, Fennema, Lamon 1990; Hyde, Linn 1988). Mówiąc prościej, informacja, czy badana osoba jest mężczyzną czy kobietą, nie powie nam zbyt wiele o jej osiągnięciach w czterech wymienionych sferach.

Obrazowo przedstawia to rycina 2.6 (za: Sapolsky 1987, na podstawie danych Benbow, Stanley 1983). Chłopcy z siódmej klasy mieli przeciętnie lepsze wyniki w matematycznej części testu SAT* niż dziewczęta z klasy siódmej. Choć rze-

badania poprzeczne

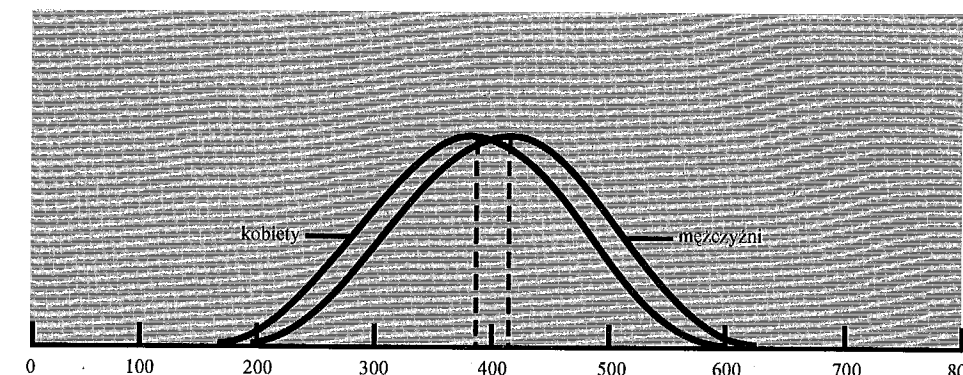
Badania, w których grupy osób w różnym wieku są porównywane w tym samym czasie.

badania podłużne

Badania, w których obserwacja odbywa się przez dłuższy czas z kilkakrotną okresową kontrolą.

metaanaliza

Procedura statystyczna polegająca na łączeniu i analizowaniu danych z wielu badań; określa w jakim stopniu zmienność wyników we wszystkich badaniach można wyjaśnić za pomocą określonej zmiennej.



Ryc. 2.6

czywiście zachodzi pewna różnica, jednak jest ona nieznaczna, a oceny chłopców i dziewcząt w dużej części pokrywają się. Nie można zatem przewidzieć, czy konkretny uczeń okaże się lepszy niż konkretna uczennica.

Metaanaliza stwarza jednak pewne problemy. Między innymi połączenie badań różnej jakości, wykonywanych z zastosowaniem różnych procedur, z udziałem osób w różnym wieku i o różnym pochodzeniu, stwarza niebezpieczeństwo pominięcia jakichś istotnych czynników. Oto przykład: badania sprawności językowej opierają się zazwyczaj na wynikach uzyskanych przez uczniów szkół średnich. Tymczasem uczniowie, którzy mają trudności z czytaniem, częściej rezygnują z ukończenia szkoły, a wśród rezygnujących jest więcej mężczyzn niż kobiet. Metaanaliza może zatem nie wykazać ogólnej przewagi kobiet w testach językowych (Halpern 1989). Podobnie metaanaliza łącząca badania wszystkich sprawności matematycznych pomija przewagę mężczyzn, jeśli chodzi o wyobraźnię przestrzenną (McGuinness 1993). Metaanalizy, które nie stwierdzają ogólnych różnic płciowych w sprawnościach poznawczych, mają wprawdzie dla wielu ludzi atrakcyjną wymowę społeczną, lecz ostatecznie okazują się nieprzydatne, jak sądzą krytycy, ponieważ błędnie interpretują różnice, o których powinni wiedzieć i nauczyciele, i uczniowie, jeśli chcemy, by wszystkie dzieci osiągały dobre wyniki w nauce. Zdaniem krytyków metaanaliza jest jak połączenie jabłka z pomarańczą; krzyżówka jest ciekawa, ale mdła, nie wyczuwa się w niej odrębnych smaków poszczególnych składników. Mimo wszystko w odniesieniu do zagadnień, które były przedmiotem kilkudziesięciu, a nawet kilkuset badań, metaanaliza jest najbardziej obiecującym sposobem stwierdzenia prawdziwości, jednak pod warunkiem uwzględnienia jej ograniczeń.

OCENA METOD BADAWCZYCH: NAUKA POD LUPĄ

Ścisłe metody badawcze są samym jądrem nauki, nic więc dziwnego, że psychologowie spędzają mnóstwo czasu na dyskusjach i rozważaniach na temat procedur gromadzenia, oceny i przedstawiania danych. W ostatnich latach dyskusje te wykroczyły daleko poza kwestie techniczne. Nie zastanawiamy się już na przykład, kiedy posługiwać się kwestionariuszem, a kiedy wywiadem, albo jak najlepiej analizować dane.

* Stanfordzki test osiągnięć (Stanford Achievement Test) — przyp. tłum.

Niektóre z dyskusji dotyczą stosowania testów psychologicznych do niewłaściwych celów. Antropolog F. Allan Hanson (1993) uważa, że testy są szkodliwą ingerencją w prywatność człowieka, „instrumentem rozwijającego się systemu instytucjonalnej dominacji, którego działanie narusza wolność i godność jednostki”. Twierdzi na przykład, że testy uzdolnień i inteligencji hamują motywację i stwarzają samospełniające się proroctwa. Cenią wyżej hipotetyczny „potencjał” jednostki od jej rzeczywistych osiągnięć, stąd bierze się dziwny zwyczaj pracodawców, którzy domagają się od absolwentów college’u okazania wyników testu SAT, nawet gdy kandydat przedstawił ważne świadectwo szkolne.

Hanson i współpracownicy obawiają się ponadto, że testów, które są próbą oceny charakteru badanego lub jego przyszłych zachowań, zbyt często się używa do dyscyplinowania pracowników obawiających się o opinię, pracę i poziom życia. Choć zabroniono rutynowego stosowania wykrywaczy kłamstw (ponieważ, jak mówiliśmy, test jest mało rzetelny i mało trafny), pracodawcy nie zmienili wcale przekonania, że możliwe jest rozpoznanie osoby, która kłamie. Wielu pracodawców wypróbowało teraz nowe metody wykrywania kłamstwa, starając się rozpoznać wśród zatrudnianych osób takie, które mogą oszukiwać, kraść albo używać narkotyków. Jednym z tych kierunków jest rozpowszechnianie „testów uczciwości”, które ponoć wykazują, czy przyszły pracownik będzie kradł w miejscu pracy. Co roku testom tym są poddawane miliony ludzi, choć nie ma pewności, na ile są rzetelne i trafne (Saxe 1991). Specjaliści od testów zauważają z niepokojem, że wiele (może nawet większość) osób, które uzyskują ocenę negatywną, wcale nie jest nieuczciwych (Cameron, Schneider 1994). Wyobraź sobie, że zapisując się do college’u, masz wypełnić test, który podobno ma wykazać, czy jesteś zdolny do oszustwa. Jak byś to przyjął? Na ile rzetelny powinien okazać się test, zanim jego rozpowszechnianie będzie można uznać za uprawnione? Czy można go zaakceptować, jeśli niesłusznie oskarża „tylko” jedną osobę niewinną na dziesięć? na sto? na tysiąc?

Obrońcy testów odpowiadają na te zarzuty, że testy są jedynie narzędziami, a wszystkie narzędzia mogą być używane błędnie. Zauważają, że istnieje różnica pomiędzy ostrożnym wykorzystywaniem testów do określonych celów badawczych lub do diagnozowania problemów psychicznych a używaniem ich rutynowo, bezmyślnie lub w złej wierze, aby nakłonić niewinnych ludzi do zeznań, próbować przewidzieć, kto w przyszłości popełni przestępstwo lub ingerować w prywatne życie pracowników. W każdym razie wydaje się oczywiste, że każdy, kto poddaje się testom, powinien czynić to w sposób świadomy, wiedząc, dlaczego przeprowadza się test i do jakich celów jest on przeznaczony.

Etyka badań naukowych

Kolejna grupa zagadnień wiąże się z etyką badań psychologicznych. Jeden z głównych zarzutów dotyczy posługiwania się oszustwem. W wielu eksperymentach, zwłaszcza w dziedzinie psychologii społecznej, istnieje konieczność, aby badacze oszukiwali badanych co do celu badania przed jego zakończeniem. Gdyby uczestnicy doświadczali od początku wiedzieli, czemu badanie ma służyć, trafność wyników byłaby wątpliwa: gdyby wiedzieli, że chodzi o gotowość pomocy innym, mogliby jako do-

brzy samarytanie próbować przewyższyć Matkę Teresę; gdyby wiedzieli, że idzie o sprawność językową, wykazywaliby elokwencję sprawozdawców sportowych. Oczywistym rozwiązaniem jest nieujawnianie celu badania, a mówiąc wprost — kłamstwo.

Co zrobić jednak, gdy oszustwo wywołuje u badanych niepokój i zaniepokojenie? Czy można wprowadzać ich w błąd, aby myśleli, że wywołują wstrząs elektryczny u innej osoby? Czy można pozwolić, aby byli obrażani albo poniżani przez pomocnika eksperymentatora, który bada ich reakcje? Ze względu na rosnące w ciągu ostatnich kilkudziesięciu lat zastrzeżenia co do moralności procedur oszukiwania zasady etyczne APA wymagają obecnie, aby badacze, którzy zamierzają posłużyć się oszustwem, zastanowili się, czy jest to uzasadnione potencjalną wartością naukową, edukacyjną lub praktyczną badania, i by rozważyli alternatywne możliwości, nie wymagające przemilczeń. Badacze muszą również chronić uczestników eksperymentu przed dolegliwościami lub zagrożeniami fizycznymi i psychicznymi, a w przypadku jakiegokolwiek ryzyka — uprzedzić o jego istnieniu.

Kolejnym zagadnieniem etycznym jest wykorzystywanie do badań zwierząt. Zwierzęta są używane w 7-8 procentach badań psychologicznych, przy czym 95 procent stanowią gryzonie (American Psychological Association, 1984), czasami jednak psychologowie wykorzystują również gołębie, koty, małpy i inne gatunki. Większość badań nie sprawia zwierzętom bólu ani przykrości (np. obserwacja parzenia się chomików), niektóre doświadczenia bywają jednak dotkliwe (np. oddzielenie młodych małp od matek i badanie rozwoju nienormalnych wzorców zachowań w następstwie rozłąki). Są także badania, które wymagają nawet śmierci zwierzęcia, na przykład w celu zbadania specyficznych zmian zachodzących w mózgu uśmierca się szczury hodowane w ubogim bądź w urozmaiconym środowisku.

Warto zauważyć, że badania na zwierzętach dają wiele korzyści samym zwierzętom. Stosując zasady behawioryzmu, farmerzy mogą, na przykład, ograniczyć zniszczenia upraw przez ptaki i zwierzynę płową bez uciekania się do tradycyjnej metody zabijania zwierząt. Badania na zwierzętach przyczyniają się również w wielu przypadkach do poprawy stanu zdrowia i samopoczucia człowieka. Wyniki tych badań pomagają psychologom i innym badaczom w opracowywaniu metod terapii moczenia nocnego, w nauczaniu dzieci upośledzonych porozumiewania się z otoczeniem, w przeciwdziałaniu groźnemu dla życia niedożywieniu niemowląt wskutek przewlekłych wymiotów, w rehabilitacji pacjentów z chorobami neurologicznymi i zaburzeniami czuciowymi. Są również przydatne przy poszukiwaniu ulepszonych sposobów zwalczania przewlekłego bólu, w tresurze zwierząt dla osób niepełnosprawnych oraz w śledzeniu mechanizmów powodujących utratę pamięci i starzenie się — oto tylko niektóre korzyści (Feeney 1987; Greenough, 1991; N. Miller 1985).

W ostatnich latach pojawiły się jednak ostre protesty w obronie zwierząt, a nawet spory o to, czy w ogóle powinno się wykonywać badania na zwierzętach. Po obu stronach spotykamy skrajne poglądy. W jednym z sondaży 85 procent zdeklarowanych obrońców praw zwierząt (17 procent w grupie porównawczej) poparło sformułowanie: „Gdyby ode mnie zależało, wykluczyłbym wszystkie badania z wykorzystaniem zwierząt” (Plous 1991). Z drugiej jednak strony nie wszyscy psychologowie skłonni są przyznać, że warunki w laboratoriach mogą mieć szkodliwy wpływ na psychikę lub stan fizyczny niektórych zwierząt. Niestety, różnica zdań często przeradza się w złośliwe pomówienia lub wywołuje jeszcze poważniejsze konsekwencje. Niektórzy obrońcy praw zwierząt demolują laboratoria, grożą i szantażują

badaczy i ich rodziny, ci z kolei niesłusznie traktują wszystkich obrońców praw zwierząt jak terrorystów.

Jeśli chodzi o skutki pozytywne, to konflikt skłonił wielu badaczy do szukania sposobów lepszego traktowania zwierząt doświadczalnych. Kodeks etyczny APA zawsze zawierał zasady humanitarnego traktowania zwierząt, a w ostatnich latach sformułowano bardziej spójne wytyczne. Uchwalono też nowe ustawy federalne, które wprowadzają surowsze przepisy dotyczące opieki nad zwierzętami laboratoryjnymi. Każdy projekt badania, w którym mają być wykorzystane kręgowce, musi być teraz rozpatrzony przez komisję złożoną z przedstawicieli instytucji naukowej i społeczeństwa.

Większość badaczy sprzeciwia się propozycjom znacznego ograniczenia badań na zwierzętach lub rezygnacji z nich. APA i inne organizacje zawodowe uważają, że ustawy o ochronie zwierząt są pożądane, ale nie powinny uniemożliwiać prowadzenia prac badawczych, które przyczyniają się do pełniejszego wyjaśnienia naukowego niektórych faktów i do poprawy sytuacji życiowej człowieka. Trudnym zadaniem — z którym będziemy się zmagać jeszcze przez wiele lat — jest zrównoważenie oceny licznych korzyści badań na zwierzętach z uznaniem dokonywanych w przeszłości nadużyć i współczującą postawą wobec innych gatunków.

Znaczenie wiedzy

Ciągłe spory wokół zagadnień etycznych wskazują, że metody badawcze w psychologii mogą wywoływać tyle samo kontrowersji ile wyniki badań. Zastrzeżenia dotyczą nie tylko sposobu przeprowadzania doświadczeń, ale również tematów, jakimi badania powinny lub nie powinny się zajmować. Zresztą wszystkie nauki ścisłe i humanistyczne zostały dotknięte głębokimi zmianami poglądów na temat samej istoty wiedzy.

W ciągu ostatnich trzystu lat odpowiedź wydawała się dość jasna. Wiedza była odkrywaniem pewnej rzeczywistości istniejącej na zewnątrz, a sposób jej poznawania miał być obiektywny, nie wartościujący i niezależny. Celem teorii było przedstawienie owej rzeczywistości. Uważano, że między badaczem z jednej strony a poznawanym zjawiskiem z drugiej istnieje wyraźna granica, której badacz nie może przekroczyć.

Obecnie wielu naukowców kwestionuje to fundamentalne przeświadczenie. Zwolennicy postmodernizmu twierdzą, że niezależna obiektywność, długo uważana za kamień węgielny zachodniej nauki, jest mitem. W ujęciu postmodernistycznym wartości, sądy i pozycja społeczna obserwatora nieuchronnie wpływają na sposób badania zjawisk, sposób ich wyjaśniania, a nawet na sam ich przebieg. Zgodnie z tym poglądem nauczyciele i uczeni nie są wolni od subiektywności. Ponieważ wykonują swoją pracę w określonym czasie i w określonej kulturze, wnoszą do niej wspólne przeświadczenia i światopoglądy, które określają, co należy uznać za istotne, jakie elementy rzeczywistości są warte uwagi i jakie obowiązują hierarchie. Jeśli kiedykolwiek uczestniczyłeś w dyskusji, jakie książki należy zaliczyć do klasyki — znane dzieła literatury zachodniej czy współczesne utwory Murzynów amerykańskich, kobiet czy pisarzy pozaeuropejskich — miałeś do czynienia z konfliktem postmodernistycznym.

W naukach społecznych postmodernistyczna teoria — zwana **konstruktywizmem społecznym** — głosi, że wiedza jest nie tyle odkrywaniem, ile konstruowaniem

albo wynajdowaniem (Gergen 1985; Guba 1990; Hare-Mustin, Maracek 1990; Rosaldo 1989; Watzlawick 1984). Nasze rozumienie rzeczywistości nie tylko odzwierciedla to, co na zewnątrz, ale również organizuje i porządkuje. George Lakoff (1987) podaje przykład z nauk przyrodniczych. W roku 1735 szwedzki botanik Karol Linneusz wprowadził sposób klasyfikacji roślin, który używany jest do dzisiaj jako obiektywna norma. System Linneusza opiera się przede wszystkim na kształcie owocu. Dlaczego uczony nie posłużył się kształtem liścia lub kolorem owocu, albo jakąkolwiek inną cechą, którą można by wyróżnić? Według jego syna kierował się raczej wygodą niż wymaganiami jakiejś zewnętrznej rzeczywistości: kształt owocu był „wyraźnie zaznaczony, bez trudu rozpoznawalny, łatwy do opisanie w słowach”.

Tak jak konstruujemy sposoby klasyfikacji roślin, tak postępujemy i wobec ludzi. Weźmy pojęcie rasy. Często mówi się o czarnych, Azjatach lub białych, jakby granice pomiędzy tymi grupami były oczywiste na podstawie różnic fizycznych. Tymczasem istnieje wiele definicji rasy. Niektóre społeczeństwa opierają swoje definicje na pochodzeniu, a nie na jakichś oczywistych cechach fizycznych, jak kolor skóry; ktoś może wyglądać jak biały, lecz zostać uznany za czarnego, jeśli ma choć jednego czarnego przodka. W 1970 roku postanowiono w Luizjanie, że za czarnego uważany będzie każdy, kto ma zaledwie jedną trzydziestą dziewiątą murzyńskiej krwi (Jones 1991). W Afryce Południowej stosuje się kryterium pochodzenia, aby odróżnić czarnych od kolorowych (osób pochodzenia mieszanego) dla celów politycznych i prawnych. W Ameryce Południowej, której mieszkańcy stanowią wszelkie odcienie i kombinacje pochodzenia europejskiego, indiańskiego i afrykańskiego, ten podział nie jest możliwy. Inne społeczeństwa zaliczają ludzi do określonej rasy, jeśli przyznają się oni do pewnej wspólnoty religijnej lub kulturowej. Niektórzy mówią, na przykład, o rasie żydowskiej, jakby nie dostrzegali europejskich Żydów o jasnych włosach i czarnych Żydów z Etiopii.

Jak pokazują przykłady, podziały rasowe pomijają fakt, że prawie nikt z nas nie jest czystej rasy; fakt ten przysparza najróżniejszych problemów amerykańskiemu Urzędowi Statystycznemu, który wytrwale stara się utworzyć nowe pojęcia. Klasyfikacje rasowe wyolbrzymiają różnice pomiędzy grupami; o wiele więcej różnic genetycznych i fizycznych można spotkać w obrębie grup niż pomiędzy nimi (Betancourt, López 1993; Jones 1991; Zuckerman 1990). Oczywiście, sposób określania i identyfikowania siebie i innych ma ogromny wpływ na zachowania człowieka, o czym przekonamy się, omawiając podejście społeczno-kulturowe. Rasizm — nienawiść lub nietolerancja wobec innych z powodu rozmaicie pojmowanej rasy — stanowi prawdziwy bolesny problem ludzkości. Ponieważ pojęcie rasy błędnie zakłada istnienie fizycznie jednolitych i odrębnych grup, z naukowego zaś punktu widzenia wydaje się problematyczne, w naszej książce staramy się unikać pojęcia rasy. W zamian posługujemy się terminem „grupa etniczna” w odniesieniu do osób mających wspólną kulturę, religię lub język.

Spojrzenie społecznego konstruktywizmu na naukę wywołuje wiele refleksji i dyskusji we wszystkich dziedzinach badań. Każda dziedzina musi weryfikować swoje założenia i zastanawiać się, w jaki sposób owe założenia odzwierciedlają się w języku (Minnich 1990). Renesans definiuje się, na przykład, jako okres artystycznego i intelektualnego odrodzenia po średniowieczu, jednak był to czas odrodzenia wyłącznie dla ludzi bogatych. Jak zauważył jeden z historyków, nie było renesansu dla kobiet — w każdym razie, nie w okresie renesansu. Podobnie dopóki historię definiuje się jako wiedzę o królach, bitwach i przywódcach politycznych, dopóty pomijane są doświadczenia zwykłych ludzi, biedaków i kobiet.

konstruktywizm społeczny
Pogląd, że nie istnieje uniwersalna prawda na temat ludzkiej natury, ponieważ ludzie rozmaicie konstruują rzeczywistość, w zależności od kultury, momentu historycznego i układu sił w społeczeństwie.

W psychologii debata konstruktywistów społecznych stanowi szczególne wyzwanie. Celem psychologów było zawsze zrozumienie zachowań i procesów psychicznych człowieka. Teraz każe się im sprawdzać, w jaki sposób ich własne wartości, płeć, pozycja społeczna i doświadczenia kulturowe wpływają na ich teorie. Zadanie to mogą realizować za pomocą różnorodnych metod naukowych, w szczególności wykorzystując osiągnięcia dwóch podejść psychologicznych: poznawczego i społeczno-kulturowego. Psychologowie poznawczy wykazują, że emocje i zachowania człowieka często zależą bardziej od interpretacji zdarzeń niż od jakiegoś rzeczywistego stanu rzeczy. Psychologowie społeczno-kulturowi twierdzą, że głęboki wpływ na sposób widzenia świata i wnioski na temat rzeczywistości mają role społeczne i kulturowe.

Wielu ludzi, również sami psychologowie, z niepokojem przyjmuje poglądy postmodernistyczne. Obawiają się, że ów krytycyzm oznacza, że nigdy nie poznamy prawdy o czymkolwiek i równie dobrze możemy zrezygnować z nauki, jednolitych standardów krytycyzmu i prób wprowadzenia obiektywnych metod naukowych. Obawiają się, że uzasadnione zastrzeżenia wobec ograniczeń badań naukowych przerodzą się w atak na wszystkie tradycyjne metody badawcze, co w rezultacie stanowi wylewanie dziecka (pożytecznych osiągnięć) z kąpielą (wypaczeniami i krótkowzrocznością) (Peplau, Conrad 1989; Smith 1994). Niepokoi ich również, że filozofia postmodernistyczna może doprowadzić do skrajnego relatywizmu w osądach moralnych i etycznych.

Jest to poważna i ciekawa debata, która długo jeszcze nie zostanie zakończona, dlatego przywołujemy ją w tym miejscu. Nasze stanowisko, któremu dajemy wyraz w podejściu do psychologii w niniejszej książce, sytuuje się pomiędzy skrajnym tradycjonalizmem a skrajnym postmodernizmem. Uważamy, że zrozumienie, w jaki sposób nauka i wiedza jest konstruowana przez wykładowców i uczonych, ma zasadnicze znaczenie dla studiowania psychologii, co staramy się pokazać w podsumowaniu każdego z podejść. Nowe spojrzenie na wiedzę i badania naukowe daje możliwość rozszerzenia i wzbogacenia naszego rozumienia zachowań (Gergen 1994). Akcentowanie w historii społecznej doświadczeń zwykłych ludzi nie oznacza, że dawniejsze zainteresowanie królami i politykami było błędne — było jedynie niekompletne i mylące. Rozwój kryteriów oceny wielkiej sztuki może jedynie zwiększyć nasz podziw i uznanie dla osiągnięć ludzi w każdej epoce i kulturze. Zakwestionowanie dawnych założeń umożliwia stworzenie nowego i pełniejszego obrazu.

Dla nas jednak, jak dla wszystkich badaczy, pewne aspekty badań psychologicznych pozostaną niezmienione. Chodzi przede wszystkim o wymagania wobec dowodów naukowych, opieranie się na weryfikowalnych wynikach i akcentowanie krytycznego myślenia. Dlatego mamy nadzieję, że czytając kolejne rozdziały, czytelnik nie ulegnie pokusie pomijania opisów wykonania badań. Jeśli założenia i metody są nieprawidłowe, to nieprawidłowe są również wyniki i oparte na nich wnioski. To, co wiemy o zachowaniu człowieka, nie da się ostatecznie odzielić od tego, skąd o tym wiemy.