

Kazimierz Czerwiński

Uniwersytet Kazimierza Wielkiego w Bydgoszczy

Metodologiczne aspekty pomiaru postaw jako wypadkowej systemu wartości

Uwagi wstępne

Przedmiotem tego tekstu jest analiza możliwości zastosowania opisanych w literaturze metodologicznej sposobów pomiaru postaw podczas prowadzenia badań empirycznych, ich ewentualnej adaptacji czy modyfikacji, jak również możliwości tworzenia przez badaczy własnych, oryginalnych narzędzi do pomiaru postaw, wykorzystujących zasady użyte przy konstrukcji narzędzi niejako „wzorcowych”.

Pomiar postaw wynikających z systemu wartości napotyka na te same problemy, co pomiar postaw w ogóle, wydaje się jednak, że w przypadku tych pierwszych pojawiać się mogą dodatkowe trudności. Postawy dotyczące kwestii określanych jako „wartościujące” silniej niż jakiegokolwiek inne bywają uzależnione od aktualnej sytuacji społecznej, materialnej, emocjonalnej, a ta potrafi zmieniać się – niejednokrotnie w sposób zaskakujący. W przypadku tego typu postaw tak zwany komponent poznawczy (o czy będzie mowa dalej) może być w wysokim stopniu kształtowany i modyfikowany przez rozpowszechnione w danym środowisku potoczne opinie i przekonania, niejednokrotnie wręcz stereotypy. Stwarza to trudność w oddzieleniu tego co „własne” od tego, co jest tylko bezrefleksyjnie zapożyczony – bo „tak wypada” sądzić. Komponent poznawczy postawy bywa po prostu nieadekwatny. Percepcja sytuacji społecznych silniej oddziałuje na zachowania niż obiektywna rzeczywistość (Zimbardo, Leippe, s. 63).

Na ten aspekt sprawy zwraca też uwagę J. Chmielewski w internetowym artykule poświęconym kształtowaniu postaw prospołecznych. „W wychowaniu wyróżnić można dwa nurty: instrumentalny (umysłowy), przekazujący atrybuty wiedzy (społecznej, politycznej, prawnej, gospodarczej) i budowanych na jej bazie umiejętności, oraz nurt kierunkowy (moralny) – kształtujący postawy. Nurt umysłowy tworzy fundament rozwoju nurtu moralnego. Wiedza stanowi bowiem podłoże do kształtowania postaw – zasadniczego celu wychowania obywatelskiego. Swoista triada postaw – od ‘ja’ (moje szczęście) przez ‘także ty’ (twoje pra-

wo) i 'my' (wspólne dobro) stanowi istotę życia człowieka w jego człowieczeństwie. Zadaniem wychowania jest przede wszystkim kształcenie postaw prospołecznych, czyli postawy 'także ty' oraz, w szczególności, postawy 'my'" (Chmielewski, s. 1). Określony poziom wiedzy jest niezbędnym warunkiem, by w ogóle przejawiała się jakakolwiek postawa, chodzi jednak głównie o wiedzę adekwatną. Proces kształtowania postaw odczytywanych jako wartościowe musi być ściśle powiązany z nabywaniem przez podmiot określonych – adekwatnych – informacji.

Podstawowy problem z pomiarem jakichkolwiek postaw polega na tym, że próbuje się uczynić mierzalnym to, co z natury swej jest „niemierzalne”, a przede wszystkim zaś – ukryte przed bezpośrednią obserwacją. Ambicją każdego badacza dokonującego pomiaru jest, by zmienna którą mierzy była rejestrowana na jak najwyższej skali pomiarowej. Tymczasem, jak to wynika z dalszych partii tekstu, postawy da się mierzyć z reguły co najwyżej na poziomie porządkowym. Stanowić to może pewne utrudnienie podczas analizy i interpretacji wyników.

Postawa i jej składniki

Postawa bywa definiowana rozmaicie. Oto kilka definicji: „trwała umysłowa reprezentacja ludzi, miejsc lub obiektów powodująca reagowanie pozytywnymi lub negatywnymi emocjami i wpływająca na zachowanie” (Rathus, s. 769); „trwała ocena – pozytywna lub negatywna – ludzi, obiektów i idei” (Aronson, Wilson, Akert, s. 313); „ogólne ukierunkowanie, które może sprawiać, że dana osoba – gdy pojawią się określone bodźce – zachowuje się lub reaguje w określony sposób” (Frankfort-Nachmias, Nachmias, s. 269); „wartościujące nastawienie oparte na elementach poznawczych, reakcjach uczuciowych oraz intencjach co do przyszłości i na zachowaniu” (Zimbardo, Leippe, s. 52); „gotowość, skłonność do reagowania w określony sposób na określone bodźce” (Oppenheim, s. 203); „względnie trwała, dynamiczna organizacja, która wyznacza sposób zachowania się jednostki względem innych ludzi lub przedmiotów, struktur i procesów poznawczych, emocjonalnych oraz schematów zachowania, którą charakteryzuje różny stopień złożoności, zwartości, siły, trwałości, natężenia, ważności i adekwatności” (Holly, s. 215).

Z pobieżnego choćby przeglądu powyższych definicji wynika, że różni autorzy – kładąc nacisk na odmienne aspekty postawy oraz inaczej rozpoczynając *definiens*: „trwała umysłowa reprezentacja”, „trwała ocena”, „ogólne ukierunkowanie”, „wartościujące nastawienie”, „skłonność do reagowania”, „względnie trwała organizacja” – zgadzają się przynajmniej w

tym, że postawa jest czymś, co jest względnie trwałe i co w określonych sytuacjach determinuje zachowanie.

Na postawę składają się trzy komponenty: emocjonalny, poznawczy i behawioralny. Komponent emocjonalny to emocje i uczucia kojarzone z obiektem postawy, komponent poznawczy to myśli i przekonania na temat właściwości obiektu postawy, zaś komponent behawioralny oznacza dające się zaobserwować działania wobec obiektu postawy (Aronson, Wilson, Akert, s. 314). Inaczej można wyrazić to w ten sposób, że postawy ulegają wzmocnieniu przez przekonania (składnik poznawczy), związane są najczęściej z silnymi uczuciami (składnik emocjonalny) i generują określone zachowania (składnik behawioralny, inaczej: składnik skłonności do działania) (Oppenheim, s. 204).

Nieco inaczej ujmują to P.G. Zimbardo i M.R. Leippe, którzy – oprócz wymienionych wyżej trzech komponentów – dodają intencje zachowań. Autorzy ci używają szerszego pojęcia, jakim jest system postawy (albo inaczej: system reakcji, charakteryzujących daną jednostkę), na który składają się: zachowania, intencje zachowań, elementy poznawcze, reakcje uczuciowe i wreszcie sama postawa, będąca swojego rodzaju sumą pozostałych elementów. „[...] po pierwsze mamy samo zachowanie [...]. Po drugie są intencje zachowań, oczekiwania lub plany zachowań w określony sposób, zanim rzeczywiście dojdzie do zachowania – czasami takie plany nigdy nie są wdrażane [...]. Po trzecie są idee kierujące naszym działaniem, nasze przekonania lub – szerzej – elementy poznawcze obejmujące zarówno przekonania, jak i fragmenty wiedzy o obiektach oraz wskazania, jak ‘powinno się’ zachować wobec nich. Po czwarte są reakcje uczuciowe, czyli emocje, czy też raczej uczucia ‘w głębi serca’, odzwierciedlające nasze postawy na płaszczyźnie fizycznego pobudzenia (odczuwanie przyjemności, smutku itp.). I po piąte, jest też sama postawa, a więc ogólna, łączna ocena uwzględniająca inne jej elementy składowe” (Zimbardo, Leippe, s. 51–53).

Poszczególne składniki owego systemu są ze sobą ściśle powiązane. Wymienieni autorzy używają przy tym pojęcia „umysłowa reprezentacja obiektu” na oznaczenie kompleksu składającego się z samej postawy i związanych z nią elementów poznawczych. Ważniejszą zaś sprawą jest to, że zmiana w dowolnym komponencie całego systemu postaw zazwyczaj prowadzi do zmian w pozostałych składnikach (s. 54).

Postawy mogą kształtować procesy percepcyjne i poznawcze. „Chociaż prawdą jest, że to nasze myślenie kształtuje postawy, prawdą będzie też stwierdzenie, iż proces ten przebiega w przeciwnym kierunku” (s. 55). Zdarza się, że postawy, które pierwotnie miały wąty komponent poznawczy, mogą wpłynąć na potrzebę pozyskania wiedzy na dany temat i ukształtowanie przekonań, co wypełni istniejącą lukę.

Postawy są łatwo dostępne. „Jako uogólnione oceny relatywnie łatwo przychodzą nam na myśl. Ludzie nie mają czasu ani możliwości umysłowych, by starannie przemyśleć sobie niezliczone bodźce i sytuacje, jakie napotykają w życiu. W obliczu decyzji dotyczących obiektów społecznych, szczególnie decyzji o znaczeniu drugorzędym, raczej nie dokonujemy gruntownego przeglądu każdego przekonania i faktu odnoszącego się do tegoż obiektu. Poszukujemy za to jakiegoś konkretnego, wierzymy w uogólnioną postawę i dajemy jej sobą kierować” (s. 56).

Postawy mają znaczenie autodefinicyjne. Informują otoczenie, a także poniekąd nas samych, jakie mamy stanowisko oraz w jaki sposób jesteśmy skłonni się zachować w stosunku do wielu ważnych kwestii. Mają charakter znaku rozpoznawczego, odróżniającego nas od innych. W dużej więc mierze określają kim jesteśmy. „Jesteśmy sumą wszystkich naszych postaw” (s. 56).

Autor hasła „postawa” i haseł powiązanych, zamieszczonych w *Słowniku psychologicznym*, Romuald Holly szczegółowiej charakteryzuje podstawowe kategorie opisu i pomiaru postaw. Wymienia następujące: treść przedmiotowa, stopień ogólności, ważność, kierunek, siła, trwałość, natężenie, adekwatność, złożoność oraz zwartość.

Treść przedmiotowa postawy to cecha, która wskazuje, czego postawa dotyczy. Określenie „przedmiot” jest rozumiane szeroko. Może to być osoba, grupa osób, idea, pojęcie, teoria, sytuacja itp., ważne z punktu widzenia zaspokajania ludzkich potrzeb. W tym kontekście wielu autorów używa szerokiego określenia: obiekt postawy. Stopień ogólności postawy oznacza ilość dotyczących jej desygnatów. Można zatem mówić o postawach wobec osób, grup osób (np. kibiców), narodów, ras – każda kolejno wymieniona kategoria jest bardziej ogólna. Ważność to znaczenie danej postawy w życiu konkretnego człowieka. Z tego punktu widzenia wyróżnia się postawy znaczące (centralne) i nieznaczące (peryferyjne). Kierunek postawy to inaczej jej „znak”: pozytywny lub negatywny. Siła określa, w jakim stopniu postawa jest na przykład pozytywna: „zdecydowanie pozytywna”, „pozytywna”, „umiarkowanie pozytywna” – to tylko przykładowe etykiety na określenie siły postawy. Siła może też być określana liczbowo – uzależnione to jest od zastosowanego sposobu pomiaru danej postawy. Trwałość to niezmiennosc przekonań, reakcji emocjonalnych i zachowań wobec przedmiotu postawy. Postawy mogą więc być mniej lub bardziej trwałe. Natężenie postawy, inaczej intensywność, to częstotliwość występowania charakterystycznych dla danej postawy zachowań, treści poznawczych oraz związanych z nimi emocji. Z tego punktu widzenia można mówić o postawach „gorących” – gdy częstotliwość wymienionych reakcji jest względnie duża, oraz o „zimnych” – w przypadku małej ich częstotliwości. Adekwatność to cecha po-

zwalająca (dychotomicznie) klasyfikować postawy jako adekwatne i nieadekwatne – w oparciu o arbitralnie przyjęte kryterium. Zrozumiałe jest, że takiej klasyfikacji dokonuje raczej zewnętrzny obserwator, bowiem dla „posiadacza” postawy jest ona zazwyczaj adekwatna. Jako przykład postawy nieadekwatnej można wskazać negatywne postawy niektórych dzieci wobec nauki szkolnej. Złożoność postawy to cecha, która informuje o ilości tworzących ją składników. Zwartość zaś to stopień wzajemnej zgodności pomiędzy komponentami tworzącymi postawę (Holly, s. 216–217).

Podstawy pomiaru

Pomiar jest czynnością (lub zespołem czynności), która przyporządkowuje jakiejś cenniejsze liczbę – według z góry określonego kryterium (Mayntz, Holm, Hübner, s. 60). Mówiąc inaczej, efektem pomiaru są dane, które stanowią wynik pojęciowo ustrukturyzowanych, systematycznych i kontrolowanych obserwacji określonych cech osób badanych (s. 46). Sprawa jest prosta, jeśli przedmiotem pomiaru mają być cechy dające się względnie łatwo obserwować, inaczej wszakże ma się rzecz z postawami. Postawy z natury rzeczy są „ukryte” przed zewnętrznym obserwatorem. By je zmierzyć, potrzebujemy wskaźników.

Według rozpowszechnionej na gruncie polskiej metodologii definicji, autorstwa S. Nowaka: „Wskaźnik zdarzenia (własności) *Z* to takie zdarzenie (taka własność) *W*, że stwierdzenie jego (jej) istnienia, pojawienia się lub stopnia intensywności bądź faktycznie jest wykorzystane jako przesłanka, bądź zasadnie nadaje się na przesłankę wnioskowania, iż w określonych przypadkach z pewnością, z określonym prawdopodobieństwem lub przynajmniej z prawdopodobieństwem wyższym niż przeciętne wystąpiło zdarzenie (własność) *Z*” (Nowak, s. 165). Dalej autor pisze: „[...] na wskaźniki w badaniach nadają się przede wszystkim zjawiska, stany rzeczy i zdarzenia (zachowania) łatwo dostępne obserwacji, i to takie, których właściwa psychologiczno-kulturowa interpretacja rozumiejąca jest stosunkowo niezawodna” (s. 166).

Związek między wskaźnikiem *W* a zjawiskiem wskaźnikowanym *Z* (w naszym przypadku – postawą) może mieć różny charakter. Najczęściej wymienia się trzy przypadki:

- *W* jest konsekwencją (skutkiem) *Z*,
- *W* jest warunkiem (przyczyną) *Z*,
- między *W* i *Z* nie zachodzi zależność przyczynowo-skutkowa, ma miejsce jedynie statystyczna korelacja (Mayntz, Holm, Hübner, s. 54).

W przypadku, gdy zjawiskiem wskaźnikowanym Z jest postawa, może mieć miejsce pierwszy z przypadków, a to wtedy, gdy jako wskaźniki będziemy traktować określone zachowania. Częściej wszakże jako wskaźników postaw używa się wypowiedzi osób badanych, na przykład pod postacią ustosunkowania się przez nie do określonych stwierdzeń. W tym przypadku należy przyjąć, że między wskaźnikiem a postawą ma miejsce związek o charakterze statystycznym.

Wskaźniki dzielą się na definicyjne i rzeczowe (realne), te drugie zaś na empiryczne i inferencyjne (Nowak, s. 166–170). Wskaźnik definicyjny jest po prostu definicją zjawiska wskazywanego. „Kiedy powiadamy, że ilość posiadanych pieniędzy jest dla nas wskaźnikiem ‘zamożności’, to w gruncie rzeczy myślimy, iż termin ‘zamożność’ znaczy ‘ilość posiadanych pieniędzy’. Tak samo wskaźnikiem ‘przestępczości’ bywa liczba rocznie popełnianych przestępstw w jakimś kraju czy regionie” (s. 166). Jeśli zaś związek między wskaźnikiem a cechą wskaźnikowaną nie jest oparty tylko na konwencji terminologicznej, lecz ma charakter realny, podlegający (bezpośredniej lub pośredniej) empirycznej kontroli, to mamy do czynienia ze wskaźnikami rzeczowymi. Jeśli oprócz wskaźnika również zjawisko wskaźnikowane jest obserwowalne, to mówimy o wskaźniku empirycznym, zaś gdy cecha wskaźnikowana jest nieobserwowalna, wskaźnik określamy jako inferencyjny. W tym ostatnim przypadku jedyne co możemy zrobić, to na podstawie wskaźnika (wskaźników) wnioskować (inferować) o interesującej nas cesze psychicznej. Postawa jest taką właśnie cechą, która nie poddaje się bezpośredniej obserwacji, możemy jedynie o niej wnioskować na podstawie obserwowalnych wskaźników; będą więc one miały charakter wskaźników inferencyjnych.

Analizując wymienione przez R. Holly’ego kategorie opisu postaw można zauważyć, że nie wszystkie z nich są jednakowo łatwo „uchwytny” dla pomiaru. Treść przedmiotowa postawy bez wątpienia jest ściśle związana z narzędziem do jej pomiaru – to, czy dane narzędzie mierzy to, do czego zostało skonstruowane, jest przedmiotem troski takiego aspektu narzędzia, jakim jest jego trafność. Jeśli narzędzie pomiarowe jest trafne, to mierzy postawę wobec tego obiektu, o który chodziło badaczowi. Podobnie jest ze stopniem ogólności. Siła postawy i jej znak są tymi wymiarami, co do których względnie łatwo uzyskamy informacje w wyniku pomiaru. Gorzej rzecz się ma z ważnością postawy dla konkretnego respondenta, z jej natężeniem, złożonością i zwartością. Typowe narzędzia ilościowe do pomiaru postaw – indeksy i skale – ignorują te wymiary. Wydaje się, że dla ustalenia tych wymiarów wskazane jest uzupełnianie tradycyjnych pomiarów ilościowych jakimiś badaniami zorientowanymi „jakościowo”, na przykład wywiadami pogłębionymi. Istnieje natomiast możliwość uchwycenia trwałości postawy – ale jedynie na podstawie kilkakrotnych pomiarów, przeprowadzo-

nych w jakichś odstępach czasu. Dzięki takim badaniom „podłużnym” możliwe jest zresztą coś więcej – uchwycenie dynamiki postaw.

Indeksy i skale

W badaniu postaw, jak zresztą w ogóle podczas pomiaru większości zmiennych w naukach społecznych, w zasadzie nie występują sytuacje, w których dla danej zmiennej potrafilibyśmy znaleźć pojedynczy, jednoznaczny wskaźnik. Wręcz przeciwnie – typowa sytuacja pomiarowa wymaga utworzenia jakiegoś miernika, na który składa się większa liczba elementarnych wskaźników. Dzieje się tak z dwóch powodów. Po pierwsze, badane zjawisko (na przykład postawa) jest wieloaspektowe i każdy z aspektów wymaga odrębnego potraktowania. Po drugie, jeśli nawet w pojedynczych przypadkach dałoby się znaleźć jedno trafne pytanie, na które odpowiedź miałaby być wskaźnikiem danej zmiennej, to zazwyczaj ma ono niewystarczający zakres zmienności; inaczej mówiąc: nie ma wystarczającej mocy różnicowania osób różniących się siłą postawy. Obie te niedogodności próbują przezwyciężyć tak zwane mierniki złożone, to jest indeksy i skale (Zob. np. Babbie, s. 173; Nowak, s. 182–189; Mayntz, Holm, Hübner, s. 52–55).

Zarówno indeksy, jak i skale są porządkowymi miernikami zmiennych. Wynik danej osoby stanowi jedynie wskazanie co do poziomu danej zmiennej względem innych osób – i tylko tak można interpretować otrzymaną wartość liczbową, nie zaś w sensie absolutnym. Kolejne podobieństwo polega na tym, że (zdaniem E. Babbie’go) każdy z tych mierników powinien odzwierciedlać tylko jeden wymiar mierzonego pojęcia. Poza tym jednak oba rodzaje mierników różnią się od siebie. Indeks tworzy się w zasadzie przez proste zsumowanie wyników przypisanych odpowiedziom na poszczególne pytania, zaś skalę – przypisując określony wynik punktowy pewnym układom odpowiedzi, przy założeniu, że niektóre pytania świadczą o relatywnie niższym, inne zaś o wyższym poziomie danej zmiennej (w naszym przypadku – postawy) (Babbie, s. 174).

Różnicę w podejściu podczas konstrukcji indeksu i skali dobrze jest rozpatrzeć na przykładzie. Taki przykład zamieszcza E. Babbie. Załóżmy, że chcemy stworzyć miernik aktywności politycznej, który dzieliłby badane osoby według stopnia aktywności: od bardzo aktywnych aż po osoby w ogóle nie biorące udziału w życiu politycznym. Podczas konstruowania indeksu bierzemy pod uwagę pewną liczbę zachowań, z których każde – mniej więcej w jednakowym stopniu – świadczy o aktywności politycznej. W omawianym przykładzie są to: napisanie listu do urzędnika państwowego, podpisanie petycji politycznej, dawanie pienięd-

dzy w sprawie politycznej, dawanie pieniędzy na kampanie wyborcze, napisanie listu o treści politycznej do redakcji gazety, przekonywanie kogoś do zmiany preferencji wyborczych. Przyznając każdej badanej osobie po jednym punkcie za każde z podjętych działań dokonuje się klasyfikacji badanych na siedem grup: od bardzo aktywnych politycznie (6 punktów) po zupełnie nie interesujących się polityką (zero punktów). Logika budowy skali jest natomiast taka, że przewidziane jako wskaźniki działania (zachowania) nie w jednakowym stopniu świadczą o aktywności politycznej, wręcz przeciwnie, mają charakter kumulatywny. W omawianym przykładzie są to: głosowanie w wyborach, dawanie pieniędzy na kampanię wyborczą, praca przy organizacji kampanii wyborczej, kandydowanie w wyborach. „[...] założenie istnienia pewnych prawidłowości działań wydaje się tutaj zupełnie bezpieczne. Na przykład wszyscy, którzy dali pieniądze, najprawdopodobniej również głosowali. Ci, którzy pracowali przy kampanii wyborczej, prawdopodobnie dali jakieś pieniądze i głosowali” (s. 174). Omawiany przykład jest po pierwsze „zanurzony” w realiach USA, po drugie zaś dla celów „dydaktycznych” jest mocno uproszczony, chodziło w nim tylko o pokazanie zasady; rzeczywiste mierniki aktywności politycznej byłyby zapewne bardziej złożone i chwytające więcej aspektów owej aktywności.

Trochę inne rozumienie indeksu – częściowo w opozycji do powyższego stanowiska – znajdziemy w książce autorów niemieckich: R. Mayntz, K. Holma i P. Hübnera. Jako przykład podają zmienną: „zadowolenie robotników z miejsca pracy”. Założenie jest takie, że uznajemy tę zmienną za dwuwymiarową, gdzie tymi wymiarami (składnikami) są: „zadowolenie ze stanowiska pracy” oraz „zadowolenie z płacy”. Zakładamy dalej, że każdy z tych wymiarów byłby mierzony na trzystopniowej skali porządkowej, określonej jako zadowolenie: niskie, średnie, wysokie. Następnie w każdym z wymiarów przyporządkowuje się (umowne) wartości punktowe poszczególnym poziomom zadowolenia: „niskie” = 0 punktów, „średnie” = 1 punkt oraz „wysokie” = 2 punkty. Następnie sumuje się te zmienne, redukując niejako oba wyróżnione wymiary do jednego – zadowolenia z miejsca pracy. Otrzymuje się więc następującą punktację sumaryczną:

- „niskie”, „niskie”: 0 punktów;
- „niskie”, „średnie”: 1 punkt (bez rozróżniania, z którego wymiaru zadowolenie jest niskie, a z którego średnie);
- „niskie”, „wysokie” albo „średnie”, „średnie”: 2 punkty;
- „średnie”, „wysokie”: 3 punkty;
- „wysokie”, „wysokie”: 4 punkty.

W omawianym przykładzie były tylko dwa wymiary, zasada pozostaje jednak niezmienną przy większej ich liczbie, chociaż „technicznie” może stać się to nieco skomplikowane. Kolejne założenie, jakie tu uczyniono jest takie, że wskazane wymiary są jednakowej wagi. Jeśli są jakieś przesłanki teoretyczne, by móc uznać wyróżnione wymiary za nierównoważne, należy nadać im odpowiednie „wagi” liczbowe. Na przykład, jeśli można uzasadnić, że na zadowolenie z pracy w danym zakładzie dwukrotnie większy wpływ ma zadowolenie z płacy, niż zadowolenie z konkretnego miejsca pracy, to umowne wartości dotyczące zadowolenia z płacy można (przed dokonaniem sumowania) pomnożyć przez 2 i zamiast liczb 0, 1 i 2 użyć do budowania indeksu liczb 0, 2 i 4. Dodatkowo milcząco zakłada się przy tym, że wskazane wymiary są niezależne (s. 56–59). „Przy tworzeniu indeksu trzeba zatem rozwiązać dwa zadania. Po pierwsze, rozstrzygnąć, dane dotyczące których cech cząstkowych mają wejść w skład indeksu. Ważne jest przy tym uwzględnienie wszystkich komponentów badanej cechy wielowymiarowej, ale zarazem tylko tych komponentów. Po drugie, wskazać, jak dane te należy ze sobą powiązać (przyporządkowanie wartości liczbowych, ważenie, kombinacje)” (s. 59). Ponieważ E. Babbie używa na swój przykład określenia „prosty indeks”, przez analogię propozycję opisaną przez trójkę autorów niemieckich można nazwać „indeksem złożonym”.

W odróżnieniu od – tak czy inaczej rozumianych – indeksów typowe skale, jak wspomniano, konstruuje się z założeniem, że poszczególne jednostki (pytania, stwierdzenia) skali mogą mieć różne wagi, ponieważ w niejednakowym stopniu wskazują natężenie danej postawy. Tym samym każda elementarna jednostka skali powinna być inaczej punktowana. Największą trudnością zaś jest sposób ustalenia owej zróżnicowanej punktacji. Istnieją pewne propozycje, związane z nazwiskami ich autorów, które odmiennie rozwiązują ów problem.

Skala Thurstone’a (Oppenheim, s. 220–226; Mayntz, Holm, Hübner, s. 69–70; Babbie, s. 191) postuluje rozwiązanie tego problemu za pomocą „sędziów kompetentnych”. Procedura jest wieloetapowa i rozpoczyna się od wyboru – na podstawie studiów z literatury przedmiotu oraz wywiadów pogłębionych – dużej liczby stwierdzeń, które mają stanowić „materiał” do budowy skali. Tych stwierdzeń powinno być znacznie więcej, niż przewidywana ich ostateczna liczba. Następnie te stwierdzenia prezentuje się dużej grupie (kilkudziesięciu, rzadziej kilkuset) osób, którzy pełnią tutaj rolę sędziów kompetentnych. Mogą to być eksperci z danej dziedziny, lecz stosuje się również takie rozwiązanie, że ci „sędziowie” to po prostu losowo dobrane osoby, nie różniące się istotnie od tych, dla których przeznaczona jest skala.

Każda osoba otrzymuje w porządku losowym zestaw kartek z poszczególnymi stwierdzeniami, jest również poinformowana, co mają mierzyć te stwierdzenia. Następnie każda z nich jest proszona, by każde stwierdzenie usytuowała – wedle swego przekonania – w jednym

z przedziałów przygotowanej wcześniej (liczącej zazwyczaj 11 odcinków) skali. Do pierwszego przedziału trafić mają stwierdzenia najbardziej przychylne, do drugiego – nieco mniej przychylne i tak dalej. W przedziale środkowym (szóstym) mają się znaleźć stwierdzenia neutralne, w siódmym – lekko negatywne, aż do jedenastego, gdzie trafią stwierdzenia najbardziej negatywne. Po zakończeniu procesu oceniania każdy sędzia numeruje każde stwierdzenie według numeru przedziału, do którego je zaliczył. Rola „sędziów” na tym się kończy, dalsze czynności należą do badacza – konstruktora skali.

Po pierwsze ocenia (oblicza za pomocą jednej z miar: wariancja, odchylenie standardowe, odchylenie środkowe itp.) rozrzut, czy inaczej rozproszenie numeracji każdego ze stwierdzeń. Im jest on większy, tym bardziej wieloznaczne (niejednako rozumiane przez poszczególne osoby) jest dane stwierdzenie. Stwierdzenia charakteryzujące się największym rozrzutem zostają oczywiście wykluczone, pozostają te, co do których sędziowie okazali się być w miarę zgodni. Kolejnym krokiem jest uporządkowanie pozostałych stwierdzeń według średnich (według innej wersji – median) przyznanych przez sędziów numeracji. Spośród uporządkowanych stwierdzeń wybiera się pewną ich liczbę, kierując się zasadą, by po pierwsze różnica między skrajnymi wartościami była możliwie największa (aby stwierdzenia „pokrywały” jak największą część *continuum* postawy), a po drugie – by interwały (różnice) między kolejno po sobie następującymi stwierdzeniami były (w miarę możliwości) równe. Stąd konstrukcja tej skali bywa też w literaturze zwana metodą równych interwałów Thurstone’a.

Pomiar za pomocą tej skali polega na tym, że respondentom prezentuje się wydrukowany zestaw losowo uporządkowanych stwierdzeń skali z prośbą o wyrażenie aprobaty bądź dezaprobaty dla każdego z nich. Następnie bierze się pod uwagę tylko te stwierdzenia, z którymi respondent się zgodził oraz oblicza się średnią punktów (tych samych punktów, które zostały obliczone jako średnie wskazań sędziów w procesie konstruowania skali). Im większa liczba, tym bardziej negatywna postawa i na odwrót – im mniejsza, tym postawa bardziej pozytywna.

Otrzymanych liczb nie można oczywiście traktować w sposób absolutny, nie ma sensu na przykład sądzić, że osoba otrzymująca 10 punktów ma dwa razy bardziej negatywną postawę niż osoba z pięcioma punktami. Otrzymane punkty powinny być traktowane jedynie jako podstawa do uporządkowania przebadanych osób – od najbardziej pozytywnej do najbardziej negatywnej. Trzeba tu dodać, że o ile większość skal tak właśnie jest traktowana przez ich twórców oraz użytkowników – jako skale porządkowe, o tyle L. L. Thurstone stał na stanowisku, że zaproponowana przez niego skala – a to dzięki sposobowi konstrukcji zapewniającemu „równe interwały” – powinna być traktowana jako skala interwałowa. Tak więc, jeśli na

przykład osoba A otrzymała 2,6 punktu, osoba B: 3,8 punktu, osoba C: 6,3 punktu oraz osoba D: 7,5 punktu, to twórca skali byłby skłonny sądzić, że dystans (różnica natężenia postawy) między osobami D i C jest taki sam, jak między osobami B i A, ponieważ w obu przypadkach różnica jest ta sama (i wynosi 1,2 punktu). Jednakże wielu autorów krytykuje ten „optymizm metodologiczny”, wskazując chociażby nie na faktyczną równość interwałów, lecz na subiektywną ich równość. Unikając głębszego rozpatrywania tych sporów, bezpieczniej będzie przyjąć, że jest to skala porządkowa.

Skala Likerta (Oppenheim, s. 226–232; Mayntz, Holm, Hübner, s. 70–75; Frankfort-Nachmias, Nachmias, s. 479–481; Babbie, s. 191–192) konstruowana jest w sposób, który nie wymaga „sędziów kompetentnych”. Podobnie jak przy skali Thurstone’a przygotowuje się długą listę stwierdzeń, które posłużą jako materiał do sporządzenia ostatecznej wersji. Osoby badane (w liczbie kilkudziesięciu – kilkuset) proszone są o ustosunkowanie się do każdego z nich poprzez wybór jednej z kategorii odpowiedzi, zazwyczaj są to: „zdecydowanie się zgadzam”, „zgadzam się”, „nie mam zdania”, „nie zgadzam się”, „zdecydowanie się nie zgadzam” (w zależności od brzmienia stwierdzeń mogą to być również inne wersje, na przykład: „tak”, „raczej tak”, „nie mam zdania”, „raczej nie”, „nie”, albo też: „prawie zawsze”, „często”, „czasami”, „rzadko”, „prawie nigdy” itp.). Najczęściej stosuje się 5 kategorii, choć możliwa jest inna ich liczba, na przykład 3 lub 7; z reguły jest to liczba nieparzysta, z kategorią „obojętną” pośrodku.

Do każdego stwierdzenia przygotowana jest wcześniej umowna punktacja za wybór określonej kategorii odpowiedzi. Może to być na przykład punktacja od 1 do 5, albo też od 0 do 4, równie dobrze od –2 do 2 (ten ostatni wariant może jednak mylnie sugerować, że wskazanie odpowiedzi środkowej typu „nie mam zdania” zawsze oznacza ambiwalentną postawę). Warto również pamiętać, że wśród przygotowanych stwierdzeń jedna część powinna być tego typu, że identyfikacja z nimi oznacza pozytywną postawę, druga zaś część przeciwnie – brak identyfikacji oznacza postawę pozytywną. Oba rodzaje stwierdzeń muszą być punktowane w przeciwnych kierunkach, jeśli za poszczególne kategorie odpowiedzi dla stwierdzeń „pozytywnych” przyznamy odpowiednio: 1, 2, 3, 4 i 5 punktów, to dla stwierdzeń „negatywnych” punktacja musi być odwrócona: od 5 do 1. Oblicza się następnie łączną liczbę punktów uzyskanych przez każdą osobę.

Kolejny krok to eliminacja tych stwierdzeń, które nie są zgodne z „duchem” pozostałych, które zakłócają wewnętrzną zgodność skali. Dla R. Likerta podstawową kwestią była jednowymiarowość skali – chodziło mu o to, by wszystkie stwierdzenia we w miarę jednokowym stopniu dotyczyły tego samego obiektu, a nie jakiegoś innego. Można to wykonać na

kilka sposobów, jednym z nich jest tak zwana analiza mocy dyskryminacyjnej każdego stwierdzenia skali. Osoby badane porządkuje się według uzyskanej liczby punktów, następnie wyróżnia się spośród nich górny i dolny kwartył (inaczej: 25% respondentów, którzy uzyskali najwyższe wyniki oraz 25% tych, którzy uzyskali wyniki najniższe; niektórzy autorzy zalecają raczej 27%). Następnie dla każdego stwierdzenia osobno oblicza się średnie wyniki odpowiednio: w górnym i dolnym kwartyłu. Jeśli różnica tych średnich miałaby znak ujemny, świadczyłoby to o tym, że dane stwierdzenie idzie „pod prąd” kierunkowi pozostałych stwierdzeń i na pewno należy je usunąć. Stwierdzenia eliminuje się kolejno, zaczynając od tych, dla których różnica jest najmniejsza. Procedura ta wykonywana jest niejako metodą kolejnych przybliżeń, ponieważ po eliminacji już pojedynczego stwierdzenia wyniki poszczególnych osób należy obliczyć na nowo oraz na nowo ustalić górny i dolny kwartył. Znakomicie przydaje się w tym miejscu technika komputerowa. Po wyeliminowaniu stwierdzeń nie pasujących do pozostałych otrzymamy względnie jednolitą, wewnętrznie spójną skalę. Wyniki poszczególnych osób są obliczane już tylko w oparciu o ich ustosunkowanie się do tych stwierdzeń, które ostatecznie pozostały.

To jednak jeszcze nie koniec. Na żadnym etapie konstruowania tej skali nie ma pewności, że poszczególne stwierdzenia mają taką samą „wagę”, że w jednakowym stopniu wskazują na natężenie postawy. Tak jak w przypadku skali Thurstone’a do ustalenia wag służyli sędziowie kompetentni, tak w przypadku skali Likerta tę rolę spełniają same osoby badane. Polega to na tym, że ustalana jest średnia liczba punktów dla każdego stwierdzenia, następnie zaś ta średnia służy jako waga. Jeśli na przykład przy przyjętej początkowo punktacji od 1 do 5 jakieś stwierdzenie uzyskało średnią 1,3, to w ostatecznym rozrachunku poszczególne kategorie odpowiedzi ustosunkowujące się do tego stwierdzenia będą punktowane odpowiednio: 1,3, 2,6, 3,9, 5,2 oraz 6,5. Po wykonaniu tego ważenia oczywiście na nowo oblicza się wyniki poszczególnych osób badanych. Jeśli taką procedurę wykonać na bardzo dużej próbie osób, to oczywiście do każdych kolejnych badań można przyjąć wagi ustalone w pierwszym badaniu, aczkolwiek można i warto każdorazowo upewniać się, że nie uległyby zmianie, gdyby je ustalić z uwzględnieniem wyników kolejnego badania.

Podobnie jak w przypadku poprzednio omówionej skali, otrzymane wyniki punktowe należy traktować jedynie jako podstawę do uporządkowania badanej grupy osób ze względu na kierunek i natężenie postawy – od najbardziej pozytywnej do najbardziej nieprzychylniej. Niektórzy teoretycy są wszakże zdania, że dzięki wyrafinowanej procedurze, wspartej przecież o zaawansowane „sztuczki” statystyczne, można tę skalę traktować jako interwałową.

Inne opisane w literaturze rodzaje skal w mniejszym stopniu są przeznaczone do pomiaru postaw. Mowa tu o skali dystansu społecznego Bogardusa (patrz: Oppenheim, s. 219–220; Babbie, s. 190–191), która służy przede wszystkim do mierzenia natężenia uprzedzeń etnicznych, oraz o skali Guttmana (patrz: Oppenheim, s. 233–237; Mayntz, Holm, Hübner, s. 75–81; Babbie, s. 193–196; Frankfort-Nachmias, Nachmias, s. 482–485), która do pewnego stopnia przypomina skalę Likerta, lecz ma charakter kumulatywny – jej szczegółowe omówienie Czytelnik znajdzie w wymienionych pozycjach.

Opisane dość szczegółowo skale: Thurstone’a i Likerta są pewnymi wzorcami z punktu widzenia konstruowania skal. Jak każde wzorce, niekoniecznie są osiągalne w praktycznym działaniu. Charakterystyczne jest, że nawet niektórzy autorzy uznanych pozycji metodologicznych „nie doprowadzają do końca” prezentacji sposobu konstruowania skali Likerta, brakuje zazwyczaj etapu ostatniego, polegającego na ustaleniu owych „wag”. Zdaniem E. Babbie’go (s. 192) brak tego etapu sprawia, że nie będziemy mieli do czynienia z prawdziwą skalą, lecz z prostym indeksem – według rozróżnienia uczynionego przez tego autora (s. 173–176). Można jedynie mówić w tym przypadku, że poszczególne stwierdzenia mają „format Likerta”.

Warto zwrócić uwagę, że ów format, polegający na udzielaniu odpowiedzi według wzorca: „zdecydowanie zgadzam się”, „zgadzam się” itd., albo: „zawsze”, „często”, ..., „nigdy” i tym podobne jest nadzwyczaj rozpowszechniony. Spotyka się go na przykład jako sposób odpowiadania na pytania ankiety. Używa się zresztą takiego języka, że wśród pytań zamkniętych dysjunktywnych występuje szczególny typ: „pytanie-skala”.

Zwłaszcza w badaniach pedagogicznych używa się tego (wygodnego) formatu dość powszechnie. Wystarczy zauważyć, że jeśli M. Łobocki pisze o „metodzie szacowania”, używając zresztą zamiennie określenia: skale ocen, to omawiając poszczególne typy skal (numeryczne, graficzne, przymiotnikowe, opisowe i inne) przy większości z nich stosuje pięciostopniowy format odpowiedzi, pochodzący od Likerta (s. 81–104). Podobnie rzecz się ma w podręczniku autorstwa T. Pilcha i T. Bauman (s. 176–183).

Uwagi końcowe

Wypadałoby jeszcze omówić zagadnienie standaryzacji zarówno indeksów, jak i skal. Na to składają się: trafność, rzetelność i normalizacja. Jeden z elementów trafności, jakim jest badanie mocy dyskryminacyjnej był już zresztą wspomniany. Przyjmując jednak założenie, że ten tekst ma dostarczyć praktycznych wskazań ewentualnym konstruktorom mierników zło-

zonych do pomiaru postaw, które to mierniki nie będą narzędziami „szerokiego użytku” lecz będą przeznaczone w zasadzie do jednokrotnego zastosowania, należy zejść na ziemię. Przeprowadzana z pełnymi sztykami standaryzacja jest zresztą procesem skomplikowanym, długotrwałym, wieloetapowym, w zasadzie zespołowym – są to warunki rzadko osiągalne w wielu typowych badaniach empirycznych. Zresztą – czy jest się o co bić? Jeśli mimo wielu wyrafinowanych zabiegów konstrukcyjnych, wspartych niejednokrotnie zaawansowanymi procedurami statystycznymi ostateczny efekt jest taki, że pozwala otrzymane liczby traktować jedynie w kategoriach porządkowych, to być może zamiast konstruować którąś z „idealnych” skal, lepiej zadowolić się indeksem. Zresztą, jeśli bliżej przyjrzeć się opisywanym przez licznych metodologów „skalom”, to w istocie są to indeksy.

Wydaje się, że można sformułować na koniec następujące wskazania praktyczne:

- pomiar postaw nie jest możliwy na podstawie jednego czy kilku wskaźników; przeciwnie – należy posłużyć się miernikiem złożonym (oddalając dyskusję, czy należy go nazwać skalą, czy też indeksem), składającym się z większej liczby stwierdzeń, część z nich powinna mieć kierunek zgodny z postawą, część – przeciwny;
- warto zadbać o sprawdzenie wewnętrznej zgodności całego narzędzia poprzez wyeliminowanie tych stwierdzeń, które „nie pasują” do większości pozostałych;
- wygodnym formatem ustosunkowania się osób badanych do tych stwierdzeń jest „format Likerta” – pięcioelementowy lub inny;
- należy przyjąć jakąś (umowną) punktację, w miarę posiadanych możliwości dokonawszy wprzód „ważenia” poszczególnych stwierdzeń;
- otrzymane sumaryczne wartości punktowe można traktować nie jako „prawdziwe” wartości liczbowe („wartości postawy”), lecz jedynie jako podstawę do uporządkowania zbadanych osób ze względu na znak i natężenie postawy.

Postawa jest „z natury rzeczy” zmienną porządkową, właściwymi do jej opisu kategoriami pomiarowymi są sformułowania typu: „bardziej”, „mniej” itp. Uporządkowanie grupy zbadanych osób na umownym *continuum*: od postawy najbardziej (w danej grupie osób) negatywnej do najbardziej pozytywnej jest zupełnie wystarczające jako punkt wyjścia do dalszych analiz, w tym także badania zależności postawy od innych zmiennych. To fakt, że jeśli wyniki pomiaru nie są na skali interwałowej, to nie możemy użyć (dla wielu badaczy najbardziej wygodnego) współczynnika korelacji Pearsona czy analizy wariancji. W zasięgu pozostają jednak współczynniki korelacji rangowej (Spearmana albo Kendalla) czy też test niezależności chi-kwadrat i jeden ze współczynników zbieżności.

Literatura

- Aronson E., Wilson T.D., Akert R.M. (1997), *Psychologia społeczna: serce i umysł*, tłum. W. Domachowski i inni, Poznań: Zysk i S-ka.
- Babbie E. (2003), *Badania społeczne w praktyce*, tłum. W. Betkiewicz i inni, Warszawa: PWN.
- Chmielewski J. (2005), *Życie w konkretnej społeczności*, „Edukacja i Dialog”, nr 2 [online:] <http://www.vulcan.edu.pl/eid/archiwum/2005/02/zyc.html>.
- Frankfort-Nachmias C., Nachmias D. (2001), *Metody badawcze w naukach społecznych*, tłum. E. Hornowska, Poznań: Zysk i S-ka.
- Holly R. (1985), *Postawa* (i hasła korespondujące), [w:] W. Szewczuk (red.), *Słownik psychologiczny*, Warszawa: „Wiedza Powszechna”, s. 215–217.
- Łobocki M. (2000), *Metody i techniki badań pedagogicznych*, Kraków: „Impuls”.
- Mayntz R., Holm K., Hübner P. (1985), *Wprowadzenie do metod socjologii empirycznej*, tłum. W. Lipnik, Warszawa: PWN.
- Nowak S. (1985), *Metodologia badań społecznych*, Warszawa: PWN.
- Oppenheim A.N. (2004), *Kwestionariusze, wywiady, pomiary postaw*, tłum. S. Amsterdamski, Poznań: Zysk i S-ka.
- Pilch T., Bauman T. (2001), *Zasady badań pedagogicznych: strategie ilościowe i jakościowe*, Warszawa: „Żak”.
- Rathus S.A. (2004), *Psychologia współczesna*, tłum. B. Wojciszke, Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- Zimbardo P.G., Leippe M.R. (2004), *Psychologia zmiany postaw i wpływu społecznego*, tłum. P. Kwiatkowski, Poznań: Zysk i S-ka.