

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/310162688>

Skale szacunkowe

Chapter · January 2011

CITATIONS
2

READS
2,662

2 authors, including:



Anna Izabela Brzezińska
Adam Mickiewicz University in Poznań

144 PUBLICATIONS 761 CITATIONS

SEE PROFILE

skale szacunkowe w badaniach diagnostycznych¹

Psycholog korzysta z wielu źródeł informacji w procesie diagnostycznym tak dla celów naukowych, jak i praktycznych. Często zanim spotka się z osobą badaną (pacjentem, klientem) i zaprosi ją do wykonania określonych zadań, zbiera na jej temat informacje, przeprowadzając rozmowy z osobami ją znającymi, np. pracodawcą, konsultantem, nauczycielem, opiekunami, kuratorem sądowym, rodzicami. Często także rozmawia z nią, a potem dopiero podejmuje decyzję, czy i jakie narzędzia diagnostyczne zastosować. Nie zawsze konieczne i możliwe jest stosowanie narzędzi spełniających jednocześnie wszystkie kryteria psychometryczne, a więc: (a) wystandaryzowanych, (b) o znanej — i wysokiej — rzetelności oraz (c) trafności, (d) znormalizowanych, (e) poddanych procedurom adaptacji kulturowej w przypadku stosowania narzędzi powstałych w innych krajach, a także (f) realnych psychologicznie i życiowo².

¹ Niniejszy rozdział stanowi znacząco zmienioną i poszerzoną wersję rozdziału naszego autorstwa pt. *Skale szacunkowe*, który ukazał się w pracy: Brzeziński J. (1978 — wyd. 1; 1980 — wyd. 2; 1984 — wyd. 3). *Elementy metodologii badań psychologicznych* (część II, rozdział 4, s. 204-220). Warszawa: Państwowe Wydawnictwo Naukowe oraz dwóch rozdziałów autorstwa Jerzego Brzezińskiego: *Skala z wymuszonym wyborem* (s. 221-231) i *Technika Q* (s. 271-286), które ukazały się w tej samej pracy.

² Korzystając z koncepcji podwójnego realizmu eksperymentu psychologicznego wg Elliota Aronsona i in. (1994, s. 58), możemy powiedzieć, że realizm

Bardzo często zdarza się, że „narzędziami z wyboru”, podstawowymi z uwagi na charakter problemu diagnostycznego (badawczego lub praktycznego) albo nawet jedynymi, jako źródło informacji o funkcjonowaniu czy właściwościach osoby badanej są:

- a) bezpośrednia obserwacja zachowania się w warunkach naturalnych i/lub eksperymentalnych albo obserwacja zachowania nagranego na wideo, np. w czasie wykonywania jakiegoś zadania w zespole i samodzielnie; w trakcie procesu terapii (monitorowanie zmian); przed i po zakończeniu cyklu szkoleniowego,
- b) wywiad (ustrukturuwany lub nieustrukturuwany) przeprowadzany z osobami z otoczenia danej jednostki, np. z ojcem i matką oraz nauczycielami dziecka z zespołem nadpobudliwości psychoruchowej; ze współpracownikami danej osoby, gdy placówka, którą kieruje, jest poddawana procedurze akredytacji; z kuratorem sądowym oraz wychowawcami w zakładzie poprawczym przed podjęciem decyzji, co do dalszych losów nieletniego,
- c) rozmowa (ustrukturuwana lub nieustrukturuwana) przeprowadzana z osobą badaną, np. z pracownikiem tracącym pracę z powodu restrukturyzacji firmy; z osobą starającą się o pracę; z uczniem sprawiającym trudności wychowawcze; kilkakrotnie

psychologiczny (ang. *psychological realism*) w odniesieniu do treści narzędzia badawczego oznacza, iż osoba badana jest w stanie zaangażować się emocjonalnie w to, czego to narzędzie od niej wymaga (odpowiedzi na pytania, rozwiązywanie zadań), nawet jeżeli wydaje się jej to dziwne (np. wypełnianie tabel, stawianie kropek na czas, zaznaczanie czegoś w określonym miejscu, wycinanie, spoglądanie przez dziurkę w kartce, skakanie na jednej nodze, tasowanie kart, układanie puzzli). Natomiast realizm życiowy (ang. *mundane realism*) oznacza, iż to, czego wymaga od niej dane narzędzie badawcze, jest w jakimś przynajmniej stopniu podobne do tego, co ją spotyka w różnych sytuacjach w życiu codziennym, co jest podobne do tego, co przytrafia się ludziom, co może się im przytrafić, nawet jeżeli jej się nie przytrafiło i raczej nie przytrafi. Wysoki poziom realizmu psychologicznego powoduje, że osoba badana angażuje się emocjonalnie w wykonywane zadania, wysoki poziom realizmu życiowego zaś wpływa na to, jakie swoje zasoby będzie aktualizowała, wykonując dane zadanie. Wysokiemu poziomowi realizmu psychologicznego (*to jest ciekawe, choć niezupełnie rozumiem, po co to robię*) nie zawsze towarzyszy wysoki poziom realizmu życiowego (*mnie to nie dotyczy, mnie to się nie przytrafiło i raczej nie przytrafi; to jest pozbawione sensu*) i odwrotnie (por. Brzeziński 1996, s. 308-309).

- z pacjentem w trakcie procesu terapii; z rehabilitowanym przez specjalistów dzieckiem; z wolontariuszem,
- d) analiza poziomu wykonania i charakterystycznych cech wytworów osoby badanej, takich jak listy, dzienniki, pamiętniki, wypracowania szkolne, rysunki, raporty, nagrania wypowiedzi itp.,
- e) sądy i opinie wydawane przez osobę badaną o samej sobie (ocena siebie pod jakimś względem) i o innych ludziach (w tym o ich wytworach).

Przy dokonywaniu porównań wyżej wymienionych źródeł informacji z testami psychometrycznymi stwierdzamy występowanie różnic na niekorzyść tych pierwszych. Najważniejsze z nich to: (1) wysoki poziom subiektywizmu przy ocenianiu uzyskanych informacji, (2) niska trafność oszacowań związana z działaniem licznych zmiennych zakłócających, (3) brak możliwości przeprowadzania ścisłych porównań interindywidualnych oraz (4) brak możliwości wiarygodnego porównania informacji zebranych od tej samej osoby w kolejnych odstępach czasu (np. w procesie uczenia się, w procesie terapii, w okresie adaptowania się do nowego miejsca pracy, przed i po wprowadzeniu zmian w zasadach funkcjonowania zespołu zadaniowego).

Wady te można sprowadzić do jednej generalnej. Jest nią fakt, że informacje uzyskane z powyższych źródeł mają charakter jakościowy, a nie ilościowy. Gdyby dać badaczowi do ręki narzędzie, za pomocą którego mógłby przełożyć uzyskane dane jakościowe na dane ilościowe, równałoby się to zrealizowaniu postulatu zwiększenia obiektywności uzyskiwanych informacji oraz możliwości dokonywania porównań interindywidualnych (tzw. porównania poprzeczne — por. Brzezińska 2000, s. 171-172) oraz intraindywidualnych, jak zmiany w czasie, monitorowanie przebiegu jakiegoś procesu (tzw. porównania podłużne — por. ibidem, s. 172-173). Tym narzędziem, naszym zdaniem, są skale szacunkowe (ang. *rating scales*).

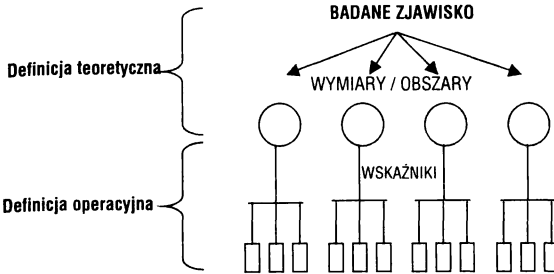
Skale szacunkowe są narzędziami bardzo często stosowanymi w praktyce badawczej psychologów. Najogólniej rzecz ujmując, skalą szacunkową jest „każde narzędzie stosowane przez sędziego przy dokonywaniu ocen” (Reber 2000, s. 673). Według Freda N. Kerlingera (1986, s. 494) skala szacunkowa to „narzędzie pomiaru wymagające, by osoba obserwująca lub szacująca oceniła dany obiekt ze względu na podane kategorie (lub kontinuum), którym przypisano odpowiednie wartości liczbo-

we". Oznacza to, że ten, kto dokonuje oszacowania, musi „przełożyć” to, co obserwuje (np. zachowanie pracownika w jakiejś sytuacji, zachowanie tej samej osoby w sytuacji nowej i sytuacji znanej, wypowiedzi ucznia w czasie pracy w zespole nad rozwiązywaniem zadania, cechy rysunków wykonanych przez pacjenta przed i po zakończeniu terapii) lub to, co wie o danym obiekcie (np. matka ocenia właściwości temperamentu swego dziecka na podstawie własnego doświadczenia, a nie obserwacji w konkretnej sytuacji, nauczycielka przedszkola ocenia ryzyko dysleksji u dzieci, korzystając z posiadanych o nich informacji, mąż ocenia umiejętności komunikacyjne swej żony i *vice versa*, czy też osoba badana sama ocenia swoje właściwości, np. w kwestionariuszu samowiedzy, poczucia kontroli, poczucia jakości życia) na kategorie opisowe i/lub liczbowe.

Badacz stosujący skale szacunkowe do przekładania danych jakościowych na dane ilościowe musi jednak najpierw spełnić dwa warunki. Po pierwsze, musi dysponować wywiedzioną z jakiejś teorii definicją badanego zjawiska (tj. rodzaju postępowania lub cechy osobowości), z której wynika, czy interesujące go „zjawisko” ma charakter jednorodny czy zróżnicowany; jeżeli zróżnicowany, to musi określić obszary lub wymiary (dymensje) opisu tego zjawiska. Po drugie, musi z tej definicji wywieść wskaźniki szacowanego zachowania się (cechy osobowości) dla każdego wyłonionego obszaru (wymiaru), na podstawie których będzie mógł określić w miarę dokładnie natężenie tego zachowania się (cechy osobowości) występujące u badanej jednostki czy w badanym obiekcie (por. rys. 1a i rys. 1b). Te dwa kroki wymagają nie tylko dobrej znajomości teorii, ale także dobrej orientacji w literaturze przedmiotu co do tego, kto i w jaki sposób badał dotąd to lub podobne zjawisko, na jakie problemy metodologiczne lub techniczne natrafił, jak je rozwiązywał oraz czy były to rozwiązania skuteczne, czy też nie.

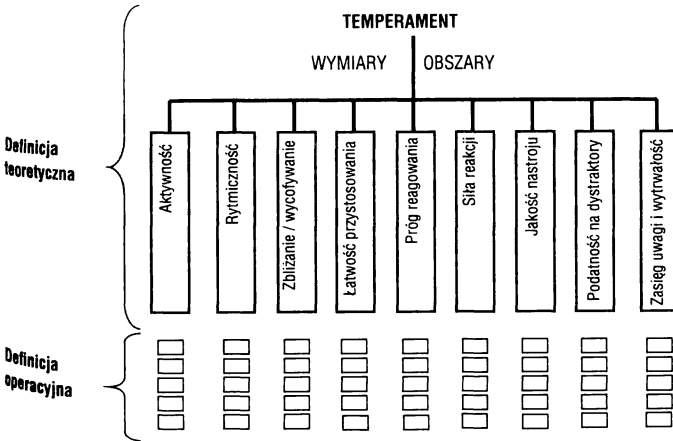
D. Peabody (1987) w przeglądowym artykule dotyczącym posługiwania się listami przymiotników, pochodzących z języka naturalnego, do badania różnych cech osobowości pisze, że skonstruowanie trafnego narzędzia wymaga określenia zbioru reprezentatywnych zmiennych „składających się” na ową cechę oraz określenia zbioru reprezentatywnych dla każdej zmiennej przymiotników. Definicja zmiennej pozwala bowiem nie tylko na wybranie z całej puli — jakkolwiek określonej — przymiotników ją charakteryzujących, ale także na uporządkowanie tych przymiotników wg stopnia bliskości znaczeniowej.

Rysunek 1a. Definicja i wskaźniki szacowanego zjawiska



Źródło: opracowanie własne.

Rysunek 1b. Definicja i wskaźniki szacowanego zjawiska — przykład: koncepcja temperamentu wg A. Thomasa i S. Chess (1977)



Wskaźniki — po pięć pozycji dla dziewięciu wymiencji temperamentu w Kwestionariuszu do szacowania cech temperamentu dzieci (Brzezińska 1995)

Źródło: opracowanie własne.

Interesującym przykładem prostej w stosowaniu i wartościowej co do zakresu dostarczanych informacji jest Skala Nieprzystosowania Społecznego (SNS) Lesława Pytki (2000, Aneks nr 1, s. 376-424). Jest to skala szacunkowa, przeznaczona do zebrania wstępnych informacji na temat wadliwego funkcjonowania dzieci i młodzieży w wieku 13-17 lat. Skala ta powstawała w kilku etapach³:

- 1) określenie punktu wyjścia: „W teoretycznym podejściu autora dominował behawioralny punkt widzenia, a zatem interpretacja wyników SNS opiera się głównie na danych pochodzących z obserwacji zachowania się jednostki będącej przedmiotem diagnozy. Skala rejestruje rozmaite kategorie zachowań i sytuacje określające kontekst psychospołeczny, w jakim one występują” (ibidem, s. 376);
- 2) zdefiniowanie kluczowego pojęcia, czyli „nieprzystosowania społecznego”;
- 3) wyodrębnienie analizowanych obszarów — Pytka mówi o trzech płaszczyznach nieprzystosowania społecznego: behawioralnej, psychologicznej i etiologicznej;
- 4) zwerbalizowanie podstawowych założeń teoretycznych (por. ibidem, s. 378);
- 5) konstrukcja skali: SNS składa się z sześciu podskal (nieprzystosowanie rodzinne, rówieśnicze, szkolne, zachowania antyspołeczne, kumulacja niekorzystnych czynników biopsychicznych, kumulacja niekorzystnych czynników socjokulturowych), z kolei każda podskala składa się z dziesięciu pozycji, a więc cała skala liczy 60 pozycji; każda pozycja oceniana jest na skali szacunkowej: 0 – 1 – 2;
- 6) dokładny opis każdej podskali: definicja, wyłonienie obszarów: dobór wskaźników, określenie, ile wynosi wynik maksymalny możliwy do uzyskania (dla każdej skali X_{\max} wynosi 20 pkt.) i co oznaczają: im wyższy wynik, tym wyższy poziom nieprzystosowania społecznego (por. ibidem, s. 380-390);
- 7) przeprowadzenie badań na próbie reprezentatywnej;

³ Przedstawimy je tutaj szczegółowo jako wręcz modelową ilustrację tego, co badacz musi zrobić, by w efekcie dysponować dobrym pod każdym względem narzędziem diagnostycznym.

- 8) psychometryczne opracowanie skali SNS (moc dyskryminacyjna pozycji, rzetelność i trafność, normalizacja) (por. *ibidem*, s. 390-404);
- 9) opis procedury badania (kto udziela informacji, jak dobierać te osoby, w jakich warunkach prowadzić badanie) oraz opis procedury obliczania wyników surowych i przeliczania ich na normy (por. *ibidem*, s. 404-410);
- 10) zasady interpretacji wyników oraz wykorzystania danych ze skali SNS w procesie diagnozy i konstruowania projektów oddziaływań profilaktycznych i resocjalizacyjnych (por. *ibidem*, s. 410-414; szczególnie schemat „Etapy i obszary działania diagnosty” s. 413).

W praktyce ze skal szacunkowych korzysta się w trzech przypadkach. Po pierwsze, gdy obiektem analiz jest jednostka — jej sposób funkcjonowania, jej właściwości — a psycholog sam dokonuje „przekładu” danych jakościowych na dane ilościowe, co ułatwia mu potem porównywanie różnych danych ze sobą. Dzieje się tak szczególnie wtedy, gdy planowane są dłuższe w czasie oddziaływania i psycholog (badacz) zamierza śledzić zakres dokonujących się zmian (porównanie podłużne) albo gdy chce uporządkować swoje obserwacje dotyczące funkcjonowania tej samej osoby, prowadzone w różnych sytuacjach w tym samym czasie (porównania poprzeczne, czyli „poziom” danej zmiennej w różnych sytuacjach czy okolicznościach).

Po drugie, gdy obiektem badań są różne osoby i psycholog chce porównywać między sobą sposób ich funkcjonowania albo postęp w jakimś zakresie; wtedy dobrze, by zebrane obserwacje czy dane z wywiadu zostały „przełożone” na dane ilościowe za pomocą skal szacunkowych przez psychologa, ale przez zespół odpowiednio do tego przygotowanych kompetentnych sędziów (piszemy o tym w końcowej części, por. Blok rozszerzający 3).

Po trzecie, gdy jednostka dokonuje samooszacowań (samoopisu), np. wypełnia skalę postaw, odpowiada na pytania kwestionariusza, ocenia swoje zachowanie czy samopoczucie w jakiejś sytuacji, np. podczas pracy w grupie lub tuż po zakończeniu szkolenia, albo też ocenia jakieś swoje właściwości, np. poziom odporności na stres w sytuacji znanej i sytuacji całkowicie dla siebie nowej, poziom gotowości do wykonania zadania, preferencje pokarmowe itp.

1. Rodzaje skal szacunkowych

Wyróżnić można kilka rodzajów skal szacunkowych. Fred N. Kerlinger wyróżnia pięć ich typów: kategoryjne, numeryczne i graficzne (op. cit., s. 494-495) oraz z wymuszonym wyborem i w postaci listy (op. cit., s. 461-463). Arthur S. Reber (op. cit., s. 673) opisuje skale szacunkowe dwubiegunowe, graficzne, itemizowane i także, jak Kerlinger, skale w postaci listy. Obaj, podobnie jak wielu innych badaczy, odwołują się do klasycznej pracy Joya P. Guilforda z roku 1954, w której wyróżnił on skale numeryczne, skale graficzne, skale skumulowanych ocen i skale z wymuszonym wyborem (por. Brzezińska, Brzeziński 1975a). Poniżej omówimy skale kategoryjne, numeryczne i graficzne, skumulowanych ocen i skale z wymuszonym wyborem. Przedstawimy także specyficzną ze względu na sposób badania skalę szacunkową zwaną „techniką Q” lub „Q-sort”.

1.1. Kryteria podziału skal szacunkowych

Skale szacunkowe można podzielić na kilka grup wg różnych kryteriów:

1) wg kryterium sposobu prezentacji skali wyróżnia się:

- a) skale kategoryjne (ang. *category rating scales*), zgodnie z terminologią Rebera (op. cit., s. 673) są to tzw. skale itemizowane (ang. *itemized rating scales*): osobie badanej prezentowana jest lista kategorii do wyboru wg podanej przez badacza zasady (por. Przykład 1a i 1b),
- b) skale numeryczne (ang. *numerical rating scales*): osobie dokonującej oszacowań prezentuje się listę zachowań (właściwości) wraz z opisami — o różnym poziomie szczegółowości — poszczególnych kategorii nasilenia lub częstości występowania, którym potem przyporządkowuje się odpowiednie wagi wg klucza albo też od razu osoba szacująca zakreśla odpowiednią cyfrę (por. Przykłady 2a, 2b, 2c, 2d),
- c) skale graficzne (ang. *graphic rating scales*): osobie dokonującej oszacowań prezentowany jest odcinek wyobrażający nasilenie lub częstość szacowanej właściwości, na którym zaznacza swoją ocenę, kierowana instrukcją badacza co do znaczenia krańców odcinka (por. Przykład 3);

- 2) wg kryterium sposobu opisu kategorii wyróżnia się:
- a) skale z opisanymi punktami krańcowymi: w postaci jednego lub dwóch wyrazów (*nie dotyczy — dotyczy; nisko — wysoko; prawie nigdy — prawie zawsze*; por. Przykład 4) albo w postaci zdań czy równoważników zdań (por. Przykład 5),
 - b) skale z opisanymi wszystkimi punktami: w postaci krótkich określeń (por. Przykład 2a, 2b, 2c) lub dłuższych opisów (por. Przykład 2d, 6);
- 3) wg kryterium charakteru skali wyróżnia się:
- a) skale jednobiegunowe (ang. *unipolar rating scales*): stosuje się je do oceny stopnia nasilenia jakiegoś wymiaru (por. Przykład 2b, 2c, 2d, 4, 8; rysunek 2) bądź do oceny częstości występowania jakiejś cechy (por. Przykład 2a),
 - b) skale dwubiegunowe (ang. *bipolar rating scales*): gdy ocenia się właściwość przyjmującą dwie różne jakościowo postaci (por. Przykład 5 oraz Blok rozszerzający 1 — pkt 2: Zachowanie uczniów i pkt 3: Właściwości nauczyciela).

1.2. Skale kategoryalne

Na podstawie przyjętej przez siebie teorii badanego zjawiska badacz przygotowuje listę kategorii, a osoba dokonująca oszacowania tego zjawiska lub jakiejś jego właściwości wybiera spośród zaprezentowanych jej opisów — tworzących wg badacza skalę — ten, który jej zdaniem stanowi najbardziej trafną charakterystykę jakiegoś zjawiska lub właściwości. Kategoriom nie przypisuje się wag liczbowych, badacz może więc analizować jedynie to, którym ocenianym właściwościom przypisano jakie kategorie.

Jeżeli badacz dokonuje porównań intraindywidualnych (np. postępy w terapii), może śledzić, czy oceniana właściwość zmienia w kolejnych badaniach swoje miejsce czy nie, a więc czy przyporządkowywane są jej te same czy inne (i jakie) kategorie. W przypadku porównań interindywidualnych może porównywać, czy ta sama kategoria została przypisana takim samym czy różnym właściwościom badanych osób, albo — jak w Przykładzie 1a — którym osobom przypisano jakie kategorie.

PRZYKŁAD 1A. SKALA KATEGORIALNA

**PORÓWNIANIA INTRAINDYWIDUALNE
(BADANIA PODŁUŻNE) — ARKUSZE DLA KOLEJNYCH OBSERWACJI**

Observacja nr 1. Data: 20.09.2003	Przedmiot: Matematyka	Klasa: IV a
1. Materiał nauczania — zgodnie z programem przedmiotu — do przerobienia wyznacza plan lekcji X	Planowanie uwzględnia dotychczasową wiedzę uczniów oraz zakładane cele, ale każda lekcja to zamknięta całość	Każda lekcja ma swój cel, ale wiąże się z poprzednimi i następnymi, kolejne lekcje tworzą spójną całość
2. Klasa jest podzielona przez nauczyciela na dwie grupy o stałym składzie	Klasa jest podzielona przez nauczyciela na trzy grupy o stałym składzie X	Planowanie obejmuje trzy lub więcej grup o zmieniającym się składzie
3. Na lekcji przerabia się fragmenty podręczników X	Planuje się fragmenty lekcji zgodnie z założonymi celami	Stosuje się różne sposoby pracy, zgodne z cechami uczniów
4. Pytania zadawane uczniom wynikają z podręcznika X	Pytania wiążą problemy poruszane na danej lekcji z poprzednimi	Przewidziany czas na samodzielną aktywność uczniów
Observacja nr 2. Data: 8.12.2003	Przedmiot: Matematyka	Klasa: IV a
1. Materiał nauczania — zgodnie z programem przedmiotu — do przerobienia wyznacza plan lekcji	Planowanie uwzględnia dotychczasową wiedzę uczniów oraz zakładane cele, ale każda lekcja to zamknięta całość X	Każda lekcja ma swój cel, ale wiąże się z poprzednimi i następnymi, kolejne lekcje tworzą spójną całość
2. Klasa jest podzielona przez nauczyciela na dwie grupy o stałym składzie	Klasa jest podzielona przez nauczyciela na trzy grupy o stałym składzie X	Planowanie obejmuje trzy lub więcej grup o zmieniającym się składzie
3. Na lekcji przerabia się fragmenty podręczników	Planuje się fragmenty lekcji zgodnie z założonymi celami X	Stosuje się różne sposoby pracy, zgodne z cechami uczniów
4. Pytania zadawane uczniom wynikają z podręcznika X	Pytania wiążą problemy poruszane na danej lekcji z poprzednimi	Przewidziany czas na samodzielną aktywność uczniów

gdzie X — kategorie zaznaczone przez badanego

Porównania interindywidualne
(badania poprzeczne) — arkusz zbiorczy dla obserwacji nr 1

1. Materiał nauczania — zgodnie z programem przedmiotu — do przerobienia wyznacza plan lekcji ① ⑤	Planowanie uwzględnia dotychczasową wiedzę uczniów oraz zakładane cele, ale każda lekcja to zamknięta całość ②	Każda lekcja ma swój cel, ale wiąże się z poprzednimi i następnymi, kolejne lekcje tworzą spójną całość ③ ④
2. Klasa jest podzielona przez nauczyciela na dwie grupy o stałym składzie ⑤	Klasa jest podzielona przez nauczyciela na trzy grupy o stałym składzie ① ②	Planowanie obejmuje trzy lub więcej grup o zmieniającym się składzie ③ ④
3. Na lekcji przerabia się fragmenty podręczników	Planuje się fragmenty lekcji zgodnie z założonymi celami ① ③ ⑤	Stosuje się różne sposoby pracy, zgodne z cechami uczniów ② ④
4. Pytania zadawane uczniom wynikają z podręcznika ①	Pytania wiążą problemy poruszane na danej lekcji z poprzednimi ② ④ ⑤	Przewidziany czas na samodzielną aktywność uczniów ③

gdzie: ① — nauczyciel ... (matematyka); ② — nauczyciel ... (język polski); ③ — nauczyciel ... (geografia); ④ — nauczyciel ... — wychowanie fizyczne i wychowawca klasy; ⑤ — nauczyciel ... — katecheta

Źródło: opracowano na podstawie arkusza obserwacji nauczyciela *Characteristics of teachers classroom observation record* D.G. Ryansa (w: Robeck, Wilson 1974, s. 630-640).

Przykładem skali kategoryjnej (tzw. itemizowanej) może być skala dystansu społecznego E.S. Bogardusa (1925; za: Błuszkowski 2003, s. 207-208). Pojęcie dystansu społecznego wprowadził do nauk społecznych R.E. Park (za: Błuszkowski, op. cit., s. 207). Oznaczało ono „odniesienie do różnych obiektów społecznych, które mogą mieścić się na skali od bieguna kontaktów ścisłych, ciepłych i zażyłych, poprzez obojętność, do bieguna antypatii, wrogości i potępienia” (ibidem). Natomiast autorem pierwszej skali do pomiaru tak rozumianego dystansu społecznego był E.S. Bogardus (ibidem). Początkowo skala ta obejmowała 60 pytań, potem poddano je ocenie 100 sędziów kompetentnych, którzy mieli owe stwierdzenia przyporządkować do siedmiu kategorii oznaczających różny poziom (stopień) dystansu społecznego. W rezultacie uzyskano

więc „skalę składającą się z siedmiu pytań tworzących kontinuum równo odległych od siebie punktów dystansu społecznego” (ibidem). Osoba badana ma stwierdzić, na podstawie swej pierwszej spontanicznej reakcji, do jakiego typu kontaktów społecznych dopuściłaby „typowego” przedstawiciela jakiejś narodowości:

- 1) do bliskiego pokrewieństwa przez małżeństwo,
- 2) do klubu, do którego sama należy, jako bliskiego przyjaciela,
- 3) na swoją ulicę jako sąsiada,
- 4) do pracy w swoim zawodzie,
- 5) jako obywatela swego kraju,
- 6) tylko jako turystę w swoim kraju,
- 7) usunęłaby tę osobę ze swego kraju.

Różne zastosowania pomysłu Bogardusa pokazuje Przykład 1b.

PRZYKŁAD 1B. SKALA KATEGORIALNA — PRZYKŁADY

Jan Błuszkowski (2003, s. 208-209) podaje kilka przykładów wykorzystania pomysłu Bogardusa w badaniach psychologicznych i socjologicznych.

Przykład 1. Badania K.W. Backa (1962, za: ibidem) nad dystansem w kontaktach interpersonalnych:

1. Chętnie zobaczyłbym go czasem w kolegium.
2. Chciałbym z nim pracować w jednej sekcji.
3. Byłoby przyjemnie z nim porozmawiać.
4. Byłoby mi przyjemnie prowadzić z nim żywą dyskusję.
5. Chciałbym z nim dyskutować o różnych poważnych sprawach ogólnych.
6. Chciałbym, żeby on do mnie przyszedł ze swoimi sprawami.
7. Chciałbym z nim porozmawiać o ważnych sprawach osobistych.

Przykład 2. Badania K. Skarżyńskiej (1976, za: ibidem) nad dystansem interpersonalnym:

1. Zdecydowanie nie chciałabym się z nią kontaktować.
2. Wolałabym się z nią nie spotykać.
3. Spotkanie z nią byłoby mi zupełnie obojętne.
4. Chętnie zobaczyłabym ją czasem na wydziale.
5. Byłoby mi przyjemnie z nią porozmawiać.

6. Byłoby mi przyjemnie spotkać się z nią poza uczelnią.
7. Chciałabym wykonywać z nią jakieś zadanie, wspólną pracę.
8. Chciałabym, żeby ona przyszła do mnie ze swoimi sprawami.
9. Chciałabym porozmawiać z nią o moich ważnych sprawach osobistych.

Przykład 3. Badania I. Kurcz (1994, za: ibidem), nad stereotypami:

Czy zgadzasz się, aby (przedstawiciel narodowości)

1. Przyjeżdżał do Polski jako turysta zwiedzający nasz kraj?
2. Został Twoim sąsiadem?
3. Był Twoim: a) współpracownikiem?
b) podwładnym?
c) bezpośrednim przełożonym?
4. Szukał w Polsce pracy jako emigrant?
5. Został Twoim lub Twojego dziecka małżonkiem?
6. Blisko się z Tobą zaprzyjaźnił?
7. Czy uważasz go za osobistego wroga?

Przykład 4. Badania nad dystansem wobec mniejszości etnicznych w Holandii (Hraba, Hagendoorn, Hagendoorn 1989, za: ibidem):

1. Nie podobałoby mi się, gdyby moimi sąsiadami byli
2. Nie podobałoby mi się, gdyby moimi kolegami w pracy byli
3. Nie podobałoby mi się, gdyby w szkole mojego dziecka było wielu
4. Nie podobałoby mi się, gdyby moim lekarzem był
5. Nie podobałoby mi się, gdyby moim przełożonym w pracy był
6. Nigdy nie poślubiłbym ani nie miał związku z
7. Nie podobałoby mi się, gdyby moje dzieci poślubiły lub miały związek z
8. Nigdy nie zaprzyjaźniłbym się z

1.3. Skale numeryczne i graficzne

Skale numeryczne i graficzne są najczęściej i najchętniej stosowane przez badaczy. Skalę numeryczną stanowi pewien zbiór kategorii opisujących poszczególne punkty określonego kontinuum psychologicznego, np. kontinuum cechy osobowości (por. Przykłady 2a, 2c.) Kategoriom tym przypisane są *a priori* cyfry od najmniejszej do największej

szej lub odwrotnie. *Implicite* przyjmuje się, że odległości między wyróżnionymi kategoriami na skali numerycznej są równe. Narzędzie badawcze może być różnie skonstruowane: (1) można na skali umieścić tylko kategorie liczbowe, (2) można umieścić tylko kategorie opisowe, którym potem przypisze się odpowiednie wagi, (3) można wreszcie umieścić i kategorie opisowe, i cyfry. Przypisywanie liczb może się niektórym badanym kojarzyć z wystawianiem ocen szkolnych i może to być czynnik istotnie obniżający trafność ich oszacowań (patrz dalej, gdzie piszemy o typowych błędach szacowania). Warto tak konstruować skale, by badani mieli do dyspozycji skalę o parzystej liczbie kategorii (2, 4, 6, 8 itd.), znika wtedy bowiem kategoria „środkowa” (dla skali 3-punktowej, np. 0 – 1 – 2, będzie to kategoria „1”; dla skali także 3-punktowej *tak – nie wiem (?)* – *nie* będzie to kategoria *nie wiem (?)*; dla skali 7-punktowej, np. 0 – 1 – 2 – 3 – 4 – 5 – 6, będzie to kategoria „3”), która może być szczególnie często używana przez osoby niepewne swej oceny.

PRZYKŁAD 2A. SKALE NUMERYCZNE

KWESTIONARIUSZ *POCZUCIE JAKOŚCI ŻYCIA* ULRIKE RAVENS-SIEBERER (1998)

Instrukcja

Witaj! Chcielibyśmy dowiedzieć się, jak czułeś się w ostatnim tygodniu, w związku z tym chcielibyśmy, żebyś odpowiedział na kilka pytań, które opracowaliśmy:

- ⇒ przeczytaj uważnie każde pytanie
- ⇒ pomyśl o tym, co wydarzyło się w ciągu ostatniego tygodnia
- ⇒ wybierz odpowiedź, która najlepiej do Ciebie pasuje, i postaw przy niej **znaczek taki, jak ten: X**

Pamiętaj, że nie ma tu ani złych, ani dobrych odpowiedzi. Napisz po prostu, co **sam myślisz**.

Przykład:

	✍	nigdy	rzadko	czasami	często	cały czas
W ostatnim tygodniu czułem się bardzo źle	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Jeśli czułeś się źle przez cały czas, postaw krzyżyk na samym końcu z prawej strony.

	✍	nigdy	rzadko	czasami	często	cały czas
W ostatnim tygodniu czułem się bardzo źle	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Jesli czuleś się źle czasami, postaw krzyżyk pośrodku.

W ostatnim tygodniu czułem się bardzo źle

nigdy	rzadko	czasami	często	cały czas
<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Jesli w ostatnim tygodniu nigdy się źle nie czuleś, postaw krzyżyk po lewej stronie.

W ostatnim tygodniu czułem się bardzo źle

nigdy	rzadko	czasami	często	cały czas
<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

W ostatnim tygodniu...

Pamiętaj, postaw znaczek: X

	nigdy	rzadko	czasami	często	cały czas
1. ... bez problemu odrabiałem lekcje	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. ... miałem wszystkiego dość	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. ... miałem mdłości, czułem się źle	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. ... rodzice byli wyrozumiali dla mnie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

PRZYKŁAD 2B. SKALE NUMERYCZNE

Podane niżej przykłady pochodzą z pracy Karoliny Appelt (2004), która badała, jak nauczyciele szkół podstawowych i gimnazjów radzą sobie w sytuacji doświadczania zmiany i od jakich czynników podmiotowych zależy styl radzenia sobie. W poniższych kwestionariuszach zastosowała 6-punktową skalę szacunkową. W zależności od klucza wskazanym odpowiedniemu przypisywała wagi: 0 - 1 - 2 - 3 - 4 - 5 lub odwrotnie.

KWESTIONARIUSZ I

Instrukcja: proszę dokładnie przeczytać poniższe stwierdzenia i zaznaczyć swoją odpowiedź, przekreślając właściwą kratkę znakiem X.

Umiem dokonywać wyborów, szczególnie w sprawach dla mnie ważnych.	ZDECYDOWANIE NIE	NIE	RACZEJ NIE	RACZEJ TAK	TAK	ZDECYDOWANIE TAK
---	------------------	-----	------------	------------	-----	------------------

Czuję się akceptowana przez innych ludzi.	ZDECYDOWANIE NIE	NIE	RACZEJ NIE	RACZEJ TAK	TAK	ZDECYDOWANIE TAK
Świat jest pełen chaosu, niczego nie można w nim przewidzieć.	ZDECYDOWANIE NIE	NIE	RACZEJ NIE	RACZEJ TAK	TAK	ZDECYDOWANIE TAK

KWESTIONARIUSZ II

Instrukcja: proszę ocenić, czy dana sytuacja jest dla Pani trudna, i zaznaczyć swoją odpowiedź, przekreślając właściwą kratkę znakiem **X**.

1. Przygotowywanie i prowadzenie zajęć w ramach wewnątrzszkolnego systemu doskonalenia nauczycieli.

ZDECYDOWANIE NIE	NIE	RACZEJ NIE	RACZEJ TAK	TAK	ZDECYDOWANIE TAK
------------------	-----	------------	------------	-----	------------------

2. Dostosowywanie wymagań do indywidualnych możliwości ucznia. Obowiązek nauczyciela, by służyć uczniowi pomocą w przewyżnianiu niepowodzeń szkolnych i mieć pełne rozeznanie potrzeb uczniów w tym zakresie.

ZDECYDOWANIE NIE	NIE	RACZEJ NIE	RACZEJ TAK	TAK	ZDECYDOWANIE TAK
------------------	-----	------------	------------	-----	------------------

KWESTIONARIUSZ III

Instrukcja: przy każdym zdaniu proszę postawić **X** przy wybranej odpowiedzi.

1. Przygotowywanie i prowadzenie zajęć w ramach wewnątrzszkolnego systemu doskonalenia nauczycieli.

• Jak Pani ocenia swoje możliwości poradzenia sobie z tą sytuacją?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	bardzo nisko	nisko	raczej nisko	raczej wysoko	wysoko	bardzo wysoko

2. Dostosowywanie wymagań do indywidualnych możliwości ucznia. Obowiązek nauczyciela, by służyć uczniowi pomocą w przewyżnianiu niepowodzeń szkolnych i mieć pełne rozeznanie potrzeb uczniów w tym zakresie.

• Jak Pani ocenia swoje możliwości poradzenia sobie z tą sytuacją?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	bardzo nisko	nisko	raczej nisko	raczej wysoko	wysoko	bardzo wysoko

PRZYKŁAD 2C. SKALE NUMERYCZNE

315

KWESTIONARIUSZ SAMOWIEDZY I SAMOOCENY PIOTRA KRZYWICKIEGO

Kwestionariusz ten powstał z potrzeby opracowania narzędzia do diagnozy samowiedzy u młodzieży z upośledzeniem umysłowym. Analiza wyników uzyskanych w badaniu pilotażowym (Brzezińska, Krzywicki 2001) oraz w badaniach porównawczych prowadzonych na grupie młodzieży z upośledzeniem umysłowym i na grupie w tzw. normie intelektualnej (Krzywicki 2002, 2003) pokazuje, że jest to rzeczywiście narzędzie trafne i rzetelne. Jako wskaźnik rzetelności przyjęto współczynnik korelacji między kolejnymi badaniami tej samej grupy osób — posłużono się metodą test-retest, czyli badano stabilność bezwzględna (stałość) testu (por. Brzeziński 1996, s. 468-469). Drugie badanie zostało przeprowadzone po upływie dwóch miesięcy od pierwszego badania. Mimo tak długiej przerwy między kolejnymi badaniami uzyskano stosunkowo wysoką wartość współczynnika korelacji: $r = 0,71$ i $p < 0,01$ dla ogólnego wskaźnika (suma wszystkich punktów) samooceny (Krzywicki 2002).

Konstrukcja narzędzia: kwestionariusz składa się z 50 krótkich stwierdzeń odnoszących się do sześciu aspektów obrazu własnej osoby (sfera fizyczna, umysłowa, działaniowa, emocjonalno-motywacyjna, społeczna, stosunku do przyszłości). Badane obszary zostały ustalone wg wskazań sędziów kompetentnych (por. Brzezińska, Krzywicki, op. cit.).

Stosowanie: osoba badana ma za zadanie ustosunkować się do 50 twierdzeń i zaznaczyć swój wybór na 5-stopniowej skali ocen. Wybór punktu na skali ocen zależy od tego, jak badany spostrzega siebie na tle swoich rówieśników. Odpowiedź na pytanie „Jak określisz siebie w porównaniu do swoich rówieśników?” osoba badana zaznacza w odpowiedniej kratce kolejno dla każdego stwierdzenia.

Ocena i interpretacja wyników: badanie pozwala ustalić, jaki jest poziom samooceny badanej osoby oraz jaka jest charakterystyka cech specyficznych dla danej jednostki. Wskaźnikiem poziomu samooceny jest suma punktów zaznaczonych przez badanego na skali od 1 do 5. Uzyskane punkty stają się podstawą do określenia niskiej (mała liczba punktów) lub wysokiej (duża liczba punktów) samooceny. W analogiczny sposób można obliczyć poziom samooceny dla wyróżnionych sześciu podskal. Cechy wyraźnie specyficzne dla danej jednostki to odpowiedzi określone przez osobę badaną na skali ocen w pozycjach 1 i 5, cechy mało specyficzne to wybory z pozycji 2 i 4, a niespecyficzne to — 3. W wyniku sumowania liczby

punktów zaznaczonych na skali ocen w przedstawionych powyżej pozycjach skali otrzymamy ilościową charakterystykę prezentującą stopień spostrzegania przez osobę badaną własnej odrębności. Analogiczna ocena może być dokonana dla każdej z sześciu podskal.

Instrukcja: przeczytaj każde zdanie i przekreśl znakiem X swoją odpowiedź.

1. Jestem zdrowy

mniej
niż
inni

raczej
mniej
niż inni

tak samo
jak
inni

raczej
bardziej
niż inni

bardziej
niż
inni

2. Odczuwam radość

mniej
niż
inni

raczej
mniej
niż inni

tak samo
jak
inni

raczej
bardziej
niż inni

bardziej
niż
inni

3. Lubię przebywać
z ludźmi

mniej
niż
inni

raczej
mniej
niż inni

tak samo
jak
inni

raczej
bardziej
niż inni

bardziej
niż
inni

PRZYKŁAD 2D. SKALE NUMERYCZNE

SKALA NIEPRZYSTOSOWANIA SPOŁECZNEGO LESŁAWA PYTKI (2000).

Podskala I. Nieprzystosowanie rodzinne

- | | |
|---|---|
| I. Formalny związek
dziecka z rodziną | 0. Wychowuje się w rodzinie pełnej.
1. Wychowuje się w rodzinie niepełnej.
2. Wychowuje się poza rodziną. |
| II. Kontrola czasu
spędzanego poza
domem | 0. Wychodząc z domu, mówi, po co i dokąd idzie.
1. Czasami mówi, dokąd i po co wychodzi.
2. Nie tłumaczy się, dokąd i po co wychodzi. |
| III. Wykonywanie
obowiązków
domowych | 0. Chętnie wykonuje obowiązki domowe.
1. Wykonuje obowiązki domowe pod przymusem.
2. Nie wykonuje obowiązków domowych. |
| IV. Spełnianie poleceń
rodziców lub
opiekunów | 0. Na ogół spełnia polecenia rodziców.
1. Niesystematycznie spełnia polecenia.
2. Na ogół nie spełnia wydawanych mu poleceń. |

Źródło: Pytka 2000, s. 418.

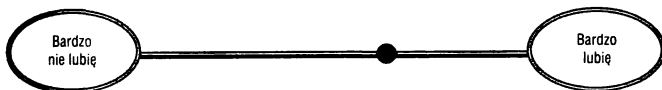
Najbardziej zbliżonymi do skal numerycznych są skale graficzne. Skalę graficzną stanowi odcinek prosty, poziomy lub pionowy. Krańcom odcinka są przyporządkowane pewne kategorie opisowe. Odcinek ten może być również podzielony na pewną liczbę równych części, a punktom podziału mogą być przypisane kategorie opisowe (por. Przykład 3).

PRZYKŁAD 3. SKALE GRAFICZNE

- Proszę ocenić swój poziom zadowolenia z pracy w zespole, stawiając znak **X** w odpowiednim miejscu



- Jak bardzo lubisz rozwiązywać różne zadania razem z kolegami (postaw znak ●):



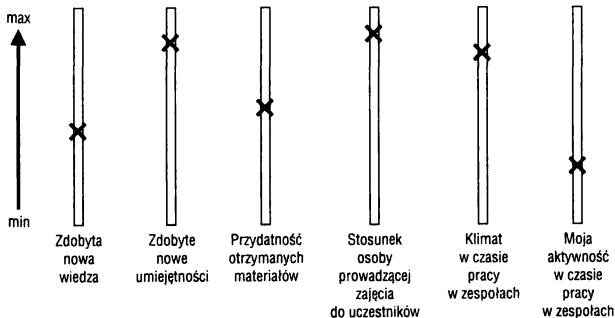
- Czy otrzymane materiały edukacyjne przydadzą się Pani (Panu) (proszę postawić **X**):

Nie widzę
żadnej ich
przydatności
w moich
działaniach



Widzę ich
różnorodne
zastosowanie
w moich
działaniach

- Oceń zajęcia, w których uczestniczyłeś (postaw znak **X** na każdej skali):



1.3.1. Wymagania techniczne stawiane skalom graficznym

Prawidłowo skonstruowaną skalę graficzną cechują pewne charakterystyczne właściwości, które niżej szczegółowo omówimy. Właściwości te wybraliśmy z listy 12 warunków, jakim powinny odpowiadać skale graficzne, podanej przez Joya P. Guilforda (1954, s. 267-268).

Optymalna długość linii stanowiącej skalę powinna wynosić 10-15 cm. Skale zbyt krótkie utrudniają lub nawet uniemożliwiają przeprowadzenie dokładnych oszacowań osób badanych, ponieważ w zbyt małym stopniu różnicują badaną grupę. Z kolei w przypadku skal zbyt długich osoby dokonujące oszacowań wykazują tendencję do gromadzenia ich na określonym odcinku skali. Jednocześnie linia ta nie powinna mieć żadnych przerw w myśl zasady, że ciągła linia wyraża ciągłość szacowanej cechy (osobowości, zachowania). Linia składająca się z kilku odcinków umieszczonych na jednym poziomie, ale w pewnych odległościach od siebie może sugerować albo brak ciągłości zmiennej, albo zupełnie różne jakościowo kategorie, a to z kolei może utrudniać dokonywanie dokładnych oszacowań. Jeżeli w badaniu posługujemy się zbiorem skal szacunkowych, to „dobre” i „złe” krańce skal powinny występować w porządku losowym, tzn. część skal powinna mieć z prawej strony kategorie dodatnie, oceniane najwyżej, a z lewej kategorie oceniane najniżej, część natomiast — odwrotnie. Takie losowe rozłożenie w pewnym stopniu przeciwdziała występowaniu błędów łagodności oraz efektu *halo* (błędy te omówimy niżej), a jednocześnie uniemożliwia automatyczne zaznaczanie pozycji badanego na skali, zmuszając każdorazowo osobę dokonującą oszacowania do zapoznania się z kierunkiem skali. Guilford (ibidem) uważa jednak, że ten sposób rozmieszczenia krańców skali stanowi pewne utrudnienie nawet dla osób wprawionych i dobrze wyćwiczonych w dokonywaniu oszacowań, a zatem może prowadzić do obniżania trafności oszacowań na skutek popełnianych błędów (także zwykłych pomyłek spowodowanych spadkiem koncentracji uwagi w trakcie szacowania). A więc, jeżeli w badaniach posługujemy się skalami szacunkowymi, a oszacowań dokonują osoby zupełnie niedoświadczone (np. rodzice, uczniowie), powinniśmy zachować ten sam kierunek dla wszystkich skal. Przy układzie pionowym kraniec „dobry” jest zwykle umieszczony u góry skali, przy układzie poziomym — z prawej strony, zgodnie z konwencjonalnym układem współrzędnych występującym w matematyce. We wszystkich innych natomiast wypadkach

należy jednak stosować losowe rozmieszczanie kategorii „dobrych” i „złych”.

W przypadku posługiwania się zbiorem skal i jednocześnie szacowania większej grupy osób powinniśmy każdą skalę umieścić na oddzielnej kartce, a szacowanie przeprowadzać kolejno dla wszystkich badanych najpierw na pierwszej skali, potem na drugiej itd. Taka kolejność szacowania umożliwi kontrolowanie efektu *halo* oraz efektu pierwszeństwa (Wojciszke 2000).

Kategorie opisowe powinny być umieszczone jak najbliżej tych punktów skali, których dotyczą, szczególnie w przypadku skal dyskretnych. Jeżeli używamy dłuższych opisów (kilka przymiotników, wielocłonowe zdania), bardziej korzystny jest pionowy układ skali. W przypadku skali poziomej konieczne jest skoncentrowanie opisów w kolumnach i posłużenie się drobnym drukiem.

Po zakończeniu badania wszystkim kategoriom opisowym przydzielamy określone wartości liczbowe (wagi). W przypadku skal dyskretnych posługujemy się specjalnym kluczem, np. skali szacunkowej o 5 kategoriach odpowiedzi (odpowiednio opisanych) możemy przypisać kolejne wagi: 0, 1, 2, 3, 4 lub 4, 3, 2, 1, 0 — zależnie od tego, czy kategoria umieszczona w lewym krańcu skali oznacza „dobry” (pożądany społecznie, wysoko ceniony) koniec danego kontinuum, czy też „zły” (nisko oceniany, społecznie niepożądany, oznaczający minimalne nasilenie danej cechy czy częstości występowania zachowań określonego rodzaju). W przypadku skal ciągłych posługujemy się podziałką milimetrową, mierząc długość odcinka: od punktu zerowego do miejsca zaznaczonego przez osobę dokonującą oszacowania. Także tu konieczne jest zwrócenie uwagi na kierunek skali. Na przykład badanej osobie dajemy 50 punktów, jeżeli długość zaznaczonego odcinka wynosi 50 mm.

1.3.2. Skale ciągłe a skale dyskretne

Skale ciągłe mają postać odcinka o określonej długości, opisanego na obu krańcach. Jeden kraniec oznacza maksymalne natężenie szacowanej cechy, a drugi minimalne. Kategorie opisowe powinny dokładnie uwzględniać różne wskaźniki behawioralne danej zmiennej, oznaczające jej krańcowe natężenie. W przypadku gdy osobą szacującą jest psycholog lub inna osoba, która dobrze zna przedmiot szacowania oraz jego

wskaźniki, możemy podać tylko nazwę zmiennej oraz stopień jej natężenia.

Skale dyskretne także mają postać odcinka umieszczonego poziomo lub pionowo, z tą różnicą, że odcinek ten podzielony jest na kilka mniejszych części o równej długości. Liczba tych części może być w zasadzie dowolna. Także i tu określonym punktom skali należy przydzielić kategorie opisowe, a przed przystąpieniem do obliczania wyników przydzielić im odpowiednie wagi.

Skala dyskretna ma jednak o wiele mniejszą wartość niż skala ciągła, ponieważ różnicuje badaną grupę tylko na kilka podgrup. Najmniej doskonała jest skala dyskretna dwustopniowa — różnicuje ona próbę tylko na dwie podgrupy. Jeżeli zatem zależy nam na dokładnym zróżnicowaniu badanej grupy, powinniśmy użyć skali ciągłej.

1.3.3. Rodzaje skal numerycznych i graficznych

Numeryczne i graficzne skale szacunkowe występują w wielu odmianach. Wyżej wymieniliśmy już dwie główne odmiany skal graficznych: ciągłe i dyskretne. Poniższy przegląd różnego typu skal będzie dotyczył w zasadzie skal numerycznych. Gdyby jednak w cytowanych tu przykładach szeregom cyfr przypisać odcinek podzielony na odpowiednią liczbę części, otrzymalibyśmy dyskretne skale graficzne.

Wszystkie skale można podzielić na dwie klasy ze względu na stopień dokładności opisu punktów krańcowych:

- a) punkty krańcowe są opisane za pomocą takich par słów, jak: *występuje* — *nie występuje*, *brak lęku* — *bardzo duży lęk*, *całkowicie się zgadzam* — *całkowicie się nie zgadzam*, *zdecydowanie tak* — *zdecydowanie nie*, *prawie nigdy* — *prawie zawsze*, *bardzo rzadko* — *bardzo często* lub też po prostu jeden kraniec skal oznaczony jest liczbą odpowiadającą minimalnemu natężeniu określonej zmiennej, a drugi liczbą oznaczającą maksimum natężenia tej zmiennej (por. Przykład 4);
- b) punkty krańcowe są bardzo dokładnie opisane, np. przez podanie określonego zachowania się, które zdaniem badacza odpowiada punktom minimum i maksimum na skali (por. Przykład 5).

PRZYKŁAD 4. SKALA NUMERYCZNA Z KRÓTKIMI OPISAMI KRAŃCÓW

SKALA SAMOOCENY POCZUCIA JAKOŚCI ŻYCIA

Skalę tę skonstruowano (Brzezińska 2000, s. 25-29; por. Brzezińska, Stolar-ska, Zielińska 2002) na podstawie pomysłu Flanagana, by wskaźnikiem poczucia jakości życia uczynić różnicę między tym, jak jednostka ocenia ważność jakiegoś obszaru swego życia, a tym, jak ocenia poczucie zaspokojenia swych potrzeb w tym obszarze.

Flanagan w swych badaniach obejmujących ludzi dorosłych odnosił się do tzw. obszarów satysfakcji życiowej. Wyróżnił 15 takich obszarów i każdemu z nich przypisał skalę 7-punktową (0-6), prosząc dwukrotnie o ich ocenę. Najpierw badani odpowiadali na pytanie: Jak ważny jest każdy obszar?, a potem na pytanie: Jaki jest poziom zaspokojenia potrzeb w każdym obszarze? Wskaźnikiem poczucia jakości życia była różnica między wagą odpowiedzi na pierwsze i drugie pytanie. Różnica wynosząca „0” oznaczała, że badany ma wysokie poczucie jakości życia, ponieważ albo (1) oceniaiany obszar jako bardzo ważny dla siebie, a zatem przypisuje mu wagę „6” i jednocześnie uznaje, że jego potrzeby w tym obszarze są w bardzo dużym stopniu zaspokojone (waga „6”), różnica wynosi: $6 - 6 = 0$, albo (2) mimo że uznaje, iż jego potrzeby w jakimś obszarze są w małym stopniu zaspokojone (waga „2”), to jednocześnie ocenia ten obszar jako mało ważny dla siebie, np. także przypisuje mu wagę „2”, więc różnica wynosi: $2 - 2 = 0$. Z kolei większa wartość tej różnicy wskazuje na niskie poczucie jakości życia. Odnosi się to przede wszystkim do sytuacji, gdy badany ocenia jakiś obszar jako bardzo ważny dla siebie (waga „6”) i jednocześnie wskazuje, że jego potrzeby w tym obszarze nie są zaspokojone (waga „0”), a zatem różnica wynosi: $6 - 0 = 6$.

Uwaga:

„0” oznacza „zupełnie nieważne”, „6” oznacza „bardzo ważne”

Jak WAŻNE są dla Pani (Pana):

Zakreśl X
swoją ocenę

Jak WAŻNE są dla Pani (Pana):	Zakreśl X swoją ocenę
1. Materialne warunki życia i bezpieczeństwo finansowe	0 1 2 3 4 5 6
2. Zdrowie i bezpieczeństwo osobiste	0 1 2 3 4 5 6
3. Związki z partnerem życiowym	0 1 2 3 4 5 6
4. Związki z rodzicami, rodzeństwem, innymi krewnymi	0 1 2 3 4 5 6
5. Związki z przyjaciółmi, znajomymi, kolegami	0 1 2 3 4 5 6
6. Posiadanie i wychowywanie dzieci, wnuków	0 1 2 3 4 5 6
7. Pomaganie innym ludziom, wspieranie ich	0 1 2 3 4 5 6
8. Działania w lokalnej społeczności i w państwie	0 1 2 3 4 5 6
9. Własny rozwój intelektualny	0 1 2 3 4 5 6

10. Twórczość i swoboda	0	1	2	3	4	5	6
11. Rozumienie tego, co się dzieje	0	1	2	3	4	5	6
12. Możliwość planowania swego życia	0	1	2	3	4	5	6
13. Aktywność zawodowa	0	1	2	3	4	5	6
14. Działalność rekreacyjna i odpoczynek	0	1	2	3	4	5	6
15. Czynne uprawianie jakiegos sportu	0	1	2	3	4	5	6

Drugie pytanie brzmiało: W jakim stopniu Pani (Pana) potrzeby są zaspokojone w każdym z obszarów? , przy czym „0” oznaczało „zupełnie nie są zaspokojone”, a „6”, że są „zaspokojone w bardzo dużym stopniu”.

PRZYKŁAD 5. SKALA NUMERYCZNA Z DOKŁADNYMI OPISAMI KRAŃCÓW

KWESTIONARIUSZ DO SZACOWANIA CECH TEMPERAMENTU DZIECI — OPIS NARZĘDZIA

Kwestionariusz został opracowany przez A. Brzezińską (1995) na podstawie koncepcji Alexandra Thomasa i Stelli Chess (1977; por. Strelau 1982, 1998). Jego zastosowanie w badaniach prowadzonych w grupie dzieci ze zdiagnozowanym zespołem nadpobudliwości ruchowej (Brzezińska 2000, s. 79-81), uczniów klas I-III (Filipiak 2002) oraz młodzieży niedostosowanej społecznie (Hetmańska-Bugajska, Brzezińska 2002) dostarczyło wyników, pozwalających na formułowanie trafnych wniosków odnośnie do związków cech temperamentu z innymi właściwościami funkcjonowania badanych osób.

Konstrukcja kwestionariusza: składa się z 45 pozycji tworzących dziewięć wymiarów. Każda pozycja ma charakter skali porządkowej z zakotwiczonymi punktami krańcowymi w postaci opisu zachowania dziecka. Poszczególne pozycje odnoszą się do zachowania będącego wskaźnikiem badanej właściwości temperamentu (tab. 1).

Sposób stosowania: arkusz testowy wypełniają rodzice lub nauczyciele badanych osób. Zaleca się, aby wypełniali go w obecności badacza, bez presji czasowej. Czas wypełniania kwestionariusza jest nieograniczony. Wypełnianie rozpoczynają od odczytania instrukcji, która brzmi: Dzieci, tak jak i dorośli, różnią się między sobą, niekiedy dość znacznie, swoimi różnymi

cechami, swoim postępowaniem w różnych sytuacjach. Proszę przeczytać zamieszczone dalej opisy zachowania i określić, w jakim stopniu odnoszą się one do dziecka. Jeśli zachowania opisane PO STRONIE LEWEJ bardzo dobrze opisują zachowanie dziecka — proszę zakreślić kółkiem cyfrę 0, ale jeżeli dotyczy to dziecka w nieco mniejszym stopniu — proszę zakreślić cyfrę 1. Jeśli zachowania opisane PO STRONIE PRAWEJ bardzo dobrze opisują zachowanie dziecka — proszę zakreślić kółkiem cyfrę 3, ale jeżeli dotyczy to dziecka w nieco mniejszym stopniu — proszę zakreślić cyfrę 2.

TABELA 1

Badane właściwości temperamentu i numery pozycji w kwestionariuszu

Skale	Symbol	Pozycje	X_{\min} = 0 pkt.	X_{\max} = 15 pkt.
1. Poziom aktywności	PA	4, 8, 10, 28, 37	niski	wysoki
2. Rytmiczność funkcji biologicznych*	RY	2, 11, 19, 20, 43	wysoka	niska
3. Zbliżanie się <i>versus</i> wycofanie*	ZW	1, 16, 25, 26, 35	zbliżanie	wycofywanie
4. Łatwość przystosowania*	LP	7, 12, 17, 34, 44	wysoka	niska
5. Próg reagowania	PR	6, 15, 24, 33, 42	na silne bodźce	na słabe bodźce
6. Siła reakcji	SR	13, 18, 22, 31, 40	mała	duża
7. Jakość nastroju*	JN	9, 27, 36, 41, 45	przewaga emocji pozytywnych	przewaga emocji negatywnych
8. Podatność na dystraktory*	PR	14, 23, 29, 32, 38	mała	duża
9. Zasięg uwagi i wytrwałość*	ZUW	3, 5, 21, 30, 39	duży zasięg i wytrwałość	mały zasięg, brak wytrwałości

* klucz odwrotny

Ocena i zasady interpretacji wyników: obliczamy — zgodnie z kluczem — wyniki dla dziewięciu podskal temperamentu (dla każdej podskali oddzielnie). Interpretacja wyników odbywa się na podstawie zaliczenia ich do określonych przedziałów, tj.: 00-02 pkt. — wynik bardzo niski; 03-05 pkt. — wynik niski; 06-08 pkt. — wynik przeciętny; 09-11 pkt. — wynik wy-

soki; 12-15 pkt. — wynik bardzo wysoki. Wynik „bardzo niski” i „niski” oznacza brak zaburzeń w obrębie danej kategorii temperamentu. Im wyższa jest liczba punktów uzyskanych przez badanego w danej skali, tym wyższe jest prawdopodobieństwo zaburzeń w zakresie danej składowej temperametu. Analiza profilu składającego się z dziewięciu wyników pozwala na wykrycie konstelacji tzw. trudnego temperamentu, temperamentu łatwo rozgrzewającego się i temperamentu łatwego (wg terminologii Thomasa i Chess, za: Strelau 1998).

KWESTIONARIUSZ DO SZACOWANIA CECH TEMPERAMENTU DZIECI

— przykłady opisów jednej pozycji dla każdej dymensji temperamentu

POZIOM AKTYWNOŚCI

podczas różnych zajęć siedzi spokojnie, nie rozmawia, nie kręci się	0 1 2 3	podczas różnych zajęć trudno mu usiedzieć na miejscu, często wierci się, zmienia pozycję ciała, rozmawia
---	---------	--

RYTMICZNOŚĆ FUNKCJI BIOLOGICZNYCH

nierregularnie zaspokajają swe potrzeby fizjologiczne, widać, że ma z tym jakiś problem	0 1 2 3	regularnie zaspokajają swe potrzeby fizjologiczne, nie ma z tym żadnego problemu
---	---------	--

ZBLIŻANIE SIĘ / WYCOFYWANIE

na nowym miejscu długo czuje się zakłopotany, nie jest sobą, jest mało mówny, zamyka się, ma problemy ze spaniem	0 1 2 3	szybko przyzwyczajają się do nowego miejsca, nie ma problemów z jedzeniem, spaniem, czuje się „jak u siebie”
--	---------	--

LATWOŚĆ PRZYSTOSOWANIA

ma wiele przyzwyczajeń odnośnie do tego, jak coś robić, niechętnie wprowadza w tym zmiany, to co nowe, jest długo „oswajane”	0 1 2 3	lubi zmiany, nowe propozycje odmieniają jego działania, nie lubi powracać do tego, co zna, co wie i umie
--	---------	--

PRÓG REAGOWANIA

nie zauważa zmian w swoim otoczeniu, trzeba mu na to zwracać uwagę, ukierunkować jego spostrzeżenie	0 1 2 3	od razu zauważa nawet bardzo drobne zmiany w otoczeniu, także zmiany w czyimś głosie, mimice, w ubraniu, w urządzeniu pokoju
---	---------	--

SILA REAKCJI

prawie po nim nie widać, że jest zły, 0 1 2 3 bardzo wyraźnie widać, kiedy
że się gniewa, trzeba się tego domy- jest zły, kiedy się złości —
ślać z całej sytuacji głóśniej mówi, zmienia mu
się twarz, czasem krzyczy lub
czymś rzuca

JAKOŚĆ NASTROJU

bardzo często wstaje rano w złym na- 0 1 2 3 wstaje rano bez problemów,
stroju, „lewą nogą”, od rana boczy się, najwyżej jest trochę zaspany,
nie wiadomo, o co mu chodzi ale jest pogodny, nie marudzi
i nie boczy się

PODATNOŚĆ NA DYSTRAKTORY

łatwo rozprasza się i nie jest w stanie 0 1 2 3 może w skupieniu coś robić,
niczego zrobić, gdy ktoś rozmawia, gra trudno go namówić na robie-
radio czy telewizor nie czegoś innego

ZASIĘG UWAGI I WYTRWAŁOŚĆ

rzadko doprowadza pracę, którą roz- 0 1 2 3 prawie zawsze doprowadza
począł, do końca, taki trochę „słomiany ogień”, porzuca wiele czynności, jest dokładny w tym, co robi,
zadań, prac skrupulatny, „jak się do cze-
goś przypnie, to nie popuści”

Niekiedy badacze posługują się skalami z dokładnie opisanymi wszystkimi punktami. Tego rodzaju skale stosuje się szczególnie wtedy, gdy zachodzi konieczność zakodowania informacji uzyskanych z wywiadu lub rozmowy. W ten sposób postąpili np. Albert Bandura i Richard H. Walters (1968) w badaniach nad agresją chłopców w okresie dorastania. Na podstawie wywiadów przeprowadzonych z badanymi chłopcami osoby dokonujące oszacowań, posługując się zestawem skal szacunkowych (kilkadziesiąt skal!), mogły przełożyć uzyskane dane jakościowe na dane ilościowe. Przykład 6 przedstawia jedno z pytań kwestionariusza wywiadu i odpowiadającą mu skalę szacunkową (Bandura, Walters, op. cit., s. 359 i 389).

Na koniec tego przeglądu przytaczamy skalę zainteresowania nauczyciela wykładanym przez siebie przedmiotem, którą podają W.J. Goode i P.K. Hatt (1965, s. 287). Jest to skala graficzna, w której „stopniowalność”

poszczególnych kategorii opisowych jest wyrażona za pomocą odpowiednich słów: *bardzo, umiarkowanie, wyraźnie, raczej* (por. rysunek 2).

PRZYKŁAD 6. SKALA NUMERYCZNA Z OPISANYMI WSZYSTKIMI PUNKTAMI SKALI (AGRESJA)

Przedstawiamy skalę 8 oraz pytanie 7 — z kwestionariusza wywiadu autorstwa A. Bandury i R. Waltersa (1968) — któremu przyporządkowana jest skala.

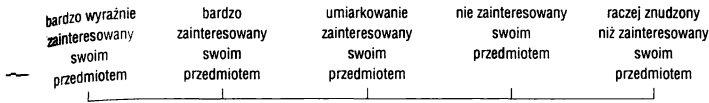
Skala 8. Opór przeciw szukaniu i przyjmowaniu pomocy lub rad nauczyciela (ad pytanie 7)

1. Żadnych objawów oporu: chociaż chłopiec może niekiedy zastanawiać się nad radą udzieloną przez nauczyciela, zanim ją zaakceptuje.
2. Nieznaczące oznaki oporu: chłopiec zdecydowanie woli sam rozwiązywać swoje sprawy, może okazywać pewne zażenowanie, jeżeli musi zwrócić się o pomoc.
3. Umiarkowany opór: pewne zażenowanie, skrępowanie, w rzadkich wypadkach miałby chęć postąpić wręcz przeciwnie, niż sugeruje nauczyciel.
4. Znaczny opór: zdecydowanie niechętnie szuka pomocy, czasami chętnie postąpiłby wręcz przeciwnie, niż sugeruje nauczyciel.
5. Silny opór: nie lubi prosić o pomoc i silnie opiera się sugestiom nauczyciela; czasami postępuje wręcz przeciwnie.

Pytanie 7:

- Co sądzisz o możliwościach pójścia do nauczyciela albo doradcy szkolnego i porozmawiania z nim na temat swoich problemów? Jak często zdarzała Ci się taka sytuacja? Z jakiego rodzaju sprawami zwróciłbyś się do niego? A jeśli chodzi o pomoc w zadaniach domowych? Czy uważasz, że tego rodzaju rozmowy z nauczycielami mogą się na coś przydać? A jeśli chodzi o doradców szkolnych? Dlaczego tak sądzisz?
- Czy zdarza Ci się, że czujesz się nieswojo, kiedy musisz zwrócić się do nauczyciela lub doradcy szkolnego o pomoc? Czemu przypisujesz to uczucie?
- Czy w przypadku, gdy nauczyciel albo doradca szkolny udziela ci jakichś rad lub coś ci proponuje, stosujesz się zazwyczaj do tych sugestii czy też raczej wolisz sam decydować o swoich sprawach?
- Czy nie odnosisz czasem wrażenia, że postępujesz w jakiejś sprawie zupełnie odwrotnie do rady nauczyciela tylko dlatego, że to właśnie on ci jej udzielił? Z jakiego powodu miałbyś tak postępować?

Rysunek 2. Graficzna skala szacunkowa pięciopunktowa



Źródło: Goode, Hatt 1965, s. 287.

1.3.4. Skala graficzna dwubiegunowa

Dobrym przykładem skali dwubiegunowej jest narzędzie zwane dyferencjałem semantycznym. Zastosowali je po raz pierwszy (do pomiaru znaczenia określeń charakteryzujących różne grupy narodowościowe) Ch.E. Osgood, G.J. Suci i P.H. Tannenbaum (1957). Według J. Błuszkowskiego (2003, s. 169): „Metoda dyferencjału semantycznego zakłada, że każda para opozycyjnych określeń wartościujących wyznacza biegunowy schemat analizy stereotypów”. W *Karcie dyferencjału semantycznego*, jaką posłużył się Błuszkowski (op. cit., s. 171) w swoich badaniach nad ewaluatywnymi treściami stereotypów narodowych (wobec 22 narodów, w tym Polaków), każda badana cecha była dwubiegunowa, czyli ujęto ją jako zaletę bądź wadę:

cechy negatywne (wady) 1 2 3 4 5 cechy pozytywne (zalety)

Poszczególnym punktom na skali szacunkowej przypisano odpowiednie opisy, przy czym pkt „1”, wyrażający ocenę negatywną, opisano jako „cecha z lewej strony właściwie określa typowych przedstawicieli danego narodu”, pkt „3”, czyli środkowy, wyrażający ocenę neutralną, opisano jako „trudno powiedzieć”, a pkt „5”, wyrażający ocenę pozytywną, opisano jako „cecha z prawej strony właściwie określa typowych przedstawicieli danego narodu”. Badani mieli ustosunkować się do 17 par określeń, wśród których były właściwości psychiczno-charakterologiczne (4 pary), właściwości moralne (3 pary), cechy mentalności (2 pary), cechy dotyczące kultury i obyczaju (2 pary), właściwości socjopolityczne (1 para), socjoekonomiczne (3 pary), cechy relacji międzyludzkich (2 pary) (op. cit., s. 170). Przykład 7 pokazuje skonstruowaną przez Błuszkowskiego (op. cit., s. 171) *Kartę dyferencjału semantycznego* oraz

rysunek z naniesionymi — jako przykład — profilami uzyskanymi dla 4 spośród 22 poddanych badaniu narodów: Polaków, Amerykanów, Niemców i Rosjan.

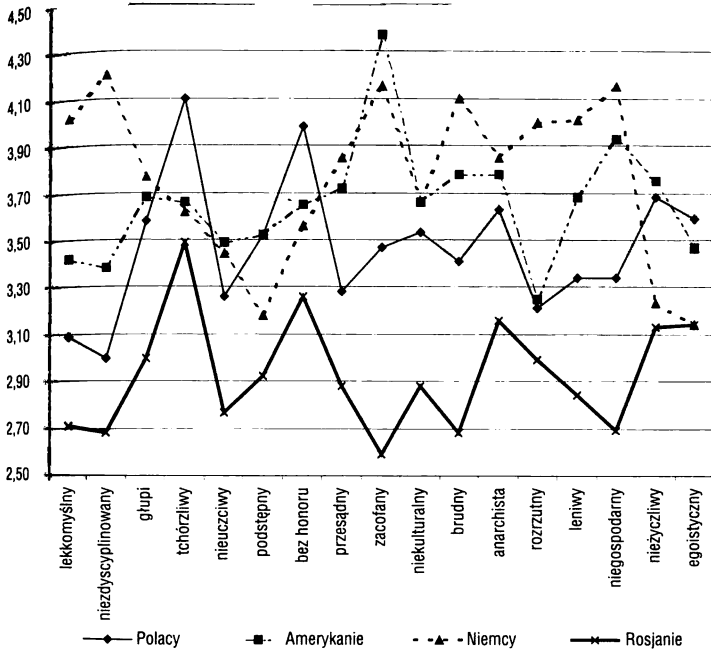
PRZYKŁAD 7. PRZYKŁAD SKALI DWUBIEGUNOWEJ

KARTA DYFERENCJAŁU SEMANTYCZNEGO J. BŁUSZKOWSKIEGO (2003, s. 171)

Poniżej przedstawiono pary przeciwstawnych cech. Po lewej stronie znajdują się cechy negatywne (wady), a po prawej cechy pozytywne (zalety). Którą i w jakim stopniu z każdej pary cech był(a)by Pan(i) skłonny(a) przypisać typowym przedstawicielom różnych narodów, posługując się pięciostopniową skalą?

	1	2	3	4	5	
Cechy negatywne (wady)	Cecha z lewej strony właściwie określa typowych przedstawicieli danego narodu	Cecha z lewej strony raczej właściwie określa typowych przedstawicieli danego narodu	Trudno powiedzieć	Cecha z prawej strony raczej właściwie określa typowych przedstawicieli danego narodu	Cecha z prawej strony właściwie określa typowych przedstawicieli danego narodu	Cechy pozytywne (zalety)
lekkomyślny	1	2	3	4	5	rozważny
niezdyscyplinowany	1	2	3	4	5	zdyscyplinowany
głupi	1	2	3	4	5	mądry
tchórzliwy	1	2	3	4	5	odważny
nieuczciwy	1	2	3	4	5	uczciwy
podstępny	1	2	3	4	5	szczerzy
bez honoru	1	2	3	4	5	honorowy
przesądny	1	2	3	4	5	racjonalny
zacořany	1	2	3	4	5	nowoczesny
niekulturalny	1	2	3	4	5	kulturalny
brudny	1	2	3	4	5	czysty
anarchista	1	2	3	4	5	lojalny
rozrzutny	1	2	3	4	5	oszczędny
leniwy	1	2	3	4	5	pracowity
niegospodarny	1	2	3	4	5	gospodarny
nieżyczliwy	1	2	3	4	5	życzliwy
egoistyczny	1	2	3	4	5	uczynny

Rysunek 3. Spolaryzowany profil cech ewaluatywnych Polaków, Amerykanów, Niemców i Rosjan



Źródło: opracowanie własne na podstawie pracy Błuszkowskiego (2003, s. 173, 188, 190, 191). Uwaga: w górnej części wykresu nie umieszczono nazw cech pozytywnych, por. Przykład 1.

1.4. Skale skumulowanych ocen

Skala skumulowanych ocen (ang. *cumulated points rating scales*) składa się zwykle z dużej liczby (nawet kilkudziesięciu) mniejszych skal, z których każda odnosi się do innego aspektu (wymiaru, obszaru) szacowanej zmiennej. Wynik ogólny jest sumą wag przypisanych kategoriom odpowiedzi zaznaczonym przez sędziów na poszczególnych skalach cząstkowych.

wych. Taki sposób obliczania wyników ogólnych jest charakterystyczny także dla innych narzędzi, wykorzystujących metodę kumulowanych punktów. Na przykład C.H. Leeds i W.W. Cook (za: Remmers 1967, s. 339; por. Leeds, Furman 1969) przeprowadziły badania, w których uczniowie dokonywali oszacowań za pomocą kwestionariusza „Mój nauczyciel” (*Minnesota Teacher Attitude Inventory*), składającego się z 50 pytań. Na każde pytanie można było odpowiedzieć przez zakreślenie jednej z trzech podanych odpowiedzi. Odpowiedzi punktowano po zakończeniu badania zgodnie z kluczem 0, 1, 2, a następnie zsumowano wagi uzyskane dla 50 odpowiedzi. Wynik ten stanowił oszacowanie nauczyciela dokonane przez danego ucznia. Oto przykłady pytań z tego kwestionariusza:

- | | |
|--|-------------|
| 4. Czy nauczyciel dużo krzyczy na uczniów? | tak, nie, ? |
| 6. Czy nauczyciel objaśnia zadania w szkole tak, że możesz to zrozumieć? | tak, nie, ? |
| 49. Czy lubisz tego nauczyciela? | tak, nie, ? |

Takiego kumulowania ocen można jednak dokonywać jedynie wtedy, gdy badane (szacowane) zjawisko jest jednorodne, czyli innymi słowy, gdy nie wyróżniamy żadnych jego składowych, wymiarów czy obszarów. Duża liczba pytań lub zdań daje wtedy okazję do rzetelnego zbadania jakiejś opinii, jak np. w badaniach oceniających nauczyciela wspomnianych wyżej.

Jeżeli jednak zjawisko jest złożone, niejednorodne (patrz rysunek 1a i 1b), w żadnym wypadku nie należy posługiwać się skalą skumulowanych ocen. W skrajnym przypadku różne osoby mogą bowiem uzyskać takie same lub zbliżone sumy punktów, podczas gdy w rzeczywistości ich oszacowania znacznie się od siebie różniły. Na przykład taką samą sumę punktów (wag) można otrzymać dla osoby, która używała skrajnie negatywnych i skrajnie pozytywnych kategorii, jak i dla osoby, która wykorzystywała głównie kategorie pośrednie.

Decyzja badacza zatem nie może być pochopna, nie warto też czasami kierować się zasadą ekonomiczności. Poza tym większa liczba wyników odpowiadających różnym dymensjom pozwala na sporządzenie profilu obrazującego natężenie badanych dymensji względem siebie. Wtedy jednak, gdy poszczególnym dymensjom odpowiada różna liczba podskal, konieczne jest dokonanie transformacji wyników, choćby na

najprostsze wskaźniki procentowe. Kolejne czynności badacza są tutaj następujące (por. tabela 2):

- 1) dla każdego szacowanego obiektu oblicza sumę wag przypisanych każdej pozycji wchodzącej w skład danej podskali — jeżeli część pozycji była kodowana np. 0–1–2–3, a część odwrotnie, tj. 3–2–1–0, badacz musi posłużyć się wcześniej przygotowanym kluczem;
- 2) sumę tę dzieli przez wynik maksymalny możliwy do uzyskania w tej podskali, a następnie mnoży przez 100;
- 3) wykreśla profil obrazujący wyniki uzyskane dla wszystkich podskal (por. rysunek 4) — może teraz te wyniki między sobą porównywać; może też obliczyć wartość średniej arytmetycznej dla grupy ocenianych obiektów, np. dla młodzieży z upośledzeniem umysłowym i młodzieży bez upośledzenia, jak w badaniach P. Krzywickiego (2003).

TABELA 2

Jak postąpić, gdy skala składa się z podskal o różnej liczbie pozycji?

Kwestionariusz samowiedzy P. Krzywickiego (2003)

Wyodrębnione obszary

Sfera fizyczna	Sfera umysłowa	Sprawność w działaniu	Sfera emocjonalno-motywacyjna	Sfera społeczna	Stosunek do przyszłości
----------------	----------------	-----------------------	-------------------------------	-----------------	-------------------------

Liczba stwierdzeń

9	10	6	9	10	6
---	----	---	---	----	---

Wynik maksymalny X_{\max} możliwy do uzyskania (skala ocen: 1–2–3–4–5)

$9 \times 5 = 45$	$10 \times 5 = 50$	$6 \times 6 = 36$	$9 \times 5 = 45$	$10 \times 5 = 50$	$6 \times 5 = 30$
-------------------	--------------------	-------------------	-------------------	--------------------	-------------------

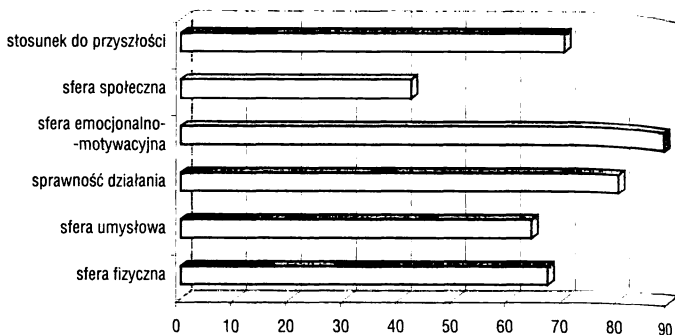
Osoba nr 12 uzyskała następujące wyniki surowe (suma zaznaczonych wag)

30 pkt.	32 pkt.	24 pkt.	40 pkt.	21 pkt.	21 pkt.
---------	---------	---------	---------	---------	---------

Osoba nr 12 uzyskała następujące wyniki przeliczone wg wzoru: $(\Sigma \text{pkt}/X_{\max}) \times 100$

$(30:45) \times 100 = 67\%$ wyniku maksymalnego	$(32:50) \times 100 = 64\%$ wyniku maksymalnego	$(24:30) \times 100 = 80\%$ wyniku maksymalnego	$(40:45) \times 100 = 89\%$ wyniku maksymalnego	$(21:50) \times 100 = 42\%$ wyniku maksymalnego	$(21:30) \times 100 = 70\%$ wyniku maksymalnego
--	--	--	--	--	--

Rysunek 4. Profil osoby nr 12 (dane w tabeli 1) — poziom samowiedzy



Źródło: opracowanie własne na podstawie Krzywicki (2003).

Skalą skumulowanych ocen posłużyli się E.S. Schaefer i R.Q. Bell (za: Rembowski 1972, s. 78; por. Przetacznik-Gierowska, Włodarski 1994, s. 132) przy konstrukcji skali postaw rodziców wobec zagadnień wychowawczych i ról rodzinnych zwanej w skrócie PARI (ang. *Parental Attitude Research Instrument*). Narzędzie to ma postać kwestionariusza składającego się ze 115 stwierdzeń. Badacze wyróżnili 23 skale postaw i każdej przypisali 5 stwierdzeń występujących w kwestionariuszu w sposób cykliczny. Obok każdego stwierdzenia znajdują się oznaczenia: **A a d D**, którym po badaniu nadaje się wartości liczbowe (wagi): 4, 3, 2, 1. Suma wag dla 5 stwierdzeń składających się na jedną skalę (podtest) odzwierciedla postawę cząstkową. Badani wyrażają swoją opinię poprzez zakreślenie odpowiedniej litery: **A** — *całkowicie się zgadzam*, **a** — *raczej się zgadzam*, **d** — *raczej się nie zgadzam*, **D** — *całkowicie się nie zgadzam*. Przykład 8 przedstawia jedną ze skal PARI.

Innym przykładem zastosowania skali skumulowanych ocen jest kwestionariusz samooceny dla dzieci — uczniów klas IV, V i VI szkoły podstawowej — skonstruowany przez Helenę Smirnową (2003) (patrz Przykład 9).

PRZYKŁAD 8. SKALA SKUMULOWANYCH OCEN (PARI)

KWESTIONARIUSZ PARI E.S. SCHAEFERA I R.Q. BELLA

(w przekładzie Józefa Rembowskiego)

Skala nr 1. Zachęcanie dziecka do wypowiadania własnych ocen i sądów

- Jeżeli dzieci uważają swe poglądy za słuszne, wolno im nie zgadzać się z poglądami rodziców A a d D
- Dzieci należy zachęcać do wypowiadania swojej opinii na temat współżycia w rodzinie, jeżeli uważają, że jest ono niewłaściwe A a d D
- Dziecko powinno mieć własne poglądy i moc je swobodnie wyrażać A a d D
- Przy podejmowaniu ważnych decyzji rodzinnych powinno się brać pod uwagę zdanie dziecka A a d D

PRZYKŁAD 9. SKALA SKUMULOWANYCH OCEN — NUMERYCZNA

KWESTIONARIUSZ SAMOOCENY POZIOMU REALIZACJI ZADAŃ ROZWOJOWYCH

DLA UCZNIÓW KLAS IV-VI SZKOŁY PODSTAWOWEJ

Helena Smirnowa (2003) — autorka kwestionariusza — przygotowała go, wykorzystując koncepcję zadań rozwojowych Roberta J. Havighursta (1981). Zadania przypisane przez Havighursta wiekowi szkolnemu i wczesnej fazie dorastania zostały podzielone na cztery grupy ze względu na to, jakich kompetencji dziecka dotyczą:

- 1) kompetencje poznawcze: (1) opanowanie umiejętności czytania, (2) opanowanie umiejętności pisania, (3) opanowanie umiejętności liczenia,
- 2) kompetencje emocjonalne: (4) socjalizacja emocji i kontrola nad emocjami,
- 3) kompetencje osobowościowe: (5) kształtowanie się samooceny, (6) kontrola moralna nad własnym postępowaniem, (7) uzyskiwanie osobistej niezależności od dorosłych, (8) kształtowanie się tożsamości płciowej,
- 4) kompetencje społeczne: (9) doskonalenie się współpracy w grupie, (10) budowanie związków z rówieśnikami, (11) wchodzenie w kontakty z osobami obcymi, (12) kształtowanie się postaw wobec instytucji społecznych.

Każdemu z dwunastu zadań rozwojowych przypisano różną liczbę stwierdzeń wybranych jako najbardziej reprezentatywne przez sędziów kompetentnych. Każde dziecko indywidualnie w obecności badacza wypełniało

kwestionariusz. Następnie zsumowano wagi przypisane każdemu zdaniu zaliczonemu do danego zadania rozwojowego, wreszcie obliczono cztery sumy odpowiednio do czterech obszarów kompetencji. Obliczone sumy i w pierwszym, i w drugim przypadku przeliczano na wskaźniki procentowe, by możliwe było porównywanie poziomu realizacji tak kolejnych zadań rozwojowych, jak i poziomu rozwoju czterech rodzajów kompetencji. Wskaźniki procentowe obliczono, dzieląc wartość średniej arytmetycznej przez wynik maksymalny, tj. możliwy do uzyskania w danej kategorii. Wynik maksymalny to liczba zdań przypisanych danej kategorii $\times 6$ pkt. Po tej transformacji wynik minimalny wynosił 0% i oznaczał, że dziecko bardzo nisko ocenia siebie w danym obszarze, wynik maksymalny wynosił 100% i oznaczał, że dziecko bardzo wysoko ocenia siebie w danym obszarze.

Przykłady stwierdzeń w kwestionariuszu:

(znakiem „*” oznaczono te, które punktowano odwrotnie, tj. waga „6” oznaczała niską samoocenę dziecka)

Dobrze się czuję z moimi rodzicami.	①	②	③	④	⑤	⑥
Jest mi obojętne, w co się ubieram.*	①	②	③	④	⑤	⑥
Widać, jak się złościę.*	①	②	③	④	⑤	⑥
Chętnie spędzam czas z moimi kolegami.	①	②	③	④	⑤	⑥
Nie lubię czytać.*	①	②	③	④	⑤	⑥
Sam zaczynam rozmowę z osobami nieznanymi.	①	②	③	④	⑤	⑥
Chętnie rozwiązuję zadania razem w grupie.	①	②	③	④	⑤	⑥

1.5. Skale z wymuszonym wyborem

Interesującą odmianą skali szacunkowej — ze względu na sposób konstruowania pozycji skali oraz sposób udzielania odpowiedzi — jest skala z wymuszonym wyborem zaprojektowana przez B.J. Cosgrove'a (1959) i przeznaczona przez niego do pomiaru umiejętności nauczyciela (*The descriptive ranking form for teachers*). Narzędzie to dostarcza rezultatu w postaci profilu, który jest odzwierciedleniem relatywnej siły czynników (cech, wymiarów), wg których dokonuje się oceny lub samooceny.

1.5.1. Charakterystyka skal szacunkowych z wymuszonym wyborem

Skala szacunkowa z wymuszonym wyborem (ang. *forced-choice rating scale*) jest narzędziem o wiele bardziej skomplikowanym od omówionych wcześniej typów skal szacunkowych. Jej konstrukcja wymaga od badacza znajomości podstawowych danych z zakresu psychometrii oraz przeprowadzenia wstępnych badań — na dość licznej próbie — zmierzających do wystandaryzowania skali. Dzięki specyficznej konstrukcji w skali tej nie występuje problem błędu łagodności (patrz dalej).

Skala szacunkowa z wymuszonym wyborem składa się z pewnej liczby pozycji, na które przypada taka sama liczba stwierdzeń (najczęściej od 2 do 5). Twierdzenia te opisują zachowanie się osoby badanej. Zadaniem osoby dokonującej oszacowań jest albo uporządkowanie tych stwierdzeń w obrębie każdej pozycji: od stwierdzenia najlepiej do najgorzej opisującego zachowanie, albo też wybór jednego stwierdzenia najlepiej opisującego i jednego najgorzej opisującego dane zachowanie. Wybrany lub uporządkowany stwierdzeniem przypisuje się następnie punkty wg określonego klucza i na ich podstawie oblicza wyniki dla poszczególnych osób badanych.

Pozycje mogą przybierać różną postać. R.W. Highland i J.R. Berkshire (za: Guilford 1954, s. 276) przeanalizowali sześć różnych typów pozycji na podstawie wyników badań przeprowadzonych na grupie osób dokonujących oszacowań. Oto one:

- a) pozycja składa się z dwóch stwierdzeń (opisujących „dobre” lub „złe” zachowanie), a osoba dokonująca oszacowań wybiera jedno stwierdzenie najlepiej (lub najgorzej) opisujące zachowanie się badanego (lub swoje),
- b) pozycja składa się z trzech stwierdzeń: wszystkie opisują albo „dobre”, albo „złe” zachowanie, a osoba dokonująca oszacowań wybiera jedno stwierdzenie najlepiej opisujące i jedno najgorzej opisujące zachowanie się badanego (lub swoje),
- c) pozycja składa się z czterech stwierdzeń opisujących „dobre” zachowanie, osoba szacująca wybiera dwa stwierdzenia najlepiej opisujące zachowanie badanego (lub swoje),
- d) pozycja składa się z czterech stwierdzeń opisujących „dobre” zachowanie, osoba szacująca wybiera jedno stwierdzenie najlepiej opisujące i jedno najgorzej opisujące zachowanie badanego (lub swoje),

- e) pozycja składa się z dwóch stwierdzeń opisujących „dobre” zachowanie i dwóch opisujących „złe” zachowanie, osoba szacująca wybiera jedno stwierdzenie najlepiej opisujące i jedno najgorzej opisujące zachowanie badanego (lub swoje),
- f) pozycja składa się z pięciu stwierdzeń (dwa opisują „dobre” zachowanie, dwa „złe”, jedno stwierdzenie jest neutralne), osoba dokonująca oszacowań wybiera jedno stwierdzenie najlepiej opisujące i jedno najgorzej opisujące zachowanie badanego (lub swoje).

Highland i Berkshire (za: ibidem) przanalizowali wymienione typy pozycji pod kątem ich rzetelności, trafności, podatności na stronniczość i popularności u osób dokonujących oszacowań. Przeprowadzona analiza pokazała, że:

- 1) pozycje o postaci: (e) i (f) dostarczają najwyższych wskaźników rzetelności,
- 2) pozycje o postaci: (c) i (d) dają najwyższe wskaźniki trafności,
- 3) pozycje o postaci: (c) i (b) wykazują najmniejszą podatność na stronniczość,
- 4) pozycje o postaci: (c) i (a) były oceniane jako najbardziej popularne,
- 5) pozycje o postaci: (b) i (d) jako najmniej popularne.

W rezultacie typ (c) został uznany za najlepszy ze względu na wszystkie analizowane kryteria.

Czy „długość” skali ma wpływ na jej trafność? Na to pytanie odpowiedzieli E.K. Taylor, D.E. Schneider i N.A. Symonds (por. Remmers 1967, s. 343). Porównali oni trafność dwóch skal, z których jedna składała się z 12 tetrad, a druga z 32 tetrad. Trafność była taka sama w obu przypadkach.

1.5.2. Wybór czynników i dobór stwierdzeń do pozycji

Konstrukcję skali z wymuszonym wyborem musimy rozpocząć od wyboru czynników, które będą podstawą do przeprowadzania oceny lub samooceny osób badanych. Decyzję w tej sprawie podejmujemy na podstawie analizy teoretycznej danego zagadnienia. Z określonej teorii osobowości, motywacji itp. wynika zarówno liczba, jak i jakość czynników, które należy uwzględnić przy budowie skali. Do każdego czynnika dobieramy stwierdzenia będące jego wskaźnikami.

Ile stwierdzeń powinny liczyć owe listy? Otóż „wyjściowa” długość każdej listy powinna być większa o około 50-100% od planowanej

długości końcowej. Jest to konieczne, gdyż w wyniku przeprowadzanej w dalszych etapach analizy psychometrycznej poszczególnych stwierdzeń część z nich odpadnie, jako że nie będzie spełniała określonych wymagań.

D.J. Cosgrove (op. cit.) przyjął czteroczynnikową koncepcję swojej skali do oceny efektywności pracy nauczyciela w szkole średniej. Oto czynniki, które wyróżnił:

- a) znajomość i umiejętność organizowania podległych sobie spraw,
- b) poprawność stosunków z uczniami w klasie,
- c) poprawność planowania i stosowanych metod pracy,
- d) zapał w pracy z uczniami.

Stwierdzenia do skali wybrał ze zbioru 900 zdań opisujących specyficzne elementy pracy nauczyciela. Zbiór ten powstał w wyniku przeanalizowania opisów „dobrego” nauczyciela sporządzonych przez studentów psychologii w college’u. Z listy tej wybrał do dalszej analizy 200 zdań, które stanowiły krótkie opisy zachowania się nauczyciela w różnych sytuacjach szkolnych. Pięciu wykładowców psychologii oraz studenci studiów psychologicznych sortowali niezależnie od siebie stwierdzenia zapisane na odrębnych kartkach do pięciu kategorii:

- 1) opanowanie i zorganizowanie podległych sobie spraw,
- 2) zdolność kontroli i utrzymywania dyscypliny wśród uczniów,
- 3) racjonalność wymagań stawianych uczniom w zakresie czasu i wysiłku przy okazywanej pomocy i wskazówkach,
- 4) efektywność i konsekwencja procedur kierowania klasą,
- 5) zdolność do motywowania, inspirowania, stwarzania zaufania u uczniów.

Dalszej analizie zostało poddanych 150 zdań, co do których sędziowie byli zgodni w maksymalnym stopniu. Po przeprowadzeniu analizy czynnikowej na pięciu zbiorach oszacowań zdań, dokonanych przez grupę 100 studentów w obrębie każdej z pięciu kategorii, wyróżniono jeden ogólny czynnik, dwa mniej ogólne i cztery specyficzne czynniki efektywności pracy nauczyciela. Ze zdań, które w najwyższym stopniu korelowały z czynnikami specyficznymi, Cosgrove utworzył cztery listy.

W zależności od tego, ile czynników badacz wyróżni — tyle stwierdzeń powinna liczyć każda pozycja skali. W przypadku dwóch czynników pozycje przyjmują postać diad, w przypadku

trzech czynników — postać triad, czterech czynników — postać tetrad itd. Cosgrove formułował pozycje swej skali w postaci tetrad: stwierdzenia zaliczone do poszczególnych pozycji miały charakter pozytywny, tzn. opisywały „dobre” zachowanie się nauczyciela. Miejsce każdego stwierdzenia w danej pozycji ustala się w sposób losowy. Skala Cosgrove’a w końcu składała się z 10 w podobny sposób zbudowanych pozycji. Zadaniem badanego było uporządkowanie stwierdzeń — w obrębie każdej pozycji oddzielnie — od najlepiej opisującego danego nauczyciela (ranga 1) do najgorzej opisującego (ranga 4). Oto, tytułem przykładu, pozycja zaczerpnięta ze Skali Cosgrove’a. Litery w nawiasach oznaczają przyporządkowanie stwierdzenia do danego czynnika.

Pozycja 5:

- cieszy się autorytetem opartym na wiedzy (a)
- przyjaźnie ustosunkowany wobec uczniów (d)
- ocenia bardzo sprawiedliwie (c)
- nigdy nie krytykuje w sposób obraźliwy (b)

1.5.3. Obliczanie wskaźników preferencji i różnicowania dla poszczególnych stwierdzeń

Po sporządzeniu list stwierdzeń dla poszczególnych czynników przystępujemy do obliczenia — dla każdego stwierdzenia oddzielnie — dwóch wskaźników: (1) wskaźnika preferencji (ang. *preference index*) i (2) wskaźnika różnicowania (ang. *discriminative index*). Obliczenia tych wskaźników dokonujemy, opierając się na opinii grupy sędziów kompetentnych (mogą nimi być studenci wyższych lat studiów psychologicznych). Grupa sędziów powinna być dość liczna. Na przykład Cosgrove posłużył się grupą 100-osobową. Sędziowie wg odpowiednich instrukcji oceniają odpowiedniość poszczególnych stwierdzeń dla danej skali.

Powróćmy znów do prezentacji procedury zastosowanej przez Cosgrove’a. Listę 150 stwierdzeń podał 100 sędziom — studentom wyższych lat psychologii wychowawczej uniwersytetu w Ohio. Sędziów podzielono na trzy grupy. Każdego sędziego poproszono o przypomnienie sobie znanego mu nauczyciela, przy czym:

- 1) w przypadku grupy pierwszej jej członków proszono o przypomnienie sobie takiego nauczyciela, którego można by ocenić w zakresie od 1 do 10 pkt. na skali 20-punktowej (od „1” — zły do „20” — dobry),
- 2) w przypadku grupy drugiej chodziło o przypomnienie sobie takiego nauczyciela, którego można by ocenić na tej samej skali, ale w zakresie od 1 do 20 pkt.,
- 3) trzecią grupę sędziów poproszono o przypomnienie sobie takiego nauczyciela, którego można by również ocenić na tej samej skali — w zakresie od 9 do 16 pkt.

Scharakteryzowana wyżej procedura została tak zaplanowana, aby w efekcie można było otrzymać rozkład rang zbliżony do normalnego. Sędziowie z każdej grupy oceniali każde ze 150 stwierdzeń pod kątem odpowiedniości (ang. *applicability*) na 5-punktowej numerycznej skali szacunkowej. Sędzia powinien przyznawać 1 pkt tym stwierdzeniom, które w niskim stopniu opisywały danego nauczyciela, a 5 pkt. tym, które tego nauczyciela opisywały w stopniu wysokim. Stwierdzeniom o pośrednich stopniach odpowiedniości powinni sędziowie przypisywać od 2 do 4 pkt. Tak więc np. sędzia uznając zdanie „Udziela wyczerpujących odpowiedzi na pytania uczniów” za opisujące danego nauczyciela w najwyższym stopniu, przyznawał temu zdaniu 5 pkt. Wskaźnik preferencji, wyrażający stopień gotowości sędziów studentów do użycia danego stwierdzenia do opisu nauczyciela obliczono w następujący sposób: (1) zliczono wszystkie punkty przypisane danemu stwierdzeniu przez poszczególnych sędziów (razem: 100 ocen), (2) obliczono średnią ocen dla każdego stwierdzenia.

Wskaźnik różnicowania, wyrażający stopień, w jakim dane stwierdzenie było odpowiednie do opisu dobrego nauczyciela, obliczono przez skorelowanie oceny odpowiedniości tego stwierdzenia (od 1 do 5 pkt.) z oceną nauczyciela (od 1 do 20 pkt.) wydaną przez każdego sędziego. A zatem dla każdego stwierdzenia otrzymano współczynnik korelacji będący wskaźnikiem różnicowania tego stwierdzenia (por. Blok rozszerzający 1, tabela 2).

OBLICZENIE WSKAŹNIKÓW PREFERENCJI I RÓŻNICOWANIA WG E.D. SISSONA

Odminną procedurą obliczania wskaźników preferencji i różnicowania dla poszczególnych stwierdzeń posłużył się E.D. Sisson (por. Ferguson 1952, s. 317-318). Każde stwierdzenie oceniane jest na pięciopunktowej numerycznej skali szacunkowej (0 – 1 – 2 – 3 – 4) przez grupę sędziów (w badaniach Sissona grupa sędziów liczyła 300 osób). Każdy z sędziów ma wyobrazić sobie dobrze znaną mu osobę spełniającą dane kryterium, ze względu na które dobierano stwierdzenia (u Sissona sędziami byli oficerowie armii, którzy oceniali znanych sobie oficerów ze względu na sukcesy, jakie odnosili w armii). Jeżeli dane stwierdzenie bardzo dobrze opisuje daną osobę — sędzia powinien przypisać mu 4 pkt. na skali szacunkowej, a jeżeli źle — 0 pkt. Po uzyskaniu oszacowań od całej grupy sędziów dzieli się je na trzy zbiory (wg ogólnego sumarycznego wyniku przypisanego ocenianej osobie) w następujący sposób:

- dołączamy do siebie punkty przypisane przez każdego sędziego poszczególnym stwierdzeniem opisującym wyobrażoną osobę modelową; w ten sposób dla każdej osoby ocenianej przez danego sędziego otrzymamy jeden sumaryczny wynik,
- porządkujemy zbiór wyników od najwyższego do najniższego (wynik najwyższy otrzymuje osoba, dla której stwierdzenia skali uzyskały najwyższą ocenę od sędziego; wynik najniższy otrzymuje osoba, dla której stwierdzenia skali uzyskały najniższą ocenę od sędziego),
- wydzielamy trzy równoliczne grupy wyników: górną, środkową i dolną (jeżeli w badaniach brało udział np. 300 sędziów — jak u Sissona — to grupy będą liczyły po 100 wyników).

Następną czynnością, jaką badacz musi wykonać, jest sporządzenie — dla każdego stwierdzenia oddzielnie — pomocniczych tabel do obliczania wskaźnika preferencji i wskaźnika różnicowania (por. tab. 3). W pola tabeli wpisujemy w kolejne wiersze odnoszące się do trzech wyłonionych grup (górną, środkową i dolną) liczbę sędziów, którzy przypisali danemu stwierdzeniu 0 pkt., liczbę sędziów, którzy przypisali danemu stwierdzeniu 1 pkt., 2 pkt., 3 pkt. i 4 pkt. Suma w ostatniej kolumnie zawsze powinna być taka sama. Podane w tabeli sumy „491” i „50” wstawiamy do odpowiednich wzorów (por. Ferguson 1952, s. 318):

$$\text{wskaźnik preferencji} = \frac{(1/3)N\Sigma(f_{og}X)}{N} = \frac{(1/3)(150)(491)}{150} = 163,6,$$

$$\text{wskaźnik różnicowania} = \Sigma(f_g - f_d) = 1 + 5 + 11 + 8 + 25 = 50.$$

TABELA 3

Tabela pomocnicza do obliczania wskaźników preferencji i różnicowania dla stwierdzenia X (ogólna liczba sędziów $N = 150$)

Punkty skali szacunkowej	0	1	2	3	4	Suma
Liczebność (f)						
Górna grupa f_g	1	0	3	3	43	50
Środkowa grupa f_s	2	2	6	8	32	50
Dolna grupa f_d	2	5	14	11	18	50
f_{og}	5	7	23	22	93	150
$f_{og} \times X$	0	7	46	66	372	491
$f_g - f_d$	1	5	11	8	25	50

1.5.4. Konstrukcja pozycji skali

Po zakończeniu procedury obliczania wskaźników preferencji i różnicowania rozdzielamy stwierdzenia, tworząc listy przyporządkowane wyróżnionym wcześniej czynnikom (wymiarom). Następnie łączymy stwierdzenia (po jednym z każdej listy) w bloki — w diady, triady, tetrazy itd. Łącząc stwierdzenia w bloki (pozycje), kierujemy się wartością wskaźników preferencji i różnicowania. I tak stwierdzenia należące do jednego bloku muszą charakteryzować się tą samą wysokością wskaźnika preferencji. Stanowi to gwarancję, że badani porządkując owe stwierdzenia w ramach każdego bloku, nie będą kierowali się ich „popularnością” w społeczeństwie, ale stopniem ich odpowiedniości do opisu konkretnej osoby. Oprócz podobnych wskaźników preferencji wymaga się od stwierdzeń zaliczonych do jednego bloku (pozycji), aby charakteryzowały się podobną wysokością wskaźników różnicowania. Ten zabieg z kolei ma dostarczyć pewności, że wszystkie stwierdzenia zaliczone do jednego bloku są mniej więcej w tym samym stopniu zrelatywizowane do poziomu oceny danego badanego. Jasne jest zatem, dlaczego wyjściowe listy stwierdzeń powinny być znacznie dłuższe od list projektowanych. Pozwalają one bowiem na odrzucenie tych stwierdzeń, które nie „sprawdziły się”, i dobrane stwierdzeń o lepszych parametrach.

Kierując się wyżej przedstawionymi regułami, Cosgrove wybrał ze 150 stwierdzeń 40 i ułożył z nich 10 pozycji (bloków) w postaci tetrad.

Każdemu stwierdzeniu w obrębie danej tetrady przypisuje się litery (lub inne znaki) oznaczające jego przyporządkowanie do danego czynnika. W celu określenia trafności poszczególnych pozycji skali oceniamy za jej pomocą grupę osób o znanych właściwościach (ocena trafności przez porównanie z kryterium zewnętrznym).

Przykład 10 pokazuje zastosowanie pomysłu Cosgrove'a, polegającego na „wymuszeniu” od osób badanych wyboru jednej z podanych — odmiennych jakościowo — odpowiedzi.

PRZYKŁAD 10

SKALA Z WYMUSZONYM WYBOREM DO BADANIA OCENY WŁASNYCH MOŻLIWOŚCI PORADZENIA SOBIE W SYTUACJI ZMIANY

Autorką kwestionariusza wykorzystującego skalę z wymuszonym wyborem jest Karolina Appelt (2004). Ocena poziomu własnych możliwości poradzenia sobie w sytuacji zmiany została odniesiona do doświadczanej od kilku lat przez nauczycieli reformy edukacji w trzech obszarach: 1) formalno-prawnym, 2) związanym z procesem kształcenia i 3) związanym z procesem wychowania i opieki nad uczniami. Kwestionariusz opracowano w dwóch wersjach: dla kobiet i dla mężczyzn, przy czym wersje te różniły się jedynie formą zwracania się do osoby badanej (Pani/Pan oraz żeńskimi i męskimi rodzajami użytych czasowników).

Wstępny zestaw opisów sytuacji trudnych, związanych z reformą systemu edukacji powstał w efekcie studiowania literatury (m.in. zeszytów wydawanych przez Ministerstwo Edukacji Narodowej w serii „Biblioteczka reformy”), dokumentów dotyczących reformy, Karty Nauczyciela, rozporządzeń Ministra Edukacji Narodowej itp. oraz własnych doświadczeń autorki związanych z pracą w szkole jako psycholog szkolny. W każdym obszarze wyróżniono, zgodnie z podziałem zaproponowanym przez Lazarusa i Folkmana (1984), trzy typy sytuacji trudnych: krzywda/strata, zagrożenie, wyzwanie.

Aby uzyskać informacje o sytuacjach, które przez samych nauczycieli są uznawane za trudne i związane są z nowymi stawianymi im wymaganiami, przeprowadzono anonimowe badania wśród nauczycieli chemii odbywających studia podyplomowe na Wydziale Chemii UAM w roku akademickim 2000/2001 (16 osób), nauczycieli pracujących w IV Liceum Ogólnokształcącym w Szczecinie (22 osoby) oraz w Zespole Szkół z Oddziałami Integracyjnymi nr 2 w Poznaniu. Badania polegały na wypisaniu, na przy-

gotowanych w tym celu arkuszach, sytuacji uznawanych za trudne w związku ze zmianami, jakie niesie reforma systemu edukacji.

W efekcie tych badań, po usunięciu pozycji powtarzających się oraz nadaniu spójnej formy gramatycznej powstała pula licząca 54 opisy sytuacji trudnych. Aby uzyskać narzędzie o wysokim stopniu trafności, poproszono siedmiu sędziów kompetentnych (czworo z nich to nauczyciele Szkoły Podstawowej nr 78 w Poznaniu, a troje jest nauczycielami Gimnazjum nr 55 w Poznaniu) o wskazanie tych sytuacji, które mogą odnosić się do nauczycieli ze wszystkich typów szkół, w różnym wieku, różnej płci. Ważne było również, aby wybierać te opisy, które wiązały się z rzeczywistością nowymi wymaganiami, wynikającymi z reformy. Ten etap badań miał miejsce na przełomie maja i czerwca 2002 roku, czyli w czasie związanym z pisaniem sprawozdań z przebiegu stażu, przystępowaniem do egzaminów kwalifikacyjnych itp. i to prawdopodobnie wpłynęło na fakt, że wybory sędziów kompetentnych oraz wysoki stopień zgodności dotyczyły sytuacji związanych głównie z aspektem formalno-prawnym działalności nauczycielskiej.

Narzędzie — poprawione w wyniku badań pilotażowych — składa się z 27 krótkich opisów sytuacji związanych z nowymi wymaganiami stawianymi przed nauczycielami.

Zadanie osoby badanej polegało na przeczytaniu opisów sytuacji i udzieleniu odpowiedzi na dwa pytania. Odpowiedź na pierwsze pytanie: Jak ocenia Pani swoje możliwości poradzenia sobie z tą sytuacją? polegała na wybraniu, poprzez zaznaczenie znakiem ×, jednej z możliwych sześciu odpowiedzi: bardzo nisko — nisko — raczej nisko — raczej wysoko — wysoko — bardzo wysoko. Odpowiedź na drugie pytanie: Jak sobie Pani najczęściej radzi w tej sytuacji? następowała poprzez wybór jednego z trzech opisów sposobu radzenia sobie.

1. Dostosowywanie wymagań do indywidualnych możliwości ucznia. Obowiązek nauczyciela, by służyć uczniowi pomocą w przezwyciężaniu niepowodzeń szkolnych i mieć pełne rozeznanie potrzeb uczniów w tym zakresie.

• Jak ocenia Pani swoje możliwości poradzenia sobie z tą sytuacją?

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bardzo nisko	nisko	raczej nisko	raczej wysoko	wysoko	bardzo wysoko

• Jak sobie Pani najczęściej radzi w tej sytuacji?

Radzę sobie sama, innych nie chcę obarczać moimi problemami, mają przecież swoje.	Zwracam się do kompetentnych osób z prośbą o pomoc, chcę wykorzystać ich doświadczenie.	Próbuję sama, ale gdy nie najlepiej sobie radzę, szukam pomocy innych osób.
---	---	---

2. Tworzenie własnych planów i programów nauczania przeznaczonych dla konkretnego oddziału danej szkoły. Nauczyciel nie realizuje ogólnych decyzji programowych, sam decyduje o swej pracy i odpowiada za jej rezultaty.

- Jak Pani ocenia swoje możliwości poradzenia sobie z tą sytuacją?

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
bardzo nisko	nisko	raczej nisko	raczej wysoko	wysoko	bardzo wysoko
- Jak sobie Pani najczęściej radzi w tej sytuacji?

Radzę sobie sama, innych nie chcę obarczać moimi problemami, mają przecież swoje.	Zwracam się do kompetentnych osób z prośbą o pomoc, chcę wykorzystać ich doświadczenie.	Próbuję sama, ale gdy nie najlepiej sobie radzę, szukam pomocy innych osób.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

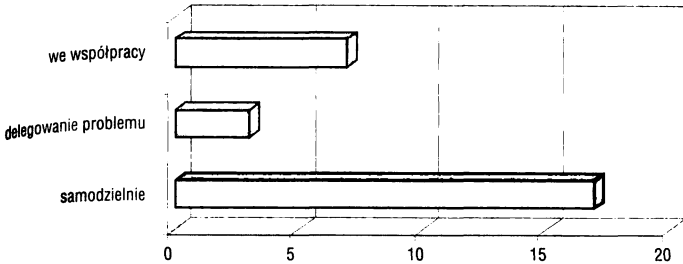
Ocena wyników polega na zliczeniu, ile razy został wybrany przez osobę badaną dany sposób korzystania z pomocy innych w każdej z trzech wyróżnionych sfer działalności nauczyciela: 1) formalno-prawnej, 2) dydaktycznej, 3) wychowawczo-opiekuńczej. Każda badana osoba uzyskuje więc trzy sumy, zawsze składające się na sumę ogólną = 27 pkt. Można sporządzić dwa profile (dane z tabeli 4): jeden obrazujący ogólnie styl radzenia sobie, a drugi pokazujący radzenie sobie w trzech badanych obszarach (por. rysunek 5a i 5b.).

TABELA 4
Rozkład odpowiedzi nauczyciela X

OBSZAR	formalnoprawny: 9 pozycji	procesu kształcenia: 8 pozycji	procesu wychowania i opieki: 10 pozycji	Suma punktów
Sposób zachowania A: samodzielnie	6	7	4	17
Sposób zachowania B: delegowanie problemu	2	0	1	3
Sposób zachowania C: we współpracy	1	1	5	7
Suma punktów	9	8	10	27

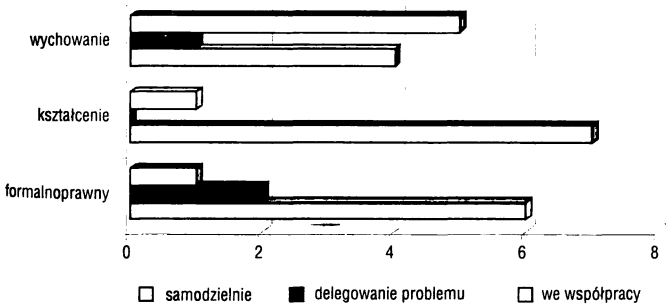
Źródło: opracowanie własne.

Rysunek 5a. Styl radzenia sobie nauczyciela X



Źródło: opracowanie własne.

Rysunek 5b. Styl radzenia sobie nauczyciela X w zależności od rodzaju problemu



Źródło: opracowanie własne.

1.5.5. Przebieg badania i analiza wyników

Osoby badane otrzymują listę pozycji. Obok każdego stwierdzenia **znaczamy** miejsce (w postaci kratki, okienka), w które mają wpisać swoją odpowiedź. W instrukcji prosimy badanych, aby np. uporządkowali stwierdzenia w ramach każdej pozycji (uwaga: dla każdej pozycji oddzielnie) od najlepiej opisującego zachowania danej osoby (lub samego siebie) przez przypisanie danemu stwierdzeniu rangi 1 do najgorzej

opisującego to zachowanie — przez przypisanie danemu stwierdzeniu rangi 4. Układ pozycji w skali może być następujący (pozycje zaczerpnięto ze skali Cosgrove'a):

Pozycja 1.

- | | |
|---|--|
| <input type="checkbox"/> Zawsze punktualny (c) | <input type="checkbox"/> Pracuje zawsze zgodnie z planem (c) |
| <input type="checkbox"/> Uprzejmy dla uczniów (b) | <input type="checkbox"/> Zajęcia łączą się ze sobą „na zakładkę” (a) |
| <input type="checkbox"/> Szczery w rozmowach z uczniami (d) | <input type="checkbox"/> Umie zainteresować danym zagadnieniem (d) |
| <input type="checkbox"/> Odczytany w literaturze (a) | <input type="checkbox"/> Wykazuje zrozumienie problemów uczniów (a) |

Po zakończeniu badania punktujemy rangi przypisane stwierdzeniom każdej pozycji wg klucza:

RANGI	1	2	3	4
PUNKTY	3	2	1	0

Następnie sumujemy punkty przypisane wszystkim stwierdzeniom typu (a), potem (b), (c) i (d). W ten sposób otrzymujemy cztery wyniki cząstkowe, które ukazują względną wartość danego czynnika na tle pozostałych. Im wyższy jest dany wynik cząstkowy, tym większa jest jego ważność w stosunku do pozostałych czynników.

Zauważmy, że bezpodstawne jest obliczanie jednego ogólnego wyniku (czyli sumy wszystkich punktów) dla danej osoby. Jest on taki sam dla każdej badanej osoby, gdyż rangowanie jest przykładem pomiaru o charakterze ipsatywnym (Cattell 1944), tzn. dla każdej badanej osoby suma punktów jest taka sama. W przypadku oryginalnej skali Cosgrove'a wynosi ona 60 pkt. (15 pozycji × 4 pkt.).

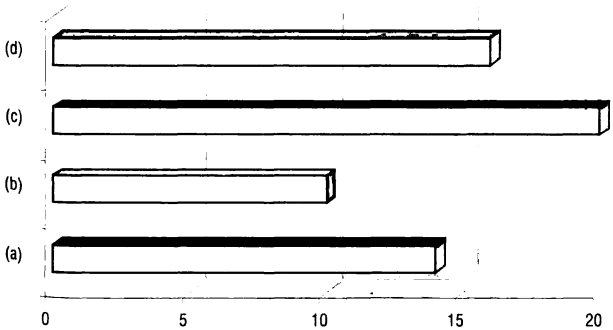
Z uzyskanych dla każdej osoby wyników cząstkowych możemy sporządzić profil psychometryczny, który nanosimy na kartę profilową wg wzoru przedstawionego na rysunku 5. Z profilu przedstawionego na rysunku 6 (dane z tabeli 5) widać, że w ocenie efektywności nauczyciela największy udział miał czynnik (c), czyli „poprawność planowania i stosowanych metod pracy”, a najniższy czynnik (b), czyli „poprawność stosunków z uczniami w klasie”.

TABELA 5

Karta wyników dla skali Cosgrove'a (wyniki pochodzą od jednej osoby badanej)

Pozycja skali	Badany czynnik			
	a	b	c	d
1	0	1	2	3
2	0	2	1	3
3	1	0	3	2
4	2	1	3	0
5	1	0	2	3
6	2	1	3	0
7	1	0	2	3
8	3	2	1	0
9	3	1	0	2
10	1	2	3	0
Suma	14	10	20	16

Źródło: opracowanie własne.

Rysunek 6. Profil wyników w skali Cosgrove'a
(z naniesionymi przykładowo wynikami pochodzącymi od jednej osoby)

Źródło: opracowanie własne na podstawie Cosgrove 1959.

Podkreślmy raz jeszcze, że profil odzwierciedla tylko względną wagę danego czynnika (na tle pozostałych) w ocenie danej osoby. Jeżeli dla jednego czynnika osoba oceniana uzyskała maksymalny wynik (u Cosgrove'a $10 \times 3 = 30$ pkt.), to dla żadnego z pozostałych czynników już nie może ona otrzymać tak wysokiego wyniku. Mówiąc inaczej, wyniki cząstkowe powstają w efekcie rozdzielenia sumy ogólnej na sumy składowe. Dla skali Cosgrove'a suma ta wynosi:

$$(10 \times 0 \text{ pkt.}) + (10 \times 1 \text{ pkt.}) + (10 \times 2 \text{ pkt.}) + (10 \times 3 \text{ pkt.}) = \\ = 0 + 10 + 20 + 30 = 60 \text{ pkt.}$$

W przykładzie przedstawionym w tabeli 4 owe 60 pkt. rozdzieliło się w następujący sposób:

$$60 \text{ pkt.} = 14 \text{ pkt.} + 10 \text{ pkt.} + 20 \text{ pkt.} + 16 \text{ pkt.} \\ \text{(a) \quad (b) \quad (c) \quad (d)}$$

Jak pisze Cosgrove (1959, s.203): „Podejście profilowe przedstawia bardziej zorganizowaną i użyteczną analizę oszacowania silnych i słabych stron jego [tj. nauczyciela — A.B. i J.B.] pracy. Narzędzie z wymuszonym wyborem minimalizuje również stronniczość, która często występuje w typowo graficznych schematach szacowania [...]”. Badając taką skalą grupę osób o znanych właściwościach, możemy sporządzić profile na podstawie wartości średnich arytmetycznych i przeprowadzić porównania intergrupowe (np. porównać opinie uczniów z różnych typów szkół, opinie chłopców i dziewcząt, opinie uczniów i ich rodziców odnośnie do pracy nauczycieli oraz nauczycieli różnych przedmiotów na swój temat).

1.6. Zastosowanie skal szacunkowych wg tzw. techniki Q-sort

1.6.1. Ogólna charakterystyka techniki Q-sort

Istota techniki Q polega na zastosowaniu procedury sortowania kartek z wypisanymi na nich stwierdzeniami czy przymiotnikami do kilku rozłącznych kategorii, rozłożonych wzdłuż k -punktowego kontinuum (por. Paszkiewicz 1974). Krańce owego kontinuum oznaczają odpowiednio:

- a) lewy kraniec: całkowity brak zgodności danego stwierdzenia z kryterium sortowania — przypisujemy mu najniższą punktację, tj. 0,
- b) prawy kraniec: całkowita zgodność danego stwierdzenia z kryterium sortowania — przypisujemy mu najwyższą punktację w przyjętej skali.

Przesuwanie ocen wzdłuż kontinuum od jego lewego krańca do prawego oznacza wzrost stopnia zgodności ocenianego przez badanego stwierdzenia z kryterium sortowania.

Kryterium, wg którego badany będzie porządkował stwierdzenia (sortowanie kartek z wypisanymi na nich stwierdzeniami do rozłącznych kategorii oznaczających różny stopień zgodności danego stwierdzenia z kryterium prowadzi w efekcie do uporządkowania całego zbioru kartek), może być własne zachowanie badanego (ocenie podlega wtedy stopień, w jakim dane stwierdzenie trafnie opisuje to zachowanie), system norm moralnych (ocenie podlega stopień, w jakim badany akceptuje stwierdzenie opisujące jakąś normę moralną) itp. Za pomocą techniki Q można dokonywać nie tylko samooceny, ale także oceny innych osób czy obiektów. Można też tę samą osobę badać kilkakrotnie — zmieniając tylko kryterium, wg którego ma ona porządkować podane stwierdzenia. Na przykład w pierwszym badaniu może ona oceniać stopień zgodności poszczególnych stwierdzeń z kryterium: *jaki jestem* (ocena tzw. ja-realnego), w drugim badaniu może oceniać stopień zgodności stwierdzenia z kryterium *jaki chciałbym być* (ocena tzw. ja-idealnego).

W uwagach wstępnych należy jeszcze podkreślić, że dużą zaletą techniki Q jest to, iż z góry określona jest liczba kartek, jaką badany może przydzielić do danej kategorii. Jest ona najczęściej zgodna z charakterem rozkładu normalnego. W ten sposób eliminuje się błąd grupowania kartek wokół paru kategorii i niewykorzystywania przez badanego innych. Na przykład badany może umieszczać wszystkie kartki tylko w środku kontinuum, nie wykorzystując w ogóle jego krańców (tzw. błąd tendencji centralnej) lub też umieszczać kartki na krańcach kontinuum, nie zapewniając jego środkowej części (tzw. błąd skrajnych ocen). O tych błędach będzie mowa dalej.

Nazwa „technika Q” pochodzi od jej twórcy W. Stephensona, który opisał ją w podstawowej pracy traktującej o technice i metodologii Q *The study of behavior. Q-technique and its methodology* (1953), natomiast zbiór pozycji nazywa się często Q-sortem.

1.6.2. Kryterium i dobór pozycji

Stosując technikę Q, możemy porządkować (przez rozsortowanie kartek do rozłącznych kategorii) albo stwierdzenia, pojedyncze słowa, np. przymiotniki, obrazki, albo inne obiekty, które chcemy poddać ocenie z punktu widzenia stopnia ich zgodności z kryterium sortowania. Kryterium sortowania może być:

- a) zgodność opisu wyrażonego danym stwierdzeniem (wyrazem) z faktycznym wyglądem fizycznym, faktycznie przejawianymi postawami i wyrażanymi przekonaniem, z faktycznie akceptowanym systemem wartości itp. badanej osoby,
- b) zgodność opisu, wyrażonego danym stwierdzeniem (wyrazem) z pożądanym przez daną osobę wyglądem fizycznym, postawami, przekonaniem, które osoba ta chciałaby u siebie wykształcić,
- c) zgodność wytworów sztuki plastycznej (rysunek, rzeźba, obrazy itp.) z akceptowanym przez daną osobę kryterium piękna,
- d) zgodność zachowania się ocenianej osoby z zachowaniem „osoby wzorcowej”,
- e) zgodność opisu wyrażonego danym zdaniem (wyrazem) z faktycznymi warunkami pracy w jakimś zakładzie pracy itp.

Od pozycji, którymi posługujemy się, stosując technikę Q, wymaga się, by dały się one uporządkować (rozsortować) wg danego kryterium wzdłuż *k*-punktowego kontinuum. Zbiór pozycji powinien być jednorodny. Jeżeli np. prosimy badanego, by dokonał oceny własnego wyglądu fizycznego wg instrukcji *jaki jestem*, to w zbiorze pozycji nie mogą znaleźć się takie stwierdzenia, jak *rozumiem postępowanie innych ludzi czy można mi powierzać odpowiedzialne zadania*, ale powinny znaleźć się np. następujące:

- *cehuje mnie drobna budowa ciała,*
- *mam wysportowaną sylwetkę,*
- *moja twarz może się podobać,*
- *mam piękne gęste włosy* itp.

Gdy badany ma ocenić własne możliwości intelektualne wg instrukcji *jaki jestem*, to w zbiorze pozycji mogą znaleźć się następujące stwierdzenia:

- *łatwo mnie nabrać,*
- *wszystko trzeba mi tłumaczyć kilka razy,*
- *szybko uczę się nowych rzeczy,*

- *zrozumienie wielu spraw sprawia mi trudności,*
- *jestem osobą utalentowaną itp.*

Możemy także pytać badanego o to, jak powinien zachować się wobec innych osób. Badany może więc porządkować zbiór np. przymiotników wedle instrukcji: *jaki powinienem być, jak powinienem się zachować.* Oto przykłady odpowiednich przymiotników:

natrętny	towarzyski	sprawiedliwy	wyrozumiały
swawolny	bezinteresowny	stanowczy	uprzejmy

Badany może oceniać także warunki pracy w zakładzie pracy wg instrukcji *jak jest.* Odpowiednie do tego celu mogą być następujące stwierdzenia:

- *hałas uniemożliwia porozumiewanie się z innymi pracownikami,*
- *stanowisko pracy jest w wystarczającym stopniu oświetlone,*
- *nie odczuwam wibracji na stanowisku pracy.*

Wyżej przedstawione próbki zbiorów pozycji cechowały się tylko tym, że były homogenicznymi zbiorami stwierdzeń (pojedynczych słów), dających się uporządkować wzdłuż pewnego k -punktowego kontinuum. Technikę Q z zastosowaniem takich zbiorów pozycji nazywa się nieustrukturowaną techniką Q (por. badania B. Ziółkowskiej — Przykład 11). Jest ona bardzo popularna m.in. ze względu na łatwość doboru pozycji, które mogą być czerpane z bardzo rozpowszechnionych wśród psychologów kwestionariuszy osobowości i skal postaw, a także ze względu na niski stopień trudności, jeżeli chodzi o stosowanie jej i obliczanie wyników. Z techniki Q w postaci nieustrukturowanej korzystał znany psycholog Carl Rogers i jego uczniowie, oceniając postępy pacjentów w psychoterapii.

Przez nałożenie dodatkowych warunków na zbiór pozycji można otrzymać tzw. ustrukturowaną technikę Q. Najprostszy sposób ustrukturowania zbioru pozycji polega na takim doborze stwierdzeń (pojedynczych wyrazów), aby połowa z nich oddawała wysokie natężenie zmiennej, stanowiącej kryterium sortowania, a połowa z nich oddawała niskie natężenie. Taką procedurą posłużyli się Butler i Haigh (por. Byrne 1966, s. 442-443) przy konstruowaniu tzw. *Chicago Q-sort*, który był później często stosowany (m.in. przez C. Rogersa) do oceny efektyw-

ności psychoterapii. Za punkt wyjścia wspomniani badacze przyjęli protokoły z posiedzeń terapeutycznych, z których wylosowali 100 stwierdzeń. Po niezbędnej językowej „obróbce” twierdzeń przekazano je do oceny dwóm psychologom klinicznym. Oceniali oni każde stwierdzenie ze względu na to, czy jest ono wskaźnikiem dobrego przystosowania, czy też złego, jeżeli pacjent uzna, że w wysokim stopniu opisuje ono jego zachowanie. W wyniku przeprowadzenia tej oceny zbiór 100 stwierdzeń rozbił się na dwa podzbiory:

- 1) podzbiór stwierdzeń będących dobrymi wskaźnikami przystosowania pacjenta (74 stwierdzenia),
- 2) podzbiór stwierdzeń nie mających nic wspólnego z przystosowaniem (26 stwierdzeń).

Z kolei czterech nowych sędziów kompetentnych rozdzieliło podzbiór (1) na dwie grupy po 37 stwierdzeń każda:

- a) na grupę stwierdzeń „pozytywnych”, wskazujących na dobre przystosowanie, jeżeli pacjent uznał je za opisujące jego zachowanie w wysokim stopniu,
- b) na grupę stwierdzeń „negatywnych”, wskazujących na brak przystosowania, jeżeli pacjent uznał je za opisujące jego zachowania w wysokim stopniu.

Kontinuum, na którym badani mieli oceniać poszczególne stwierdzenia jako podobne lub niepodobne do nich (*like me — unlike me*), było 9-punktowe:

niepodobne do mnie 0 1 2 3 4 5 6 7 8 podobne do mnie

Pacjenta uznawano za w pełni przystosowanego, jeżeli stwierdzenia z grupy (a) umieszczał na prawym krańcu kontinuum, tzn. przydzielał im punkty 5, 6, 7, 8, a stwierdzenia z grupy (b) umieszczał na lewym krańcu, tzn. przydzielał im punkty 0, 1, 2, 3. Punkt środkowy kontinuum „4” oznaczał, że dane stwierdzenie nie dotyczy badanej osoby. Jeżeli pacjent sortował stwierdzenia wg odwrotnej reguły, uznawał go za w pełni nieprzystosowanego. Poniżej podajemy przykładowe stwierdzenia zaczerpnięte z *Chicago Q-sort* (por. Byrne, op. cit., rozdz. 12).

Stwierdzenia świadczące o dobrym przystosowaniu:

- Sam jestem sprawcą swoich kłopotów.
- Nie mam kłopotów z opanowaniem się.
- Swoje uczucia wyrażam bez skrępowania, swobodnie.
- Jestem atrakcyjny seksualnie.
- Różnię się od innych.
- Czuję się dojrzały emocjonalnie.
- Rozumiem siebie, swoje postępowanie.
- Jestem odprężony i właściwie nic mnie nie martwi.

Stwierdzenia świadczące o braku przystosowania:

- Mam poczucie beznadziejności.
- Często czuję się upokorzony.
- Trudno mi kontrolować własną agresywność.
- Jestem zdeorientowany.
- Jestem nierzetelny.
- Gardzę sobą.
- Przybieram fałszywą maskę.
- Czuję się apatycznie.

Oczywiście, badani otrzymują stwierdzenia w porządku losowym, a nie najpierw grupę stwierdzeń pozytywnych, a potem negatywnych.

W badaniach A. Brzezińskiej i J. Brzezińskiego (1975b) posłużono się Q-sortem (Skala Samoakceptacji SQ) składającym się z 80 przymiotników, przy czym wyróżniono cztery 20-przymiotnikowe grupy (szerzej o samej procedurze dalej). Każda grupa przymiotników dotyczyła jednego wymiaru obrazu własnej osoby, a mianowicie:

- I właściwości funkcjonowania poznawczego,
- E właściwości funkcjonowania w sferze emocji i motywacji,
- F właściwości funkcjonowania fizycznego i wyglądu,
- S właściwości funkcjonowania społecznego.

Badany otrzymywał zbiór 80 kartek — w porządku losowym — z przymiotnikami, który musiał uporządkować na 9-punktowym kontinuum:

nie opisuje mnie 0 1 2 3 4 5 6 7 8 opisuje mnie

Badany dokonywał dwukrotnego sortowania: raz wg instrukcji *jaki jestem*, a drugi raz wg instrukcji *jaki chciałbym być*, za każdym razem postępował tak samo, tj. musiał przyporządkować podaną liczbę kartek do poszczególnych kategorii:

nie opisuje mnie 0 1 2 3 4 5 6 7 8 opisuje mnie
Liczba kartek 4 6 9 13 16 13 9 6 4 Liczba kartek

I z pierwszego, i z drugiego badania uzyskano informacje o względnym natężeniu poszczególnych wymiarów obrazu własnej osoby u danego badanego. Z porównania wyników pochodzących z obu badań uzyskano informację o poziomie zgodności (czy rozbieżności) ja-idealnego i ja-realnego, czyli o poziomie samoakceptacji. Oto 14 przymiotników zaczerpniętych ze Skali Samoakceptacji SQ wraz z ich przyporządkowaniem do danego wymiaru:

1	Aktywny	EM	46	Ponury	EM
2	Bezczelny	S	54	Słabowity	F
11	Dokuczliwy	S	61	Tępy	I
12	Drażliwy	EM	63	Troskliwy	S
15	Elegancki	F	73	Wyrozumiały	S
27	Męski	F	78	Zrównoważony	EM
32	Omylny	I	79	Zwinny	F

Na tym można już zakończyć prezentację różnych form, jakie może przybierać Q-sort, jeżeli chodzi o charakter zbioru pozycji oraz o charakter samych pozycji. Czytelnik sam powinien zdecydować, w jakiej postaci językowej mają być sformułowane pozycje. Powinien wziąć przy tym pod uwagę takie dane, jak płeć badanej populacji, wykształcenie, wiek, kryterium sortowania. Przy doborze pozycji do Q-sortu należy kierować się takimi samymi względami jak w przypadku wybierania pozycji do kwestionariusza osobowości czy skali postaw.

1.6.3. Charakterystyka rozkładu pozycji

Charakterystyczną cechą techniki Q jest to, że wymaga ona od badacza określenia liczby kartek (z wypisanymi na nich stwierdzeniami lub pojedynczymi słowami), jaka może być przez osobę badaną zaliczona do danej kategorii. Najczęściej tak dobiera się liczbę kartek dla poszczególnych punktów kontinuum, aby otrzymać rozkład zbliżony do normalnego, a więc symetryczny i jednomodalny. Jeżeli chodzi o wielkość zbioru kartek i liczbę punktów kontinuum, to w zasadzie badacz ma nieograniczone możliwości, chociaż ze względów praktycznych nie zaleca się, by liczba kartek w Q-sorcie przekraczała 140 ze względu na zmęczenie bądź znudzenie badanych procedurą sortowania. Natomiast względy natury statystycznej (stabilność, rzetelność) przemawiają za tym, by liczba ta nie była mniejsza niż 60. Optymalna wielkość zbioru kartek zawiera się w granicach od 60 do 120 (Kerlinger 1986, s. 508) — chociaż sam Ker-

linger podaje, że uzyskał zadowalające rezultaty przy jedynie 40 pozycjach wybranych ze znacznie większego zbioru. To, że badacz „wymusza” normalny rozkład pozycji na kontinuum, ma tę zaletę, że pozwala na stosowanie odpowiednich wskaźników w analizie rezultatów badania. Przy porównywaniu dwóch zbiorów uporządkowań pochodzących z dwukrotnie przeprowadzonych badań można się posłużyć np. współczynnikiem korelacji r -Pearsona. Przyjrzyjmy się teraz różnym rozwiązaniom, jeżeli chodzi o liczbę punktów kontinuum i liczbę sortowanych pozycji.

Dymond (por. Byrne 1966, s. 443), stosując *Chicago Q-sort* Butlera i Haigha, posłużył się 9-punktowym kontinuum (por. tabela 6a i rysunek 7a). Także Akeret (1965, s. 492) zastosował 100-kartkowy Q-sort przy 9-punktowym kontinuum i rozkładzie liczebności pozycji jak u Dymonda. Z kolei nasza Skala Samoakceptacji SQ (1975a) jest oparta na rozkładzie podanym w tabeli 6b i na rysunku 7b. W tabeli 6c i na rysunku 7c-1 i 7c-2 przedstawiono — za Stephensonem (1953) — kilka rozkładów liczebności pozycji dla różnej ich liczby i dla kontinuum o różnej liczbie punktów. Także Kerlinger (1964, s. 583) przytacza kilka typowych rozkładów liczebności kartek. Przedstawiają je tabela 6d oraz rysunek 7d-1 i 7d-2.

TABELA 6a

Rozkład pozycji w *Chicago Q-sort* Butlera i Haigha

$n = 100$	Najmniej podobne do mnie								Najbardziej podobne do mnie	
	0	1	2	3	4	5	6	7	8	
Kontinuum										
Liczba pozycji	1	4	11	21	26	21	11	4	1	

Źródło: opracowanie własne na podstawie Butler i Haigh 1966.

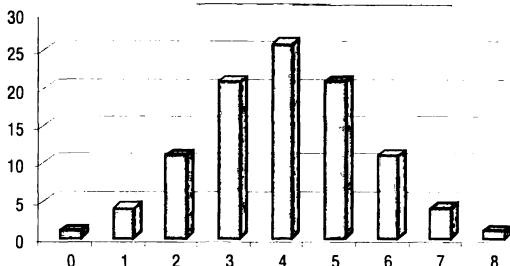
TABELA 6b

Rozkład pozycji w Skali Samoakceptacji SQ Brzezińskiej i Brzezińskiego

$n = 80$	W najmniejszym stopniu mnie opisuje								W największym stopniu mnie opisuje	
	0	1	2	3	4	5	6	7	8	
Kontinuum										
Liczba pozycji	4	6	9	13	16	13	9	6	4	

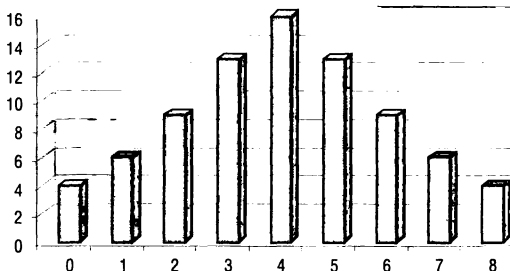
Źródło: opracowanie na podstawie Brzezińska, Brzeziński 1975a.

Rysunek 7a. Rozkład pozycji w *Chicago Q-sort* Butlera i Haigha



Źródło: opracowanie na podstawie Brzezińska, Brzeziński 1975a.

Rysunek 7b. Rozkład pozycji w *Skali Samoakceptacji SQ* Brzezińskiej i Brzezińskiego



Źródło: opracowanie na podstawie Brzezińska, Brzeziński 1975a.

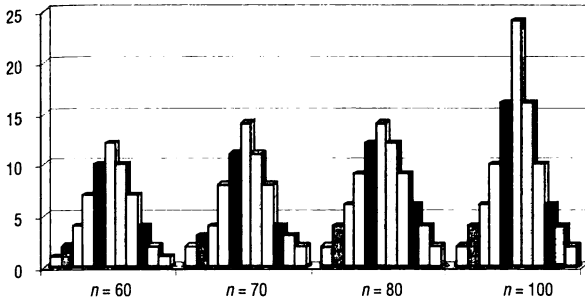
TABELA 6c

Przykładowe rozkłady pozycji w Q-sorcie wg Stephensona

Kontinuum Liczba pozycji	Niepodobne								Podobne		
	0	1	2	3	4	5	6	7	8	9	10
<i>n</i> = 48	2	4	5	8	10	8	5	4	2		
<i>n</i> = 100	2	4	6	10	16	24	16	10	6	4	2
<i>n</i> = 60	1	2	4	7	10	12	10	7	4	2	1
<i>n</i> = 160	8	12	20	24	32	24	20	12	8		
<i>n</i> = 80	2	4	6	9	12	14	12	9	6	4	2
<i>n</i> = 70	2	3	4	8	11	14	11	8	4	3	2

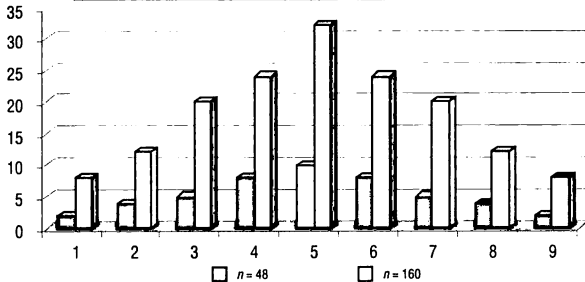
Źródło: Stephenson 1953, s. 17, 20, 9, 115, 72 i 277.

Rysunek 7c-1. Przykładowe rozkłady pozycji w Q-sorcie wg Stephensona



Źródło: opracowanie własne na podstawie Stephenson 1953.

Rysunek 7c-2. Przykładowe rozkłady pozycji w Q-sorcie wg Stephensona



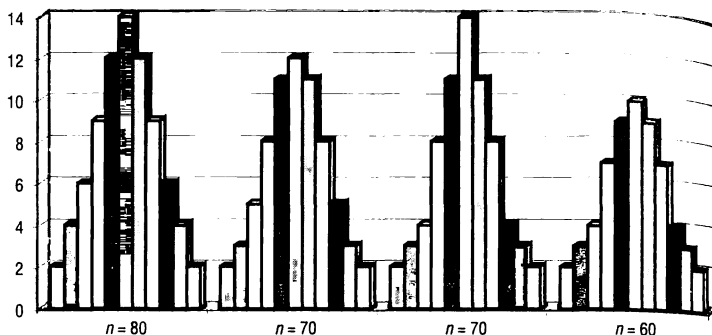
Źródło: opracowanie własne na podstawie Stephenson 1953.

TABELA 6d
Przykładowe rozkłady pozycji w Q-sorcie wg Kerlingera

Kontinuum Liczba pozycji	Najmniej aprobowane					Najbardziej aprobowane					
	0	1	2	3	4	5	6	7	8	9	10
n = 80	2	4	6	9	12	14	12	9	6	4	2
n = 80	4	6	9	13	16	13	9	6	4		
n = 80	4	6	10	12	16	12	10	6	4		
n = 70	2	3	5	8	11	12	11	8	5	3	2
n = 70	2	3	4	8	11	14	11	8	4	3	2
n = 60	2	3	6	11	16	11	6	3	2		
n = 60	2	3	4	7	9	10	9	7	4	3	2

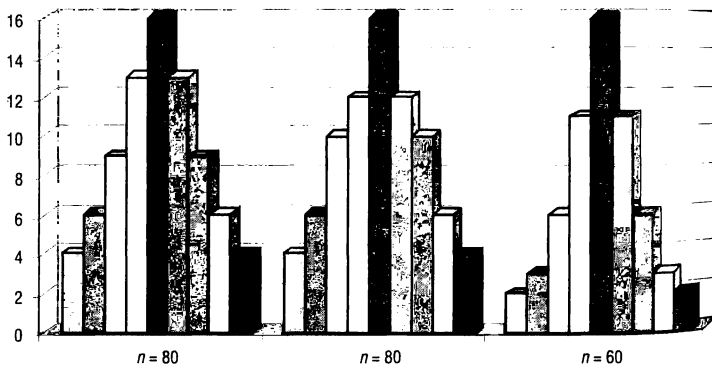
Źródło: na podstawie Kerlinger 1986, s. 509, Table 32.3.

Rysunek 7d-1. Przykładowe rozkłady pozycji w Q-sorcie wg Kerlingera



Źródło: opracowanie własne na podstawie Kerlingera 1964.

Rysunek 7d-2. Przykładowe rozkłady pozycji w Q-sorcie wg Kerlingera



Źródło: opracowanie własne na podstawie Kerlingera 1964.

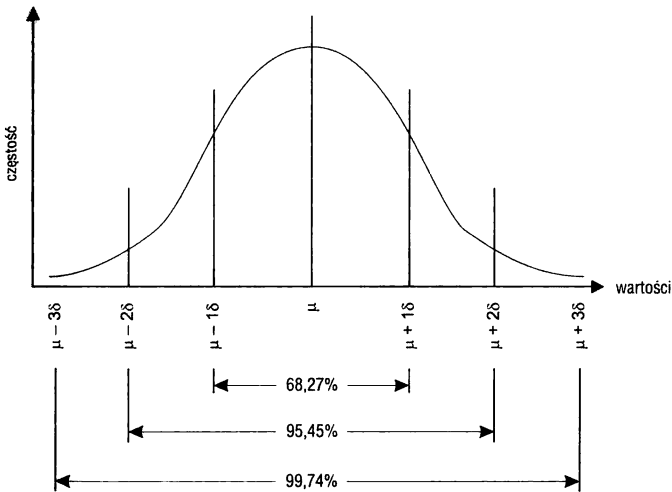
BLOK ROZSZERZAJĄCY 2

ANALIZA ROZKŁADU POZYCJI Q-SORT

Jeżeli Czytelnikowi nie będzie odpowiadał żaden z przedstawionych przykładów rozkładów kartek do sortowania, może sporządzić samodzielnie rozkład liczebności kartek na dowolnym kontinuum. Pomocne tu będą ta-

blisce dystrybuanty rozkładu normalnego (inaczej: tablice powierzchni pod krzywą normalną). Dokładne tablice podaje np. J. Greń (1975, s. 307, tab. 4). Przy ustalaniu rozstępu między sąsiednimi punktami kontinuum należy kierować się wielkością odchylenia standardowego. Na mocy reguły „trzech sigm” prawie 100% powierzchni pod krzywą normalną leży w granicach $\langle -3$ odchylenia standardowe od średniej, $+3$ odchylenia standardowe od średniej \rangle — por. rysunek 8a.

Rysunek 8a. Powierzchnia pod krzywą normalną



Źródło: opracowanie własne.

Założmy, że badacz chce skonstruować Q-sort o 100 pozycjach i kontinuum 11-punktowym. Niech odstęp między punktami kontinuum wynosi $0,4 \delta$. Wobec tego otrzymujemy podział powierzchni pod krzywą normalną na 11 części (por. rys. 8b). Odejmując kolejno mniejsze powierzchnie od większych, otrzymujemy wielkość powierzchni leżącej między dwoma punktami założonego kontinuum:

pkt 10 obejmuje całą powierzchnię leżącą na prawo od $1,8 \delta$, czyli $3,59\% \approx 4\%$

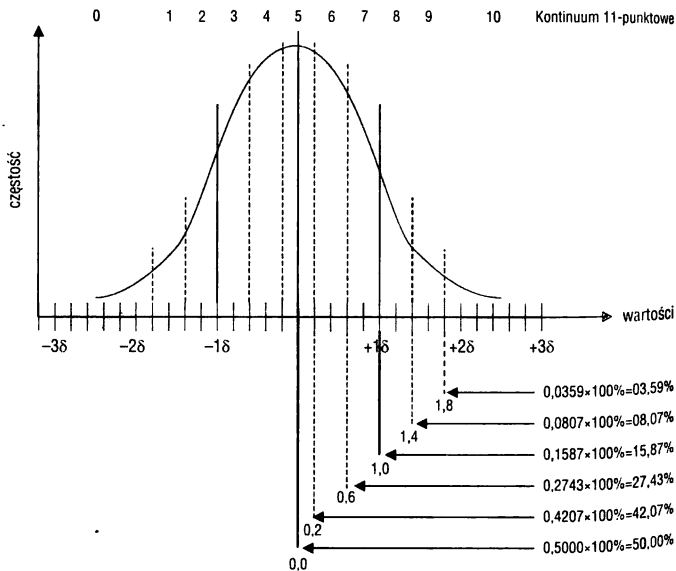
pkt 9 obejmuje całą powierzchnię leżącą między $1,4 \delta$ i $1,8 \delta$, czyli $8,07 - 3,59 = 4,46\% \approx 5\%$

pkt 8 obejmuje całą powierzchnię leżącą między $1,0 \delta$ i $1,4 \delta$, czyli $15,87 - 8,07 = 7,80\% \approx 8\%$

pkt 7 obejmuje całą powierzchnię leżącą między $0,6\delta$ i $1,0\delta$, czyli $27,43 - 15,87 = 11,56\% \approx 11\%$
 pkt 6 obejmuje całą powierzchnię leżącą między $0,2\delta$ i $0,6\delta$, czyli $42,07 - 27,43 = 14,64\% \approx 14\%$
 pkt 5 obejmuje całą powierzchnię leżącą między $-0,2\delta$ i 0 oraz między 0 i $0,2\delta$, czyli $50,00 - 42,07 = 7,93\%$; $7,93 \times 2 = 15,86\% \approx 16\%$

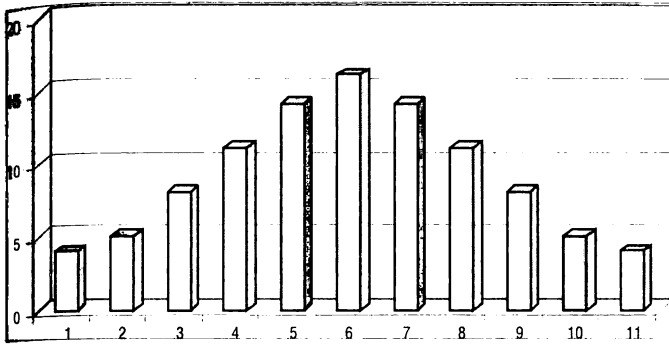
Ponieważ krzywa normalna jest symetryczna (tzn. po obu stronach średniej leży 50% powierzchni), więc nie ma potrzeby obliczania powierzchni odpowiadającej pozostałym punktom kontinuum. Należy tylko przypisać powierzchnie obliczone dla punktów 6, 7, 8, 9 i 10 — odpowiednio — punktom 0, 1, 2, 3 i 4. Zaokrąglone wyniki procentowe wskazują, ile kartek badany może przypisać danemu punktowi kontinuum. Otrzymany rozkład 100 kartek na 11-punktowym kontinuum prezentuje rysunek 8c.

Rysunek 8b. Rozkład powierzchni pod krzywą normalną na kontinuum 11-punktowym (odległość między punktami kontinuum wynosi $0,4\delta$)



Źródło: opracowanie własne.

Rysunek 8c. Rozkład kartek na kontynuum 11-punktowym



Źródło: opracowanie własne.

1.6.4. Przebieg badania

Za pomocą techniki Q można przeprowadzić badania jednorazowe i wielokrotne (najczęściej dwukrotne). Przy jednokrotnym badaniu interesuje badacza ustosunkowanie się badanych osób do pozycji, które są wskaźnikami określonej zmiennej (zmiennych). Można np. posłużyć się zbiorem pozycji, których część wyraża pozytywną motywację do pracy, a część negatywną. Zakładamy, że badani, których cechuje pozytywna motywacja, będą przydzielali pozycje pozytywne do najwyższych kategorii, a badani, których cechuje motywacja negatywna, będą przydzielali pozycje negatywne do najwyższych kategorii. Analizując zatem otrzymany rozkład kartek, możemy określić charakter motywacji do pracy każdej badanej osoby.

Przy dwukrotnym badaniu tej samej osoby tym samym zestawem pozycji, ale przy różnych instrukcjach możemy określić stopień podobieństwa dwóch zbiorów ocen. Ponieważ „wymuszaliśmy” każdorazowo rozkład normalny liczebności kartek, więc jako wskaźnika podobieństwa możemy użyć współczynnika korelacji r -Pearsona. Przy badaniu stopnia zgodności tzw. realnego obrazu samego siebie (ja-realne, ang. *real self*) z tzw. idealnym obrazem samego siebie (ja-idealne, ang. *ideal*

sealf) także korzysta się ze współczynnika korelacji. Im wyższą wartość przyjmuje ten współczynnik, tym większy jest stopień zgodności obu elementów obrazu samego siebie, czy inaczej: wyższy jest poziom samoakceptacji i odwrotnie. Jak pokazały liczne badania (por. Reykowski 1970), u osób z zaburzeniami emocjonalnymi współczynnik korelacji jest bardzo niski, a nawet zerowy. Osoby poprawnie funkcjonujące społecznie otrzymują współczynniki rzędu 0,50-0,60. Natomiast u osób z zaburzeniami psychicznymi (psychozy, psychopatia) zauważono bardzo wysoką zbieżność ja-idealnego i ja-realnego.

Przytaczamy teraz, dla przykładu, dwuczęściową instrukcję dla osoby badanej, jaka dołączona jest do Skali Samoakceptacji SQ (wersja dla mężczyzn), przy czym część A dotyczy badania ja-realnego, a część B — badania ja-idealnego.

Część A

Ma Pan przed sobą zestaw 80 kartek. Na każdej z nich wypisane jest jedno słowo. Słowa te opisują w różnym stopniu Pana cechy. Zadaniem Pana jest rozdzielenie tych kartek do 9 przegródek. Cyfry od 0 do 8 oznaczają, w jakim stopniu posiada Pan daną cechę:

„0” oznacza, że danej cechy w ogóle Pan nie posiada,

„8” oznacza, że posiada Pan tę cechę w maksymalnym stopniu.

Pozostałe cyfry oznaczają stopnie pośrednie. A zatem kartki zawierające słowa najmniej odnoszące się do Pana należy umieścić w przegródkach o najniższych cyfrach. Kartki ze słowami, które w dużym stopniu odnoszą się do Pana, należy umieścić w przegródkach o wyższych cyfrach.

Do każdej przegródki należy włożyć określoną liczbę kartek. Liczby te wypisane są u dołu każdej przegródki.

Słowa na kartkach można interpretować w dowolny sposób, a podczas rozdzielania kartek można je przekładać z jednej przegródki do drugiej dowolną liczbę razy.

Po zakończeniu rozdzielania kartek należy sprawdzić, czy w każdej przegródce jest odpowiednia liczba kartek — zgodna z liczbą wypisaną pod przegródką.

Proszę dokonać rozdzielenia kartek do 9 przegródek, aby odpowiedzieć na pytanie: **jaki jestem?**

Część B

Proszę dokonać rozdzielenia kartek w podobny sposób jak poprzednio. Teraz jednak proszę odpowiedzieć na pytanie: **jaki chciałbym być?** Do przegródek o niskich cyfrach włożyć Pan kartki z tymi cechami, których nie

chciałby Pan posiadać. Natomiast do przegródek o wysokich cyfrach — kartki z cechami, które chciałby Pan posiadać. Proszę rozdzielić kartki.

Niezależnie od tego, w jakim celu będzie się stosowało technikę Q, w instrukcji dla osoby badanej powinna znaleźć się informacja o tym, że zaklasyfikowanie przez badanego danej pozycji do określonej kategorii może być tyle razy zmieniane, aż uzna on, iż przeprowadzone sortowanie jest zgodne z jego przekonaniem na wskazany instrukcją temat. Celem tej instrukcji jest zminimalizowanie stopnia wpływu przydzielenia danej pozycji do określonej kategorii na przydzielanie innych pozycji do innych kategorii. Mówiąc inaczej, chodzi o zapewnienie jak największej niezależności kolejnych sortowań. Niemniej badanym nie wolno przydzielić ani mniejszej, ani większej liczby kartek do poszczególnych kategorii, niż to wynika z podanego im w instrukcji rozkładu liczebności.

Dla ułatwienia sortowania kartek wskazane jest posługiwanie się w badaniach indywidualnych specjalnie sporządzonym sorterem. Funkcję sortera może pełnić pudełko z przegródkami, po jednej przegródce na każdą kategorię. Na widocznej dla badanego ściance przegródki piszemy odpowiadającą jej liczebność kartek oraz dany punkt kontinuum (por. badania B. Ziółkowskiej — Przykład 11 i rysunek 7).

PRZYKŁAD 11

ZASTOSOWANIE TECHNIKI Q-SORT W DIAGNOZIE ZACHOWAŃ ANOREKTYCZNYCH

Beata Ziółkowska (2001, s. 153-154; 2003) w swoich badaniach nad uwarunkowaniami zachowań anorektycznych u młodych kobiet postawiła m.in. pytanie „Jakiego rodzaju zachowania uznać można za diagnostyczne dla rozpoznania *Syndromu gotowości anorektycznej* u kobiet w okresie przejścia od adolescencji do wczesnej dorosłości?”. Na podstawie analizy literatury przedmiotu za diagnostyczne dla ekspresji *Syndromu gotowości anorektycznej* uznała: odmawianie przyjmowania pokarmów; liczenie kalorii; zwiększenie aktywności ruchowej; przesadną dbałość o wygląd; wyrzucanie, gromadzenie, chowanie jedzenia; wydzielanie sobie racji pokarmowych i zachęcanie innych do jedzenia; częste kontrolowanie wagi i wymiarów ciała, czemu towarzyszy strach, niepokój, napięcie; tendencję do przecenia-

nia rozmiarów swojego ciała oraz wagi; przyjmowanie środków przeczyszczających i moczopędnych; sporadyczne odczuwanie głodu i ssania w żołądku.

Dla zweryfikowania hipotezy zakładającej istnienie określonego dla *Syndromu gotowości anorektycznej* zespołu diagnostycznych zachowań anorektycznych wykorzystano technikę Q-sort. Przed badaną osobą rozkładano sześć pudełek (por. rysunek 9) opatrzonych odpowiednią cyfrą i nazwami kategorii z towarzyszącą instrukcją słowną, że lewy kraniec (lewe pudełko oznaczone jako „0”) oznacza całkowity brak zgodności z określonym twierdzeniem, prawe zaś (oznaczone jako „5”) opisuje pełną z nim zgodność. Z kolei wartości „1” oznacza *nie zgadzam się*, „2” — *raczej się nie zgadzam*, „3” — *raczej się zgadzam*, „4” — *zgadzam się*. Następnie osoba badana otrzymała zestaw kartek z wypisanymi na nich twierdzeniami. Jej zadaniem było odczytywanie kolejnych zdań i zdecydowanie, poprzez umieszczenie kartki w odpowiednim pudełku, na ile opisuje ono jej sposób funkcjonowania. Zdania — w liczbie 35 — badana odczytywać mogła w dowolnej kolejności. Twierdzenia obejmują sześć kategorii definiujących zachowania anorektyczne:

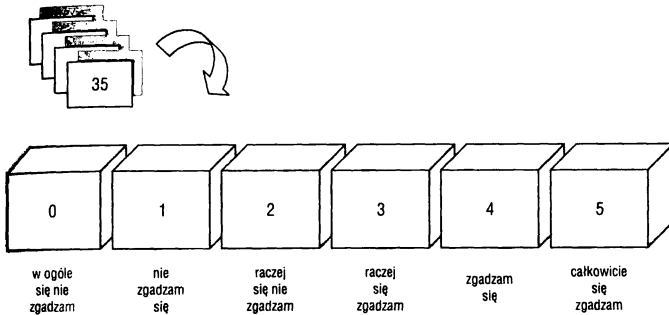
- | | |
|-------------------------------------|-------------------------------------|
| 1) emocje związane z jedzeniem: | stwierdzenia nr 2, 13, 19, 21, 23 |
| 2) ograniczanie jedzenia: | stwierdzenia nr 4, 6, 10, 22, 26 |
| 3) ocena i kontrola atrakcyjności: | stwierdzenia nr 5, 7, 12, 15, 31 |
| 4) wpływ reklamy i kultury masowej: | stwierdzenia nr 1, 16, 18, 28, 32 |
| 5) formy „zbijania wagi”: | stwierdzenia nr 3, 8, 9, 11, 24, 25 |
| 6) jedzenie a stosunki społeczne: | stwierdzenia nr 14, 17, 20, 33, 34 |
| 7) perfekcjonizm i rywalizacja: | stwierdzenia nr 27, 29, 30, 35 |

Zadanie uznano za zakończone w chwili, gdy wszystkie twierdzenia znalazły się w kartonikach. Za odpowiedź (1) przyznawano 5 punktów, za (2) — 4 punkty, za (3) — 3 punkty, za (4) — 2 punkty, za (5) — 0 punktów. Zasady punktacji polegały na tym, że im większa ostateczna suma punktów uzyskana przez badaną osobę, tym wyraźniejszy wskaźnik manifestowania zachowań anorektycznych.

Badanie prowadziły dwie odpowiednio do tego celu przeszkolone osoby, młode kobiety z wykształceniem psychologicznym. Ten etap procedury miał postać jednorazowej 2-godzinnej, indywidualnej sesji. Zadaniem osoby badającej było krótkie przedstawienie idei spotkania i nawiązanie kontaktu. W większości wypadków sesję rozpoczynało badanie techniką Q-sort. Z jednej strony technika ta angażowała osobę badaną w konkretną, samodzielną pracę w obecności badacza, z drugiej zaś pozwalała na oswajanie się z sytuacją i rozładowanie napięcia. Następnie inicjowano rozmowę na temat przechowywanego w pamięci obrazu rodziny. W trakcie rozmowy rejestrowano wypowiedzi osoby badanej, ale także systematycznie anga-

zowano ją w uzupełnianie zdań niedokończonych oraz zaznaczanie oceny różnych swych właściwości na przygotowanych numerycznych i graficznych skalach szacunkowych.

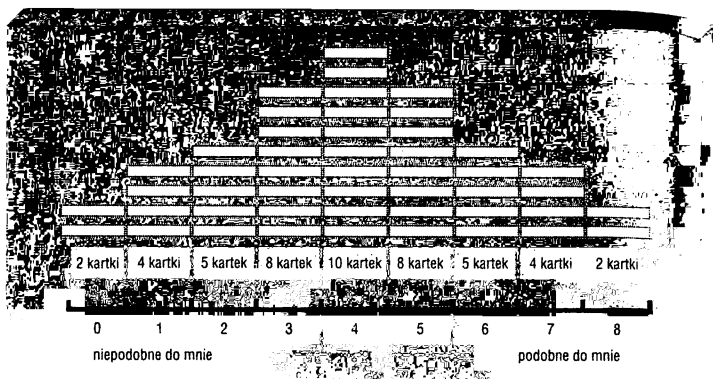
Rysunek 9. Sposób sortowania kartek w badaniach B. Ziółkowskiej



Źródło: opracowanie własne na podstawie Ziółkowska 2001.

Zamiast pudełka można użyć planszy z nacięciami. Nad każdą kategorią umieszczamy tyle nacięć, ile kartek powinno się w tej kategorii znaleźć. Jeżeli chcemy prowadzić badania dwukrotnie, dobrze jest posłużyć się dwiema planszami i dwoma analogicznymi zestawami pozycji. Skróci nam to znacznie czas badania, gdyż nie będziemy musieli w obecności badanego wyjmować z sortera kartek, nanosić punktacji na kartę wyników oraz ponownie układać kartek w pierwotnym, standardowym porządku. Jeżeli mamy dwa identyczne zestawy kartek, to po rozsortowaniu ich zgodnie z pierwszą instrukcją prosimy badanego o rozsortowanie drugiego zestawu przy zmienionej instrukcji i przy wykorzystaniu drugiego sortera. Sorter w postaci planszy można wykonać wg wzoru podanego na rysunku 10. Poszczególne pozycje wypisujemy na odrębnych kartkach z zaznaczonym numerem ułatwiającym identyfikację kartki na *Karcie wyników*.

Rysunek 10. Sorter w postaci planszy z nacięciami na kartki



Źródło: opracowanie własne.

1.6.5. Zasady obliczania wyników

Należy wyraźnie podkreślić, iż w technice Q nie można obliczać jednego ogólnego wyniku, jak to ma miejsce w kwestionariuszach osobowości czy skalach postaw typu Likerta. Jeżeli np. przebadamy 20 osób danym Q-sortem i obliczymy ogólne wyniki przez zsumowanie punktów przypisanych każdej pozycji, to otrzymamy 20 takich samych „wyników ogólnych”. Dla Q-sortu przedstawionego na rysunku 10 (rysunek sortera z podaną liczbą kartek dla każdej kategorii) suma ta wynosi:

$$[(2 \times 0) + (4 \times 1) + (5 \times 2) + (8 \times 3) + (10 \times 4) + (8 \times 5) + (5 \times 6) + (4 \times 7) + (2 \times 8)] = 192 \text{ pkt.}$$

Można natomiast obliczać wskaźniki cząstkowe, jeżeli postępuje się ustrukturowaną techniką Q.

Powróćmy raz jeszcze do przykładu Skali Samoakceptacji SQ. Składa się ona z 80 pozycji podzielonych na 4 grupy po 20 przymiotników. Grupy te odpowiadają czterem wymiarom obrazu własnej osoby. Symbolicznie oznaczono je jako: I, EM, F i S. Na każdej kartce zawierającej daną pozycję jest umieszczony numer określający jej przynależność do jednej

z czterech grup. Po wpisaniu do Karty wyników (por. rys. 11a) obok numeru każdej pozycji jej punktacji (tzn. punktu kontinuum, któremu osoba badana przypisała tę kartkę), sumujemy oddzielnie punkty dla pozycji należących do grupy I, EM, F i S. W ten sposób otrzymujemy cztery sumy — wyniki cząstkowe (przy czym ΣA oznacza sumę dla ja-realnego, a ΣB sumę dla ja-idealnego). Sumy te umieszczamy w tabeli (por. tab. 7) i wykreślamy profil (por. rys. 11b). Profil ten interpretujemy wg takich samych reguł jak profil uzyskany z badania skalą z wymuszonym wyborem Cosgrove'a. Jeżeli badanie przeprowadzamy dwukrotnie — otrzymujemy dwa profile. Karta wyników jest tak skonstruowana, że jeden blok dotyczy jednego wymiaru. Każdy blok składa się z czterech wierszy oznaczonych kolejno symbolami: I (EM, F, S), A, B i D^2 , gdzie: A oznacza punkty przypisane przymiotnikom w badaniu I (ja-realne), B — punkty przypisane przymiotnikom w badaniu II (ja-idealne), D^2 to kwadrat różnicy ($A - B$).

TABELA 7

Tabela sum dla ja-realnego i ja-idealnego (wyniki surowe na rysunku 11a)

	Sfera I	Sfera EM	Sfera F	Sfera S
Ja-realne	68	78	81	81
Ja-idealne	75	87	79	79

Jako miary podobieństwa dwóch zbiorów wyników pochodzących z dwukrotnie przeprowadzonych badań tej samej osoby Skalą Samoakceptacji SQ przy dwóch różnych instrukcjach (*jaki jestem?* i *jaki chciałbym być?*) użyto współczynnika korelacji r -Pearsona. Można go obliczyć wg bardzo prostego wzoru, który podaje Kerlinger (1964, s. 585):

$$r = \frac{\Sigma ab}{\sqrt{\Sigma a^2 \Sigma b^2}},$$

gdzie: $a = A - \bar{A}$; $b = B - \bar{B}$; \bar{A} i \bar{B} to średnie wartości A i B wg punktacji Q-sortu; oznaczenia literowe (A, B) dostosowano od oznaczeń użytych na Karcie wyników. Ponieważ $\Sigma A^2 = \Sigma B^2$, gdyż są to te same zbiory wyników, wystarczy obliczyć jedynie iloczyn $A \times B$ dla każdej pozycji oraz wyrażenie A^2 dla każdej pozycji.

Rysunek 11a. Karta wyników dla Skali Samoakceptacji SQ

I	4	10	18	20	25	26	32	38	44	45	51	52	56	61	64	66	67	70	72	76	ΣA	ΣB	ΣD ²	
A	4	5	1	3	1	1	2	1	6	4	0	5	4	0	4	5	5	5	6	6	68			
B	2	3	1	3	2	0	3	0	6	6	1	5	4	0	7	7	4	6	7	8		75		
D ²	4	4	0	0	1	1	1	1	0	4	1	0	0	0	9	4	1	1	1	4			37	
EM	1	8	9	12	17	22	23	24	36	39	40	43	46	48	53	55	59	68	75	78	ΣA	ΣB	ΣD ²	
A	4	2	8	3	3	4	5	6	3	4	7	5	2	2	6	7	3	8	4	4	78			
B	6	4	4	3	4	4	4	2	8	5	4	4	2	4	5	5	6	5	3	5		87		
D ²	4	4	16	0	1	0	1	16	25	1	9	1	0	4	1	4	9	9	1	1			107	
F	13	15	19	27	28	29	31	33	34	37	42	49	50	54	58	60	69	74	77	79	ΣA	ΣB	ΣD ²	
A	8	6	3	4	3	7	1	4	0	5	2	4	6	3	4	5	2	5	7	2	81			
B	3	5	2	7	6	7	1	3	0	4	1	4	4	2	3	4	6	4	6	7		79		
D ²	25	1	1	9	9	0	0	1	0	1	1	0	4	1	1	1	16	1	1	25			98	
S	2	3	5	6	7	11	14	16	21	30	35	41	47	57	62	63	65	71	73	80	ΣA	ΣB	ΣD ²	
A	5	5	6	3	0	2	3	1	7	3	4	2	6	7	8	3	3	4	4	5	81			
B	3	4	5	3	1	3	3	1	8	2	3	2	8	5	6	5	5	2	5	5		79		
D ²	4	1	1	0	1	1	0	0	1	1	1	0	4	4	4	4	4	4	1	0			36	
																					Σ	320	320	278
Wskaźnik samoakceptacji: $1 - \Sigma D^2 / 672 = 1 - 278 / 672 = 1 - 0,41 = 0,59$																								

Źródło: opracowanie własne.

Cohen (1957; Paszkiewicz 1974, s. 110-111) podaje daleko idącą modyfikację wzoru na r , która wymaga jeszcze mniejszej liczby obliczeń. Ostateczna postać wzoru jest więc następująca:

$$r = 1 - (\Sigma D^2 / 2n\sigma^2),$$

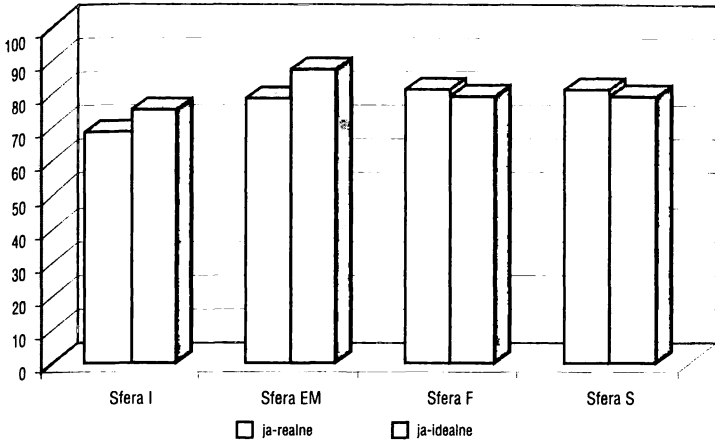
gdzie:

$$D = B - A,$$

σ^2 — typowa dla danego rozkładu liczebności kartek wartość wariancji, wystarczy raz ją obliczyć i podać w charakterystyce Q-sortu (dla Skali Samoakceptacji SQ wartość ta wynosi 4,20),

n — liczba kartek w Q-sortcie.

Rysunek 11b. Profil dla ja-realnego i ja-idealnego (wyniki w tabeli 7 i na rysunku 11a)



Źródło: opracowanie własne.

A zatem dla Skali Samoakceptacji SQ mamy:

$$2 \times n \times \sigma^2 = 2 \times 80 \times 4,20 = 672.$$

Wzór na σ^2 jest następujący:

$$\sigma^2 = \left[n \sum f \times X^2 - (\sum f \times X)^2 \right] : n(n-1),$$

gdzie:

X — punkty kontinuum,

f — liczba kartek przypadających na daną kategorię kontinuum.

Do obliczenia σ^2 można posłużyć się tabelą pomocniczą (por. tab. 8).

Skala Samoakceptacji SQ pozwala zatem na uzyskanie dwójakiego rodzaju informacji:

- pokazuje względną ważność poszczególnych wymiarów obrazu siebie samego dla danej osoby, dla jej ja-realnego i ja-idealnego — analiza profilu,
- pokazuje stopień zgodności (samoakceptacji) dwu elementów obrazu siebie, tj. ja-realnego i ja-idealnego — oblicza się

współczynnik korelacji, którego wartość wskazuje stopień zgodności obu elementów obrazu samego siebie.

TABELA 8

Tabela pomocnicza do obliczania wariancji rozkładu pozycji w Q-sortcie (na przykładzie Skali Samoakceptacji SQ)

X	f	$f \times X$	X^2	$f \times X$
0	4	0	0	0
1	6	6	1	6
2	9	18	4	36
3	13	39	9	117
4	16	64	16	256
5	13	65	25	325
6	9	54	36	324
7	6	42	49	294
8	4	32	64	256
suma	80	320		1614

Gdzie: X — punkty kontinuum, f — liczba kartek; $(\sum f \times X)^2 = 320^2 = 102\,400$

1.6.6. Procedura budowy Skali Samoakceptacji SQ

Na przykładzie Skali Samoakceptacji SQ pokażemy, w jaki sposób badacz otrzymuje zbiór pozycji. Oczywiście, badacz może zastosować rozwiązanie, które będzie się różniło od tu pokazanego. Niemniej jednak ukaże mu ono, jakiego rodzaju problemy musi rozwiązać, przystępując do konstruowania Q-sortu.

Wyróżniono — jak wspomniano — cztery wymiary odpowiadające różnym obszarom obrazu własnej osoby. Wymiary te oznaczono jako I, EM, F i S, gdzie:

I — właściwości funkcjonowania poznawczego, czyli poziom ogólnej sprawności umysłowej, zdolności, umiejętności,

- EM — właściwości funkcjonowania w sferze emocji i motywacji, czyli cechy reagowania emocjonalnego, poziom aspiracji,
- F — właściwości funkcjonowania fizycznego, czyli wygląd, sprawność i wydolność fizyczna, sfera seksualna,
- S — właściwości funkcjonowania społecznego, czyli kontakty z innymi ludźmi, zależność od innych, stosunek do otoczenia społecznego.

Każdemu wymiarowi (I, EM, F, S) przypisano zestaw 20 przymiotników. Ogółem Skala Samoakceptacji SQ zawiera 80 pozycji. Zestawy opracowano w następujący sposób:

- 1) ze *Słownika języka polskiego* Witolda Doroszewskiego (1958-1969) wybrano przymiotniki określające właściwości człowieka ($n = 568$); po wstępnej analizie językowej, której celem było wyeliminowanie synonimów, określeń dwu- i wieloznacznych oraz zbyt trudnych (np. słowa pochodzenia obcego) dla badanych wyeliminowano 358 przymiotników;
- 2) zbiór przymiotników w kolejności alfabetycznej ($n = 210$) podano czterem sędziom kompetentnym (psychologowie kliniczni, dwóch mężczyzn i dwie kobiety) w celu rozdzielenia całego zbioru na kategorie odpowiadające wyróżnionym wymiarom I, EM, F, S; dodano kategorię piątą „nie wiem” — dla przymiotników, których nie można było przydzielić do żadnej z czterech kategorii podstawowych;
- 3) po dokonaniu szacowania przez czterech sędziów wybrano tylko te przymiotniki, co do których wszyscy sędziowie byli zgodni, że dotyczą one danej kategorii; zbiór ten liczył 109 przymiotników (a zatem wyeliminowano 91) podzielonych na cztery kategorie;
- 4) trzech innych sędziów kompetentnych (psychologowie kliniczni, dwóch mężczyzn, jedna kobieta) dokonywało analizy wewnątrz każdej kategorii w celu wyboru tylko 20 przymiotników, spełniających dodatkowo warunek: połowa powinna mieć charakter negatywny, a połowa — pozytywny;
- 5) w rezultacie otrzymano cztery równoliczne zbiory przymiotników ($n = 20$); przymiotniki te po odpowiednim zakodowaniu (przypisanie liter oznaczających przynależność do danej kategorii) ułożono w porządku losowym i przypisano im kolejne numery (korzystano z tablic liczb losowych).

Badanie za pomocą Skali Samoakceptacji SQ przebiega w kilku etapach:

- 1) osoba badana otrzymuje dwa analogiczne zestawy kartek, które są uporządkowane wg numerów; ponadto badany otrzymuje sorter, a na nim wypisane punkty (kategorie) kontinuum od 0 do 8 oraz liczba kartek, które można włączyć do danej kategorii;
- 2) etapy badania:
 - badający czyta instrukcję — część A (*jaki jestem?*),
 - osoba badana dokonuje sortowania,
 - badający czyta instrukcję — część B (*jaki chciałbym być?*),
 - badany dokonuje ponownego sortowania;
- 3) badający nanosi wyniki na Kartę wyników kolejno dla dwóch sortowań:
 - sortowanie I: *jaki jestem?* — wyniki wpisuje się w wiersze oznaczone literą A; każda kartka ma numer odpowiadający numerowi na Karcie wyników; Karta ta uwzględnia podział kartek na cztery zbiory: I, EM, F i S,
 - sortowanie II: *jaki chciałbym być?* — wyniki wpisuje się w wiersze oznaczone literą B;
- 4) badający oblicza wskaźnik: $(A - B)^2 = D^2$ dla każdej pozycji i wpisuje w wiersze oznaczone symbolem D^2 na Karcie wyników; następnie oblicza odpowiednie sumy: ΣA , ΣB i ΣD^2 (prawa strona Karty wyników); cztery cząstkowe sumy D^2 dodaje się, a rezultat wpisuje w prawym dolnym rogu Karty wyników ($\Sigma = 278$);
- 5) obliczoną ogólną sumę D^2 wstawia się do wzoru zapisanego na Karcie wyników; jest to wzór na współczynnik korelacji r_{AB} , będący wskaźnikiem poziomu samoakceptacji badanej osoby; po wykonaniu obliczeń uzyskujemy wartość wskaźnika samoakceptacji — wynosi on: $1 - \Sigma D^2 / 672 = 0,59$;
- 6) na podstawie sum ΣA i ΣB dla poszczególnych wymiarów: I, EM, F i S wykreśla się dwa profile — jeden odpowiadający ja-realnemu, a drugi ja-idealnemu;
- 7) przeprowadza się interpretację profilów; dokonuje się także interpretacji wartości obliczonego wskaźnika samoakceptacji.

1.6.7. Ocena techniki Q-sortu

Poza oczywistymi zaletami technika Q-sort posiada także pewną wadę. Jest nią utrata informacji o wartości średniej arytmetycznej i odchylenia standardowego rozkładu wyników. Wiąże się to z faktem, że pomiar przeprowadzany za jej pomocą ma charakter ipsatywny. Bez sensu jest obliczanie ogólnego wyniku jako sumy ocen wszystkich stwierdzeń składających się na Q-sort, gdyż wszyscy badani uzyskują ten sam wynik ogólny, tak samo jak w przypadku skali z wymuszonym wyborem Cosgrove'a. Można natomiast obliczać wyniki cząstkowe, będące sumami ocen pozycji zaklasyfikowanych przez badacza do określonych grup, np. grupy pozycji dotyczących różnych wymiarów obrazu samego siebie.

Nie można zatem za pomocą techniki Q-sort porównywać grup osób pod względem ogólnych średnich wyników. Można jednak porównywać średnie wartości (i utworzone z nich profile) uzyskane przez różne grupy (np. matki i ojców; dzieci matek o różnym poziomie wykształcenia; osoby zamieszkałe na wsi, w małym mieście i w dużym mieście; pracodawcy i pracownicy: osoby uczestniczące i nieuczestniczące w eksperymencie), w przypadku gdy Q-sort składa się z kilku podzbiorów pozycji (ustrukturuwana wersja techniki Q).

Gdy zależy nam na analizie relacji intraindywidualnych (ocena indywidualnych sortowań — np. przed, w trakcie i po zakończeniu procesu terapii) czy intragrupowych (analiza macierzy korelacji między osobami — taką analizę wyników Q-sortu zastosował R.U. Akeret 1965), to technika Q-sort jest niezastąpiona. Dzięki jej zastosowaniu możemy wykryć powiązania jednostek (podobieństwa osób, obiektów) niedostępne przy powierzchniowej obserwacji.

2. Zasady konstruowania skal szacunkowych

2.1. Zasady opisywania szacowanych cech i zachowań

Podstawowym warunkiem otrzymania rzetelnych i trafnych oszacowań jest prawidłowy dobór i zdefiniowanie wskaźników dla zmiennych, które mają podlegać szacowaniu, oraz prawidłowe skonstruowanie (na podstawie tych wskaźników) kategorii opisowych dla poszczegół-

nych punktów skali. Kategorie muszą być tak sformułowane, aby sędziowie (czy też osoby dokonujące samooszacowań) po otrzymaniu instrukcji nie musieli zgłaszać różnych pytań i wątpliwości. A zatem kategorie opisowe powinny być: (a) reprezentatywne dla zmiennej, której dotyczą, (b) prawidłowo sformułowane, (c) prawidłowo rozmieszczone wzdłuż skali. Podane niżej zasady opisywania szacowanych obiektów opracowaliśmy na podstawie pracy A.W. Edwardsa (1957) oraz J.P. Guilforda (1954):

- 1) Używane kategorie opisowe powinny być krótkie. Najlepiej posługiwać się krótkimi opisami zachowań, zestawem przymiotników, imiesłów. Jeżeli w opisie występują zdania, to każde zdanie powinno zawierać tylko jedną pełną myśl; zdania złożone utrudniają dokładne dokonanie oszacowań. Poważnym uchybieniem jest pytanie w jednym zdaniu o kilka różnych kwestii.
- 2) Język sformułowań powinien być prosty, jasny, bezpośredni. Należy starać się unikać:
 - a) wszelkich zwrotów dwuznacznych i wieloznacznych, dużych kwantyfikatorów typu *wszystko, zawsze, żaden, nigdy* itp.,
 - b) sformułowań ogólnikowych typu *tylko, właśnie, zaledwie,*
 - c) sformułowań zawierających oceny estetyczne, moralne, społeczne: *dobry — zły, godny — niegodny, pożądany — niepożądany, wspaniały — okropny,*
 - d) podwójnych przeczeń,
 - e) słów i wyrażeń, których sędziowie mogą nie rozumieć, specyficznej terminologii psychologicznej, żargonu, slangu, choć czasami użycie wyrażeń potocznych ułatwia np. dzieciom, młodzieży, osobom o niższym poziomie wykształcenia dokonywanie oszacowań,
 - f) powtarzania tych samych określeń we wszystkich kategoriach opisowych, np. *zupełnie nierozmowny, raczej nierozmowny, raczej rozmowny, bardzo rozmowny.*
- 3) Opis powinien być zgodny z nazwą cechy i jej definicją i nie może zawierać elementów charakterystycznych także dla innych cech (jest to problem trafności opisu), czyli kategorie opisowe powinny być rozłączne; aby to sprawdzić, warto posłużyć się metodą sędziów kompetentnych.
- 4) Prawidłowo skonstruowana kategoria opisowa powinna leżeć jak najbliżej tego punktu na skali, którego dotyczy (dlatego też niekiedy

korzystniejszy jest — szczególnie przy dłuższych opisach — pionowy układ skali).

- (5) Opisy powinny dotyczyć zachowań możliwych do zaobserwowania aktualnie, a nie występujących w przeszłości lub mogących wystąpić w przyszłości.

D. Peabody (1987, s. 65) podkreśla, że wtedy gdy posługujemy się przymiotnikami, dobrze jest formułować opisy na zasadzie opozycji (a więc tworzyć skale dwubiegunowe) oraz dbać o to, by kategoria — składająca się z kilku skal — odpowiadająca szacowanej właściwości była homogeniczna jednocześnie pod względem wartości deskryptywnej i ewaluatywnej opisów krańców każdej skali. Możemy tworzyć pary na zasadzie odwrotności ewaluatywnej, np. *rozważny — pochopny, hojny — skąpy*, albo deskryptywnej, np. *zimny — ciepły, wysoki — niski, unikający — dążący do, skupiony — rozpraszający się*.

Także K.R. Murphy i V.A. Pardaffy (1989, s. 343) podkreślają, że badacz konstruując skale szacunkowe, musi rozważyć problem podobieństwa pojęciowego (wymiar deskryptywny) i podobieństwa ewaluatywnego. Podobieństwo pojęciowe dotyczy tego, jak „spozrzegane są relacje między szacowanymi dymensjami” (ibidem), a podobieństwo ewaluatywne — „stopnia, w jakim szacowane dymensje wnoszą swój wkład do ogólnej oceny szacowanej jednostki” (ibidem). Można z tego wysnuć wniosek, że stosowanie skal podobnych pojęciowo (nierozłącznych, o zachodzących na siebie pojęciowo kategoriach opisu) może sztucznie podwyższać trafność opisu, zatem lepiej posługiwać się skalami, które oceniane są przez sędziów kompetentnych w fazie przygotowywania narzędzia jako mało podobne do siebie pod względem tego, czego dotyczą. Będą też wtedy lepiej pokrywały cały obszar interesujący badacza, a nie skupiały się w jakimś jego fragmencie. Z kolei wysoki poziom podobieństwa ewaluatywnego zapewnia większą rzetelność całego zestawu skal szacunkowych, ponieważ każda z nich ważna jest z punktu widzenia oceny ogólnej szacowanego obiektu, każda wnosi jakąś informację.

Tytułem przykładu w Bloku rozszerzającym 3 przytaczamy fragmenty Arkusza obserwacji dla nauczyciela czytania, opracowanego przez M.C. Robeck i J.A.R. Wilsona (1974, s. 630-640).

CHARAKTERYSTYKA ARKUSZA OBSERWACJI DLA NAUCZYCIELA CZYTANIA

Interesującym przykładem zastosowania skal szacunkowych do „zobiektywizowania” danych pochodzących z obserwacji jest Arkusz obserwacji dla nauczyciela czytania opracowany przez M.C. Robeck i J.A.R. Wilsona (1974, s. 630-640). Autorzy wykorzystali opracowany przez D.G. Ryansa (1960) arkusz obserwacji nauczyciela *Characteristics of teachers classroom observation record*.

Arkusz składa się z dwóch części. Część pierwsza zawiera zestaw obserwowanych zmiennych. Są to trzy zmienne: (1) proces nauczania czytania (obserwowanych tu jest 10 cech), (2) zachowanie uczniów (obserwowane są 4 cechy), (3) właściwości nauczyciela (obserwowanych jest 8 cech). Łącznie zatem obserwacja dotyczy 22 cech. Każdej z tych cech przyporządkowana jest siedmiopunktowa numeryczna skala szacunkowa. Część drugą stanowi bardzo szczegółowo opracowana instrukcja — podzielono ją także na trzy działy.

W dziale pierwszym, Nauczanie czytania, autorzy opisują dokładnie trzy spośród siedmiu punktów skali, a mianowicie punkt 2, 4 i 6 (kolejno: 2 — ocena niska, 4 — ocena średnia, 6 — ocena wysoka). Opisy te dotyczą czynności nauczyciela oddzielnie dla każdej z 10 zmiennych. W dziale drugim, Zachowanie uczniów, dla każdej z czterech zmiennych podano szczegółowy opis krańców kontinuum (tzn. pkt 1 i pkt 7), poprzez określenie cech zachowania uczniów (najczęściej w postaci przymiotników). W dziale trzecim, Właściwości nauczyciela, zastosowano podobną procedurę jak w dziale drugim, tzn. szczegółowo opisane są krańce kontinuum (pkt 1 i pkt 7). Część pierwszą wspomnianego Arkusza prezentujemy w całości, natomiast z części drugiej wybraliśmy fragmenty ilustrujące dwa wspomniane wyżej sposoby opisywania punktów skali szacunkowej.

ARKUSZ OBSERWACJI DLA NAUCZYCIELA CZYTANIA

Szkola:

Klasa:

Nauczyciel:

Data obserwacji:

Czas obserwacji:

1. Proces nauczania czytania:

- | | | | | | | |
|--|---|---|---|---|---|---|
| 1) wyposażenie klasy | 1 | 2 | 3 | 4 | 5 | 6 |
| 2) sposób nauczania (jedno- czy wielopoziomowy) | 1 | 2 | 3 | 4 | 5 | 6 |
| 3) kształtowanie umiejętności rozpoznawania słów | 1 | 2 | 3 | 4 | 5 | 6 |

4) kształtowanie umiejętności czytania ze zrozumieniem	1	2	3	4	5	6
5) doskonalenie czytania głośnego	1	2	3	4	5	6
6) przygotowanie lekcji i jej organizacja	1	2	3	4	5	6
7) wyrównywanie braków	1	2	3	4	5	6
8) techniki nagradzania	1	2	3	4	5	6
9) wdrażanie uczniów do samokształcenia	1	2	3	4	5	6
10) korelacja nauki czytania z innymi przedmiotami	1	2	3	4	5	6

2. Zachowanie uczniów:

11) apatyczni	1	2	3	4	5	6	7	pobudzeni
12) zahamowani	1	2	3	4	5	6	7	reaktywni
13) niezdecydowani	1	2	3	4	5	6	7	pewni siebie
14) zależni od nauczyciela	1	2	3	4	5	6	7	z inicjatywą

3. Właściwości nauczyciela:

15) powściągliwy	1	2	3	4	5	6	7	żywo reagujący
16) stosujący restrykcje	1	2	3	4	5	6	7	rozumiejący uczniów
17) nudny, pośepny	1	2	3	4	5	6	7	stymulujący
18) stereotypowy	1	2	3	4	5	6	7	oryginalny
19) apatyczny	1	2	3	4	5	6	7	pobudzony, żywy
20) niezdecydowany	1	2	3	4	5	6	7	pewny siebie
21) zdezorganizowany	1	2	3	4	5	6	7	systematyczny
22) sztywny	1	2	3	4	5	6	7	elastyczny

INSTRUKCJA DLA OSOBY DOKONUJĄCEJ OBSERWACJI**1. Proces nauczania czytania**

1	2	3	4	5	6
	Ocena niska		Ocena średnia		Ocena wysoka

Pozycja nr 6: przygotowanie lekcji i jej organizacja

- | | | |
|--|---|---|
| <ul style="list-style-type: none"> ■ Sekwencja materiału nauczania do przerobienia wyznacza plan lekcji | <ul style="list-style-type: none"> ■ Planowanie opiera się na uwzględnianiu dotychczasowej wiedzy uczniów oraz na zakładanych celach | <ul style="list-style-type: none"> ■ Każda lekcja ma swój cel, kierunek, wiąże się z poprzednimi i następnymi, stanowi sensowną całość |
| <ul style="list-style-type: none"> ■ Klasa podzielona jest na 2 grupy o stałym składzie | <ul style="list-style-type: none"> ■ Klasa podzielona jest na 3 stałe grupy | <ul style="list-style-type: none"> ■ Planowanie obejmuje 3 lub więcej grup o zmieniającym się składzie |
| <ul style="list-style-type: none"> ■ Na lekcji przerabia się fragmenty podręczników | <ul style="list-style-type: none"> ■ Planuje się fragmenty lekcji zgodnie z założonymi celami | <ul style="list-style-type: none"> ■ Stosuje się różne sposoby pracy, zgodne z cechami uczniów |

- Pytania zadawane uczniom wynikają z podręcznika
- Pytania wiążą problemy poruszane na danej lekcji z poprzednimi
- Przewidziany czas na samodzielną aktywność uczniów

2. Zachowanie uczniów

Pozycja nr 14: zależni od nauczyciela 1 2 3 4 5 6 7 z inicjatywą

- | | |
|---|--|
| <ul style="list-style-type: none"> ■ Oczekuje wskazówek i wyjaśnień od nauczyciela ■ Wykazuje małą zdolność robienia czegoś „sam z siebie” ■ Niezdolny do działania, gdy wymagana jest pomysłowość ■ Nie jest skłonny do przewodzenia innym lub brania za coś odpowiedzialności | <ul style="list-style-type: none"> ■ Samorzutnie proponuje nowe rozwiązania ■ Wykazuje się pomysłowością ■ Chętnie zostaje przywódcą ■ Przyjmuje odpowiedzialność bez wykręcania się |
|---|--|

3. Właściwości nauczyciela

Pozycja nr 21: zdezorganizowany 1 2 3 4 5 6 7 systematyczny

- | | |
|--|---|
| <ul style="list-style-type: none"> ■ Nie planuje pracy uczniów w klasie ■ Nieprzygotowany do lekcji ■ Niejasne cele, niezdecydowany, co ma robić ■ Traci czas w czasie lekcji ■ Mało trafne wyjaśnienia ■ Roztargniony | <ul style="list-style-type: none"> ■ Jasno widać plan działania, choć zmienia metody pracy ■ Dobrze przygotowany do lekcji ■ Rozważny we włączaniu uczniów w planowanie ■ Systematyczny w działaniu ■ Rozumie potrzeby uczniów ■ Udziela jasnych i pełnych wyjaśnień ■ Dyskutuje z uczniami, nie gubi przy tym celu lekcji |
|--|---|

2.2. Zakotwiczenie skali

Zakotwiczenie (ang. *anchoring*) skali szacunkowej polega na użyciu takich krańcowych kategorii opisowych, aby żadna (lub prawie żadna) osoba dokonująca szacowania nie mogła ich wykorzystać. Kategorie zastosowane jako „kotwice” skali mają charakter opisów stanów idealnych, takich jak np. całkowita pewność, zawsze (bez wyjątku) stabilna samoocena itp.

Może się wydawać, że takie skrajne kategorie są bezużyteczne. Istnieją jednak dwie przyczyny uzasadniające ich stosowanie (Guilford 1954, s. 264-265). Niektóre osoby szacujące korzystają z krańcowych, „idealnych” kategorii wtedy, gdy w szacowanej grupie wyróżnią takie obiekty, które — wg nich — cechuje natężenie danej właściwości w stopniu krańcowym, nie dającym się porównać z innymi obiektami. Jeżeli brak kategorii krańcowych, obiekty te zostaną zaklasyfikowane do tej samej klasy, do jakiej będą włączane obiekty o mniejszym natężeniu tej cechy. Będzie to sprzeczne z faktycznym zróżnicowaniem obiektów. W tym przypadku zatem krańcowe kategorie traktuje się jako przedłużenie skali w obu kierunkach, z których czasami, choć raczej rzadko, korzystają osoby dokonujące oszacowań.

Druga przyczyna wiąże się z faktem zauważonym przez W.A. Hunta i J. Volkmana (za: ibidem), że dodanie kategorii krańcowych w charakterze „kotwic” rozszerza cały rozkład oszacowań w kierunku tych właśnie kategorii. Daje to w efekcie większy rozsiew wyników i zwiększa stopień zróżnicowania obiektów. U osób szacujących występuje bowiem naturalna tendencja do bardzo rzadkiego używania kategorii krańcowych, a zgodnie z tym do przesuwania wszystkich sądów nieznacznie w kierunku środka skali (błąd tendencji centralnej). Tak więc skala 7-punktowa (0-6) jest w rzeczywistości skalą 5-punktową (1-5), co prowadzi do zmniejszenia dyspersji oszacowań. Jeżeli chcemy mieć rzeczywiście skalę 7-punktową (tj. taką, która podzieli całą badaną próbę na siedem podgrup), musimy sporządzić skalę co najmniej 9-punktową (0-8).

Badania prowadzone przez K.R. Murphy’ego i jego współpracowników (por. Murphy, Constans 1987; Murphy, Pardaffy 1989) także pokazują, że skale zakotwiczone są bardziej obiektywne niż niezakotwiczone. Ponadto obiektywizm ten jest większy, gdy opisy punktów krańcowych (kotwic) dotyczą zachowań aktualnie obserwowanych i od razu szaco-

wanych — a nie zachowań, które wystąpiły kiedyś, i osoba szacująca musi odwoływać się do swej pamięci.

W badaniach eksperymentalnych K.R. Murphy i V.A. Pardaffy (1989) sprawdzali, czy sposób opisywania kategorii krańcowej jest źródłem stronniczości w dokonywaniu oszacowań. W pierwszym eksperymencie osobny badane ($n = 94$) zostały podzielone na grupy 6- do 8-osobowe. Ich zadaniem była obserwacja wykładu na temat snu nagranego — na wideo — przez aktora. Aktor wykladał na poziomie przeciętnym, ale jednocześnie popełnił kilka rażących błędów, a w innych momentach pokazał mistrzostwo. Badani mieli oszacować sposób zaprezentowania na odpowiednio przygotowanych pięciu skalach: (1) jakość wykładu, (2) wzbudzenie zainteresowania (wysoki poziom podobieństwa pojęciowego i wysoki ewaluatywnego do skali 1), (3) prezentowanie materiału na odpowiednim poziomie (niski poziom podobieństwa pojęciowego i niski ewaluatywnego do skali 1), (4) prezentowanie różnych punktów widzenia (wysoki poziom podobieństwa pojęciowego i niski ewaluatywnego do skali 1), (5) włączanie informacji z innych dziedzin (niski poziom podobieństwa pojęciowego i wysoki ewaluatywnego do skali 1). Cztery skale dotyczyły różnych aspektów wykonania zadania, natomiast piąta skala — podawana jako pierwsza — była zakotwiczona tak, że dotyczyła albo błędów popełnionych przez aktora, albo tylko epizodów mistrzowskich. Badani najpierw zapoznawali się ze skalami, potem oglądali nagranie, wreszcie dokonywali oszacowań. Połowa osób w każdej grupie jako pierwszą otrzymywała skalę z negatywnymi epizodami, a połowa z pozytywnymi. Badacze sprawdzali hipotezy mówiące, że (1) „oszacowania na czterech niemanipulowanych skalach będą w sposób stronniczy ukierunkowane przez sposób szacowania na skali, których zakotwiczeniem badacz manipulował”, (2) „efekt ten będzie silniejszy w przypadku skal o dużym stopniu podobieństwa pojęciowego do skali manipulowanej” i (3) „efekt ten będzie silniejszy w przypadku skal o dużym stopniu podobieństwa ewaluatywnego do skali manipulowanej” (op. cit., s.343). Wyniki pokazały, że rzeczywiście oszacowania na skali „jakość wykładu” były istotnie wyższe ($p < 0,01$) w grupie pierwszej (pozytywne zakotwiczenie) niż drugiej (negatywne zakotwiczenie), jednakże oszacowania na pozostałych czterech skalach nie różniły się istotnie, a więc nie wykryto efektu stronniczości. W drugim eksperymencie ($n = 91$) zastosowano taką samą procedurę jak w pierwszym. Pierwsza skala pozostała taka sama (jakość wykładu), zmieniono pozostałe cztery

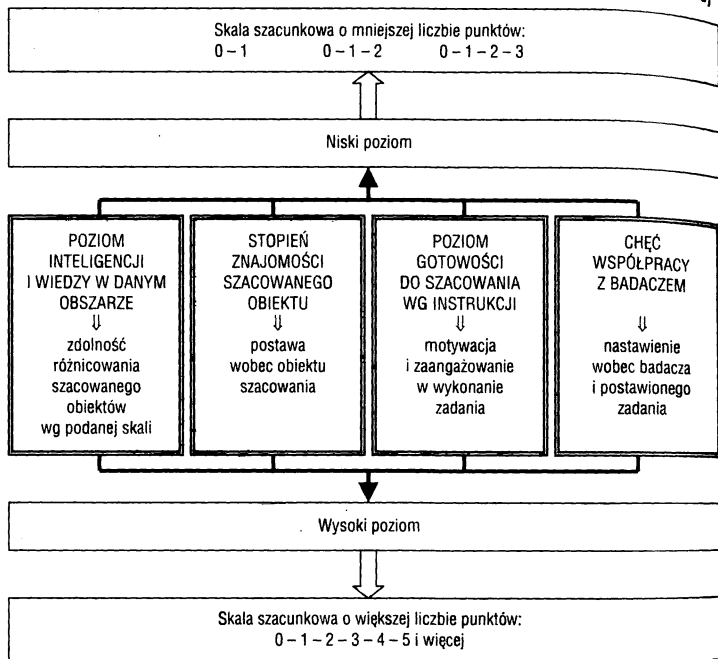
skale składające się na „styl prezentacji wykładu”. Tutaj wszystkie różnice okazały się statystycznie nieistotne, także dla skali pierwszej. Autorzy w konkluzji stwierdzają, że „stronniczość w szacowaniach jest ściśle związana z charakterem skali (ang. *scale-specific*) i nie ulega generalizacji na inne skale” (op. cit., s. 345). Różnica w wynikach eksperymentów, jakie prowadzili K.R. Murphy i J.I. Constans (1987) oraz K.R. Murphy i V.A. Pardaffy (1989) została wyjaśniona następująco: stronniczość wynikająca ze sposobu opisu zachowania stanowiącego zakotwiczenia skali szacunkowej dotyczy raczej sytuacji w laboratorium niż sytuacji życiowych. Mamy tu więc potwierdzenie tego, że narzędzie badawcze realne życiowo jest dużo trafniejsze i rzetelniejsze niż narzędzie o treściach i poleceniach odległych od codziennych doświadczeń ludzi.

2.3. Optymalna liczba punktów skali

Liczba punktów w skalach dyskretnych jest w zasadzie dowolna i zależy tylko od badacza. Różni autorzy zaopatrywali konstruowane przez siebie skale w różną liczbę punktów. Niemniej istnieje pewien związek między liczbą użytych punktów a rzetelnością skali. Wprowadzenie zbyt małej liczby punktów prowadzi do obniżenia mocy dyskryminacyjnej, a więc do obniżenia rzetelności narzędzia. Z drugiej strony, zbyt duża ich liczba może przekraczać zdolność różnicowania u osób dokonujących oszacowania oraz powodować, że będą one grupowały kilka punktów w jeden. Liczba punktów zależy zatem od kilku czynników. Prezentuje je rysunek 12.

Przeprowadzono wiele badań zmierzających do ustalenia optymalnej liczby punktów skali szacunkowej. Według A.W. Bendiga (1953) rzetelność skal 3-, 5-, 7- i 9-punktowych jest prawie taka sama, natomiast rzetelność skal 11-punktowych jest dużo mniejsza. Badania, jakie przeprowadził na grupie 225 studentów nad oceną ich wiedzy o 12 państwach, pokazały także, że rzetelność była większa w przypadku skal z opisanymi zarówno punktami krańcowymi, jak i punktem środkowym, najmniejsza zaś przy opisanu jedynie punktu środkowego. W innych badaniach przeprowadzonych na grupie 236 osób nad preferencjami pokarmowymi A.W. Bendig (1954) użył skal 2-, 3-, 5-, 7- i 9-punktowych i obliczył dwa wskaźniki rzetelności: rzetelność skal szacunkowych i rzetelność (zgodność) osób szacujących. Ta pierwsza była podobna dla

Rysunek 12. Czynniki wpływające na decyzję badacza co do liczby punktów skali szacunkowej



Źródło: opracowanie własne.

skal o różnej liczbie punktów, natomiast ta druga była podobna dla skal 5-, 7- i 9-punktowych, niższa w przypadku skal 3-punktowych, a najniższa dla skal 2-punktowych.

Samuel S. Komorita i William K. Graham (1965) przeprowadzili badania nad rzetelnością skal 2- i 6-punktowych. Wykazali, że zwiększenie liczby punktów skali z dwóch do sześciu daje wyraźny wzrost rzetelności. Ponadto z ich badań wynika, że jeżeli zbiór skal szacunkowych jest niezbyt liczny, należy zwiększyć liczbę punktów na poszczególnych skalach. Zabieg ten zapewnia większe zróżnicowanie badanej próby.

Z analizy 23 tysięcy oszacowań przeprowadzonej przez E.S. Conkлина (por. Guilford 1954, s. 290) wynika, że jeżeli sędziami dokonującymi oszacowań mają być osoby niemające odpowiedniego doświadczenia w tym zakresie, a więc nieprzygotowane poprzez ćwiczenie, to maksymalna liczba punktów, którą można się posłużyć, wynosi 5 dla skali jednobiegunowej (np. brak lęku — duży lęk) oraz 9 dla skali dwubiegunowej (np. apatyczny — ruchliwy). P.M. Symonds (por. Komorita, Graham 1965) proponuje skalę 7-punktową jako optymalną. C.E. Osgood (Osgood, Suci, Tannenbaum 1957) w opracowanym przez siebie narzędziu — dyferencjale semantycznym — posłużył się właśnie skalami 7-punktowymi. J.P. Guilford (op. cit., s. 292) jednak uważa, że liczba siedmiu kategorii podana przez N.A. Symonds jest zaniżona i że w pewnych przypadkach można użyć nawet skali 25-punktowej (!).

C. Peterson i M. Grant (2001) przeprowadziły badania na dzieciach w wieku przedszkolnym, poszukując odpowiedzi na pytanie o to, kiedy ich odpowiedzi są bardziej rzetelne. Przebadały indywidualnie 32 dzieci: 16 w grupie młodszej, czyli w wieku od 3;2 do 4;4 oraz 16 w grupie starszej, czyli w wieku od 4;8 do 5;1. Badania polegały na udzielaniu odpowiedzi na pytania, które miały różną liczbę i różny charakter odpowiedzi: (1) odpowiedzi typu „tak — nie” i odpowiedź „tak” była prawdziwa, (2) odpowiedzi typu „tak — nie” i odpowiedź „nie” była prawdziwa, (3) dwie odpowiedzi do wyboru, przy czym poprawna była odpowiedź pierwsza, (4) dwie odpowiedzi do wyboru, przy czym poprawna była odpowiedź druga, (5) dwie odpowiedzi do wyboru, przy czym żadna z nich nie była prawidłowa (por. op. cit., s. 121). Połowę dzieci poinstruowano, że mogą powiedzieć „nie wiem”, gdy nie będą znały odpowiedzi na stawiane im pytania. Badanie przebiegało w dwóch krokach. Pierwszy krok to sesja, w której każde dziecko oddzielnie uczestniczyło w różnych sytuacjach z dwoma eksperymentatorami. Jeden z nich nawiązywał różne interakcje z dzieckiem, drugi często wchodził i wychodził, rozmawiał z pierwszym eksperymentatorem i z dzieckiem. W połowie przypadków głównym eksperymentatorem była kobieta, a pomocniczym mężczyzna, a w połowie — odwrotnie. Krok drugi — po upływie tygodnia — to odpowiadanie na pytania dotyczące tego, co zdarzyło się tydzień wcześniej. Osoba, która prowadziła tę drugą sesję, nie uczestniczyła w pierwszej. Każdemu dziecku zadano dwadzieścia pytań — przyporządkowanych po cztery do wyróżnionych pięciu sposobów odpowiadania. Odpowiedzi dzieci były prawidłowe znacznie

częściej w sytuacji, gdy odpowiedź była typu „tak — nie” i odpowiedź „tak” była prawidłowa oraz gdy podano im dwie odpowiedzi do wyboru i jedna z nich była prawidłowa. Gdy prawidłowa była odpowiedź „nie” w sytuacji „tak — nie”, dzieci popełniały dużo błędów. Badaczki stwierdzają w konkluzji, że odpowiedzi typu „tak — nie” stwarzają więcej problemów niż odpowiedzi z wymuszonym wyborem (tj. gdy podaje się odpowiedzi do wyboru), ponieważ silna jest u dzieci tendencja do odpowiadania „tak”. Ponadto jeżeli odpowiedź miała charakter wymuszonego wyboru, dzieci znacznie częściej mówiły „nie wiem”, gdy nie mogły wybrać żadnej z podanych odpowiedzi. W sytuacji propozycji odpowiedzi „tak — nie” dzieci bardzo rzadko mówiły „nie wiem”, nawet gdy nie były pewne swej odpowiedzi. Według badaczek wyniki te wskazują na wyższość narzędzi typu wymuszonego wyboru oraz mają duże znaczenie dla konstruowania scenariuszy rozmowy z małymi dziećmi jako świadkami w czasie przesłuchań dla potrzeb sądu (por. op. cit., s. 122).

3. Poziom pomiaru oszacowań

Badacz, który zebrał dane szacunkowe dotyczące zachowania się członków badanej grupy staje przed pytaniem: Jakie operacje statystyczne można przeprowadzić na tym zbiorze? Jest to pytanie o poziom pomiaru, jakiemu odpowiadają oszacowania otrzymane za pomocą skal szacunkowych. Wstępna orientacja w pracach o charakterze eksperymentalnym, których autorzy obok innych narzędzi stosowali skale szacunkowe, pokazuje, że istnieją dwa stanowiska w kwestii poziomu pomiaru oszacowań. Jedno z nich traktuje skale szacunkowe jako skale porządkowe, a drugie jako skale interwałowe.

Uważamy jednak, że zgodnie z klasyfikacją skal pomiarowych wprowadzoną przez S.S. Stevensa, skalom szacunkowym przysługuje status skali porządkowej. Chcąc być w zgodzie z tą klasyfikacją, należy przeprowadzać tylko takie operacje na zbiorze danych szacunkowych i stosować tylko takie testy oceny istotności różnic oraz miary korelacji, jakie dopuszcza ten typ skali pomiarowej. Traktowanie przez niektórych badaczy (np. przez Bandurę i Waltersa (1968)) skal szacunkowych jako interwałowych jest sprzeczne z ideą klasyfikacji Stevensa. Przyjęcie proponowanego przez nas stanowiska uchroni badacza przed opaczynym

wyciąganiem daleko idących wniosków z wyników nieprawidłowo zastosowanych testów statystycznych i miar korelacji.

4. Rzetelność skal szacunkowych

Zanim zaczniemy posługiwać się jakimkolwiek narzędziem przez nas skonstruowanym musimy poznać jego rzetelność, a co za tym idzie błąd, na jaki jesteśmy narażeni przy posługiwaniu się nim. Współczesna psychometria dysponuje różnymi technikami obliczania rzetelności narzędzi pomiarowych i różnymi jej miarami (por. np. Brzeziński 1996, rozdział 15; Hornowska 2001, s. 41-79).

W przypadku skal szacunkowych sprawdzanie rzetelności może przebiegać w dwojaki sposób w zależności od tego, czy (1) badacz chce sprawdzić rzetelność tylko jednej skali czy też (2) homogenicznego ich zestawu, dla którego oblicza się jeden ogólny wynik, będący sumą wyników cząstkowych pochodzących od różnych skal szacunkowych (jak w przypadku skali skumulowanych ocen). W jaki zatem sposób obliczać rzetelność w obu tych przypadkach?

Jeżeli dwie lub więcej osób ocenia za pomocą jednej skali szacunkowej (lub zestawu skal) zachowanie się dowolnej liczby osób badanych, miarą rzetelności będzie wysokość współczynnika korelacji między dwoma zbiorami oszacowań. I tak w przypadku dwóch osób dokonujących oszacowań miarą tą może być współczynnik tau-Kendalla. Natomiast współczynnik W-Kendalla będzie miarą zgodności opinii sędziów kompetentnych w przypadku większej ich liczby (troje sędziów i więcej) (por. Brzeziński 1996, s. 500-505). Można jednak tak postąpić tylko w sytuacji, gdy (Goode, Hatt 1965, s. 285):

- a) w danej populacji istnieje jednorodny system ocen, który powinien ujawnić się podczas szacowania; miarą jednorodności tego systemu jest wariancja ocen; jeżeli wariancja jest zbyt duża, należy podejrzewać, że założenie o jednorodności ocen nie jest spełnione;
- b) wzrost liczby sędziów kompetentnych sprawia, że zwiększa się szansa, iż oszacowania będą reprezentacyjne dla danej populacji.

W przypadku skal skumulowanych ocen najbardziej efektywną miarą rzetelności będzie estymacja wewnętrznej zgodności, np. wg wzoru

²⁰ Kudera-Richardson, przystosowanego przez G.A. Fergusona (1951)

do obliczania rzetelności narzędzi składających się ze stwierdzeń wielokategorialnych.

Rozważając zagadnienie dokładności oszacowań przeprowadzonych za pomocą skal szacunkowych, należy uwzględnić fakt, że dokładność ta nie zależy tylko od samego narzędzia, ale także od umiejętności szacowania i kompetencji osób dokonujących tych oszacowań. Można zwiększyć dokładność oszacowań przez zwiększenie liczby sędziów kompetentnych dokonujących oszacowań oraz odpowiednie ich przeszkolenie w zakresie umiejętności postępowania się skalami szacunkowymi (por. Blok rozszerzający 4).

BLOK ROZSZERZAJĄCY 4

ZASADY TRENINGU SĘDZIÓW KOMPETENTNYCH

Przebieg procesu trenowania w zakresie umiejętności szacowania szczegółowo omawia D.G. Ryans (1960). Proces ten może przebiegać zarówno w formie indywidualnej, jak i grupowej. Znacznie częściej stosuje się jednak trening grupowy. Przebiega on w kilku etapach:

- 1) instruktor omawia z osobami poddanymi treningowi podstawowe zagadnienia związane z obserwacją i szacowaniem, przedstawia przykłady arkuszy obserwacyjnych i skal,
- 2) przyszli sędziowie sami zapoznają się szczegółowo z budową i rodzajami skal i arkuszy,
- 3) jednocześnie z instruktorem obserwują np. nauczyciela w klasie i niezależnie od siebie dokonują oszacowań na kopiach tego samego arkusza obserwacji,
- 4) następuje porównanie oszacowań, dyskusja, wyjaśnienie trudności,
- 5) ponowna wspólna obserwacja, porównywanie oszacowań, dyskusja,
- 6) w końcowej fazie treningu przyszli sędziowie oraz instruktor przeprowadzają (u Ryansa — 45-minutową) obserwację tych samych nauczycieli, ale w różnym czasie (w poprzednich etapach w tym samym czasie), następnie znów porównuje się oszacowania.

Podczas całego treningu oblicza się korelacje między oszacowaniami instruktora i osób trenowanych — miary te ułatwiają kontrolę przebiegu treningu oraz ocenę jego efektywności. U Ryansa po pełnym treningu korelacja mieściła się w przedziale 0,80-0,90.

5. Błędy popełniane przy szacowaniu

Oszacowania uzyskane od sędziów wykorzystuje się po odpowiednich przetworzeniach do różnych porównań i analiz intra- oraz interindywidualnych. Przyjmuje się przy tym, że pochodzą one od osób nie popełniających żadnych błędów i zdolnych do prawidłowej oceny natężenia określonej cechy u badanej osoby. W rzeczywistości jednak w wielu sytuacjach, w których korzystamy z oszacowań, tak nie jest. Można wyróżnić kilka najbardziej powszechnie występujących i typowych błędów popełnianych przez osoby dokonujące oszacowań (por. Guilford 1954, s. 278-279; Kerlinger 1986, s. 495; Manstead i in. 2001, s. 71-73 i 80). Są to: błąd łagodności, błąd tendencji centralnej, efekt *halo* (błąd *halo*, efekt aureoli), błąd bliskości, błąd kontrastu (por. rodzaje błędów w ocenianiu ludzi analizowane przez Wojciszke 2000; 2003, s. 120-123).

5.1. Błąd łagodności

Błąd łagodności (ang. *error of leniency*) wiąże się z występującą u osób dokonujących oszacowań — najczęściej nieświadomą — tendencją do dokonywania oceny osób znanych lub bliskich w sposób korzystny dla nich. Tendencja ta występuje niezależnie od rodzaju szacowanej cechy czy zachowania. Niektórzy sędziowie, aby przeciwdziałać tej tendencji, z której zdają sobie sprawę, celowo oceniają znanych sobie badanych surowiej, niż to jest uzasadnione.

Można zatem podzielić ogół sędziów na dwie grupy:

- 1) grupa szacujących łagodnie (ang. *easy raters*): osoby te nie uświadamiają sobie popełnianych przez siebie błędów ani ich przyczyn i szacują znanych sobie badanych znacznie łagodniej niż innych, przedstawiając ich w bardziej korzystnym świetle,
- 2) grupa sędziów szacujących surowo (ang. *hard raters*): ta część sędziów, uświadamiając sobie rodzaj błędów oraz ich przyczyny, świadomie szacuje bardziej surowo osoby znane sobie; efektem może być nieuzasadnione przedstawienie badanych w mniej korzystnym dla nich świetle. W przypadku zbyt surowego szacowania mówimy o ujemnej łagodności (ang. *negative leniency*).

Ten rodzaj błędów charakterystyczny jest także — choć w nieco szerszym zakresie — dla osób, które cechuje prokuratorska lub adwo-

kacka postawa wobec badania. Nazwy tych postaw pochodzą od N. Talenta (por. Płużek 1971, s. 48). Postawę prokuratorską cechuje przy opisywaniu i ocenianiu cech osobowości podkreślanie cech ujemnych, czy wręcz patologicznych, postawę adwokacką — uwypuklanie cech dodatnich. A przecież w prawidłowym opisie człowieka (zdrowego czy chorego) muszą występować i elementy dodatnie, i elementy ujemne.

Przewyciężenie dodatniego błędu łagodności polega m.in. na takim konstruowaniu skal, aby w kategoriach opisowych występowało jak najmniej określeń negatywnych. Na przykład skala do oceny zdrowia fizycznego może składać się z pięciu kategorii, z których tylko jedna zawiera określenie negatywne: *złe, nie najgorsze, dobre, bardzo dobre, wyśmienite*.

5.2. Błąd tendencji centralnej

Błąd tendencji centralnej (ang. *error of central tendency*) polega na unikaniu przez sędziów kategorii skrajnych i przesuwaniu tym samym oszacowań w kierunku kategorii środkowych. Jest to spowodowane z jednej strony naturalną tendencją do unikania kategorii skrajnych, a z drugiej strony faktem, że skrajne nasilenie danej zmiennej występuje w rzeczywistości u niewielu osób (a takich może np. nie być w badanej grupie). Skutkiem tego rodzaju błędu jest uzyskiwanie bardzo podobnych — i to najczęściej na poziomie przeciętnym — ocen dla większości szacowanych osób, a więc bardzo małe zróżnicowanie grupy. Z podobnym błędem mamy do czynienia w tzw. efekcie Barnuma (por. Płużek 1971, s. 48), kiedy to w opisie danej jednostki podaje się takie cechy, które mogą charakteryzować wiele osób.

Tendencji tej można przeciwdziałać przez:

- pozostawienie większych przerw między kategoriami środkowymi, szczególnie jeżeli posługujemy się skalami graficznymi,
- wprowadzenie większej różnicy znaczeniowej między kategoriami opisowymi dotyczącymi punktów położonych bliżej krańców skali niż między kategoriami opisującymi punkty leżące bliżej środka,
- dodanie tzw. kotwic rozszerzających w obu kierunkach zakres szacowanej cechy (zagadnienie to omówiliśmy wyżej).

5.3. Efekt halo

Efekt *halo* stanowi bardzo powszechny typ błędu. Jego nazwa pochodzi od Edwarda L. Thorndike'a. Wojciszke (2000) określa ten rodzaj błędu w ocenianiu ludzi jako efekt aureoli (op. cit., s. 220). W jego interpretacji jest on bliski błędowi określanemu jako efekt pierwszeństwa, który sprawia, że: „osoba opisana jako <inteligentna — przedsiębiorcza — impulsywna — krytyczna — uparta — zazdrosna> wyżej jest ceniona niż opisana w odwrotnej kolejności” (op. cit., s. 214).

Błąd ten przysparza wielu kłopotów nie tylko osobom dokonującym oszacowań, ale także psychologom posługującym się takimi metodami klinicznymi, jak rozmowa kliniczna, obserwacja, analiza wytworów pacjenta. Efekt *halo* można określić jako „tendencję w dokonywaniu oceny lub oszacowania jakiejś cechy jednostki do ulegania wpływom innej cechy lub własnemu ogólnemu wrażeniu odnośnie do tej osoby” (English, English, 1958, s. 236). L. Zebrowitz (w: Manstead i in. 2001, s. 71) pisze, że efekt aureoli (ang. *halo effect*) „występuje w procesie tworzenia wrażenia i polega na tym, że jedna pozytywna cecha obiektu oceny podnosi pozytywność jego pozostałych cech [...]”.

Wpływ efektu *halo* na szacowanie jest taki, że sędziowie ulegają ogólnemu wrażeniu, jakie wywierają na nich badane osoby i w rezultacie ich oszacowania stają się nietrafne. Dzieje się tak dlatego, że istnienie tej ogólnej postawy, jaką mamy wobec aktualnie badanej osoby, powoduje, że mamy podobną postawę wobec poszczególnych właściwości jej osobowości i zachowania się. Krańcowym przypadkiem występowania efektu *halo* jest taka sytuacja, kiedy wysoko oceniamy wszystkie cechy badanej osoby, ponieważ jakaś jedna lub dwie cechy jej osobowości wywierają na nas bardzo dobre ogólne wrażenie, lub też gdy oceniamy jakąś osobę bardzo nisko pod względem wszystkich zmiennych, ponieważ jakaś jej jedna właściwość wyraźnie nam się nie podoba.

Wojciszke (op. cit., s. 214-215) wyjaśniając wpływ efektu pierwszeństwa na sposób dokonywania oceny innych ludzi, formułuje trzy hipotezy. Pierwsza to hipoteza spadku u wagi: w miarę jak badacz dokonuje kolejnych oszacowań, słabnie u niego koncentracja uwagi, a więc kolejne cechy są oceniane mniej rzetelnie. Druga to hipoteza dyskredytowania sprzecznych informacji: gdy oceniany obiekt jest w jakiś sposób ważny dla osoby oceniającej, staje się ona bardziej wrażliwa na istnienie sprzeczności między informacjami i wtedy, gdy pierwsze jej

oszacowania będą wysokie, kolejne będą również wysokie, co pozwoli na uniknięcie dysonansu poznawczego między własnymi oszacowaniami. Trzecia to hipoteza ukierunkowanego rozwijania sądu o drugiej osobie: pierwsze oszacowania dostarczają kontekstu do interpretacji kolejno szacowanych cech. Jak pisze Wojciszke (ibidem): „informacje uzyskane jako pierwsze z kolei i początkowy sąd o ocenianym człowieku służą jako umysłowe narzędzie subiektywnego ujednoznaczniania danych”.

Jeżeli dokonujemy oszacowania kilku czy kilkunastu cech osobowości u tej samej osoby, to efekt *halo* może wprowadzić pewną dodatnią — w istocie iluzoryczną — korelację między szacowanymi cechami. Taka pozorna interkorelacja podwyższa wewnętrzną zgodność zestawu skal szacunkowych jako jednorodnego narzędzia. Efekt *halo* przyczynia się także do podniesienia wartości współczynnika stałości oszacowań i zgodności między osobami ich dokonującymi (Medley, Mitzel 1967, s. 306).

Łatwiej nam będzie przeciwdziałać wystąpieniu efektu *halo*, jeżeli będziemy wiedzieli, w jakich sytuacjach na pewno się on pojawi. P.M. Symonds (por. Guilford 1954, s. 279) wymienia pięć takich sytuacji:

- a) zmienną stanowiącą przedmiot szacowania jest bardzo trudno zaobserwować,
- b) szacowana zmienna rzadko występuje w badaniach, jest rzadko wyróżniana i oceniana,
- c) cecha jest niejasno zdefiniowana, brak (lub jest ich bardzo niewiele) konkretnych wskaźników świadczących zarówno o jej występowaniu, jak i natężeniu; aby zapobiec występowaniu efektu *halo* w tej sytuacji, należy szczegółowo opisać poszczególne punkty skali (najlepiej przez podanie konkretnych przykładów zachowań w różnych sytuacjach),
- d) cecha dotyczy kontaktów jednostki z innymi ludźmi i ujawnia się w tych kontaktach,
- e) cecha dotyczy sfery moralności człowieka oraz cech tzw. charakteru, czyli stosunku jednostki do siebie samej, do innych ludzi oraz do świata jako całości.

Szacowanie wszystkich osób najpierw pod względem stopnia występowania u nich jednej cechy, a potem pod względem kolejnych cech oraz umieszczanie zapisu każdej cechy (a nie osoby) na odrębnej kartce także przeciwdziała wystąpieniu efektu *halo*. Najbardziej jednak

efektywnym sposobem przeciwdziałania efektowi *halo* oraz innym błędom szacowania jest odpowiednie przygotowanie, poprzez ćwiczenie, osób szacujących. Osoba, która wie o różnego rodzaju błędach, jakie może popełnić, będzie bardziej uważna i będzie się starała ich unikać.

5.4. Błąd bliskości

Błąd bliskości (ang. *proximity error*), podobnie jak efekt *halo*, powoduje powstawanie zbyt wysokich, nieprawdziwych korelacji między szacowanymi zmiennymi. Przyczyną tego zjawiska jest w tym przypadku zbyt duża bliskość cech, zarówno przestrzenna (kiedy sędzia otrzymuje kilka skal szacunkowych, dotyczących różnych wymiarów osobowości lub zachowania, umieszczonych na jednej kartce), jak i czasowej (sędzia dokonuje szacowania różnych jakościowo cech bezpośrednio po sobie). To wszystko powoduje, że oszacowania następujące bezpośrednio po sobie są bardziej do siebie podobne niż oszacowania w dość dużym odstępście czasowym. Wykazali to w swoich badaniach L. Stockford i H.W. Bissell (por. Guilford 1954). Pokazali mianowicie, że średnia interkorelacja dla cech sąsiadujących ze sobą wynosiła 0,66. Wartość interkorelacji w miarę wzrostu oddalenia cech od siebie systematycznie malała i np. dla dwóch cech „przedzielonych” pięcioma innymi wynosiła już tylko 0,46. Podobne rezultaty otrzymali dla zmienionego w sposób losowy układu tych samych zmiennych, które wystąpiły w pierwszej fazie badania. Rezultat tych badań pokazuje, że w żadnym wypadku nie można traktować interkorelacji wyników (oszacowań) jako wskaźnika interkorelacji szacowanych zmiennych.

Istnieje kilka sposobów przeciwdziałania występowaniu błędu bliskości. Oto one:

- a) umieszczenie każdej skali szacunkowej na oddzielnej kartce,
- b) w przypadku badania nie pojedynczych obiektów (osób), ale całych ich grup (kilku czy kilkunastu osób) należy najpierw dokonać oszacowania wszystkich obiektów na pierwszej skali, a potem dopiero na drugiej, trzeciej itd.
- c) rozmieszczenie skal dotyczących podobnych cech (lub cech składających się razem na jedną bardziej ogólną zmienną) daleko od siebie, a cech zupełnie odmiennych bezpośrednio po sobie,

- d) takie zorganizowanie badań, aby sędziowie jednorazowo szacowali badaną osobę (osoby) tylko na jednej skali, po upływie pewnego czasu, np. po upływie 1-2 dni, na drugiej itd.

5.5. Błąd kontrastu

Błąd kontrastu (ang. *contrast error*) został wykryty przez H.A. Murraya (por. Guilford 1954, s. 279). Polega on na występowaniu u sędziów tendencji do przypisywania osobom podlegającym badaniu takich cech, które stanowią przeciwieństwo cech posiadanych przez nich samych. Na przykład sędziowie, którzy są bardziej zdyscyplinowani — i sami też się za takich uważają — mogą oceniać osoby badane jako mało systematyczne i mało zdyscyplinowane. Tendencja ta ujawnia się w takim przypadku w nadmiarze zbyt niskich oszacowań. Może też zdarzyć się sytuacja odwrotna, kiedy to sędziowie szacują pewne cechy badanych (te, których sami nie posiadają) bardzo wysoko i wtedy wystąpi zbyt wiele oszacowań w nieuzasadniony sposób wysokich.

W przypadku niektórych cech może zachodzić zjawisko zupełnie odwrotne. Osoba szacująca, którą cechuje np. wyjątkowa gotowość do współpracy i dobrego współżycia z innymi ludźmi, może nie zauważać przejawów niechęci do współpracy występujących u badanych. Będzie ona wtedy dawała znacznie wyższe oszacowania, które nie będą zgodne z rzeczywistym nasileniem szacowanej zmiennej. Ten rodzaj błędu wiąże się z występującą u ludzi tendencją do oczekiwania, że inni ludzie są podobni do nich. Zjawiska te można wyjaśnić, odwołując się do wpływu mechanizmów obronnych na zachowanie człowieka, a szczególnie wpływu projekcji oraz tzw. formowania reakcji przeciwnych (ang. *reaction formation*).

6. Typy postaw wobec badania

Błąd kontrastu wiąże się z zagadnieniem szerszym dotyczącym ogólnych postaw wobec badania. W naszym przypadku będą to postawy dwójakiego rodzaju:

- a) jeżeli mamy grupę osób i oszacowań dokonuje sędzia, interesuje nas postawa sędziego wobec badania i wobec osób, które ma oszacować na skalach,

b) jeżeli dana jednostka sama dokonuje oszacowania swoich cech na skalach, interesuje nas jej postawa wobec samej siebie.

Zajmiemy się tu bardziej szczegółowo zagadnieniem drugim ze względu na to, że w sytuacji tej jest znacznie trudniej zapewnić sobie wysoką rzetelność wyników (samooszacowań). Najczęściej wymienia się dwa typy zniekształcenia odpowiedzi występujące w sytuacji, kiedy badani mają sami siebie opisywać czy oceniać. Zniekształcenia te występują niezależnie od tego, jakim rodzajem narzędzia posłużymy się do zebrania potrzebnych informacji: skalą szacunkową, rozmową czy kwestionariuszem osobowości. Pierwszy rodzaj zniekształceń nazwany przez A.L. Edwardsa (1957; także Nederhof 1991) „zmienną aprobaty społecznej” zależy od ogólnej treści pytania czy skali. Polega on na udzielaniu przez badanego takich informacji o sobie, aby były zgodne ze spostrzeganymi przez niego wymaganiami społeczeństwa, stawianymi jednostkom dobrze przystosowanym. Drugi rodzaj zniekształceń to „skłonność do zgadzania się”: wiąże się ona ze sposobem postawienia pytania i sformułowania kategorii opisowych na skali. Występuje on głównie wtedy, gdy osoba badana obawia się wyników badania. Udziela wtedy odpowiedzi w sposób bardzo ostrożny, np. daje potwierdzające odpowiedzi na pytania kwestionariusza, jeżeli nie wiąże się to dla niej z żadnym ryzykiem, na skalach szacunkowych zaznacza pozycje środkowe, „neutralne” lub kategorie bliższe dodatniemu krańcowi skali.

Postawę wobec samego siebie, bezpośrednio wpływającą na sposób udzielania informacji o sobie, wiąże się z niektórymi właściwościami samooceny: z jej poziomem (wysoka *versus* niska), adekwatnością (trafna *versus* nietrafna), stabilnością (stabilna *versus* niestabilna) (por. Brzezińska 1973; Paszkiewicz 1974).

W. Sanocki (1970) wyróżnił trzy rodzaje postaw wobec badania. Postawy te można uporządkować wzdłuż pewnego kontinuum. Na lewym krańcu znajduje się postawa obronna. Przejawia się ona w skłonności do ukrywania swoich wad i przedstawiania się w jak najlepszym świetle. Jest to postawa charakterystyczna dla osób o bardzo wysokiej, ale nieadekwatnej, czyli zawyżonej samoocenie. Pośrednie miejsce na kontinuum zajmuje postawa otwarta, cechuje ją szczerść i krytyczność odpowiedzi (samoocena obiektywna, realistyczna, trafna). Prawy kraniec to postawa nadmiernie krytyczna. Osoby, u których występuje ten rodzaj postawy, cechują się znacznie zaniżoną samooceną i przejawiają skłonność do przypisywania sobie różnych wad. W przypadku badania z użyciem skali

szacunkowej będą one zaznaczały kategorie bliskie „negatywnym” krańcom skali. W przypadku korzystania z samooszacowań w zasadzie nie mamy możliwości wyeliminowania wpływu nieadekwatnej — czyli zawyżonej lub zaniżonej — samooceny na wyniki. Powinniśmy jednak czynnik ten uwzględniać przy interpretowaniu wyników, wyciąganiu wniosków oraz — co najważniejsze — formułowaniu uogólnień.

7. Podsumowanie

Skale szacunkowe są czasami jedynym narzędziem, jakiego może użyć psycholog, by uzyskać dane ilościowe, na podstawie których będzie mógł dokonywać porównań inter- i intraindywidualnych. Nie zawsze przecież konieczne czy możliwe jest posługiwanie się „twardymi” narzędziami diagnostycznymi. Obserwacja, rozmowa, analiza wytworów, techniki projekcyjne — te wszystkie narzędzia dostarczają wielu wartościowych informacji, jeżeli tylko dysponujemy doświadczeniem badawczym i odpowiednią teorią pozwalającą na wyjaśnienie, a nie tylko opisanie badanego zjawiska.

Badacz decyduje też o tym, kto będzie dokonywał oszacowania — on sam, osoba badana czy też odpowiednio przygotowani sędziowie kompetentni. Kiedy osoba badana nie jest w stanie dokonać samooszacowania (dotyczy to szczególnie dzieci — por. Kamphaus, Petoskey, Rowe 2000 — ale także osób o niskim poziomie wykształcenia, generalnie osób, u których niska jest zdolność dokonania refleksji nad sobą i wyrażenia opinii o sobie w formie proponowanej przez badacza), należy posłużyć się sędziami kompetentnymi, a kiedy jest to technicznie trudne lub wręcz niemożliwe albo budzi wątpliwości natury etycznej, wtedy sam badacz musi „przełożyć” zebrane przez siebie informacje na dane ilościowe. Zawsze jednak o tym, jaka będzie procedura gromadzenia i porządkowania informacji diagnostycznych, decyduje badacz, biorąc pod uwagę cel badania, możliwości osoby badanej i warunki zewnętrzne, w tym czas i liczbę możliwych z nią spotkań.

Na koniec raz jeszcze podkreślmy, że skale szacunkowe dostarczają cennych informacji, a współcześnie — jak piszą Kamphaus, Petoskey i Rowe (op. cit.) — ich stosowanie to wyraźna i nasilająca się tendencja wśród badaczy i praktyków zajmujących się zaburzeniami zachowania u dzieci. Możliwe zatem staje się, po pierwsze, wyłanianie grup dzieci

podobnych do siebie ze względu na charakter ujawnianych problemów dzięki zastosowaniu analizy skupień, np. metodą *k*-średnich, przeprowadzonej na przypadkach/osobach (patrz: Marek 1989; por. wyniki analizy skupień na osobach w pracy Hetmańska-Bugajska, Brzezińska 2002 oraz Smirnowa 2003). Po drugie, wewnątrz tych w miarę jednorodnych grup można przeprowadzić analizę skupień na zmiennych, np. wg metody aglomeracyjnej (Marek, op. cit.) lub analizę czynnikową (Zakrzewska 1994), co pozwoli na połączenie szacowanych zachowań w wiązki, wg przyjętych za podstawę diagnozy konstruktów teoretycznych (dymensji), a to znacznie się różni od postępowania w ramach tzw. diagnozy różnicowej. Tak wyodrębniane typy dzieci, jak i wiązki zmiennych cechuje wysoki poziom trafności prognostycznej (Ferguson, Horwood 1995; za: Kamphaus, Petoskey, Rowe, op. cit., s. 161), co nie jest bez znaczenia dla konstruowanych programów edukacyjnych, profilaktycznych czy terapeutycznych.

Jeżeli zatem badacz zadba o poprawne skonstruowanie zestawu skal szacunkowych, skorzysta z pomocy i rad sędziów kompetentnych w fazie przygotowania narzędzia, przeprowadzi badania pilotażowe, sprawdzi ich rzetelność i trafność, to otrzyma w rezultacie wiarygodne narzędzie dostarczające mu wielu cennych informacji diagnostycznych, przede wszystkim jednak wartościowych i użytecznych w procesie planowania różnego typu oddziaływań.

Literatura

- Akeret R.U. (1965). Interrelationships among various dimensions of the self-concept. W: G. Lindzey, C.S. Hall (red.), *Theories of personality: primary sources and research*. New York, NY: J. Wiley.
- Appelt K. (2004). *Kompetencja twórczej adaptacji u nauczycieli w warunkach zmiany społecznej*. Poznań: Instytut Psychologii Uniwersytetu im. Adama Mickiewicza (maszynopis niepublikowany).
- Aronson E., Wilson T.D., Akert R.M. (1994). *Social psychology. The heart and the mind*. New York, NY: Harper Collins.
- Bandura A., Walters R. (1968). *Agresja w okresie dorastania*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Bendig A.W. (1953). The reliability of self-ratings as a function of the amount of verbal anchoring and of the number of categories of the scale. *Journal of Applied Psychology*, 37, 38-41.

- Bendig A.W. (1954). Reliability and the number of rating-scale categories. *Journal of Applied Psychology*, 38, 38-40.
- Błuszkowski J. (2003). *Stereotypy narodowe w świadomości Polaków*. Warszawa: Dom Wydawniczy ELIPSA.
- Brzezińska A. (1973). Struktura obrazu własnej osoby i jego wpływ na zachowanie. *Kwartalnik Pedagogiczny*, 3, 87-97.
- Brzezińska A. (1995). *Kwestionariusz do szacowania cech temperamentu dzieci*. Poznań: Instytut Psychologii Uniwersytetu im. Adama Mickiewicza w Poznaniu (maszynopis niepublikowany).
- Brzezińska A. (2000). *Spoleczna psychologia rozwoju*. Warszawa: Wydawnictwo Naukowe „Scholar”.
- Brzezińska A., Brzeziński J. (1975a). Konstrukcja i stosowanie skal szacunkowych do oceny zachowania się. *Psychologia Wychowawcza*, 18, 202-227.
- Brzezińska A., Brzeziński J. (1975b). *Skala Samoakceptacji SQ-75/02*. Poznań: Instytut Psychologii (maszynopis niepublikowany).
- Brzezińska A., Krzywicki P. (2001). Właściwości samowiedzy na przełomie późnej adolescencji i wczesnej dorosłości. *Czasopismo Psychologiczne*, 7 (1), 87-94.
- Brzezińska A., Stolarska M., Zielińska J. (2002). Poczucie jakości życia w okresie dorosłości. W: K. Appelt, J. Wojciechowska (red.). *Zadania i role społeczne w okresie dorosłości* (s. 103-126). Poznań: Wydawnictwo Fundacji Humaniora.
- Brzeziński J. (1996). *Metodologia badań psychologicznych*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Byrne D. (1966). *An introduction to personality*. New Jersey, NJ: Prentice Hall.
- Catell R.B. (1944). Psychological measurement: normative, ipsative, interactive. *Psychological Review*, 51, 291-303.
- Cohen J. (1957). An aid in the computation of correlations based on Q sorts. *Psychological Bulletin*, 54, 138-139.
- Cosgrove D.J. (1959). Diagnostic rating of teacher performance. *Journal of Educational Psychology*, 50, 200-204.
- Edwards A.L. (1957). *Techniques of attitude scale construction*. New York, NY: Appleton-Century-Crofts Inc.
- English H.B., English, A. C. (1958). *A comprehensive dictionary of psychological and psychoanalytical terms*. London: Longmans.
- Ferguson G.A. (1951). A note on the Kuder-Richardson formula. *Educational and Psychological Measurement*, 11, 612-615.
- Ferguson L.W. (1952). *Personality measurement*. New York, NY: McGraw-Hill.
- Filipiak E. (2002). *Konteksty rozwoju aktywności językowej dzieci w wieku wczesnoszkolnym*. Bydgoszcz: Wydawnictwo Akademii Bydgoskiej im. Kazimierza Wielkiego.
- Goode W.J., Hatt P.K. (1965). Skale szacunkowe. W: S. Nowak (red.), *Metody badań socjologicznych* (s. 284-290). Warszawa: Państwowe Wydawnictwo Naukowe.

- Greń J. (1975). *Statystyka matematyczna. Modele i zadania*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Guilford J.P. (1954). *Psychometric methods*. New York, NY: McGraw-Hill.
- Havighurst R.J. (1981). *Developmental tasks and education*. New York, NY: Longman.
- Hetmańska-Bugajska K., Brzezińska A. (2002). Temperament jako czynnik ryzyka zaburzeń zachowania w okresie adolescencji. W: A. Brzezińska, M. Bardziejewska, B. Ziółkowska (red.), *Zagrożenia rozwoju w okresie dorastania* (s. 11-46). Poznań: Wydawnictwo Fundacji Humaniora.
- Hornowska E. (2001). *Testy psychologiczne. Teoria i praktyka*. Warszawa: Wydawnictwo Naukowe „Scholar”.
- Horrocks J.E., Schoonover T.J. (1968). *Measurement for teachers*. Columbus, OH: Ch.E. Merrill Publishing Company.
- Kamphaus R.W., Petoskey M.D., Rowe E.W. (2000). Current trends in psychological testing of children. *Professional Psychology: Research and Practice*, 31 (2), 155-164.
- Kerlinger F.N. (1986). *Foundations of behavioral reserach*. Fort Worth: Harcourt Brace Jovanovich College Publishers.
- Komorita S.S., Graham W.K. (1965). Number of scale points and the reliability of scales. *Educational and Psychological Measurement*, 25, 987-995.
- Krzywicki P. (2002). *Właściwości struktury Ja u młodzieży z upośledzeniem umysłowym w stopniu lekkim*. Poznań: Instytut Psychologii Uniwersytetu im. Adama Mickiewicza (niepublikowany maszynopis rozprawy doktorskiej).
- Krzywicki P. (2003). Właściwości samowiedzy młodzieży z upośledzeniem umysłowym w stopniu lekkim oraz młodzieży nieupośledzonej. *Roczniki Psychologiczne*, 6, 91-114.
- Lazarus R., Folkman S. (1984). *Stress appraisal and coping*. New York: Springer-Verlag.
- Leeds C.H., Furman U. (1969). Predictive validity of Minnesota Teacher Attitude Inventory. *Journal of Teacher Education*, 20 (1), 51-56.
- Manstead A.S.R., Hewstone M., Fiske S.T., Hogg M.A., Reis H.T., Semin G.R. (2001). *Encyklopedia „Blackwella”. Psychologia społeczna*. Warszawa: Jacek Santorski & Co Wydawnictwo.
- Marek T. (1989). *Analiza skupień w badaniach empirycznych*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Medley D.M., Mitzel H.E. (1967). Measuring classroom behavior by systematic observation. W: N.L. Gage (red.), *Handbook of research on teaching* (s. 247-328). Chicago: Rand McNally.
- Murphy K.R., Constans J.I. (1987). Behavioral anchors as a source of bias in rating. *Journal of Applied Psychology*, 72, 573-579.
- Murphy K.R., Pardaffy V.A. (1989). Bias in behaviorally anchored rating scales. *Journal of Applied Psychology*, 74, 343-346.

- Nederhof A.J. (1991). Metody korygowania zniekształcającego wpływu zmiennej aprobaty społecznej. W: J. Brzeziński; J. Siuta (red.). *Społeczny kontekst badań psychologicznych i pedagogicznych* (s. 211-233). Poznań: Wydawnictwo Naukowe UAM.
- Osgood C.E., Suci G.J., Tannenbaum P.H. (1957). *The measurement of meaning*. Urbana, Ill: University of Illinois Press.
- Paszkievicz E. (1974). Technika Q-sort. W: E. Paszkiewicz (red.), *Materiały pomocnicze do ćwiczeń z metod badania osobowości (dla I i III roku studiów)* (s. 103-124). Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Peabody D. (1987). Selecting representative trait adjectives. *Journal of Personality and Social Psychology*, 52 (1), 59-71.
- Peterson C., Grant M. (2001). Forced-choice: are forensic interviewers asking the right questions? *Canadian Journal of Behavioural Sciences*, 33 (2), 118-127.
- Płużek Z. (1971). *Wartość testu WISKAD-MMPI dla diagnozy różnicowej w zakresie nozologii psychiatrycznej*. Lublin: Wydawnictwo Katolickiego Uniwersytetu Lubelskiego.
- Przetacznik-Gierowska M., Włodarski Z. (1994). *Psychologia wychowawcza* (tom 2). Warszawa: Państwowe Wydawnictwo Naukowe.
- Pytka L. (2000). *Pedagogika resocjalizacyjna. Wybrane zagadnienia teoretyczne, diagnostyczne i metodyczne*. Warszawa: Wydawnictwo Akademii Pedagogiki Specjalnej.
- Ravens-Sieberer U. (1998). *Mierzenie związanej ze zdrowiem jakości życia dzieci i młodzieży: Kwestionariusz KINDL. Podręcznik*. Hamburg: Universitäts-Krankenhaus Eppendorf Medizinische Klinik, Abteilung für Medizinische Psychologie — Forschungsgruppen (maszynopis niepublikowany; przekład za zgodą autorki z języka niemieckiego; Marta Stolarska, opracowanie za zgodą autorki: Anna Brzezińska).
- Reber A.S. (2000). *Słownik psychologii*. Warszawa: Wydawnictwo Naukowe „Scholar”.
- Rembowski J. (1972). *Więzi uczuciowe w rodzinie*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Remmers H.H. (1967). *Rating methods in research on teaching*. W: N.L. Gage (red.). *Handbook of research on teaching* (s. 329-378). Chicago: Rand McNally.
- Reykowski J. (1970). Obraz własnej osoby jako mechanizm regulujący postępowanie. *Kwartalnik Pedagogiczny*, 15 (3), 45-57.
- Robeck M.C., Wilson J.A.R (1974). *Psychology of reading: foundations of instruction*. New York, NY: John Wiley.
- Ryans D.G. (1960). *Characteristics of teachers*. Washington DC: American Council of Education.
- Sanocki W. (1970). Postawa wobec badania a wyniki inwentarza osobowości. *Przegląd Penitencjarny*, 27, 49-72.

- Smirnowa H. (2003). Poziom realizacji zadań rozwojowych w okresie późnego dzieciństwa. W: A. Brzezińska, S. Jabłoński, M. Marchow (red.), *Niewidzialne źródła. Szanse rozwoju w okresie dzieciństwa* (s. 191-211). Poznań: Wydawnictwo Fundacji Humaniora.
- Stephenson W. (1953). *The study of behavior. Q-technique and its methodology*. Chicago: The University of Chicago Press.
- Strelau J. (1982). *O temperamentie i jego poznaniu*. Warszawa: Nasza Księgarnia.
- Strelau J. (1998). *Psychologia temperamentu*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Thomas A., Chess S. (1977). *Temperament and development*. New York, NY: Brunner and Mazel.
- Wojciszke B. (2000). Psychologia oceniania: mechanizmy, pułapki, środki zaradcze. W: A. Brzezińska, J. Brzeziński (red.), *Ewaluacja procesu kształcenia w szkole wyższej* (s. 211-237). Poznań: Wydawnictwo Fundacji Humaniora.
- Wojciszke B. (2003). *Człowiek wśród ludzi. Zarys psychologii społecznej*. Warszawa: Wydawnictwo Naukowe „Scholar”.
- Zakrzewska M. (1994). *Analiza czynnikowa w budowaniu i sprawdzaniu modeli psychologicznych*. Poznań: Wydawnictwo Naukowe UAM.
- Ziółkowska B. (2001). *Ekspresja syndromu gotowości anorektycznej u dziewcząt w stadium adolescencji*. Poznań: Wydawnictwo Fundacji Humaniora.
- Ziółkowska B. (2003). Specyfika stadium adolescencji jako kontekst ekspresji *anorexia nervosa*. W: A. Brzezińska, M. Bardziejewska, B. Ziółkowska (red.), *Zagrożenia rozwoju w okresie dorastania* (s. 153-171). Poznań: Wydawnictwo Fundacji Humaniora.