



ISiD
ISiD

Zeszyty naukowe

Working papers

**Operacjonalizacja i skalowanie
w ilościowych badaniach społecznych**

Dorota Węziak-Białowolska

Zeszyty naukowe
Instytut Statystyki i Demografii SGH

Nr 16, rok 2011

Streszczenie

Niniejszy tekst jest efektem prac badawczych mających na celu zebranie i uporządkowanie pojęć i metod związanych z początkowymi etapami przygotowania naukowego badania ilościowego. Jednocześnie stanowi punkt wyjścia do kolejnych zadań badawczych, które obejmą zebranie i uporządkowanie metod (1) zbierania danych niezbędnych do realizacji założonych celów badawczych, (2) ich odpowiedniego przygotowania oraz (3) analizy.

Tekst został w całości poświęcony zagadnieniom związanym z operacjonalizacją badanych pojęć oraz metodom ich skalowania. Ze względu na charakter opracowania (część większej pracy badawczej) tematyka została celowo zawężona do tych obszarów. Tekst został przygotowany w ten sposób, aby nie tylko prezentować podstawy teoretyczne przedstawionych zagadnień, ale również ilustrować je skrótowymi przykładami zastosowań w badaniach naukowych wraz ze wskazaniem, gdzie można znaleźć pełen opis. Ponadto w sposób zamierzony nie prezentowano szczegółowo zagadnień metodologicznych omawianych w wielu podręcznikowych pozycjach polskojęzycznych.

Tekst składa się z dwóch głównych części: operacjonalizacja oraz pomiar i skalowanie. W części poświęconej operacjonalizacji pokrótce przedstawiono etapy projektowania badania naukowego ze szczególnym naciskiem na etap budowy modelu koncepcyjnego oraz jego operacjonalizację.

W części poświęconej zagadnieniom pomiaru i skalowania najpierw omówiono zagadnienia związane z definicjami pomiaru i skalowania, a następnie przedstawiono wybrane metody skalowania.

Spis treści

1. Wprowadzenie.....	3
2. Operacjonalizacja.....	4
3. Skalowanie i pomiar.....	17
3.1. Skalowanie cech obserwowalnych	25
3.2. Skalowanie cech ukrytych	25
Literatura	47

1. Wprowadzenie

Niniejsze opracowanie jest efektem prac badawczych mających na celu zebranie i uporządkowanie pojęć i metod związanych z początkowymi etapami przygotowania naukowego badania ilościowego. Jednocześnie stanowi punkt wyjścia do kolejnych zadań badawczych, które obejmą zebranie i uporządkowanie metod (1) zbierania danych niezbędnych do realizacji założonych celów badawczych, (2) ich odpowiedniego przygotowania oraz (3) analizy.

Opracowanie zostało w całości poświęcone zagadnieniom związanym z operacjonalizacją badanych pojęć oraz metodom ich skalowania. Ze względu na charakter opracowania (część większej pracy badawczej) tematyka została celowo zawężona do tych obszarów. Tekst został przygotowany w ten sposób, aby nie tylko prezentować podstawy teoretyczne przedstawionych zagadnień, ale również ilustrować je skrótowymi przykładami zastosowań w badaniach naukowych wraz ze wskazaniem, gdzie można znaleźć pełen opis. Ponadto w sposób zamierzony nie prezentowano szczegółowo zagadnień metodologicznych omawianych w wielu podręcznikowych pozycjach polskojęzycznych.

Tekst składa się z dwóch głównych części: operacjonalizacja oraz pomiar i skalowanie. W części poświęconej operacjonalizacji pokrótce przedstawiono etapy projektowania badania naukowego ze szczególnym naciskiem na etap budowy modelu konceptualnego oraz jego operacjonalizację.

W części poświęconej zagadnieniom pomiaru i skalowania najpierw omówiono zagadnienia związane z definicjami pomiaru i skalowania, a następnie przedstawiono wybrane metody skalowania. Metody przedstawiono przyjmując podział na skalowanie cech obserwowalnych bezpośrednio oraz skalowanie cech latentnych, przy czym skoncentrowano się głównie na skalowaniu cech latentnych. Wynikało to z faktu, że skalowanie cech obserwowalnych związane jest ze stosowaniem zwykle odpowiednio sformułowanych pojedynczych pytań, do których formułuje się odpowiednią skalę odpowiedzi. Na temat możliwych do zastosowania w takim przypadku skal odpowiedzi (nominalna – dwu- bądź więcej kategoryjna, porządkowa różnego typu, np. Likerta, Stapela, dyferencjału semantycznego, przedziałowa i ilorazowa) w polskojęzycznej literaturze napisano już bardzo dużo. Z tego względu zagadnienie to nie było rozwijane.

W ramach metod skalowania cech ukrytych wyróżniono dwa podejścia: (1) oparte na czynnikach i (2) czynnikowe. Skalowanie oparte na czynnikach przedstawiono w kontekście

budowy skal sumarycznych, natomiast przy skalowaniu czynnikowym największy nacisk położono na zaakcentowanie możliwości tworzenia skal wielowymiarowych. Ponadto w ramach drugiego z podejść pokrótce scharakteryzowano podstawy teoretyczne, a następnie zaprezentowano przykłady zastosowań w dwóch obszarach: eksploracyjnym oraz confirmacyjnym.

2. Operacjonalizacja

Każdemu badaniu naukowemu¹, już na samym jego początku, powinien zostać postawiony cel. Taki cel powinien zostać wyraźnie sformułowany. Może to być jeden cel, bądź też celów takich może być kilka. W przypadku sformułowania kilku celów mogą to być cele równoważne, bądź też badacz wskazać może cel główny oraz cele szczegółowe, których realizacja składa się na cel główny.

W sytuacji gdy badane zagadnienie jest dosyć nowe (np. brak jest na jego temat literatury), zaś badanie służy wyjaśnieniu jego istoty oraz związków między nim a światem zewnętrznym, celem badania może być określenie charakteru zjawiska, jego struktury oraz metody jego pomiaru. W takiej sytuacji cel badania sformułować można tak, jak zrobiła to Węziak-Białowolska: „za cel badania przyjęto zdefiniowanie modelu konceptualnego kapitału intelektualnego regionu, zaproponowanie jego modelu pomiarowego oraz w kontekście zaproponowanych rozwiązań wykazanie, że może on stanowić dodatkowy wymiar w analizie porównawczej regionów Polski”²

Tak sformułowany cel jest celem głównym. W tym konkretnym przypadku jego realizacja obejmowała realizację szeregu celów cząstkowych. Ich sformułowanie oraz związki z celem głównym przedstawiono w tabeli 1.

¹ Postulat ten dotyczy również badań nienaukowych – biznesowych

² Węziak-Białowolska D., 2008, Model kapitału intelektualnego regionu – koncepcja pomiaru i jej zastosowanie w analizie porównawczej regionów, rozprawa doktorska, Kolegium Analiz Ekonomicznych, Szkoła Główna Handlowa, Warszawa, s. 7

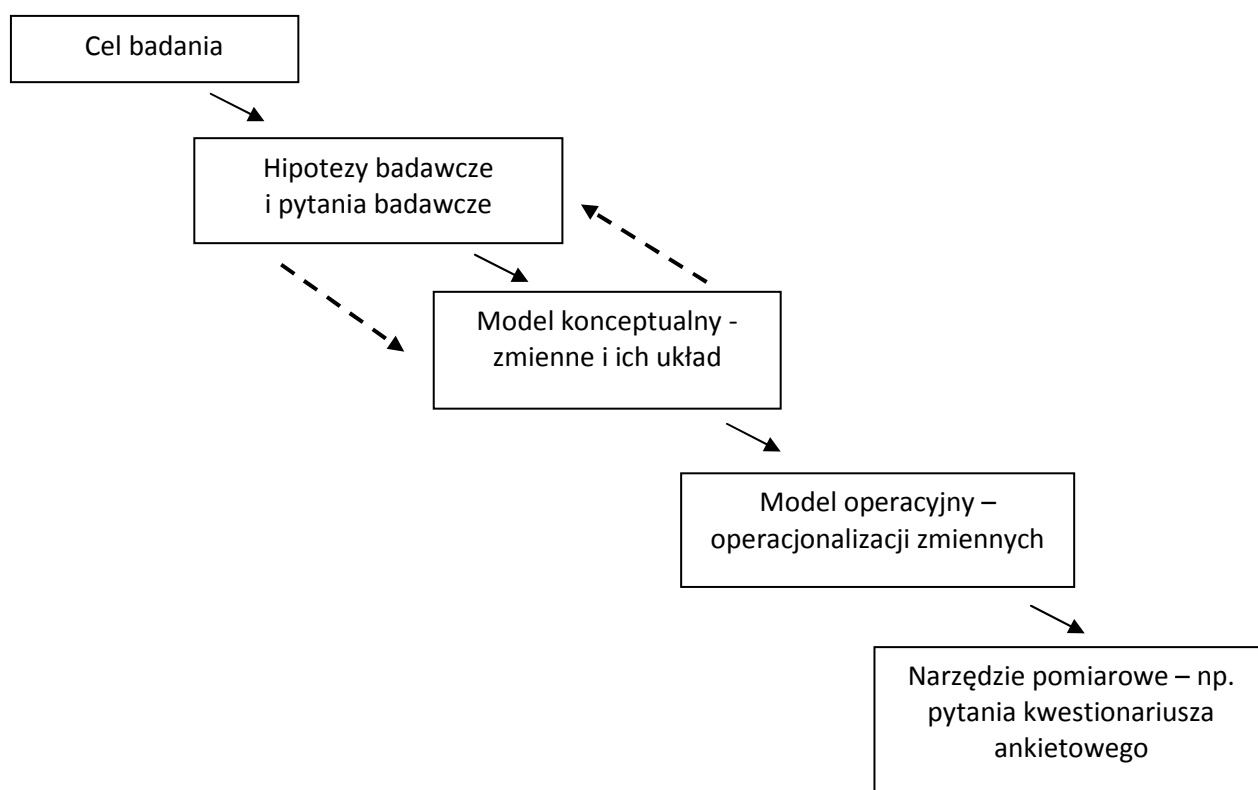
Tabela 1. Przykładowe cele szczegółowe i ich związek z celem głównym

Cel szczegółowy	Element celu głównego
1. Uporządkowanie pojęć, definicji i klasyfikacji związanych z koncepcją kapitału intelektualnego oraz w konsekwencji przedstawienie definicji kapitału intelektualnego regionu.	„zdefiniowanie modelu konceptualnego kapitału intelektualnego regionu”
2. Sformułowanie modelu konceptualnego kapitału intelektualnego regionu.	
3. Operacjonalizacja modelu konceptualnego kapitału intelektualnego regionu.	„zapropozowanie jego modelu pomiarowego”
4. Pomiar poszczególnych elementów kapitału intelektualnego zgodnie z modelem konceptualnym oraz ocena ich trafności i rzetelności.	
5. Weryfikacja związków założonych w hipotezach badawczych	„wykazanie, że może on stanowić dodatkowy wymiar w analizie porównawczej regionów Polski”
6. Wykazanie, że wskaźnik kapitału intelektualnego może służyć do oceny kondycji gospodarki i poziomu rozwoju gospodarczego.	

Źródło: opracowanie własne na podstawie Węziak-Białowolska D., *Model kapitału intelektualnego regionu. Koncepcja pomiaru i jej zastosowanie*, Oficyna Wydawnicza SGH, Warszawa 2010

Sformułowanie celów cząstkowych wyznacza niejako kolejne etapy postępowania badawczego mającego na celu zaprojektowanie badania naukowego (rysunek 1). Postępowanie to obejmuje oprócz ustalenia celu bądź celów badawczych, sformułowanie hipotez badawczych oraz pytań badawczych, przedstawienie modelu konceptualnego, a następnie modelu operacyjnego, które to prowadzić powinny do zaprojektowania narzędzia pomiarowego (np. kwestionariusza ankietowego lub eksperymentu).

Rysunek 1. Etapy projektowania badania naukowego.



Źródło: opracowanie własne

Umieszczenie na powyższym rysunku etapów *hipotezy badawcze i pytania badawcze* oraz *model konceptualny* nie jest stałe. To od badacza, celu badania i jego specyfiki zależy, czy najpierw sformułuje się model konceptualny a potem hipotezy i pytania badawcze, czy też odwrotnie³. Po określeniu celów badawczych konieczne jest sformułowanie hipotez badawczych, a więc pewnych przypuszczeń dotyczących badanych zagadnień, które będą weryfikowane w procesie badawczym. Przypuszczenia te powinny zostać sformułowane w języku dyscypliny, której dotyczy badanie. Warto podkreślić, że hipotezy badawcze mogą zostać potwierdzone lub odrzucone w toku procesu badawczego. Oba te rezultaty są równie ważne i dostarczają wiedzy o badanym zagadnieniu. Przykładowo w odniesieniu do celów zestawionych w tabeli 1 sformułowano następujące hipotezy badawcze⁴:

³ Częstokroć to model konceptualny poprzedza etap hipotez i pytań badawczych, ponieważ hipotezy stanowią słowny opis wszystkich zależności (ilustrowanych w postaci strzałek) występujących w modelu konceptualnym.

⁴ Węziak-Białowolska D., Model kapitału intelektualnego regionu. Koncepcja pomiaru i jej zastosowanie, Oficyna Wydawnicza SGH, Warszawa 2010, s. 16

- H1. Zmienne wskaźnikowe nie pełnią równoważnej roli w ujawnianiu kapitału intelektualnego.
- H2. Między składnikami kapitału intelektualnego regionu, tj. kapitałem ludzkim, kapitałem społecznym, kapitałem strukturalnym oraz kapitałem rozwoju, występują związki korelacyjne o dodatnim kierunku.
- H3. Kapitał intelektualny dobrze identyfikuje stan gospodarki i poziom rozwoju gospodarczego regionu obok innych zmiennych ekonomicznych.

Hipoteza H1 odnosi się do wskaźników kapitału intelektualnego, które na podstawie studiów literatury zostały wyselekcjonowane do mierzenia kapitału intelektualnego. Stanowi zatem przypuszczenie co do poprawności modelu operacyjnego i metody pomiaru. Hipoteza H2 odnosi się do związków występujących między kapitałem intelektualnym i jego składnikami. Jest to zatem przypuszczenie co do poprawności modelu konceptualnego oraz związków występujących między elementami tego modelu. Hipoteza H3 dotyczy związków występujących między wskaźnikiem kapitału intelektualnego i podstawowymi wskaźnikami makroekonomicznymi opisującymi kondycję gospodarki i poziom rozwoju gospodarczego. Jej weryfikacja pozwala zatem na sprawdzenie, czy związki występują, a zatem prowadzi do realizacji jednego z celów badawczych.

Obok określenia hipotez badawczych badacz może sformułować również pytania badawcze. Mówi się czasem, że hipotezy badawcze są proponowanymi przez badacza odpowiedziami, jakich można udzielić na pytania badawcze. Ich treść zatem powinna odpowiadać treści hipotez badawczych. Pytania badawcze powinny mieć formę zdań pytających, natomiast hipotezy badawcze wyraża się w formie zdań oznajmujących, czasem wyrażonych w formie przypuszczającej (*im ... , tym...; jeśli ... , to ...*).

Przykłady różnych sformułowań hipotez badawczych oraz odpowiadających im pytań badawczych zestawiono w tabeli 2.

Tabela 2. Przykładowe sformułowania hipotez badawczych i pytań badawczych.

Przykład hipotezy badawczej	Przykład pytania badawczego
H2. Zmienne wskaźnikowe nie pełnią równoważnej roli w ujawnianiu kapitału intelektualnego.	P1. Czy zmienne wskaźnikowe pełnią równoważną rolę w ujawnianiu kapitału intelektualnego?
H2. Między składnikami kapitału intelektualnego regionu, tj. kapitałem ludzkim, kapitałem społecznym, kapitałem strukturalnym oraz kapitałem rozwoju, występują związki korelacyjne o dodatnim kierunku.	P2. Czy między składnikami kapitału intelektualnego regionu, tj. kapitałem ludzkim, kapitałem społecznym, kapitałem strukturalnym oraz kapitałem rozwoju, występują związki korelacyjne? P3. Czy związki występujące między składnikami kapitału intelektualnego regionu, tj. kapitałem ludzkim, kapitałem społecznym, kapitałem strukturalnym oraz kapitałem rozwoju, mają kierunek dodatni? P4. Jaki jest kierunek związku między składnikami kapitału intelektualnego regionu, tj. kapitałem ludzkim, kapitałem społecznym, kapitałem strukturalnym oraz kapitałem rozwoju?
H3. Kapitał intelektualny dobrze identyfikuje stan gospodarki i poziom rozwoju gospodarczego regionu obok innych zmiennych ekonomicznych.	P5. Czy występuje związek między wskaźnikiem kapitału intelektualnego z wskaźnikami stanu gospodarki i poziomu rozwoju gospodarczego regionu? P6. Jaki jest kierunek związku między wskaźnikiem kapitału intelektualnego z wskaźnikami stanu gospodarki i poziomu rozwoju gospodarczego regionu?
H4. Im wyższy jest poziom rozwoju gospodarczego, tym wyższy jest poziom kapitału intelektualnego	P7. Czy związek między wskaźnikiem kapitału intelektualnego, a wskaźnikami stanu gospodarki i poziomu rozwoju gospodarczego regionu ma kierunek dodatni?
H5. Lepsza samoocena zdrowia współwystępuje z pozytywnymi zachowaniami zdrowotnymi	P8. Czy występuje dodatni związek między lepszą samooceną zdrowia i pozytywnymi zachowaniami zdrowotnymi?
H6. Częstsze używanie substancji psychoaktywnych współwystępuje z negatywnymi zachowaniami zdrowotnymi	P9. Czy występuje ujemny związek samooceną zdrowia i częstszym używaniem substancji psychoaktywnych?

Źródło: opracowanie własne na podstawie Węziak-Białowolska D., *Model kapitału intelektualnego regionu. Koncepcja pomiaru i jej zastosowanie*, Oficyna Wydawnicza SGH, Warszawa 2010 (hipotezy H1 – H4) oraz Hildt-Ciupińska K., Węziak D., Woynarowska-Sołdan M., *Skala pozytywnych zachowań zdrowotnych dla młodzieży. Wstępna analiza psychometryczna*, „Medycyna Wieku Rozwojowego”, Nr 2 część I tom XII 2008, s.577-585. (hipotezy H5 – H7).

Model konceptualny

Sformułowanie hipotez i pytań badawczych zmusza badacza do zdefiniowania wszystkich elementów badanego zjawiska oraz do zastanowienia się nad potencjalnymi związkami, które mogą między nimi wystąpić. Ponadto każe również zastanowić się, jakie związki występują między badanym zjawiskiem (często z uwzględnieniem jego poszczególnych elementów) a światem zewnętrznym. Przedstawienie tych wszystkich elementów (pojęć i zjawisk związanych z badanym zjawiskiem, jak również ze światem zewnętrznym względem niego) oraz występujących między nimi związków określa się mianem budowy modelu konceptualnego. Model konceptualny jest modelem badawczym, według którego badacz analizować będzie badane zjawisko. Do jego sformułowania niezbędne jest zapoznanie się z literaturą na temat badanego zjawiska. Prześledzić należy: (1) teorię badanego zjawiska, (2) pojęcia, które są z nim związane oraz (3) ich klasyfikację, (4) związki, jakie między tymi pojęciami mogą wystąpić oraz (5) związki, jakie mogą wystąpić między badanym pojęciem a światem zewnętrznym. Nie wszystkie z tych elementów muszą wystąpić w modelu konceptualnym. Decyzję o tym, co powinno zostać uwzględnione, podejmuje badacz na podstawie analizy celów badawczych, które sobie wytyczył. Jeśli pewne elementy potencjalnego modelu konceptualnego (np. związki między wybranymi zjawiskami, czy procesami) wydają się ciekawe, ale nie przewidziano ich analizy w celach badawczych, nie powinno się ich uwzględniać w modelu konceptualnym. Ich analiza, choć może dostarczyć ciekawych informacji, nie będzie możliwa do wykorzystania. Oczywiście alternatywą jest zawsze rozszerzenie listy celów badawczych.

W przypadku gdy analizowany problem ma charakter nowatorski, model konceptualny może w całości opierać się na hipotezach badawczych sformułowanych przez badacza wyłącznie na podstawie wiedzy własnej i intuicji. Bez względu na to, jakie podstawy merytoryczne leżały u podstaw konstrukcji modelu konceptualnego, ważne jest, aby model ten ujmował problem w sposób całościowy i kompletny.

Często do przedstawienia modelu konceptualnego korzysta się z formy graficznej i buduje się schemat ilustrujący wszystkie elementy badanego zjawiska oraz wskazuje się na powiązania między nimi. Przy ilustrowaniu powiązań ważne jest nie tylko wskazanie tych, których występowania badacz się spodziewa, ale należy również pokazać, jaki jest kierunek tych związków⁵. Możliwe jest przedstawienie kilku wariantów takich powiązań. Podkreślić należy, że schemat taki powinien odzwierciedlać problem badawczy w sposób całościowy.

⁵ związki przyczynowo-skutkowe ilustruje się za pomocą strzałki z grotem skierowanym od przyczyny do skutku; natomiast związki korelacyjne (bez wskazania przyczyny i skutku) ilustruje się strzałkami o grotach na obu końcach;

Aby zilustrować mechanizm budowy modelu konceptualnego poniżej przedstawiono przykład takiego postępowania na przykładzie pomiaru kapitału społecznego na podstawie Diagnozy Społecznej (tabela 3).

Tabela 3. Przykład budowy modelu konceptualnego – kapitał społeczny na podstawie Diagnozy Społecznej.

Etap budowy modelu konceptualnego	Przykład
1. Definicja pojęcia na podstawie badań literatury	Kapitał społeczny zdefiniowano jako zespół norm społecznych oraz wspólnie podzielanych wartości i zwyczajów kształtujących świat relacji społecznych i ekonomicznych, a wyrażających się przede wszystkim w poziomie zaufania społecznego oraz zaufania do instytucji, a także skali formalnych i nieformalnych międzyludzkich sieci powiązań odpowiedniej jakości.
2. Klasyfikacja pojęć występujących w definicji na podstawie badań literatury	Kierując się pracami van Oorschota, Artsa i Gelissena [2006] wyróżniono trzy wymiary kapitału społecznego: zaufanie, sieci oraz normy. W ramach dwóch z trzech wymiarów wyróżniono po dwa podwymiary. Były to: (1) Zaufanie: (a) do instytucji, (b) do innych osób; (2) Sieci: (a) aktywność obywatelska, (b) kontakty towarzyskie;
3. Określenie powiązań między pojęciami	<p>„Kapitał społeczny to ... wyrażających się przede wszystkim w poziomie zaufania społecznego oraz zaufania do instytucji, a także skali formalnych i nieformalnych międzyludzkich sieci powiązań odpowiedniej jakości.”</p> <p>Ze względu na nieobserwowalny bezpośrednio charakter kapitału społecznego oraz jego wymiarów przyjęto, że kapitał społeczny ujawnia się w swoich trzech wymiarach. Natomiast o istnieniu wymiarów kapitału społecznego świadczą ich symptomy.</p>
4. Model konceptualny – forma graficzna	<pre> graph LR KS(kapitał społeczny) --> N(normy) KS --> S(sieci) KS --> Z(zaufanie) S --> KT(kontakty towarzyskie) S --> AO(aktywność obywatelska) Z --> ZI(zaufanie do instytucji) Z --> ZSL(zaufanie w stosunku do ludzi) </pre>

Źródło: opracowanie własne na podstawie Węziak-Białowolska D., 2011, Kapitał społeczny w Polsce w świetle Diagnozy Społecznej – pomiar i wyniki, *Studia Demograficzne* (w druku).

Operacjonalizacja modelu konceptualnego

Naturalną konsekwencją sformułowania modelu konceptualnego jest jego operacjonalizacja. Operacjonalizacja to proces służący stworzeniu definicji operacyjnej zarówno dla pojęć, jak i związków między nimi. Jest to takie opisanie wszystkich elementów modelu konceptualnego (zarówno pojęć, jak i związków między nimi), aby:

- ich rozumienie nie tylko przez badawczą, ale i inne osoby było takie samo; oznacza to, że wszystkie abstrakcyjne pojęcia powinny zostać wyrażone jak najbardziej jednoznacznie oraz w formie zdarzeń, które można zmierzyć,
- umieć określić charakter zjawisk opisywanych przez pojęcia w modelu konceptualnym; oznacza to, że należy podjąć decyzję, czy dane zjawisko można obserwować i mierzyć wprost (np. płeć), czy też jego charakter jest bardziej złożony (np. wielowymiarowy i/lub nieobserwowalny bezpośrednio),
- umieć odpowiedzieć na pytanie: jak konkretnie mierzyć wszystkie elementy występujące, w modelu konceptualnym czyli:
 - w przypadku projektowania badania ankietowego – jakie pytanie kwestionariusza zastosować, czy wystarczy jedno pytanie, czy też powinno być ich kilka, jeśli kilka, to jak je ze sobą powiązać, czy też należy je analizować niezależnie, jakiego słownictwa w nim (nich) użyć, jakie formy odpowiedzi zastosować,
 - w przypadku korzystania z danych zastanych – które dane można wykorzystać do pomiaru określonego pojęcia, jakie zmienne/pytania, czy wystarczy jedno pytanie, czy też powinno być ich kilka, jeśli kilka, to jak je ze sobą powiązać, czy też należy je analizować niezależnie,
 - określić metodę pomiaru związków występujących między pojęciami występującymi w modelu konceptualnym (np. współczynnik korelacji liniowej Pearsona, współczynnik zbieżności V-Kramera itd.).

Niedokładne przeprowadzenie operacjonalizacji lub co gorsza pominięcie jej, uniemożliwia całkowicie wyciągnięcie wiarygodnych wniosków co do relacji zachodzących między mierzonymi zmiennymi, a także interpretację wyników analiz przeprowadzonych na danych. W celach ilustracyjnych w tabeli 4 przedstawiono operacjonalizację kapitału społecznego według modelu konceptualnego przedstawionego w tabeli 3.

Tabela 4. Operacjonalizacja kapitału społecznego na podstawie Diagnozy Społecznej.

Element procesu operacjonalizacji	Przykład
Definicja	Kapitał społeczny zdefiniowano jako zaufanie (do instytucji oraz do innych osób), sieci (aktywność obywatelska i kontakty towarzyskie) oraz normy społeczne. Zaufanie, sieci i normy określono jako wymiary kapitału społecznego.
Wybór danych/zmiennych	<p>Do każdego z wymiarów kapitału społecznego wyselekcjonowano odpowiednie wskaźniki⁶. Były to wybrane pytania kwestionariuszowe z badania Diagnoza Społeczna. Pytania te brzmiały następująco:</p> <p><u>Zaufanie:</u></p> <p>1. Zaufanie do instytucji (3 pytania): <i>Czy ma Pan/Pani zaufanie do:</i></p> <ul style="list-style-type: none"> – banków, – funduszy inwestycyjnych, – funduszy emerytalnych; <p><i>odpowiedzi: tak/nie/nie mam zdania;</i></p> <p>2. Zaufanie do innych ludzi (1 pytanie): <i>Ogólnie rzecz biorąc, czy uważa Pan/Pani, że można ufać większości ludzi, czy też sądzi Pan, że w postępowaniu z ludźmi ostrożności nigdy za wiele?</i> <i>odpowiedzi: większości ludzi można ufać/ostrożności nigdy za wiele/trudno powiedzieć;</i></p> <p><u>Sieci:</u></p> <p>1. Aktywność obywatelska (2 pytania)</p> <ul style="list-style-type: none"> – <i>Czy jest Pan/Pani członkiem jakichś organizacji, stowarzyszeń, partii, komitetów, rad, grup religijnych, związków lub kół?</i> <i>odpowiedzi: tak/nie;</i> – <i>Czy w ciągu ostatnich dwóch lat zdarzyło się, że angażował/a się Pan/Pani w działania na rzecz społeczności lokalnej (gminy, osiedla, miejscowości, w najbliższym sąsiedztwie)?</i> <i>odpowiedzi: tak/nie</i> <p>2. Kontakty towarzyskie (3 pytania):</p> <ul style="list-style-type: none"> – <i>Ile osób zalicza Pan/Pani do grona swoich przyjaciół?</i> – <i>W jakim stopniu jest Pan/Pani zadowolony/a z kontaktów z przyjaciółmi</i> <i>odpowiedzi: 6 – bardzo zadowolony, 5 – zadowolony, 4 – dosyć zadowolony, 3 – dosyć niezadowolony, 2 – niezadowolony, 1 – bardzo niezadowolony;</i> – <i>W jakim stopniu jest Pan/Pani zadowolony/a z kontaktów z rodziną?</i> <i>odpowiedzi: 6 – bardzo zadowolony, 5 – zadowolony, 4 – dosyć zadowolony, 3 – dosyć niezadowolony, 2 – niezadowolony, 1 – bardzo niezadowolony;</i> <p><u>Normy (6 pytań):</u></p>

⁶ Dla zachowania jasności wywodu opuszczono uzasadnienie takiego a nie innego wyboru pytań; niemniej jednak ważne jest, aby takie uzasadnienie zostało przez badacza podane; Uzasadnienia dla wyborów prezentowanych w przykładzie z tabeli 4 zainteresowany Czytelnik znajdzie w Węziak-Białowska D., 2011, *Kapitał społeczny w Polsce w świetle Diagnozy Społecznej – pomiar i wyniki*, *Studia Demograficzne* (w druku).

	<p><i>Jak bardzo Pana/Panią obchodzi, czy:</i></p> <ul style="list-style-type: none"> – ktoś płaci podatki mniejsze niż powinien, – ktoś unika płacenia za korzystanie z transportu publicznego, – komuś udaje się nie płacić za światło, – ktoś pobiera nieustannie zasiłek dla bezrobotnych, – ktoś nie płaci (choć może) czynszu za mieszkanie, – ktoś sprowadza towary z zagranicy i nie płaci cła; <p>odpowiedzi: 1 – w ogóle mnie nie obchodzi, 2 – mało mnie obchodzi, 3 – trochę mnie obchodzi, 4 – bardzo mnie obchodzi</p>
Model operacyjny – schemat	<pre> graph LR KS((kapitał społeczny)) --> N((normy)) KS --> S((sieci)) KS --> Z((zaufanie)) N --> N1[normy1] N --> N6[normy6] S --> O[organizacje] S --> Sp[społeczność] S --> L[l_przyjaciele] S --> Zp[z_przyjaciele] S --> Zr[z_rodzina] Z --> Zs(zaufanie w stosunku) Z --> Zi(zaufanie do instytucji) Zs --> Z1[zaufanie_1] Zi --> Zb[z_bank] Zi --> ZFI[z_FI] Zi --> ZFE[z_FE] </pre>
Związki między pojęciami – kierunek strzałek	<p>Ze względu na poniższe dwa zdania (z opisu modelu konceptualnego):</p> <ol style="list-style-type: none"> 1. Kapitał społeczny wyraża się w poziomie zaufania społecznego oraz zaufania do instytucji, a także skali formalnych i nieformalnych międzyludzkich sieciach powiązań oraz normach społecznych; 2. Kapitał społeczny ujawnia się w swoich trzech wymiarach. Natomiast o istnieniu wymiarów kapitału społecznego świadczą ich symptomy; <p>Kierunek strzałek na schemacie modelu operacyjnego ustalono w taki sposób, aby odzwierciedlały one znaczenie słów „wyraża się”, „ujawnia się” oraz „symptomy”.</p> <p>Oznacza to, że kapitał społeczny jest źródłem (przyczyną) pojawiania się zaufania społecznego oraz zaufania do instytucji, a także formalnych i nieformalnych międzyludzkich sieci powiązań oraz norm społecznych, a więc jest przyczyną pojawiania się swoich trzech wymiarów.</p> <p>Każdy z wymiarów kapitału społecznego ma swoje symptomy występowania, jest zatem źródłem zjawisk operacjonalizowanych przez odpowiednie symptomy.</p>
Metoda analizy związków między zmiennymi	Konfirmacyjna analiza czynnikowa

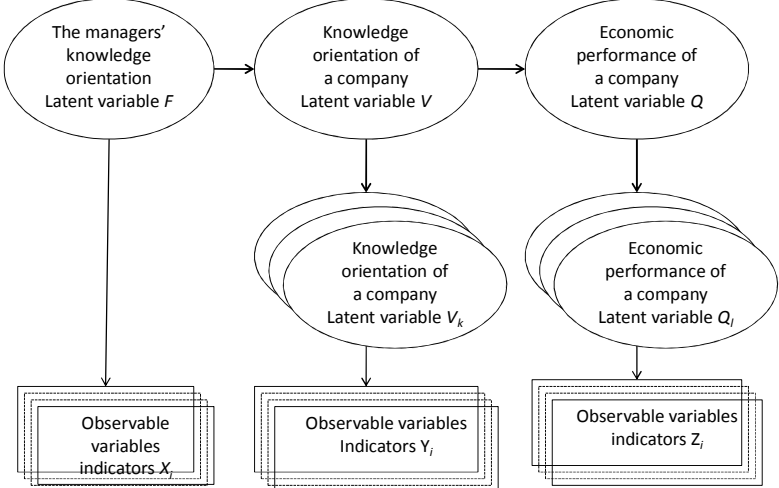
Źródło: opracowanie własne na podstawie Węziak-Białowolska D., 2011, *Kapitał społeczny w Polsce w świetle Diagnozy Społecznej – pomiar i wyniki*, *Studia Demograficzne* (w druku).

Przedstawiona w tabeli 4 operacjonalizacja kapitału społecznego została przeprowadzona w taki sposób, aby pomiar był możliwy do przeprowadzenia. Przeprowadzono ją po uprzednim wyborze źródła danych, stąd wybór zmiennych do mierzenia poszczególnych elementów modelu konceptualnego mógł być bardzo precyzyjny. Niemniej jednak operacjonalizację modelu konceptualnego można przeprowadzić w formie bardziej ogólnej, jeszcze przed przeprowadzeniem badania ankietowego. Przykład taki przedstawiono w tabeli 5.

Tabela 5. Operacjonalizacja modelu orientacji przedsiębiorstwa na wiedzę.

Etap procesu badawczego	Przykład
Definicja na podstawie badań literatury	Mazur, Rószkiewicz, Strzyżewska ⁷ zaproponowały model konceptualny wiążący wyniki ekonomiczne przedsiębiorstw ze stopniem zorientowania na wiedzę zarówno kadry menedżerskiej jak i przedsiębiorstwa rozumianego jako organizacji. Zaproponowany model teoretyczny zakłada występowanie dwóch elementarnych dla całej koncepcji kategorii, tj. postawy wobec wiedzy, będącej cechą kadry kierowniczej oraz orientacji na wiedzę przedsiębiorstwa, rozumianej jako właściwość prowadzonej działalności ekonomicznej.
Model konceptualny	<p>The diagram illustrates the conceptual model. On the left, a box labeled 'Orientacja na wiedzę kadry kierowniczej' (Orientation of management staff) has a solid arrow pointing to a larger box 'Orientacja przedsiębiorstwa na wiedzę' (Orientation of the company towards knowledge). This larger box contains three sub-sections: 'Orientacja przedsiębiorstwa na wiedzę w obszarze pozyskiwania i upowszechniania danych' (data acquisition and dissemination), 'Orientacja przedsiębiorstwa na wiedzę w obszarze tworzenia i upowszechniania wiedzy' (knowledge creation and dissemination), and 'Orientacja przedsiębiorstwa na wiedzę w obszarze wykorzystania i transferu wiedzy poza przedsiębiorstwo' (knowledge use and transfer outside the company). A solid arrow points from this large box to a box on the right labeled 'Wyniki ekonomiczne przedsiębiorstwa' (Economic results). Below the 'Management staff orientation' box is an oval labeled 'Wskaźniki orientacji na wiedzę kadry kierowniczej' (Indicators of management staff orientation). Below the 'Company orientation' box is a large oval labeled 'Wskaźniki działalności przedsiębiorstwa odnoszące się do pozyskiwania i upowszechniania danych, tworzenia i upowszechniania wiedzy oraz wykorzystywania i transferu wiedzy poza przedsiębiorstwo' (Indicators of company activity related to data acquisition, knowledge creation, and transfer). Below the 'Economic results' box is an oval labeled 'Wskaźniki wyników ekonomicznych' (Economic result indicators). Dotted arrows indicate the relationship between the indicators and the boxes they measure.</p>
Operacjonalizacja	Zarówno postawa kadry kierowniczej wobec wiedzy jak i orientacja przedsiębiorstwa na wiedzę są konstruktami intelektualnymi, czyli kategoriami bezpośrednio nieobserwowalnymi, które mogą podlegać pomiarowi tylko za pośrednictwem innych zmiennych, tzw. wskaźników postawy lub orientacji. Założenia przyjęte w zdefiniowanym modelu teoretycznego wskazują, że mamy do czynienia z jedną zmienną latentną egzogeniczną, opisującą postawę kadry kierowniczej (F) oraz wektorem powiązanych z nią zmiennych mierzalnych, które stanowią deskryptory kadry kierowniczej (X_i). Ponadto w modelu występują endogeniczne zmienne latentne, opisujące aktywność przedsiębiorstwa w obszarach

⁷ Mazur J., Rószkiewicz M., Strzyżewska M., 2005, Orientacja na wiedzę a wyniki ekonomiczne przedsiębiorstwa. Konceptualizacja badań, „Współczesne Zarządzanie”, nr 1/2005, p. 5-23.

	orientacji na wiedzę (V_k) (przyjęto występowanie trzech obszarów, czyli $k = 1, 2, 3$) oraz zmienna opisująca orientację na wiedzę przedsiębiorstwa w ogóle (V). Zmienne te są powiązane bezpośrednio lub pośrednio ze zmiennymi mierzalnymi, które dotyczą konkretnych działań w obszarach orientacji na wiedzę przedsiębiorstwa (Y_i). Ostatnią grupę zmiennych w zdefiniowanym modelu zależności tworzą zmienne mierzalne odnoszące się do wyników ekonomicznych przedsiębiorstwa (Z_p).
Schemat modelu operacyjnego	
Metoda analizy związku między pojęciami	Rodzaje zmiennych i zależności między nimi wskazują, że zdefiniowany model ma własności modelu strukturalnego

Źródło: Mazur J., Rószkiewicz M., Strzyżewska M., 2005 oraz M. Rószkiewicz, D. Węziak, Model orientacji przedsiębiorstwa na wiedzę – próba operacjonalizacji, Prace i Materiały Wydziału Zarządzania Uniwersytetu Gdańskiego, Nr 5(2007), Wydział Zarządzania Uniwersytetu Gdańskiego, Fundacja Rozwoju Uniwersytetu Gdańskiego, Sopot 2007, s. 287-302

Operacjonalizacja modelu orientacji przedsiębiorstwa na wiedzę Mazur, Rószkiewicz i Strzyżewskiej⁸ obejmowała następujące kroki. Najpierw na podstawie badań literatury opisana została koncepcja orientacji na wiedzę przedsiębiorstwa. Na tej podstawie sformułowano i przedstawiono w formie graficznej model konceptualny tej koncepcji. Obejmował on wszystkie pojęcia w niej występujące, jak również definiował występujące między nimi związki. Następnie przeprowadzono operacjonalizację tych pojęć. Przyjęto, że wszystkie mają charakter nieobserwowalny bezpośrednio. Należało zatem określić, jakie są symptomy ich występowania, czyli jakie zjawiska świadczą o tym, że dane pojęcie występuje. To następnie pozwoliło określić, za pomocą jakich zmiennych powinny być one mierzone. Ponadto okazało się, że orientacja na wiedzę przedsiębiorstwa jest pojęciem złożonym – trójwymiarowym. Składają się na nie: (1) orientacja przedsiębiorstwa na wiedzę w obszarze pozyskiwania i upowszechniania danych, (2) orientacja przedsiębiorstwa na wiedzę w obszarze tworzenia wiedzy i jej upowszechniania, (3) orientacja przedsiębiorstwa na wiedzę

⁸ Mazur J., Rószkiewicz M., Strzyżewska M., 2005, Orientacja na wiedzę a wyniki ekonomiczne przedsiębiorstwa. Konceptualizacja badań, „Współczesne Zarządzanie”, nr 1/2005, p. 5-23.

w obszarze wdrażania wiedzy i jej transferu poza przedsiębiorstwo. Należało zatem przeprowadzić również proces operacjonalizacji każdego z wymiarów orientacji przedsiębiorstwa na wiedzę. W stosunku do każdego z nich trzeba było określić, jakie są symptomy ich występowania. W ten sposób zdecydowano, że symptomy te zostaną zmierzone w specjalnie zaprojektowanym badaniu ankietowym, którego poszczególne części zawierać będą zestawy pytań bądź stwierdzeń opisujących każdy w wymiarów orientacji na wiedzę przedsiębiorstwa oraz orientację na wiedzę kadry kierowniczej.

Ponadto w modelu operacyjnym określono związki między pojęciami. Kierunek strzałek pokazuje, że założono (a później weryfikowano), że postawa kadry kierowniczej w stosunku do wiedzy (orientacja kadry kierowniczej na wiedzę) wpływa na orientację na wiedzę całego przedsiębiorstwa, zaś orientacja na wiedzę przedsiębiorstwa wpływa na jego wyniki ekonomiczne. Warto podkreślić, że zarówno w modelu konceptualnym, jak i w modelu operacyjnym wpływ postawy kadry kierowniczej na wyniki ekonomiczne przedsiębiorstwa jest jedynie pośredni – przez orientację na wiedzę przedsiębiorstwa. Choć autorki badania zaproponowały tylko jeden model konceptualny, możliwe było zaproponowanie konceptualizacji alternatywnej.

Przy okazji omawiania kwestii operacjonalizacji pojęć warto przedstawić obecną od niedawna w literaturze i praktyce koncepcję konceptu lub konstruktów postulowanych przez teorię (ang. *concept-by-postulation*) oraz konceptu lub konstruktów intuicyjnych (ang. *concept-by-intuition*)⁹. Koncepcja konceptu lub konstruktów intuicyjnych mówi, że konstrukt taki może być postrzegany przez ludzkie zmysły, zaś jego znaczenie jest natychmiast jednoznacznie rozumiane zarówno przez badacza, jak i respondenta. Natomiast w przypadku konstruktów postulowanych przez teorię rozumienie nie jest oczywiste i jednoznaczne. Określenie znaczenia konstruktów postulowanych przez teorię wymaga tej właśnie teorii. Na jej podstawie wnioskuje się o znaczeniu konstruktów i tworzy się ich definicję operacyjną. Innymi słowy, o ile konstrukty intuicyjne nie wymagają dodatkowych objaśnień, to konstrukty postulowane przez teorię nie będą właściwie i jednoznacznie zmierzone bez definicji operacyjnej.

Przykładowo, kapitał społeczny – zgodnie z konceptualizacją oraz operacjonalizacją przedstawioną powyżej – nie tylko jest bezpośrednio nieobserwowalny, ale również ma wielowymiarową strukturę oraz niejednoznaczny charakter. Kwantyfikuje się go zatem korzystając ze zmiennych obserwowalnych będących wskaźnikami jego występowania,

⁹ Saris, W. E., Gallhofer, I. N., 2007, *Design, Evaluation, and Analysis of Questionnaires for Survey Research*. Hoboken, NJ: Wiley

zgodnie z podejściem opartym na koncepcie postulowanym przez teorię. W przypadku definiowania kapitału społecznego lub jego wymiarów jako źródeł występowania ich odpowiednich wskaźników wykorzystuje się do tego narzędzia modelowania zmiennych ukrytych, w tym confirmacyjną analizę czynnikową, które pozwalają na przeprowadzenie kwantyfikacji zmiennych nieobserwowalnych bezpośrednio. W sytuacji gdy kapitał społeczny lub któryś z jego wymiarów jest jednak definiowany jako efekt obserwowanych wskaźników, to wtedy do pomiaru wykorzystuje się metody bazujące na wskaźnikach formatywnych¹⁰.

3. Skalowanie i pomiar

Waga pomiaru w naukach społecznych jest ogromna, ale często niedoceniana przez badaczy. Chociaż wiarygodny pomiar leży u podstaw uzyskania godnych zaufania wyników, to często zbyt mało uwagi poświęca się uzyskaniu narzędzia pomiarowego wysokiej jakości. Podstawowym warunkiem uzyskania dobrej jakości wyników jest wysoka trafność i rzetelność narzędzi pomiarowych (pytań kwestionariusza, skal odpowiedzi i in.) zastosowanych do zmierzenia odpowiednich zjawisk (postaw, zachowań etc.).

Zastosowaniem praktycznym miar i skal zajmuje się teoria pomiaru. Jej podstawowymi problemami są¹¹:

- a) zagadnienie istnienia reprezentacji,
- b) zagadnienie jednoznaczności tej reprezentacji,
- c) zagadnienie znaczenia (sensowności),
- d) problem skalowania.

Zagadnienie istnienia reprezentacji opiera się na odpowiedzi na pytanie, czy wszystkie cechy można zmierzyć bezpośrednio, a jeśli nie, to jakie warunki należy spełnić, aby zachować relacje między stanem empirycznym, a zastosowanym do jego operacjonalizacji systemem numerycznym (teoretycznym).

Zagadnienie jednoznaczności to określenie zakresu jednoznaczności danej procedury pomiarowej. Lub inaczej – określenie zbioru wszystkich możliwych izomorfizmów.

Zagadnienie znaczenia (sensowności) związane jest z odpowiedzią na pytanie, jakie wnioski można wyciągnąć z danych uzyskanych za pomocą określonej skali pomiarowej, czy i kiedy dane te znaczą cokolwiek.

¹⁰ Więcej na temat tych metod znaleźć można w Saris, W. E., Gallhofer, I. N., 2007, Design, Evaluation, and Analysis of Questionnaires for Survey Research. Hoboken, NJ: Wiley..

¹¹ C. H. Coombs, R. M. Dawes, A. Tversky, *Wprowadzenie do psychologii matematycznej*, PWN, Warszawa 1977, s. 28, a także *Pomiar statystyczny*, red. W. Ostasiewicz, , Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2003, s.17.

Problem skalowania, czyli jak konstruuje się skale pomiarowe¹², w jaki sposób tworzy się reguły przyporządkowania symboli właściwościom empirycznym, a także jak postępuje się z błędami pomiaru.

Czym natomiast jest pomiar? Zgodnie z definicją podaną przez Stevensa, pomiar to procedura przyporządkowania wartości liczbowych¹³ – cyfr lub innych symboli – obiektom lub zdarzeniom zgodnie z określonymi zasadami¹⁴ lub inaczej – za Tversky’em – to „postępowanie służące badaczowi do reprezentowania właściwości empirycznych za pomocą liczb”¹⁵. Rezultatem pomiaru powinna być skala pomiarowa.

Najbardziej istotnymi elementami obu definicji są słowa: procedura i postępowanie. To one decydują o jakości pomiaru poprzez poprawne, odpowiednie i przede wszystkim odzwierciedlające rzeczywistość powiązanie ze sobą właściwości empirycznych i symboli (cyfr). Innymi słowy, mierząc (dokonując pomiaru) należy zawsze powtarzać sobie pytanie, czy struktura wykorzystywanego systemu liczbowego odzwierciedla strukturę mierzonych zjawisk i pojęć, a więc czy sposób przyporządkowania jest izomorficzny względem badanych zjawisk¹⁶.

W sensie naukowym skala pomiarowa powinna być obiektywną abstrakcją posiadającą równe jednostki. Słowo „obiektywna” oznacza, że powinna być ona stała bez względu na dobór próby badawczej (celowy lub losowy), a także niezależna od zestawu pozycji/stwierdzeń wybranych do pomiaru określonej cechy. Innymi słowy skala pomiarowa powinna gwarantować, aby jej stosowanie za każdym razem było identyczne (niezależne od osoby przeprowadzającej badanie). Słowo „abstrakcja” odnosi się do liniowego kontinuum będącego idealną skalą pomiarową każdego mierzonego pojęcia i zjawiska. Jak twierdzi Thurstone, kontinuum to ma zawsze charakter abstrakcyjny, bo każdy pomiar związany jest z przetworzeniem mierzonego konstruktów w abstrakcyjną formę liniową.¹⁷

Poza tym, użyteczna skala pomiarowa nie powinna być skomplikowana. Powinna dostarczać danych wysokiej jakości, co jest konsekwencją jej wysokiej rzetelności i trafności.

¹² W literaturze polskojęzycznej zamiast określenia skala pomiarowa stosuje się również pojęcia miary lub testu. Pojęcie testu częściej występuje w psychologii i psychometrii, miara najczęściej kojarzona jest z naukami fizykalnymi, zaś skala pomiarowa to pojęcie charakterystyczne dla nauk społecznych, badań sondażowych i marketingowych.

¹³ Należy podkreślić, że liczby uzyskane poprzez przypisanie cyfr poszczególnym poziomom mierzonych cech nie posiadają własności addytywności i odtwarzalności. Są to dane surowe i nie wolno traktować ich jako miar. Niemniej jednak w praktyce często zdarza się, że traktuje się je jako miary. Przykładowo sztucznie wzmacnia się dane uzyskane na skali Likerta (skala porządkowa) w celu uzyskania danych typu interwałowego.

¹⁴ Ch. Frankfort-Nachmias, D. Nachmias, *Metody badawcze w naukach społecznych*, Zysk i S-ka Wydawnictwo, s. 170.

¹⁵ Coombs, Dawes, Tversky, op. cit., s. 25.

¹⁶ Ch. Frankfort-Nachmias, D. Nachmias, op. cit., s. 172.

¹⁷ Wright B. D., *Fundamental Measurement in Social Science and Education*, Research Memorandum No. 33a, MESA Psychometric Laboratory <http://www.rasch.org/memo33a.htm>, strona odwiedzona 15 VII 2004.

Dodatkowo Frankfort-Nachmias i Nachmias twierdzą, że skala pomiarowa powinna charakteryzować się jednowymiarowością, tzn. że zbiór pozycji składających się na skalę powinien odzwierciedlać tylko jeden wymiar. Innymi słowy wszystkie pozycje, z których składa się narzędzie pomiarowe, mogą być umieszczone na kontinuum odnoszącym się do jednego i tylko jednego pojęcia¹⁸.

Decyzję o tym, w jaki sposób wyskalować dane zjawisko lub pojęcie, badacz podejmuje na podstawie jego konceptualizacji i operacjonalizacji. To wtedy badacz decyduje, czy:

- badane zjawisko lub pojęcie ma charakter obserwowalny bezpośrednio, czy też nieobserwowalny bezpośrednio (ukryty, latentny),
- badane zjawisko lub pojęcie ma charakter prosty, czy też charakter złożony (wtedy mówi się, że składa się z wielu elementów składowych lub wymiarów).

W przypadku pomiaru cech obserwowalnych bezpośrednio zwykle do pomiaru stosuje się pojedyncze pytania lub skale proste. W przypadku zjawisk nieobserwowalnych bezpośrednio należy określić, jakie zjawiska, które można zaobserwować, są symptomami lub wskaźnikami występowania zjawisk nieobserwowalnych. Ponadto trzeba określić również strukturę badanego zjawiska – czy jest to zjawisko jednowymiarowe, czy też wielowymiarowe. Innymi słowy, trzeba określić, za pośrednictwem jakiego zestawu zmiennych wskaźnikowych, zwanych również wskaźnikami lub indykatorami, można mierzyć zjawisko latentne. Każda taka zmienna wskaźnikowa skalowana może być za pomocą pojedynczego pytania lub skali prostej lub też za pomocą jednej (w przypadku jednowymiarowych zjawisk nieobserwowalnych) lub kilku skal złożonych¹⁹ (w przypadku wielowymiarowych zjawisk nieobserwowalnych).

Jeśli zjawisko lub pojęcie latentne ma charakter prosty, również do jego pomiaru stosuje się pojedyncze pytania lub skale proste. Natomiast jeśli zjawisko latentne ma charakter złożony, czyli wielowymiarowy zastosowanie mają metody skalowania wielowymiarowego stosowane do zestawu kilku lub kilkunastu pytań bądź skal prostych.

Skale proste to pojedyncze pytania lub stwierdzenia o określonej skali odpowiedzi. Zwykle skala odpowiedzi może być: (1) nominalna, (2) porządkowa, (3) przedziałowa lub (4) ilorazowa, co odpowiada klasyfikacji skal zaproponowanej przez Stevensa²⁰. O skali

¹⁸ Ch. Frankfort-Nachmias, D. Nachmias, op. cit., s. 471.

¹⁹ Skala złożona to zestaw określonej liczby skal prostych lub zestaw kilku bądź nawet kilkunastu pytań lub stwierdzeń.

²⁰ Rószkiewicz M., *Metody ilościowe w badaniach marketingowych*, Wydawnictwo Naukowe PWN, Warszawa 2002; Sagan A., *Badania marketingowe. Podstawowe kierunki*. Wydawnictwo Akademii Ekonomicznej w

nominalnej mówi się, gdy pozwala ona wyłącznie na identyfikację kategorii odpowiedzi lub typów mierzonych obiektów. O skali porządkowej mówi się, gdy – oprócz możliwości identyfikacji kategorii odpowiedzi lub typów mierzonych obiektów – pozwala ona również na ustalenie ich porządku. Skala przedziałowa służy zarówno do identyfikacji kategorii odpowiedzi lub typów mierzonych obiektów, ustalania porządku rodzajów odpowiedzi, ale również porównywania ocenianych obiektów między sobą. Porównania te przeprowadzane są w formule absolutnej, tj. określić można, o ile dany obiekt jest lepszy lub gorszy od innego obiektu. Skala ilorazowa posiada natomiast (oprócz własności wszystkich skal wymienionych wcześniej) możliwość dokonywania porównań w formie względnej. Oznacza to, że dla obiektów zmierzonych za pomocą takiej skali można określić, ile razy każdy z nich jest lepszy lub gorszy od innego²¹.

W praktyce skale przedziałowa i ilorazowa traktowane są równoważnie, natomiast najczęściej wątpliwości budzi stosowanie skal porządkowych. Biorąc pod uwagę, że skale porządkowe są bardzo często stosowane (np. do pomiaru postaw) warto poświęcić im kilka słów. Mówiąc o skali porządkowej ma się myśleć o sposobie gradacji wariantów odpowiedzi na zadane pytanie lub przedstawione stwierdzenie, przy czym wariantów tych powinno być co najmniej trzy. Najpowszechniej stosowanymi skalami porządkowymi są skale Likerta, dyferencjału semantycznego i Stapela²². Skale te stosuje się nie tylko do pojedynczych pytań lub stwierdzeń, ale przede wszystkim w przypadku gdy pytań lub stwierdzeń w kwestionariuszu ankietowym jest kilka lub kilkanaście i dotyczą one tego samego zagadnienia. W takim przypadku mówi się nawet o skali Likerta w odniesieniu do zestawu (baterii) stwierdzeń, z których każde skalowane jest za pomocą odpowiedzi definiowanych właśnie na skali Likerta. Stosowanie tego typu określeń, choć używane w praktyce, może być mylące. Określenie „skala Likerta” występuje bowiem w dwojakim znaczeniu: (1) zestaw (bateria) stwierdzeń, (2) sposób gradacji odpowiedzi.

Skala porządkowa (jak również skala nominalna) określana jest mianem skali słabej bądź niemetrycznej, ponieważ na danych mierzonych na nominalnym poziomie pomiaru nie

Krakowie, Kraków 2004. Ostasiewicz W. Istota pomiaru statystycznego, [w:] Pomiar statystyczny, red. W. Ostasiewicz, Wydawnictwo Akademii Ekonomicznej im. Oskara Lanego we Wrocławiu, Wrocław 2003.

²¹ Więcej na temat klasyfikacji skal wg Stevensa znaleźć można w Rószkiewicz M., *Metody ilościowe w badaniach marketingowych*, Wydawnictwo Naukowe PWN, Warszawa 2002; Sagan A., *Badania marketingowe. Podstawowe kierunki*. Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004. Ostasiewicz W. Istota pomiaru statystycznego, [w:] Pomiar statystyczny, red. W. Ostasiewicz, Wydawnictwo Akademii Ekonomicznej im. Oskara Lanego we Wrocławiu, Wrocław 2003.

²² Zob. Rószkiewicz M., *Metody ilościowe w badaniach marketingowych*, Wydawnictwo Naukowe PWN, Warszawa 2002; Sagan A., *Badania marketingowe. Podstawowe kierunki*. Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004; Ch. Frankfort-Nachmias, D. Nachmias, *Metody badawcze w naukach społecznych*, Zysk i S-ka Wydawnictwo, Kaczmarczyk S., *Badania marketingowe. Metody i techniki*, PWE, Warszawa 2002.

powinno się wykonywać żadnych operacji matematycznych²³. Niemniej jednak w praktyce sumowanie i uśrednianie danych porządkowych stosuje się (np. gdy dane mierzone są na skali Likerta, wykorzystuje się je do budowy skali skumulowanych ocen). Warto jednak pamiętać, że liczby porządkowe wskazują jedynie na sposób uporządkowania obiektów, zaś odległości między nimi nie muszą być takie, na jakie wskazują zastosowane liczby. Liczby te, zastosowane jako skala porządkowa, są tylko symbolami, bądź kodami dla poszczególnych typów odpowiedzi. Odległości między nimi przedstawiają „skoki” między odpowiedziami w sensie psychologicznym. Kategorie odpowiedzi mogą w percepcji respondentów nie być od siebie oddzielone w sposób ostry. W przypadku gdy dwa stopnie skali są do siebie zbliżone w sensie psychologicznym, należy zadać sobie pytanie, czy jednego z nich nie należy wyeliminować. Natomiast gdy dwa stopnie skali są od siebie oddalone, należy rozważyć, czy nie należałoby jednego dodać. Faktem jest, że nie istnieje optymalna liczba stopni skali porządkowej, która byłaby poprawna w każdej sytuacji. Decyzja leży w gestii badacza i powinna być poparta badaniami empirycznymi.

W celach ilustracyjnych rozważono sytuację, gdy respondent proszony jest o udzielenie odpowiedzi na pytanie „*W jakim stopniu zadowolona/y jest Pani/Pan ze swoich stosunków z najbliższymi w rodzinie*”²⁴ i ma do wyboru jedną z następujących możliwości: 1–bardzo zadowolona/y, 2–zadowolona/y, 3–dosyć zadowolona/y, 4–dosyć niezadowolona/y, 5–niezadowolona/y, 6–bardzo niezadowolona/y. W takiej sytuacji pomiar jego zadowolenia ze stosunków z najbliższymi w rodzinie przeprowadzany jest na skali porządkowej, zaś udzielona odpowiedź w bazie danych zapisywana jest za pomocą określonego kodu, czyli w tym przypadku symbolu liczbowego. Niemniej jednak, gdy porówna się odpowiedzi udzielone przez dwie pary respondentów (A1 i A2 oraz B1 i B2), takie jak pokazano w tabeli 6, to nie można mieć żadnej pewności, że znaczeniowa różnica w ich odpowiedziach jest taka sama, choć wskazują na to liczby (różnica w obu przypadkach równa 2).

²³ Przykładowo nie powinno się takich danych sumować ani uśredniać. Z możliwych do przeprowadzenia na tego typu danych analiz statystycznych warto wymienić: budowę rozkładu odpowiedzi oraz tylko dla danych porządkowych – obliczenie miar pozycyjnych położenia, zróżnicowania i asymetrii. Więcej na temat możliwych do wykonania operacji matematycznych i statystycznych na danych z poszczególnych poziomów pomiaru znaleźć można w Rószkiewicz M., *Metody ilościowe w badaniach marketingowych*, Wydawnictwo Naukowe PWN, Warszawa 2002;

²⁴ Diagnoza Społeczna 2011 – kwestionariusz ankietowy dostępny na stronie <http://www.diagnoza.com/>

Tabela 6. Przykładowe odpowiedzi na pytanie: „W jakim stopniu zadowolona/y jest Pani/Pan ze swoich stosunków z najbliższymi w rodzinie”

Numer respondenta	Odpowiedź	Liczbowy kod odpowiedzi	Różnica między odpowiedziami respondentów A1 i A2	Różnica między odpowiedziami respondentów B1 i B2
A1	bardzo zadowolona/y	1	2	
A2	dosyć zadowolona/y	3		
B1	dosyć niezadowolona/y	3		2
B2	niezadowolona/y	5		

Źródło: opracowanie własne

Niemniej jednak warto wiedzieć, że w praktyce często skale porządkowo typu Likerta, dyferencjału semantycznego, czy Stapela o co najmniej pięciu kategoriach odpowiedzi (o co najmniej pięciu punktach) traktuje się jak skale ciągłe. Ponadto ze względu na fakt, że wzrost liczby punktów skali powoduje wzrost jej mocy dyskryminacyjnej, a więc możliwości różnicowania odpowiedzi, zwiększa się liczbę kategorii odpowiedzi. Trzeba jednak pamiętać, że ludzkie możliwości rozróżniania poszczególnych wariantów odpowiedzi są ograniczone. Ograniczenia te pojawiają się szybciej (czyli przy mniejszej liczbie punktów skali) wtedy, gdy nie wszystkie kategorie odpowiedzi zostały opisane słownie (etykietowane). Chociaż przyjmuje się najczęściej, że liczba punktów skali powinna zawierać się w przedziale od 3 do 9, a Stobiecka²⁵ dodaje, że „człowiek posługuje się najwyżej skalą 11 – punktową”, to już Sztabiński²⁶ podkreśla, że konieczność posługiwania się zbyt długą skalą (np. 9 – punktową) wywołuje u badanych tendencje do jej spłaszczania, czyli łączenia ze sobą poszczególnych punktów skali, a dopiero potem próbę ich interpretacji i udzielenie odpowiedzi.

Przykład zbyt długiej skali odpowiedzi przedstawiono poniżej oraz zilustrowano na wykresie 1. Wykres ten prezentuje, jak respondenci odpowiadali na zestaw dziewięciu pytań dotyczących stosunku do zabezpieczenia warunków życia na starość²⁷. Pytania sformułowano następująco: *Czy zabezpieczenie warunków na starość jest mniej ważne, tak samo ważne czy ważniejsze niż:*

²⁵ J. Stobiecka, *Porządkowy czy interwałowy charakter skal szacunkowych w naukach społecznych* [w:] *Analiza i prognozowanie zjawisk rynkowych o charakterze niemetrycznym*, red. Gatnar E., Wydawnictwo Akademii Ekonomicznej w Katowicach, Katowice 2003, s. 84

²⁶ F. Sztabiński, *Logika badacza i logika respondenta. Problem adekwatności narzędzia badawczego* [w:] *ASK Społeczeństwo Badania Metody* 12/2003, red. Domański H., Wydawnictwo Instytutu Filozofii i Socjologii PAN, Warszawa 2003, s. 165.

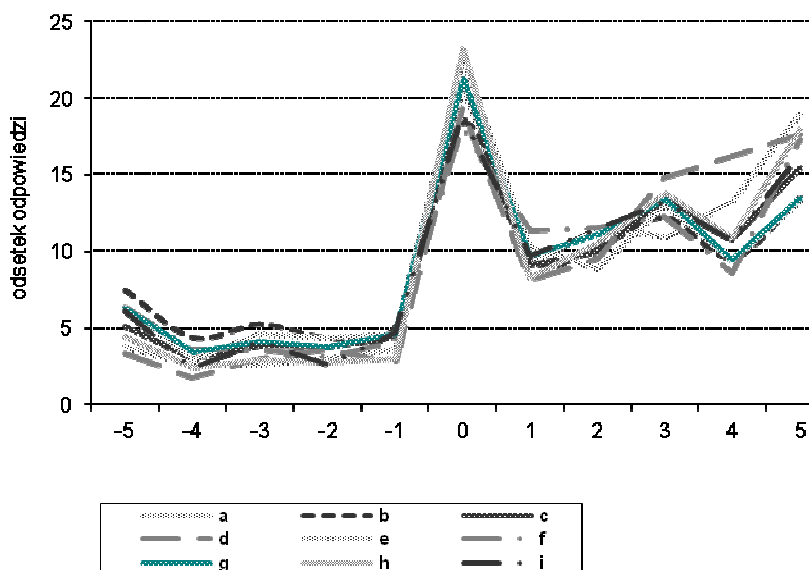
²⁷ Analiza została przeprowadzona na danych pochodzących z badania zrealizowanego przez CBOS w listopadzie 2001r. na losowej próbie 1635 Polaków. Rószkiewicz M., *Oszczędzanie. Postawy i zachowania polskich gospodarstw domowych wobec oszczędzania*, Aureus, Kraków 2008

- a. *Zdobywanie pieniędzy i osiągnięcie wyższych dochodów?*
- b. *Zdobywanie popularności?*
- c. *Zdobywanie szacunku?*
- d. *Urządzenie sobie obecnie wygodnych warunków życia?*
- e. *Zdobywanie awansu społecznego?*
- f. *Poprawienie wyglądu zewnętrznego?*
- g. *Zdobycie uznania wśród współpracowników i znajomych?*
- h. *Uzyskanie lub zachowanie dobrej kondycji fizycznej?*
- i. *Zapewnienie rozrywki i przyjemności?*

Respondenci proszeni byli o udzielenie odpowiedzi na 11-punktowej skali Stapela, przy czym etykietami opatrzone tylko trzy kategorie odpowiedzi (dwie skrajne i środkową) w następujący sposób: -5 – zdecydowanie mniej ważne, 0 – tak samo ważne, +5 – zdecydowanie ważniejsze. Niezdecydowanym respondentom pozostawiono możliwość wyboru kategorii „*trudno powiedzieć*”. Ponieważ odsetek odpowiedzi „*trudno powiedzieć*” na każde z analizowanych pytań był stosunkowo niski i zawierał się w przedziale 1,1% - 1,3%, takie odpowiedzi zostały w analizie potraktowane jako braki danych.

Przeprowadzona przez Węziak²⁸ wstępna analiza wyników badania pokazała, że w przypadku wszystkich pytań zadanych respondentom kształt rozkładu odpowiedzi jest zbliżony. Jego cechą charakterystyczną było występowanie kilku kategorii dominujących w punktach: -5, -3, 0, +3 oraz +5 z niewielkimi wyjątkami dla pozytywnej połówki skali w przypadku stwierdzeń *a* oraz *d*.

²⁸ Węziak D., *Wpływ długości skali odpowiedzi na wyniki pomiaru – zastosowanie wielorakiej analizy korespondencji*, Badanie statutowe Kolegium Analiz Ekonomicznych: Metody i techniki w analizie zachowań ekonomicznych gospodarstw domowych, Warszawa 2005

Wykres 1. Odsetki wskazań poszczególnych kategorii odpowiedzi²⁹.

Źródło: Węziak D., *Wpływ długości skali odpowiedzi na wyniki pomiaru – zastosowanie wielorakiej analizy korespondencji*, Badanie statutowe Kolegium Analiz Ekonomicznych: Metody i techniki w analizie zachowań ekonomicznych gospodarstw domowych, Warszawa 2005.

Jak pokazuje wykres 1, najczęściej wybieranymi kategoriami odpowiedzi były: 0 – *tak samo ważne*, następnie w kolejności pozytywny punkt ekstremalny +5 – *zdecydowanie ważniejsze*, kategoria +3 (wyjątek stanowią stwierdzenia a i d) oraz w negatywnej części skali punkt maksymalny -5 – *zdecydowanie mniej ważne*. Zauważyć można również, że zdecydowanie częściej wybierane były kategorie pozytywne (nawet te nie będące maksimumami lokalnymi), czyli odpowiadające większej wadze potrzeby zabezpieczenia starości nad innymi potrzebami. Pomimo stosunkowo rzadszych wskazań kategorii ujemnej części skali, również tutaj można zauważyć schemat odpowiedzi przypominający kształtem ten z dodatniej połówki skali, co przejawiało się częstszymi wskazaniami kategorii ekstremalnej, a także leżącej pośrodku tej części skali, czyli -3.

Na podstawie uzyskanych wyników Węziak³⁰ sformułowała następujące wnioski. Po pierwsze, kształt uzyskanych rozkładów sugerować może, że respondenci, chcąc udzielić odpowiedzi, dekomponowali skalę odpowiedzi na dwie części: od -5 do 0 i od 0 do +5. Po drugie, badani stosunkowo łatwo określali się w jednej z dwóch połówek, ale potem nie

²⁹ Poprawną formą graficzną prezentacji rozkładu odpowiedzi jest wykres słupkowy (nieciągły). Wynika to stąd, że skala Stapela jest skalą porządkową. W tym jednak przypadku celowo zastosowano wykres liniowy, aby pokazać, że może być ona traktowana jako skala ciągła.

³⁰ Węziak D., *Wpływ długości skali odpowiedzi na wyniki pomiaru – zastosowanie wielorakiej analizy korespondencji*, Badanie statutowe Kolegium Analiz Ekonomicznych: Metody i techniki w analizie zachowań ekonomicznych gospodarstw domowych, Warszawa 2005

zawsze byli w stanie różnicować odpowiedzi. Po trzecie, bardzo rzadko w porównaniu do innych kategorii w odpowiedniej połówce skali odpowiedzi wskazywane były punkty odpowiadające słabej preferencji: +1 – w pozytywnej połówce – oraz -1 i -2 dla negatywnej połówki. Stosunkowo rzadsze wybieranie niektórych kategorii odpowiedzi nasunęło podejrzenie, że być może skala odpowiedzi była zbyt trudna dla respondentów, a uzyskane kształty rozkładów były nie tylko wynikiem treści pytań, ale również efektem długości skali. Wysunięta została hipoteza, że o ile respondent jest w stanie wypowiedzieć się na temat znaczenia potrzeby zabezpieczania własnej starości względem innych wyróżnionych w badaniu potrzeb, to już z określeniem siły preferencji ma większy problem.

Podsumowując, skalowanie dotyczy ustalenia skali odpowiedzi na poszczególne pytania zamknięte, co przekłada się na ustalenie jednostki pomiaru dla badanych zjawisk. Polega ono na przyporządkowaniu (zgodnie z określonymi zasadami) wybranych symboli (cyfr) obserwowanym faktom lub wyrażanym opiniom i ma decydujący wpływ na sposoby opracowania danych. Natomiast pomiar zjawiska to ustalenie za pomocą jakich i ilu pytań lub stwierdzeń zmierzyć badane zjawisko. Pytania i stwierdzenia określa się mianem wskaźników³¹ zjawiska, zaś o mierzonej cesze (operacjonalizującej badane zjawisko) mówi się wtedy, że jest ona zjawiskiem wskaźnikowanym.

3.1. Skalowanie cech obserwowalnych

Cechy obserwowalne to te cechy, które można obserwować i mierzyć wprost w badaniach empirycznych o charakterze ilościowym. Jako narzędzie pomiaru wykorzystać można m.in. kwestionariusz ankietowy. Cechy te, ze względu na swój obserwowalny i jednowymiarowy charakter, występują zwykle w postaci jednej z charakterystycznych i znanych dla nich kategorii oraz mają zwykle określone jednostki miary. Z tego powodu do ich pomiaru stosuje się zwykle pojedyncze pytania kwestionariuszowe lub skale proste³².

3.2. Skalowanie cech ukrytych

³¹ wskaźnik to zdarzenie obserwowalne, które pozwala z określonym prawdopodobieństwem stwierdzić zaistnienie zjawiska wskaźnikowanego.

³² Pojedyncze stwierdzenia o odpowiedziach skalowanych na skali porządkowej, np. typu Likerta

W badaniach społecznych do badania zjawisk złożonych i bezpośrednio nieobserwowalnych wykorzystuje się skale pomiarowe złożone, zbudowane najczęściej z kilku, kilkunastu lub kilkudziesięciu skal prostych określanych również mianem stwierdzeń³³, twierdzeń³⁴, wypowiedzi³⁵, aspektów³⁶, kwestii³⁷ bądź pozycji skali³⁸. Są to na ogół krótkie zdania oznajmujące³⁹, z którymi respondent może się zgadzać lub nie zgadzać, a z których każde ma na celu objęcie innego aspektu badanego zjawiska. Ze względu na nieobserwowalny i często złożony charakter mierzonych zjawisk i pojęć, niezwykle ważne jest wykazanie przez badacza, że proponowane do jego pomiaru skale pomiarowe są wysokiej jakości. Oznacza to, że wyniki uzyskiwane za ich pomocą powinny rzeczywiście dotyczyć mierzonego zjawiska lub pojęcia latentnego oraz powinny być precyzyjne. Pierwszy aspekt dotyczy trafności skali pomiarowej, zaś drugi – jej rzetelności. Metody statystyczne dostarczają narzędzi do zweryfikowania, na ile przyjęte przez badacza rozwiązania pomiarowe zapewniają uzyskanie narzędzia wysokiej jakości.

Do pomiaru cech ukrytych stosuje się jedno z dwóch alternatywnych podejść: (1) skalowanie oparte na czynnikach oraz (2) skalowanie czynnikowe.

3.2.1. Skale oparte na czynnikach – skalowanie jednowymiarowe

W badaniach społecznych do badania złożonych i często bezpośrednio nieobserwowalnych zjawisk wykorzystuje się jednowymiarowe skale złożone. Są to skale zbudowane z szeregu stwierdzeń lub pytań określanych mianem pozycji skali. Czasem, aby zwiększyć dokładność pomiaru pozycjom tym nadaje się różną intensywność⁴⁰, nie jest to jednak konieczne. Odpowiedzi udzielane na poszczególne pozycje skali mierzy się za pomocą którejś ze skal porządkowych, np. za pomocą skali Likerta. Pozycja respondenta na takiej

³³ Ostasiewicz W. Istota pomiaru statystycznego, [w:] Pomiar statystyczny, red. W. Ostasiewicz, Wydawnictwo Akademii Ekonomicznej im. Oskara Lanego we Wrocławiu, Wrocław 2003, s. 20; Rószkiewicz M., Metody ilościowe w badaniach marketingowych, Wydawnictwo Naukowe PWN, Warszawa 2002, s.48; Sagan A., Badania marketingowe. Podstawowe kierunki. Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004, s. 89.

³⁴ Nachmias D, Frankfort-Nachmias Ch., Metody badawcze w naukach społecznych, Wydawnictwo Zysk i Spółka, Poznań 2001, s.477-480.

³⁵ Churchill G. A., Badania marketingowe. Podstawy metodologiczne, Wydawnictwo Naukowe PWN, Warszawa 2002, s. 412

³⁶ Ostasiewicz W. Istota pomiaru statystycznego, [w:] Pomiar statystyczny, red. W. Ostasiewicz, Wydawnictwo Akademii Ekonomicznej im. Oskara Lanego we Wrocławiu, Wrocław 2003, s. 20

³⁷ Ostasiewicz W. Istota pomiaru statystycznego, [w:] Pomiar statystyczny, red. W. Ostasiewicz, Wydawnictwo Akademii Ekonomicznej im. Oskara Lanego we Wrocławiu, Wrocław 2003, s. 20

³⁸ Kaczmarczyk S., Badania marketingowe. Metody i techniki, PWE, Warszawa 2002, s. 128-129; Ostasiewicz W. Istota pomiaru statystycznego, [w:] Pomiar statystyczny, red. W. Ostasiewicz, Wydawnictwo Akademii Ekonomicznej im. Oskara Lanego we Wrocławiu, Wrocław 2003, s. 20

³⁹ Kaczmarczyk S., Badania marketingowe. Metody i techniki, PWE, Warszawa 2002, s. 127

⁴⁰ Kaczmarczyk S., Badania marketingowe. Metody i techniki, PWE, Warszawa 2002

skali, a więc intensywność charakteryzującego tego respondenta zjawiska latentnego (np. postawy), określana jest najczęściej za pomocą sumy ocen⁴¹ uzyskanych dla poszczególnych stwierdzeń. Tworzy się w ten sposób skalę sumaryczną, zaś formuła jej wyznaczania jest następująca:

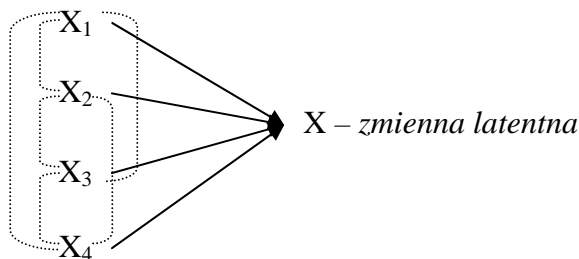
$$X_i = X_{i1} + X_{i2} + \dots + X_{ik} \quad (1)$$

gdzie: X_i – intensywność zjawiska latentnego charakteryzująca i -tego respondenta, X_{ij} – ocena wystawiona przez i -tego respondenta j -tej pozycji skali, $i = 1, \dots, n$ oraz $j = 1, \dots, k$.

Sumowanie takie jest uprawomocnione tylko dzięki przyjęciu założenia, że można traktować dane mierzone na skalach porządkowych jako mające charakter interwałowy. Przyjęcie takiego założenia oznacza, że odległości między kategoriami odpowiedzi traktowane są jako równe. Takie podejście ma zarówno swoich zwolenników jak i przeciwników popierających swoje poglądy licznymi badaniami⁴².

Jak podaje Rószkiewicz⁴³, konstruowanie skal złożonych odnoszących się do symptomów określonych zjawisk latentnych (np. postawy), polega na sprawdzeniu, na ile poszczególne stwierdzenia lub pytania, zwane pozycjami skali, wiążą się z określonym, jednym zjawiskiem latentnym (np. postawą wobec oszczędzania na przyszłość). Jeśli diagnozowane zjawisko latentne ma być mierzone za pomocą proponowanego zestawu stwierdzeń lub pytań, to między nim a udzielanymi przez respondentów odpowiedziami odnoszącymi się do tych stwierdzeń lub pytań, powinna występować zależność przyczynowo-skutkowa. Jej rezultatem powinna być z kolei korelacja między ocenami wystawionymi przez respondentów poszczególnym stwierdzeniom, co schematycznie ilustruje rysunek 2.

Rysunek 2. Zależności przyczynowo-skutkowe (linie ciągłe) między pozycjami skali X1, X2, X3 i X4 oraz zmienną latentną mierzoną (diagnozowaną) za pomocą tej skali – zastosowanie wskaźników formatywnych



zmienne wskaźnikowe – stwierdzenia lub pytania

Źródło: opracowanie własne

⁴¹ oceny te przypisuje się każdemu z wariantów odpowiedzi na etapie konstrukcji skali

⁴² Por. Górniak J., 2000, *My i nasze pieniądze*. Kraków, Aureus, s.312-313

⁴³ Rószkiewicz M., *Analiza klienta*, w druku

Oznacza to, że budowa skali sumarycznej wymaga istnienia skorelowania między pozycjami skali. Im silniejsze są te korelacje, tym wyższa jest wewnętrzna zgodność skali sumarycznej, czyli bardziej rzetelna jest proponowana skala. Zatem do weryfikacji zasadności budowy skali sumarycznej z zestawu proponowanych do tego stwierdzeń niezbędne jest sprawdzenie najpierw, czy skala ta jest rzetelna. Oznacza to, że skala sumaryczna może być zastosowana do pomiaru jednowymiarowego zjawiska ukrytego dopiero wtedy, gdy analiza rzetelności potwierdzi jednowymiarowość budowanej skali⁴⁴.

Podsumowując, warunki niezbędne do pomiaru zjawisk nieobserwowalnych bezpośrednio bazującego na skalach opartych na czynnikach są następujące. Po pierwsze, wysoka zgodność proponowanej skali z określoną koncepcją mierzonego zjawiska ukrytego. Oznacza to, że aspekty badanego zjawiska latentnego objęte zestawem stwierdzeń lub pytań do jego pomiaru powinny wynikać z teorii tego zjawiska. Realizacja tego warunku silnie związania jest z poprawnie przeprowadzonym etapem konceptualizacji badania i budową bazującego na teorii (analizie literatury) modelu konceptualnego. Po drugie, wysoka zgodność wewnętrzna proponowanej skali pomiarowej. Oznacza to, że poszczególne stwierdzenia (pozycje skali) powinny być powiązane z mierzonym zjawiskiem ukrytym (jednym).

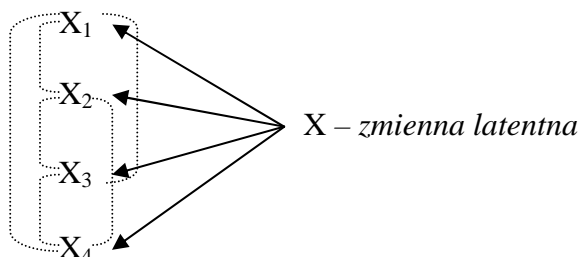
Powiązania te mogą być dwojakiego typu. Po pierwsze, poszczególne stwierdzenia tworzące skalę pomiarową, w języku teorii pomiaru zwane zmiennymi wskaźnikowymi, mogą być determinantami zjawiska latentnego. Mówimy wtedy, że badane zjawisko ukryte jest konsekwencją występowania określonych zjawisk obserwowalnych, zaś jego zmienne wskaźnikowe mają charakter wskaźników formatywnych (formujących lub budujących) (ang. *formative indicators*). W takim przypadku kierunek związku między stwierdzeniami, a mierzoną przez nie zmienną ukrytą jest taki, jak prezentuje to rysunek 2.

Po drugie, poszczególne stwierdzenia tworzące skalę pomiarową, mogą być symptomami zjawiska latentnego. Mówimy wtedy, że badane zjawisko ukryte jest źródłem występowania określonych zjawisk obserwowalnych lub przejawia się za pomocą występowania określonych zjawisk obserwowalnych. Zjawiska stanowią wtedy symptomy zjawiska latentnego. W takim przypadku zmienne wskaźnikowe mają charakter wskaźników refleksywnych (refleksyjnych lub odbijających) (ang. *reflective indicators*), zaś kierunek związku między stwierdzeniami, a mierzoną przez nie zmienną ukrytą jest taki, jak prezentuje to rysunek 3. W sytuacji gdy zmienne wskaźnikowe mają charakter refleksywny, do

⁴⁴ Rószkiewicz M. Analiza klienta, w druku

konstrukcji skali pomiarowej powinno być wykorzystywane podejście oparte na skalach czynnikowych.

Rysunek 3. Zależności przyczynowo-skutkowe (linie ciągłe) i korelacyjne (linie przerywane) między pozycjami skali X1, X2, X3 i X4 oraz zmienną latentną mierzoną (diagnozowaną) za pomocą tej skali – zastosowanie wskaźników refleksywnych



zmienne wskaźnikowe – stwierdzenia lub pytania

Źródło: Rószkiewicz M. Analiza klienta, w druku

3.2.2 Skalowanie czynnikowe - skale wielowymiarowe

Wśród metod skalowania zjawisk wielowymiarowych lub zjawisk latentnych wielowymiarowych lub jednowymiarowych za pomocą skal czynnikowych wyróżnić można dwa podejścia:

1. eksploracyjne,
2. konfirmacyjne.

W podejściu eksploracyjnym, na podstawie proponowanego zestawu stwierżeń traktowanych jako potencjalne wskaźniki mierzonego zjawiska latentnego, dąży się do określania liczby jego wymiarów, ich interpretacji oraz do ich zmierzenia (wyskalowania). Stosuje się zatem podejście polegające na wykryciu struktury badanego zjawiska oraz interpretacji uzyskanych w ten sposób wyników. Struktura ta może być złożona (wielowymiarowa) lub prosta (jednowymiarowa). W przypadku struktury złożonej, metody eksploracyjne służą do wyjaśnienia liczby wymiarów oraz ich interpretacji. Natomiast w przypadku struktury jednowymiarowej, metody eksploracyjne zastosować można w celu potwierdzenia jednowymiarowości proponowanej do pomiaru jednowymiarowego zjawiska latentnego skali pomiarowej. Do metod eksploracyjnych zalicza się między innymi eksploracyjną analizę czynnikową, analizę głównych składowych oraz skalowanie Rascha.

W podejściu konfirmacyjnym punktem wyjścia jest teoretyczny model badanego zjawiska latentnego. Model ten, zwany również modelem konceptualnym, konstruowany jest na podstawie analizy literatury tematu oraz przyjętych przez badacza celów badawczych oraz

pytań i hipotez badawczych. Obejmuje nie tylko zestaw zmiennych wskaźnikowych, ale również występujące między nimi relacje. Model ten jest następnie weryfikowany, czyli, za pomocą odpowiednich metod statystycznych, sprawdza się, na ile proponowana przez badacza w modelu konceptualnym koncepcja teoretyczna znajduje potwierdzenie w danych. Jeśli potwierdzenie to jest satysfakcjonujące, estymuje się model pomiarowy i kwantyfikuje zmiennych skalujące zdefiniowane apriorycznie konstrukty. Do metod confirmacyjnych zalicza się confirmacyjną analizę czynnikową oraz modelowanie równań strukturalnych.

3. 2.2.1. Podejście eksploracyjne

W ramach podejścia eksploracyjnego wyróżnić można wiele metod skalowania. Ich podstawowa klasyfikacja bazuje na określeniu charakteru zmiennych obserwowalnych oraz charakteru zmiennych latentnych. W obu przypadkach rozróżnia się zmienne ciągłe od zmiennych dyskretnych. W konsekwencji powstają cztery kombinacje tak, jak to zaprezentowano w tabeli 7.

Tabela 7. Klasyczna klasyfikacja metod modelowania zmiennych ukrytych

Zmienne obserwowalne	Zmienne ukryte	
	Ciągłe (<i>latent traits</i>)	Dyskretne (<i>latent classes</i>)
Ciągłe	Analiza czynnikowa	Analiza profili ukrytych
Dyskretne	Modele IRT	Analiza klas ukrytych

Źródło: opracowanie własne na podstawie Skrondal A., Rabe-Hesketh S., Generalized Latent Variable Modeling. Multilevel, Longitudinal, and Structural Equation Models Chapman&Hall/CRC, 2004

W sytuacji gdy zmienne obserwowalne mają charakter ciągły oraz zjawisko latentne kwantyfikuje się jako zmienną ciągłą właściwą metodą analityczną jest analiza czynnikowa. W sytuacji gdy zmienne obserwowalne mają charakter ciągły, ale badane zjawisko latentne kwantyfikuje się jako zmienną dyskretną, właściwą metodą analityczną jest analiza profili ukrytych. W sytuacji gdy zmienne obserwowalne mają charakter dyskretny i badane zjawisko latentne również ma charakter dyskretny (tzn. kwantyfikuje się jako zmienną latentną dyskretną), właściwą metodą analityczną jest analiza klas ukrytych. I wreszcie gdy zmienne obserwowalne mają charakter dyskretny, ale badane zjawisko latentne kwantyfikuje się jako zmienną ciągłą, właściwymi metodami analitycznymi są modele IRT, w tym model Rascha.

Przedstawiona powyżej klasyfikacja pokazuje podział klasyczny. Wraz z obserwowanym w ciągu kilkunastu ostatnich lat wzrostem mocy obliczeniowej komputerów możliwe stało się łączenie ze sobą w ramach jednej specyfikacji czterech klasycznych modeli służących do kwantyfikacji zmiennych latentnych. Ze względu na fakt, że tematyka niniejszej

publikacji koncentruje się na metodach przygotowania badania, a nie samej analizie jego wyników, poniżej przedstawiono przykłady zastosowania dwóch metod, które stosuje się do weryfikacji poprawności i trafności modelu konceptualnego oraz do konstrukcji skali pomiarowej – analizę czynnikową w wersji eksploracyjnej oraz model Rascha⁴⁵.

Eksploracyjna analiza czynnikowa – przykłady zastosowań

Eksploracyjna analiza czynnikowa może być stosowana nie tylko do wykrycia struktury związków występujących między zmiennymi wskaźnikowymi diagnozującymi określone zjawisko lub zjawiska nieobserwowalne bezpośrednio, ale również do badania wymiarowości zbioru zmiennych wskaźnikowych lub też do sprawdzenia jednowymiarowości skali pomiarowej. To pierwsze zastosowanie zaprezentowała Rószkiewicz⁴⁶, zaś drugie – Woynarowska-Sołdan i Węziak-Białowolska⁴⁷.

Rószkiewicz⁴⁸ zastosowała procedurę analizy czynnikowej w celu wygenerowania postaw wobec gromadzenia oszczędności. Analizę przeprowadziła na zestawie 7 stwierdzeń odnoszących się do motywów gromadzenia oszczędności (treść stwierdzeń zamieszczono w tabeli 8). Uzyskane przez nią wartości ładunków po rotacji⁴⁹, dla trzech czynników⁵⁰, do których ograniczono analizę zgodnie z sugestiami płynącymi z prezentowanej w literaturze tematu teorii oszczędzania, zestawiono w tabeli 8.

⁴⁵ Przykład zastosowania analizy klas ukrytych do wyskalowania zmiennej latentnej o charakterze dyskretnym znaleźć można w Białowolski P., Węziak-Białowolska D., 2011, Zastosowanie analizy klas ukrytych do identyfikacji wykluczenia gospodarstw domowych z rynku kredytowego „Roczniki Kolegium Analiz Ekonomicznych”, Nr 23/2011, s.159 - 174

⁴⁶ Rószkiewicz M., *Oszczędzanie. Postawy i zachowania polskich gospodarstw domowych wobec oszczędzania*, Aureus, Kraków 2008

⁴⁷ Woynarowska-Sołdan M., Węziak-Białowolska D., Narzędzia do badania klimatu społecznego w szkole promującej zdrowie i ich analiza psychometryczna, „Edukacja, Studia, Badania, Innowacje”, nr 3 (111), 2010, s. 71-80

⁴⁸ Rószkiewicz M., *Oszczędzanie. Postawy i zachowania polskich gospodarstw domowych wobec oszczędzania*, Aureus, Kraków 2008.

⁴⁹ Przeprowadzona rotacja typu Varimax nie zmieniła struktury zależności między wyróżnionymi motywami oraz trzema pierwszymi czynnikami wyznaczonymi metodą głównych składowych. Jednakże rotacja ta zwiększyła polaryzację wartości ładunków, co jest rozwiązaniem bardziej jednoznacznym względem rozwiązania początkowego.

⁵⁰ W ocenie adekwatności próby do założeń tej metody, mierzonej statystyką Kaisera-Maylera-Olkina uzyskano satysfakcjonujący poziom 0,791. Przy czym wariancje trzech zmiennych czynnikowych odtworzyły 72,8% wariancji zmiennych podstawowych.

Tabela 8. Wartości ładunków czynnikowych stojące przy pierwszych trzech czynnikach opisujących poszczególne motywy oszczędzania.

Motywy oszczędzania	Ładunki stojące przy:		
	Czynniku 1	Czynniku 2	Czynniku 3
S1. Chwilowy brak pieniędzy	0,822	-0,049	0,282
S2. Finansowanie poważnych zakupów i inwestycji	0,723	0,376	0,082
S3. Finansowanie niespodziewanych wydatków	0,697	0,486	0,059
S4. Zabezpieczenie przyszłości najbliższych	0,178	0,816	0,105
S5. Zabezpieczenie okresu starości	0,157	0,777	0,229
S6. Leży to w naturze ludzkiej	0,173	0,110	0,856
S7. Wyznawanie zasady wstrzemięźliwości	0,124	0,203	0,848

Źródło: Rószkiewicz M., *Oszczędzanie. Postawy i zachowania polskich gospodarstw domowych wobec oszczędzania*, Aureus, Kraków 2008, s. 107

Na podstawie zaobserwowanego układu ładunków czynnikowych Rószkiewicz wskazała trzy postawy wobec gromadzenia oszczędności. Były to:

- postawa ostrożności związana z motywem chwilowego braku pieniędzy, motywem finansowania poważnych zakupów i motywem inwestycji oraz finansowania niespodziewanych wydatków (stwierdzenia S1, S2, S3);
- postawa przezornościowa – związana z cyklem życia – odpowiadająca za motyw zabezpieczania przyszłości najbliższych oraz zabezpieczania okresu starości (stwierdzenia S4, S5);
- postawa wstrzemięźliwa związana z naturalnymi predyspozycjami natury ludzkiej oraz wyznawaniem zasady wstrzemięźliwości (stwierdzenia S6, S7);

Dla każdej z wyróżnionych postaw Rószkiewicz oceniła rzetelność mierzących je skal pomiarowych za pomocą współczynników α -Cronbacha. Dla skali pomiarowej postawy ostrożności wartość współczynnika α -Cronbacha wyniosła 0,722, dla skali pomiarowej postawy przezornościowej była na poziomie 0,629, z kolei dla skali pomiarowej postawy wstrzemięźliwej współczynnik α -Cronbacha była na poziomie 0,721. Przeprowadzona analiza pozwoliła zatem uznać, że respondenci postrzegali motywy oszczędzania przez pryzmat trzech wymiarów, tj. ostrożności, cyklu życia i wstrzemięźliwości z natury rzeczy, tak jak jest to ujmowane w literaturze tematu.

Z kolei Woynarowska-Sołdan i Węziak-Białowolska⁵¹ zastosowały analizę czynnikową⁵² do sprawdzenia wymiarowości skali pomiarowej do badania klimatu

⁵¹ Woynarowska-Sołdan M., Węziak-Białowolska D., Narzędzia do badania klimatu społecznego w szkole promującej zdrowie i ich analiza psychometryczna, „Edukacja, Studia, Badania, Innowacje”, nr 3 (111), 2010, s. 71-80

⁵² czynniki zostały wyodrębnione metodą czynników głównych z ortogonalną rotacją Varimax

społecznego w szkole promującej zdrowie (a pośrednio również do weryfikacji jej trafności) oraz analizę głównych skadowych do potwierdzenia jednowymiarowości poszczególnych podskal składających się na skalę klimatu społecznego takiej szkoły. Ze względu na fakt, że klimat społeczny szkoły jest pojęciem subiektywnym, wielowymiarowym, niejednoznacznym i trudnym do zdefiniowania, a w literaturze znaleźć można wiele jego definicji⁵³ autorki przyjęły, że dla potrzeb weryfikacji proponowanego podejścia konceptualnego zastosowane zostaną metody oparte na podejściu czynnikowym⁵⁴.

W eksploracyjnej analizie czynnikowej współczynnik K-M-O osiągnął satysfakcjonującą wartość 0,894, zaś zmienne wskaźnikowe rozlokowały się w 8 czynnikach, które wyjaśniały w sumie 67% ogólnej zmienności (wyjaśnianej wariancji). Rozwiązanie ośmioczynnikowe poddano rotacji *Varimax*. Zdecydowana większość stwierdzeń zakwalifikowała się wyraźnie do poszczególnych czynników odpowiadających poszczególnym podskalom kwestionariusza ankietowego. O miejscu kilku stwierdzeń zdecydowano na podstawie analizy wielkości ich ładunku czynnikowego oraz treści. W kolejnym kroku sprawdzono jednowymiarowość podskal, które pokazało, że:

1. w przypadku każdej z podskal, poza podskalą „satysfakcja z pracy w szkole”, potwierdzono, że mogą być one uznane za jednowymiarowe (wyłącznie jedna wartość własna przekroczyła 1); w przypadku podskali „satysfakcja z pracy w szkole” dwie wartości własne przekroczyły 1, co oznacza, że podskala ta obejmuje dwa aspekty;
2. udział wyjaśnionej wariancji dla każdej z podskal wynosił od 53% do 73%.

Również rzetelność ostatecznych podskal kwestionariusza mierzona współczynnikiem α -Cronbacha okazała się satysfakcjonująca w przypadku wszystkich podskal i wyniosła od 0,72 do 0,94. Ponadto średnie korelacje między stwierdzeniami można było uznać za wysokie (przekraczały one zalecany poziom 0,3).

⁵³ Adrian B. (2003) Klimat szkoły, w: Encyklopedia Pedagogiczna XXI wieku, T. Pilch (red.), Wydawnictwo Akademickie Żak, Warszawa, s. 609-613; Arends R (1994) Uczymy się nauczać, WSiP, Warszawa; Freiberg H.J. (red.) (1999) School Climate: Measuring, Improving and Sustaining Healthy Learning Environments, Falmer Press, London; Śliwerski B., Milerski B. (2000) Leksykon Pedagogika, Wydawnictwo Naukowe PWN, Warszawa.

⁵⁴ Analizy prowadzono na danych zebranych za pomocą kwestionariusza ankietowego, który zawierał stwierdzenia w postaci zdań oznajmujących pogrupowane w podskale odpowiadające przyjętym wymiarom klimatu społecznego szkoły. Na każde ze stwierdzeń można było udzielić odpowiedzi na pięciopunktowej skali Likerta.

Ocena jakości skali pomiarowej mierzącej stosunek do osób starszych za pomocą modelu Rascha

Węziak⁵⁵ zastosowała model Rascha do oceny jakości skali mierzącej stosunek do osób starszych. Respondentów poproszono o wyrażenie opinii na osiem następujących stwierdzeń⁵⁶:

- A1. Dzięki doświadczeniu są ciągle potrzebni
- A2. Gwarantują zachowanie tradycyjnych wartości w społeczeństwie
- A3. Młodsze generacje mogą korzystać z ich obecności, wiedzy, doświadczenia
- A4. Społeczeństwo powinno brać pod uwagę prawa starszych
- A5. Społeczeństwo powinno brać pod uwagę problemy osób starszych
- A6. Osoby starsze są nieproduktywne i tylko stanowią obciążenie dla społeczeństwa
- A7. Starsze osoby stanowią przeszkodę dla zmian
- A8. Starsze osoby stanowią ciężar dla społeczeństwa

Respondenci mieli do wyboru jedną z pięciu kategorii odpowiedzi, z których każda była etykietowana (1 – zdecydowanie się zgadzam, 2 – zgadzam się, 3 – ani się zgadzam, ani się nie zgadzam, 4 – nie zgadzam się, 5 – zdecydowanie się nie zgadzam). W celu ułatwienia interpretacji wyników dokonano przekodowania wyników w taki sposób, aby osoby z wyższym wynikiem sumarycznym charakteryzowały się bardziej pozytywnym stosunkiem do osób starszych niż osoby o niższym wyniku ogólnym. Ze względu na porządkową (Likertowską) skalę odpowiedzi analizy wykonano za pomocą porządkowego modelu Rascha (ang. *rating scale model*).

W pierwszym kroku sprawdzono, czy dane spełniają założenia modelu. Przeciętne wartości statystyki dopasowania *IMS* i *OMS* dla respondentów były na poziomie równym bądź bliskim 1, co odpowiada wartości oczekiwanej tych statystyk i świadczy o tym, że badana grupa respondentów nie powinna być uznana za nietypową⁵⁷. Natomiast statystyki

⁵⁵ Węziak D., *Ocena jakości skali ze szczególnym uwzględnieniem jej rzetelności i trafności za pomocą skalowania Rascha*, „ASK Społeczeństwo Badania Metody”, Nr. 15, 2006, s. 107-131

⁵⁶ Do ilustracji wykorzystano dane pochodzące z badania „Badanie poglądów na zagadnienia ludnościowe oraz politykę ludnościową PPA2 (Population Policy Attitudes Survey)” przeprowadzonego w IV kw. 2001 przez Instytut Statystyki i Demografii Szkoły Głównej Handlowej w Warszawie przy współpracy z Głównym Urzędem Statystycznym w ramach projektu „Population Policy Acceptance Study. The Viewpoint of Citizens and Policy Actors regarding the Management of Population Related Change DIALOG”. Badanie objęło osoby w wieku 18-64 lat z gospodarstw domowych w mieszkaniach wylosowanych do próby BAEL (Badanie aktywności ekonomicznej ludności). Wywiady przeprowadzono z 4244 respondentami z 2027 gospodarstw domowych. (Kotowska i in. 2003).

⁵⁷ Bond T.G., Fox Ch. M., *Applying The Rasch Model. Fundamental Measurement in the Human Science*, Lawrence Erlbaum Associates, Publishers Mahaw, New Jersey 2001.

SIMS i *SOMS* wyniosły -0,4, co również nie odbiegało znacząco od wartości oczekiwanych tych statystyk.

Obliczone dla wszystkich pozycji skali pomiarowej łącznie wartości *IMS* i *OMS* były na poziomie 0,99 i 0,95, a więc bardzo bliskim wartości oczekiwanej tych statystyk⁵⁸, co świadczyło o tym, że skala nie powinna być uznana za nietypową, a w dalszym kroku pozwoliło przyjąć, że badana grupa stwierdzeń mierzy stosunek do osób starszych. Powyższy wniosek potwierdziły również wartości statystyk *SIMS* i *SOMS*, które wyniosły -0,1 i -0,4.

Współczynnik α -Cronbacha, będący klasyczną miarą rzetelności skali pomiarowej, wyniósł 0,851 przy średniej korelacji między pozycjami skali na poziomie 0,42. Pozwoliło to uznać, że analizowana skala pomiarowa okazała się być rzetelna na satysfakcjonującym poziomie. Wniosek ten potwierdziła analiza wartości współczynnika rzetelności uzyskanego w toku analiz modelem Rascha. Współczynnik ten (R_p) był na poziomie 0,72, co oznaczało, że zgodnie z zaleceniami Green i Frantom⁵⁹ był on na akceptowalny.

Przeciętne oszacowanie pozycji respondentów mierzone logitem wyniosło 2,63, co przy uwzględnieniu informacji, że w skalowaniu Rascha przeciętne oszacowanie pozycji skali wynosi 0, pozwala uznać, że przeciętnie biorąc respondenci bardzo łatwo zgadzali się z analizowanymi stwierdzeniami. Oznaczało to, że ich stosunek do osób starszych był pozytywny. W następnym kroku prowadzi to jednak do konkluzji, że dla tej grupy respondentów skala nie była optymalnie dopasowana pod względem zróżnicowania intensywności stwierdzeń i sugeruje jej rozszerzenie o stwierdzenia, z którymi trudniej się zgodzić, po uprzednim przeprowadzeniu analizy poszczególnych stwierdzeń indywidualnie. W kolejnym kroku przeprowadzono analizę każdego ze stwierdzeń osobno. W tabeli 9 przedstawiono pozycje skali uszeregowane według rosnących wartości oszacowań ich intensywności.

Tabela 9. Charakterystyki pozycji skali

POZYCJE SKALI	OIPS	B.S.	STATYSTYKI DOPASOWANIA			
			IMS	SIMS	OMS	SOMS
<i>(r)A7. Starsze osoby stanowią przeszkodę dla zmian</i>	0,32	0,15	0,98	-0,1	0,94	-0,4
<i>(r)A6. Osoby starsze są nieproduktywne i tylko stanowią obciążenie dla społeczeństwa</i>	0,21	0,15	1,14	1,1	1,09	0,7

⁵⁸ Bond T.G., Fox Ch. M., Applying The Rasch Model. Fundamental Measurement in the Human Science, Lawrence Erlbaum Associates, Publishers Mahaw, New Jersey 2001.

⁵⁹ Green K. E., Frantom C. G., 2000. Survey development and validation with the Rasch Model, A paper presented at the International Conference on Questionnaire Development, Evaluation, and Testing. http://www.jpsm.umd.edu/qdet/final_pdf_papers/green.pdf

(r)A8. Starsze osoby stanowią ciężar dla społeczeństwa	0,13	0,15	1,13	1	1,11	0,9
A2. Gwarantują zachowanie tradycyjnych wartości w społeczeństwie	0,08	0,15	0,93	-0,5	0,84	-1,3
A3. Młodsze generacje mogą korzystać z ich obecności, wiedzy, doświadczenia	0,02	0,15	0,82	-1,5	0,81	-1,5
A4. Społeczeństwo powinno brać pod uwagę prawa starszych	-0,17	0,15	0,78	-1,9	0,78	-1,8
A1. Dzięki doświadczeniu są ciągle potrzebni	-0,26	0,15	1,22	1,7	1,14	1,1
A5. Społeczeństwo powinno brać pod uwagę problemy osób starszych	-0,33	0,15	0,93	-0,5	0,92	-0,6
ŚREDNIA	0,00	0,15	0,99	-0,1	0,95	-0,4
ODCHYLENIE STANDARDOWE	0,22	0,00	0,15	1,22	0,14	1,0

*OIPS – szacowanie intensywności pozycji skali; B.S. – błąd szacunku;

Źródło: Węziak D., Ocena jakości skali ze szczególnym uwzględnieniem jej rzetelności i trafności za pomocą skalowania Rascha, „ASK Społeczeństwo Badania Metody”, Nr. 15, 2006, s. 120

Uiszeregowanie badanych stwierdzeń według malejących wartości oszacowań ich intensywności należy rozumieć w następujący sposób. Respondenci najtrudniej zgadzali się z rozumianym w sposób przeciwny stwierdzeniem (r)A7. *Starsze osoby stanowią przeszkodę dla zmian*, następnie ze stwierdzeniami (również o przeciwnym znaczeniu): (r)A6. *Osoby starsze są nieproduktywne i tylko stanowią obciążenie dla społeczeństwa* i (r)A8. *Starsze osoby stanowią ciężar dla społeczeństwa*. Najłatwiejsze do zaakceptowania były stwierdzenia A1. *Dzięki doświadczeniu są ciągle potrzebni* i A5. *Społeczeństwo powinno brać pod uwagę problemy osób starszych*.

Kwestią niezwykle istotną z punktu widzenia trafności treściowej skali pomiarowej jest zastanowienie się, czy kolejność, w jakiej ustawione są pozycje skali (według oszacowań stopnia akceptowalności), ma sens. Czyli czy na przykład łatwiej jest zgodzić się ze stwierdzeniem A2. *Gwarantują zachowanie tradycyjnych wartości w społeczeństwie* niż A1. *Dzięki doświadczeniu są ciągle potrzebni*. A także czy osoba charakteryzująca się pozytywnym stosunkiem do osób starszych nie zgodzi się ze stwierdzeniem (r)A7. *Starsze osoby stanowią przeszkodę dla zmian*, podczas gdy osoba o bardziej negatywnej postawie nie zgodzi się ze stwierdzeniem A5. *Społeczeństwo powinno brać pod uwagę problemy osób*. Analiza treści stwierdzeń i wyników porządkowego skalowania Rascha pozwoliła na powyższe pytania odpowiedzieć twierdząco.

Żadnego z badanych stwierdzeń nie można było uznać za nie pasujące do skali, co pokazały statystyki IMS i OMS zawierające się w przedziałach odpowiednio $\langle 0,78; 1,22 \rangle$ i $\langle 0,78; 1,14 \rangle$. Powyższy wniosek okazał się poprawny również w przypadku

statystyk *SIMS* i *SOMS*, z których żadna nie przekroczyła zalecanego poziomu $<-2, 2>$ ⁶⁰. Pozwoliło to uznać, że skala mierząca stosunek do osób starszych charakteryzowała się trafnością teoretyczną⁶¹.

Podsumowując, przeprowadzona przez Węziak⁶² analiza skali oceniającej stosunek do osób starszych wykonana przy pomocy porządkowego skalowania Rascha dała satysfakcjonujące wyniki. Sposób uszeregowania pozycji według stopnia ich intensywności pozwolił uznać trafność treściową skali za odpowiednią. Natomiast o dobrej trafności teoretycznej narzędzia świadczyły statystyki dopasowania *SIMS*, *SOMS*, *IMS* i *OMS*, które w żadnym przypadku nie wykroczyły poza przedziały uznawane za dopuszczalne. Rzetelność skali mierzona współczynnikiem α -Cronbacha, jak i współczynnikiem rzetelności R_p była na satysfakcjonującym poziomie. Niemniej jednak analizowanej skali można zarzucić, że obejmuje zbyt wąski zakres stosunku do osób starszych, a także że składa się ze stwierdzeń, których zawartość treściowa narzuca respondentom określony typ odpowiedzi – odpowiedzi pożądanej i poprawnej społecznie. W celu ulepszenia narzędzia należałoby dodać stwierdzenia o wyższej intensywności, a więc z którymi trudniej będzie się zgadzać.

3.2.2.2. Podejście confirmacyjne

W podejściu confirmacyjnym punktem wyjścia jest model konceptualny. Oznacza to, że jeszcze przed przeprowadzeniem jakichkolwiek analiz zakłada się istnienie pewnego modelowego obrazu świata rzeczywistego w wycinku odpowiadającym postawionemu celowi badania. W toku procesu badawczego bada się zasadność przypuszczenia, że obraz ten wygląda tak, jak założono w modelu konceptualnym. Analizy confirmacyjne mają zatem na celu jedną z alternatyw: (1) potwierdzenie poprawności modelu konceptualnego, (2) odrzucenie modelu konceptualnego, (3) modyfikację modelu konceptualnego. Uzyskanie w toku analiz confirmacyjnych potwierdzenia poprawności modelu konceptualnego (również po wprowadzeniu modyfikacji) umożliwia estymację modelu pomiarowego i kwantyfikację zmiennych skalujących zdefiniowane apriorycznie konstrukty. Jak już wspomniano, do metod confirmacyjnych zalicza się confirmacyjną analizę czynnikową oraz modelowanie równań strukturalnych.

⁶⁰ Bond T.G., Fox Ch. M., *Applying The Rasch Model. Fundamental Measurement in the Human Science*, Lawrence Erlbaum Associates, Publishers Mahaw, New Jersey 2001.

⁶¹ W analogiczny sposób, jak ten przedstawiony powyżej, można dokonać analizy odpowiedzi poszczególnych respondentów.

⁶² Węziak D., *Ocena jakości skali ze szczególnym uwzględnieniem jej rzetelności i trafności za pomocą skalowania Rascha*, „ASK Społeczeństwo Badania Metody”, Nr. 15, 2006, s. 107-131.

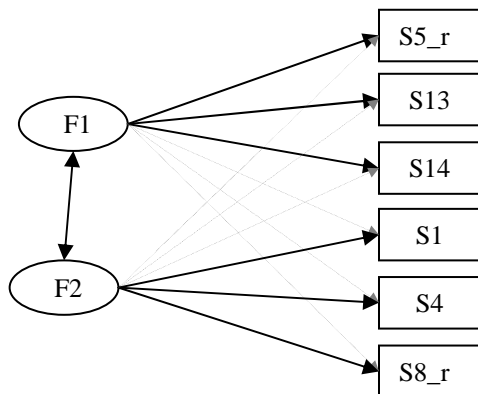
Konfirmacyjna analiza czynnikowa – przykład zastosowania

W podejściu bazującym na analizie czynnikowej (zarówno konfirmacyjnej jak i eksploracyjnej) przyjmuje się, że o występowaniu zjawiska ukrytego świadczą jego symptomy, które są kwantyfikowane przez odpowiednie zmienne wskaźnikowe.

Jednak w podejściu konfirmacyjnym, w odróżnieniu do podejścia eksploracyjnego, a priori ustala się listę tych symptomów oraz powiązania między nimi. W konsekwencji konfirmacyjna analiza czynnikowa (CFA) oznacza modelowanie związków między zmiennymi wskaźnikowymi z narzuconymi warunkami ograniczającymi (ang. *restricted factor model*), zaś eksploracyjna analiza czynnikowa to modelowanie związków między zmiennymi wskaźnikowymi bez żadnych warunków ograniczających⁶³ (ang. *unrestricted factor model*). Różnicę między oboma podejściami schematycznie przedstawiono na rysunku 4.

Rysunek 4. Porównanie modelu eksploracyjnej i konfirmacyjnej analizy czynnikowej.

Model eksploracyjnej analizy czynnikowej – rozwiązanie dwuczynnikowe, rotacja ukośna

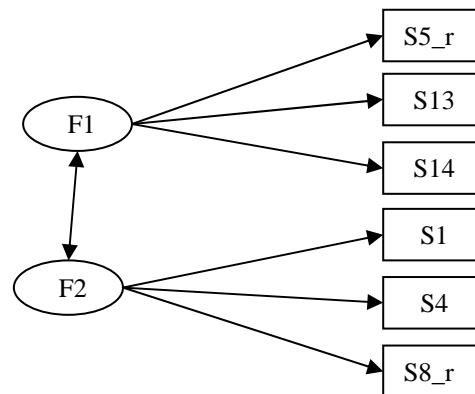


*czarne strzałki ilustrują największe (co do wartości bezwzględnej) ładunki czynnikowe;
 **szare strzałki ilustrują ładunki czynnikowe o niskich wartościach bezwzględnych
 ***strzałka o dwóch grotach symbolizuje skorelowanie głównych składowych F1 oraz F2 (efekt zastosowania rotacji ukośnej)

Źródło: opracowanie własne.

*na schematach zastosowano typowe w modelowaniu równań strukturalnych formy graniczne; zmienne latentne przedstawiono w postaci elips, natomiast zmienne obserwowalne (zmienne wskaźnikowe) – w postaci prostokątów; kierunek strzałek odpowiada kierunkowi zależności przyczynowo-skutkowych – zmienne latentne ujawniają się przez swoje symptomy, zaś zmienne wskaźnikowe mają charakter wskaźników refleksywnych.

Model konfirmacyjnej analizy czynnikowej – rozwiązanie dwuczynnikowe



*niskie ładunki czynnikowe (szare strzałki na rysunku z lewej strony) zostały ustalone na poziomie 0; oznacza to, że do modelu dodano warunki ograniczające, mówiące, że wybrane ładunki czynnikowe równe są zero.

⁶³ Kaplan D., 2009, Structural equation modeling. Foundation and Extensions, Sage, Los Angeles; Brown T. A., 2006, Confirmatory Factor Analysis for Applied Research, The Guilford Press, New York, London.

Rysunek ten jest ilustracją podejścia badawczego zaproponowanego przez Węziak-Białowolską⁶⁴ do oceny własności psychometrycznych skali motywacji testowej stworzonej na podstawie odpowiedzi udzielonych przez uczniów klas trzecich gimnazjum na stwierdzenia kwestionariusza ankietowego do samodzielnego wypełnienia⁶⁵. Proponowane podejście obejmowało przeprowadzenie analiz eksploracyjnych⁶⁶ i confirmacyjnych⁶⁷. Analizy eksploracyjne miały na celu zweryfikowanie, czy zastosowane narzędzie – stwierdzenia kwestionariusza – jest przydatne do stworzenia skali motywacji testowej. Natomiast analizy confirmacyjne posłużyły do zweryfikowania, czy stworzone narzędzie charakteryzuje się odpowiednią dokładnością pomiarową – trafnością i rzetelnością.

W toku eksploracyjnych okazało się, że zbiór 10 zmiennych wskaźnikowych może być wiarygodnie zastąpiony czterema nowymi zmiennymi (głównymi składowymi), które razem odtwarzają 65,85% wariancji zbioru zmiennych wskaźnikowych. Analiza wzoru ładunków czynnikowych pokazała, że główne składowe kwantyfikowały następujące wymiary motywacji testowej i stopnia przygotowania do sesji zrównującej: (F1) pierwsza główna składowa – zaangażowanie w rozwiązanie testu zrównującego (stwierdzenia S5_r, S13 oraz S14), (F2) druga główna składowa – subiektywna ocena stopnia aktualnie posiadanej wiedzy niezbędnej do rozwiązania testu zrównującego i właściwego egzaminu (stwierdzenia S1, S4 oraz S8_r), (F3) trzecia główna składowa – subiektywna ocena stopnia potencjalnego przygotowania do właściwego egzaminu (stwierdzenia S2_r i S9_r), (F4) czwarta główna składowa – ocena stopnia przygotowania się do rozwiązania testu zrównującego (stwierdzenia S3 i S7).

Odpowiadający powyższemu rozwiązaniu w obszarze dwóch pierwszych wymiarów schemat modelu zaprezentowano w lewej części rysunku 4. Natomiast z prawej strony rysunku 4 zamieszczono odpowiadający mu schemat modelu confirmacyjnej analizy czynnikowej, czyli model z warunkami ograniczającymi.

⁶⁴ Węziak-Białowolska D., Ocena własności psychometrycznych skali motywacji testowej, XVII Konferencja Diagnostyki Edukacyjnej, w druku

⁶⁵ Do zmierzenia poziomu motywacji testowej uczniów oraz stopnia przygotowania do sesji zrównującej zastosowanych zostało 10 stwierdzeń. Uczniowie proszeni byli o ustosunkowanie się do nich na 5-stopniowej skali Likerta.

⁶⁶ Zastosowano analizę głównych składowych z rotacją varimax – w celu rozpoznania możliwości utworzenia z zaproponowanego zestawu zmiennych mniejszego zbioru zmiennych, które dobrze wyjaśniają korelację między tymi zmiennymi i tym samym dobrze odtwarzają informację zawartą w zbiorze wyjściowym; sprawdzony został poziom zmienności wspólnej każdego ze stwierdzeń (min. 0,3) oraz poziom wyjaśnianej wariancji (min. 60%);

⁶⁷ Modele zostały zweryfikowane m. in. pod względem istotności zmiennych wskaźnikowych (stwierdzeń) oraz jakości dopasowania.

Strategia postępowania w skalowaniu zmiennych latentnych w podejściu czynnikowym

Przystępując do skalowania zjawisk nieobserwowalnych za pomocą podejścia czynnikowego, zwłaszcza w sytuacji gdy badane zjawisko jest słabo rozpoznane w literaturze tematu, warto rozważyć zastosowanie następującej strategii postępowania. Najpierw na zbiorze zmiennych wskaźnikowych wyselekcjonowanych do pomiaru zjawiska latentnego przeprowadza się eksploracyjną analizę czynnikową. Jej wyniki pozwalają na wykrycie struktury związków występujących między zmiennymi wskaźnikowymi oraz na wprowadzenie modyfikacji do tej listy (usunięcie zmiennych wskaźnikowych). Następnie za pomocą współczynnika α -Cronbacha weryfikuje się rzetelność uzyskanych skal czynnikowych. W przypadku uzyskania niesatysfakcjonujących wyników ponownie możliwe jest zmodyfikowanie zestawu zmiennych wskaźnikowych przez usunięcie tych z nich, których zgodność ze skalą nie została potwierdzona. W ostatnim kroku przeprowadza się confirmacyjną analizę czynnikową i ponownie na podstawie uzyskanych wyników przeprowadza się modyfikację zestawu zmiennych wskaźnikowych przez usunięcie tych z nich, które okazały się być nieistotne statystycznie lub nietrafne (znaki ładunków czynnikowych nie były zgodne z przewidywaniami).

Ostateczne rozwiązanie ocenia się pod względem rzetelności stosując na przykład współczynnik rho Joreskoga, którego formuła jest następująca:

a) w przypadku występowania skorelowania składników resztowych

$$\rho = \frac{(\sum \lambda)^2}{(\sum \lambda)^2 + \sum (1 - \lambda^2) + 2 \sum \theta_{ij}} \quad (2)$$

lub

b) w przypadku braku skorelowania składników resztowych

$$\rho = \frac{(\sum \lambda)^2}{(\sum \lambda)^2 + \sum (1 - \lambda^2)} \quad (3)$$

gdzie: λ – ładunek czynnikowy wystandaryzowany, θ – niezerowe kowariancje składników resztowych.

Współczynnik ten przyjmuje wartości z przedziału $<0; 1>$.

Procedurę taką zastosowała Węziak-Białowska do weryfikacji koncepcji kapitału intelektualnego regionu⁶⁸. Konstrukcja modelu do pomiaru kapitału intelektualnego regionu zrealizowana została poprzez:

⁶⁸ Węziak-Białowska D., *Model kapitału intelektualnego regionu. Koncepcja pomiaru i jej zastosowanie*, Oficyna Wydawnicza SGH, Warszawa 2010

- zidentyfikowanie skal pomiarowych do badania zmiennych o charakterze nieobserwowalnym występujących w modelu konceptualnym,
- analizę trafności i rzetelności zaproponowanych skal pomiarowych,
- opis sformułowanych w modelu konceptualnym relacji za pomocą równań strukturalnych.

Do oceny adekwatności zestawów zmiennych obserwowalnych zaproponowanych do konstrukcji jednowymiarowych skal opisujących określone składniki kapitału intelektualnego wykorzystane zostały eksploracyjna oraz confirmacyjna analiza czynnikowa. Analiza eksploracyjna miała na celu sprawdzenie, czy za obserwowanymi zmiennymi kryje się struktura, która mogła być ujawniona za pomocą mniejszej liczby wymiarów. Z kolei analiza confirmacyjna pozwoliła ocenić jakość zaproponowanych modeli pomiarowych poszczególnych składników i elementów kapitału intelektualnego regionu.

Pomiaru kapitału intelektualnego regionu, jego składników oraz ich elementów dokonano na podstawie danych z badania Diagnoza Społeczna 2005 oraz danych z Banku Danych Regionalnych Głównego Urzędu Statystycznego dotyczących 2005 roku za pomocą analizy głównych składowych oraz eksploracyjnej i confirmacyjnej analizy czynnikowej.

Konfirmacyjna analiza czynnikowa dla wielu grup – badanie zgodności pomiaru

W sytuacji gdy analiza prowadzona jest w celu porównania przeciętnych wartości zmiennych ukrytych klasyczny model confirmacyjnej analizy czynnikowej

$$x = \Lambda_x F + \varepsilon \quad (4)$$

gdzie: Λ_x - macierz ładunków confirmacyjnej analizy czynnikowej, x – zmienne bezpośrednio obserwowalne, F – zmienne nieobserwowalne bezpośrednio, ε – czynnik swoisty,

rozszerza się tak, aby zawierał wyraz wolny τ oraz uwzględniał zróżnicowanie modelu w g grupach:

$$x_g = \tau_g + \Lambda_{xg} F_g + \varepsilon_g \quad (5)$$

zakładając jednocześnie, że:

$$E(x_g) = E(\tau_g) + \Lambda_{xg} E(F_g) + E(\varepsilon_g) = \tau + \Lambda_{xg} \kappa_g \quad (6)$$

gdzie:

κ_g – wektor przeciętnych wartości zmiennej F w grupie g (ang. vector of factor means for group g).

Ponadto do modelu włącza się warunki dodatkowe gwarantujące, że:

- W1. zmienna kwantyfikująca zjawisko latentne przyjmuje formę indeksu ilustrującego intensywność tego zjawiska w odniesieniu do poziomu obserwowanego w grupie referencyjnej (metoda grupy odniesienia, ang. *reference group method*);
- W2. zmienną kwantyfikującą zjawisko latentne cechuje zgodność pomiaru⁶⁹, tzn. zgodność definicyjna, zgodność skali oraz zgodność punktu referencyjnego skali, która jest niezbędna do porównywania wartości zmiennych ukrytych między grupami.

W wyniku spełnienia warunku W1 wartość oczekiwana każdej zmiennej ukrytej w modelu w pierwszej z grup ustalana jest na poziomie zero, zaś ładunek czynnikowy jednego wskaźnika dla każdej zmiennej ukrytej lub wariancja każdej zmiennej ukrytej przyjmują wartość 1 (ang. *reference group method*)⁷⁰.

Spełnienie warunku W2 wymaga zapewnienia zgodności pomiaru, czyli jednocześnie:

1. zgodności definicyjnej,
2. zgodności skali,
3. zgodności punktu referencyjnego.

Zgodność definicyjna w literaturze przedmiotu opisywana jest jako identyczność konceptualizacji i operacjonalizacji mierzonych zjawisk. Zgodność skali oznacza, że zmienna ukryta ma to samo znaczenie w badanych grupach. Natomiast zgodność punktu referencyjnego oznacza, że obserwowane między grupami różnice w wartościach zmiennych wskaźnikowych są wynikiem różnic w wartościach mierzonej zmiennej ukrytej oraz że skwantyfikowana zmienna ukryta w każdej z grup ma nie tylko tą samą skalę, ale i ten sam punkt referencyjny (wspólne zero).

W literaturze, jak i w praktyce, wskazuje się na dwie strategie postępowania odnoszące się do sprawdzania zgodności skali oraz zgodności punktu referencyjnego. W pierwszym przypadku najpierw sprawdza się istnienie zgodności skali, a następnie zgodność

⁶⁹ Ze względu na brak w języku polskim powszechnie stosowanego odpowiednika angielskiego terminu *measurement invariance* podajemy inne terminy pojawiające się w polskojęzycznej literaturze. Są to: zgodność pomiaru, niezmiennosc pomiaru, ekwiwalencja bądź ekwiwalentność pomiaru

⁷⁰ W analizie czynnikowej dla wielu grup oprócz metody grupy odniesienia stosuje się również metodę znacznika (ang. *marker-variable method*) oraz metodę kontrastu (ang. *effect-coding method*). Metoda znacznika polega na narzuceniu w modelu pomiarowym zmiennej ukrytej dwóch warunków ograniczających, mówiących, że wyraz wolny jednego ze wskaźników jest równy zero, zaś jego ładunek czynnikowy równy 1. Wartości wyrazów wolnych oraz ładunków czynnikowych pozostałych wskaźników pozostają swobodne, przy czym spełniają warunek równości między grupami. Natomiast metoda kontrastu polega na narzuceniu warunku ograniczającego mówiącego, że suma wyrazów wolnych wszystkich wskaźników wynosi zero, zaś średnia ładunków czynnikowych wszystkich wskaźników wynosi jeden.

punktu referencyjnego. W drugim przypadku rozpoczyna się od zgodności punktu referencyjnego, a w przypadku jej braku, „schodzi się” jeden poziom niżej – do zgodności skali.⁷¹

Występowanie zgodności skali oraz zgodności punktu referencyjnego stwierdza się na podstawie analizy statystyk dopasowania w modelach confirmacyjnej analizy czynnikowej dla wielu grup. Istnieje również inny sposób badania istnienia zgodności pomiaru. Wykorzystuje się nie tyle same statystyki dopasowania modelu, ile porównania ich wartości dla modeli z „różnych poziomów” zgodności pomiaru. Jak podaje Chen⁷², jeśli próba jest mała ($n \leq 300$), zaś liczebności grup nie są takie same, za brakiem zgodności skalarnej przemawiają jednoczesne różnice w wartościach:

1. $CFI \geq 0,005$ ($\Delta CFI \geq 0,005$),
2. $RMSEA \geq 0,01$ ($\Delta RMSEA \geq 0,01$) lub $SRMR \geq 0,025$ ($\Delta SRMR \geq 0,025$),

zaś za brakiem zgodności punktu referencyjnego:

1. $CFI \geq 0,005$ ($\Delta CFI \geq 0,005$),
2. $RMSEA \geq 0,01$ ($\Delta RMSEA \geq 0,01$) lub $SRMR \geq 0,005$ ($\Delta SRMR \geq 0,005$).

Jeśli próba jest odpowiednio duża ($n > 300$) oraz grupy są równoliczne, kryteria są ostrzejsze. Za brakiem zgodności skalarnej przemawiają jednoczesne różnice w wartościach:

1. $CFI \geq 0,01$ ($\Delta CFI \geq 0,01$),
2. $RMSEA \geq 0,015$ ($\Delta RMSEA \geq 0,015$) lub $SRMR \geq 0,03$ ($\Delta SRMR \geq 0,03$),

zaś za brakiem zgodności punktu referencyjnego:

1. $CFI \geq 0,01$ ($\Delta CFI \geq 0,01$),
2. $RMSEA \geq 0,015$ ($\Delta RMSEA \geq 0,015$) lub $SRMR \geq 0,01$ ($\Delta SRMR \geq 0,01$).

W przypadku braku spójnych wyników, Chen⁷³ sugeruje kierowanie się w pierwszej kolejności wartościami ΔCFI .

Obecnie coraz częściej stosuje się jednak podejście, które głosi, że jedynie znaczące odchylenia od zgodności pomiaru powinny być brane pod uwagę. Przy czym do ich

⁷¹ Podejście pierwsze jest zwykle stosowane przez użytkowników programu AMOS, zaś podejście drugie – przez użytkowników programu Mplus.

⁷² Chen F. F., 2007, Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance, *Structural Equation Modeling*, 14(3), 464–504, s.501

⁷³ Chen F. F., 2007, Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance, *Structural Equation Modeling*, 14(3), 464–504, s.501

identyfikacji zalecana jest analiza mocy testu oraz wartości parametru EPC (ang. *Expected Parameter Change*) określającego oczekiwaną przybliżoną zmianę wartości analizowanego parametru w przypadku estymacji modelu bez żadnego warunku ograniczającego narzuconego na ten parametr.⁷⁴

Występowanie wszystkich trzech typów zgodności dla wszystkich zmiennych wskaźnikowych określane jest w literaturze przedmiotu mianem pełnej zgodności pomiaru (ang. *full measurement invariance*)⁷⁵. Warto jednak zauważyć, że w praktyce potwierdzenie pełnej zgodności bywa czasem niemożliwe. Niemniej jednak, jak pokazali Steenkamp i Baumgartner⁷⁶, do przeprowadzenia porównań zarówno przeciętnych wartości zmiennej ukrytej, jak i poziomów skorelowania między zmiennymi w modelu, wystarczy, aby wystąpiła zgodność częściowa (ang. *partial measurement invariance*). Oznacza ona, że do prowadzenia porównań wystarczy, aby równość ładunków czynnikowych i wyrazów wolnych (dla zmiennych wskaźnikowych ciągłych) lub progów skali odpowiedzi (dla zmiennych wskaźnikowych kategoryalnych) zachodziła dla dwóch zmiennych wskaźnikowych⁷⁷.

Badanie zgodności pomiaru stosuje się zarówno (1) w przypadku analiz przekrojowych, jak i (2) w przypadku analizy danych wzdłużnych. W pierwszym podejściu mówi się o porównaniach przeciętnych wartości zmiennej latentnej między wyróżnionymi grupami respondentów – np. między kobietami i mężczyznami lub między grupami wyróżnionymi ze względu na status na rynku pracy (pracujący, bezrobotny, bierny zawodowo). Przykład takiego zastosowania przedstawili Davidov, Schmidt i Schwartz⁷⁸. Autorzy ci weryfikowali, czy między krajami biorącymi udział w badaniu Europejski Sondaż Społeczny (European Social Survey – ESS) zachodzi zgodność pomiaru dla teorii podstawowych wartości ludzkich Schwartza (1992) (ang. *theory of basic human values*), która jest operacjonalizowana i mierzona w tym badaniu. Do jej operacjonalizacji stosuje się 21 stwierdzeń, które mierzą ważność 10 podstawowych wartości ludzkich. Autorzy za pomocą confirmacyjnej analizy czynnikowej dla wielu grup sprawdzali, czy znaczenie poszczególnych wartości definiowanych przez teorię Schwartza jest ekwiwalentne między krajami. Jest to konieczne, bo tylko potwierdzenie takiej ekwiwalencji umożliwia analizowanie i porównywanie między krajami związków między wartościami a postawami,

⁷⁴ Taką analizę przeprowadza się wykorzystując program JRULE.

⁷⁵ Davidov E. (2008) A Cross-Country and Cross-Time Comparison of the Human Values Measurements with the Second Round of the European Social Survey, "Survey Research Methods", Vol.2, No.1, 33-46.

⁷⁶ Steenkamp J-B., Baumgartner H., 1998, Assessing Measurement Invariance in Cross-National Consumer Research, "The Journal of Consumer Research", Vol. 25, No. 1 (Jun., 1998), 78-90.

⁷⁷ Steenkamp J-B., Baumgartner H., 1998, Assessing Measurement Invariance in Cross-National Consumer Research, "The Journal of Consumer Research", Vol. 25, No. 1 (Jun., 1998), 78-90.

⁷⁸ Davidov E., Schmidt P., Schwartz S. H., Bringing values back in the adequacy of the European Social Survey to measure values in 20 countries, Public Opinion Quarterly, Vol. 0, No. 0, 2008, pp. 1-26

zachowaniami oraz charakterystykami społeczno-demograficznymi respondentów. Wyniki uzyskane przez Davidova, Schmidta i Schwartza potwierdzają istnienie takiej porównywalności dla rundy 2002-2003 dla 14 z 25 krajów biorących udział w badaniu.

W podejściu drugim analizuje się zmiany intensywności zjawiska latentnego mierzonego za pomocą zmiennej ukrytej w czasie. Takie zastosowanie konfirmacyjnej analizy czynnikowej dla wielu grup znaleźć można m.in. u Davidova⁷⁹ oraz u Węziak-Białowolskiej⁸⁰.

Davidov⁸¹ powtórzył badania dotyczące zgodności pomiaru podstawowych wartości ludzkich Schwartza między krajami biorącymi udział w I (runda z roku 2002-2003) oraz II rundzie badania ESS (runda z roku 2004-2005). Weryfikował, czy występuje zgodność skali oraz zgodność punktu referencyjnego osobno dla każdego z 19 krajów biorących udział w obu rundach. Wyniki, które uzyskał, pokazują, że w 9 krajach zgodność taka występuje, natomiast w 10 pozostałych krajach pierwotne modele konceptualne dla teorii wartości ludzkich Schwartza wymagały modyfikacji. Niemniej jednak po wprowadzeniu niezbędnych zmian w modelach konceptualnych udało się potwierdzić istnienie zgodności pomiaru dla wszystkich 19 krajów. Wniosek z badań Davidova jest taki, że dla tych 19 krajów można analizować zmiany w wyznawanych wartościach, które zaszły między analizowanymi rundami badania.

Węziak-Białowolska⁸² podjęła próbę zweryfikowania, czy podejście do mierzenia kapitału społecznego proponowane przez van Oorschota, Artsa i Gelissena⁸³ może być zastosowane w warunkach polskich. Podejście zastosowane przez autorkę obejmowało (1) przedstawienie modelu konceptualnego kapitału społecznego, (2) weryfikację, czy wyselekcjonowane zmienne wskaźnikowe mają odpowiednie własności, aby tworzyć wskaźniki syntetyczne, (3) budowę modeli pomiarowych dla wymiarów kapitału społecznego za pomocą konfirmacyjnej analizy czynnikowej oraz (4) analizę zmian ich poziomu w okresach 2005 – 2007 oraz 2007 – 2009. Krok czwarty możliwy był do realizacji tylko dzięki potwierdzeniu istnienia pełnej zgodności skalarnej dla wszystkich wymiarów kapitału

⁷⁹ Davidov E., 2008, A Cross-Country and Cross-Time Comparison of the Human Values Measurements with the Second Round of the European Social Survey, "Survey Research Methods", Vol.2, No.1, 33-46.

⁸⁰ Węziak-Białowolska D., Kapitał społeczny w Polsce w świetle Diagnozy Społecznej – pomiar i wyniki, „Studia Demograficzne”, w druku

⁸¹ Davidov E., 2008, A Cross-Country and Cross-Time Comparison of the Human Values Measurements with the Second Round of the European Social Survey, "Survey Research Methods", Vol.2, No.1, s. 33-46.

⁸² Węziak-Białowolska D., Kapitał społeczny w Polsce w świetle Diagnozy Społecznej – pomiar i wyniki, „Studia Demograficzne”, w druku.

⁸³ van Oorschot W., Arts W., Gelissen J., 2006, Social Capital in Europe. Measurement and Social and Regional Distribution of a Multifaceted Phenomenon, "Acta Sociologica", Vol. 49(2), 149-167.

społecznego. Istnienie pełnej zgodności skalarnej oznaczało stałość w czasie konceptualizacji mierzonych zjawisk, a tym samym trafność proponowanego do ich pomiaru podejścia.

Literatura

- Adrian B., 2003, Klimat szkoły, w: Encyklopedia Pedagogiczna XXI wieku, T. Pilch (red.), Wydawnictwo Akademickie Żak, Warszawa, s. 609-613.
- Arends R., 1994, Uczymy się nauczać, WSiP, Warszawa.
- Białowolski P., Węziak-Białowolska D., 2011, Zastosowanie analizy klas ukrytych do identyfikacji wykluczenia gospodarstw domowych z rynku kredytowego „Roczniki Kolegium Analiz Ekonomicznych”, Nr 23/2011, s.159 – 174
- Bond T.G., Fox Ch. M., Applying The Rasch Model. Fundamental Measurement in the Human Science, Lawrence Erlbaum Associates, Publishers Mahaw, New Jersey 2001.
- Brown T. A., 2006, Confirmatory Factor Analysis for Applied Research, The Guilford Press, New York, London.
- Chen F. F., 2007, Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance, Structural Equation Modeling, 14(3), 464–504
- Churchill G. A., Badania marketingowe. Podstawy metodologiczne, Wydawnictwo Naukowe PWN, Warszawa 2002.
- Coombs C. H., Dawes R.M., Tversky A., Wprowadzenie do psychologii, PWN, Warszawa 1977.
- Davidov E., 2008, A Cross-Country and Cross-Time Comparison of the Human Values Measurements with the Second Round of the European Social Survey, “Survey Research Methods”, Vol.2, No.1, 33-46.
- Freiberg H.J. (red.) (1999) School Climate: Measuring, Improving and Sustaining Healthy Learning Environments, Falmer Press, London.
- Górniak J., 2000, My i nasze pieniądze. Kraków, Aureus.
- Green E. K., Applications of the Rasch Model to Evaluation of Survey Data Quality, “New Directions for Evaluation”, No. 70, Jossey-Bass Publishers, 1996.
- Green, Kathy E. i Catherine G. Frantom. 2000. Survey development and validation with the Rasch Model, A paper presented at the International Conference on Questionnaire Development, Evaluation, and Testing. http://www.jpism.umd.edu/qdet/final_pdf_papers/green.pdf
- Kaczmarczyk S., Badania marketingowe: metody i techniki, Warszawa 2002, PWE.
- Kotowska, Irena E. i in. 2003. Polityka ludnościowa – cele, rozwiązania, opinie. Raport z badania finansowanego przez KBN, 5H02B 020 20. Warszawa.
- Mazur J., Rószkiewicz M., Strzyżewska M., 2005, Orientacja na wiedzę a wyniki ekonomiczne przedsiębiorstwa. Konceptualizacja badań, „Współczesne Zarządzanie”, nr 1/2005, p. 5-23.
- Nachmias D, Frankfort-Nachmias Ch., Metody badawcze w naukach społecznych, Wydawnictwo Zysk i Spółka, Poznań 2001.
- OECD 2008, Handbook on Constructing Composite Indicators. Methodology and User Guide.

- Ostasiewicz W., Istota pomiaru statystycznego, [w:] Pomiar statystyczny, red. W. Ostasiewicz, Wydawnictwo Akademii Ekonomicznej im. Oskara Lanego we Wrocławiu, Wrocław 2003.
- Rószkiewicz M., Metody ilościowe w badaniach marketingowych, Wydawnictwo Naukowe PWN, Warszawa 2002.
- Rószkiewicz M., Oszczędzanie. Postawy i zachowania polskich gospodarstw domowych wobec oszczędzania, Aureus, Kraków 2008.
- Sagan A., Badania marketingowe. Podstawowe kierunki. Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2004.
- Saris, W. E., Gallhofer, I. N., 2007, Design, Evaluation, and Analysis of Questionnaires for Survey Research. Hoboken, NJ: Wiley
- Skrondal A., Rabe-Hesketh S., Generalized Latent Variable Modeling. Multilevel. Longitudinal, and Structural Equation Models Chapman&Hall/CRC, 2004
- Steenkamp J-B., Baumgartner H., 1998, Assessing Measurement Invariance in Cross-National Consumer Research, "The Journal of Consumer Research", Vol. 25, No. 1 (Jun., 1998), 78-90.
- Sztabiński F., Logika badacza i logika respondenta. Problem adekwatności narzędzia badawczego [w:] ASK Społeczeństwo Badania Metody 12/2003, red. Domański H., Wydawnictwo Instytutu Filozofii i Socjologii PAN, Warszawa 2003, 51-83.
- Śliwerski B., Milerski B., 2000, Leksykon Pedagogika, Wydawnictwo Naukowe PWN, Warszawa.
- van Oorschot W., Arts W., Gelissen J., 2006, Social Capital in Europe. Measurement and Social and Regional Distribution of a Multifaceted Phenomenon, "Acta Sociologica", Vol. 49(2), 149-167.
- Węziak D., Indeks stanu posiadania dóbr trwałego użytku przez gospodarstwa domowe w świetle badań kapitału intelektualnego Lubelszczyzny, "Wiadomości Statystyczne", nr 1/2007
- Węziak D., Ocena jakości skali ze szczególnym uwzględnieniem jej rzetelności i trafności za pomocą skalowania Rascha, „ASK Społeczeństwo Badania Metody”, Nr. 15, 2006, s. 107-131
- Węziak D., Zastosowanie porządkowego skalowania Rascha do optymalizacji długości skali odpowiedzi, „Zeszyty Naukowe Akademii Ekonomicznej w Poznaniu, Ilościowe i jakościowe metody badania rynku”, Nr. 71, 2006, s.137-146
- Węziak D., Zastosowanie skalowania Rascha do oceny trafności skali pomiarowej, „Statystyka w badaniach społecznych – Prace Naukowe Nr 1142 Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Nr. 1142, 2006, s.34-48
- Węziak-Białowolska D., Kotowska I. E., Pomiar kapitału ludzkiego i jego zróżnicowanie według cech demograficznych, społecznych i ekonomicznych, [w:] Diagnoza Społeczna 2009, (red.). J. Czapiński, T. Panek, 2009
- Węziak-Białowolska D., Kotowska I. E., Pomiar kapitału ludzkiego i jego zróżnicowanie według cech demograficznych, społecznych i ekonomicznych, [w:] Rynek pracy i wykluczenie społeczne w kontekście percepcji Polaków – Diagnoza Społeczna 2009; Raport tematyczny, (red.) Kotowska I. E, Warszawa 2009

- Węziak-Białowska D., Model kapitału intelektualnego regionu. Koncepcja pomiaru i jej zastosowanie, Oficyna Wydawnicza SGH, Warszawa 2010
- Węziak-Białowska D., Kapitał społeczny w Polsce w świetle Diagnozy Społecznej – pomiar i wyniki, „Studia Demograficzne”, w druku
- Wojnarowska-Sołdan M., Węziak-Białowska D., Narzędzia do badania klimatu społecznego w szkole promującej zdrowie i ich analiza psychometryczna, „Edukacja, Studia, Badania, Innowacje”, nr 3 (111), 2010, s. 71-80