

Elżbieta Sobczak

SKALE POMIARU W MIĘDZYNARODOWYCH BADANIACH MARKETINGOWYCH

1. Wstęp

Opracowanie skal pomiaru dla międzynarodowych badań marketingowych jest przedsięwzięciem bardzo trudnym i złożonym. Badacze zajmujący się tą problematyką muszą rozważyć kwestię równoważności pomiarów przeprowadzonych w wielu krajach, poprawności technik pomiarowych stosowanych do gromadzenia danych oraz związanej z tym rzetelności i trafności danych.

Istotnym celem badań międzynarodowych jest poszukiwanie podobieństw i różnicowań występujących w krajach poddanych analizie. Jest to całkowicie uwarunkowane porównywalnością uzyskanych danych. Wskaźniki marketingowe podlegające pomiarowi nie zawsze wyrażone są w ten sam sposób, ponadto nie wszystkie metody i instrumenty pomiarowe mogą być stosowane z podobną skutecznością w różnych krajach. Stąd też trudne a zarazem konieczne jest określenie, czy zaobserwowane podobieństwa lub różnice występują w rzeczywistości, czy są jedynie skutkiem niewłaściwego pomiaru. W praktyce badań międzynarodowych można zaobserwować różne podejścia do tego typu problemów.

Celem tego opracowania jest analiza typów skal pomiaru pod względem możliwości ich wykorzystania w badaniach międzynarodowych oraz sposobów oceny ich przydatności. Scharakteryzowano również alternatywne podejścia do opracowania wielowymiarowych skal pomiaru.

2. Typy skal pomiaru i możliwości ich wykorzystania w badaniach międzynarodowych

Pomiar to proces przypisywania określonych symboli lub liczb cechom obiektów, z zachowaniem relacji zachodzących między obiektami i działań dokonywa-

nych na tych obiektach [12, s. 7-9; 3, s. 51; 1, s. 226; 19, s. 19; 8, s. 27; 6, s. 403]. W sposób bardziej sformalizowany można określić pomiar jako proces homomorficznego przekształcenia empirycznego systemu relacyjnego w liczbowy system relacyjny. Homomorfizm to pewna funkcja zachowująca struktury relacyjne zbiorów obiektów i liczb [10, s. 32; 4, s. 20; 17, s. 17].

Pomiaru można dokonać, stosując określoną miarę zwaną skalą pomiaru. Wyróżnia się wiele skal pomiarowych, jednak najczęściej stosowane w badaniach marketingowych są cztery spośród nich: nominalna, porządkowa, przedziałowa i ilorazowa. Wymienione skale można podzielić na metryczne i niemetryczne w zależności tego, czy dotyczą cech niemierzalnych (jakościowych), czy mierzalnych (ilościowych). Skale nominalna i porządkowa uznawane są za skale niemetryczne, pozostałe skale są metryczne [15, s. 80].

Skale niemetryczne nazywane są słabymi, skale metryczne silnymi. Kryterium takiego podziału jest zawartość informacyjna wyników pomiarów oraz zbiór dopuszczalnych transformant. Wszystkie przekształcenia możliwe w skali słabszej są również dopuszczalne w skali silniejszej, przeciwna zależność nie zachodzi.

Skala nominalna jest najprostszym a zarazem najmniej precyzyjnym typem skali. W hierarchii skal pomiaru jest to skala najniższa, ponieważ korzystając z niej, uzyskuje się najmniej informacji o stosunkach między obiektami, których dotyczy pomiar.

Jest to skala niemetryczna służąca do pomiaru wartości cech jakościowych badanych obiektów. W skali nominalnej cechy obiektów nie są mierzone w ścisłym tego słowa znaczeniu, lecz jedynie identyfikowane. Symbole przyporządkowywane cechom obiektów mierzonych w tej skali pełnią zatem funkcję identyfikatorów umożliwiających ich rozróżnianie. Jeżeli do identyfikacji stosowane są liczby, służą one jedynie do wyróżnienia klas obiektów, nie ma żadnego związku między empirycznymi relacjami obiektów a matematycznymi stosunkami liczb oznaczających poszczególne klasy [15, s. 81; 5, s. 18].

Ze względu na dopuszczalne przekształcenia można w tej skali określać tylko relacje równości i różności, symetryczności i przechodniości [2, s. 28]. Skala nominalna nie ma własności wartościujących, niedopuszczalne jest w niej wykonywanie takich operacji arytmetycznych, jak dodawanie, odejmowanie, mnożenie czy dzielenie. W skali tej możliwe jest tylko zliczanie relacji tego samego rodzaju, tzn. równości i różności, określenie częstości ich występowania oraz sortowanie obiektów ze względu na zachodzące między nimi relacje. Jedyną miarą statystyczną, jaką można wyznaczyć, posługując się nominalną skalą pomiaru, jest dominanta, czyli relacja najczęściej występująca. Konsekwencją wymienionych właściwości skali nominalnej są ograniczone możliwości stosowania metod wielowymiarowej analizy statystycznej do wyników pomiarów dokonywanych w tej skali.

Szczególnym rodzajem skali nominalnej jest skala dychotomiczna, w której można rozróżniać i klasyfikować obiekty opisane za pomocą cech binarnych, zwa-

nych dychotomicznymi. Są to zmienne skokowe przyjmujące jedynie dwie wartości. Jeżeli wartościom zmiennych dychotomicznych przypisane zostaną symbole 0 i 1, noszą one nazwę zmiennych zero-jedynkowych.

Pomiary dokonywane z wykorzystaniem skali nominalnej stanowią najmniej skomplikowany i często stosowany sposób gromadzenia informacji w międzynarodowych badaniach marketingowych. Skala nominalna stanowi najmniejsze obciążenie dla respondenta, ponieważ jest najprostszym rodzajem pomiaru, właściwym dla respondentów o niskim poziomie wykształcenia bądź nieumiejących czytać lub pisać. Skala ta wymaga jednak, by definicja określonej cechy była jednoznaczna i znana respondentom pochodzącym z różnych krajów. Wówczas respondent określa jedynie, czy dana właściwość go dotyczy, czy też nie [7, s. 258].

Kolejną skalą niemetryczną jest skala porządkowa. Umożliwia ona pomiar na poziomie porównywania względnego, wprowadza bowiem uporządkowanie obiektów, polegające na rangowaniu, tj. przypisywaniu obiektom kolejnych liczb naturalnych w porządku rosnącym lub malejącym, wg stopnia nasilenia danej cechy. Skala ta umożliwia porównywanie obiektów typu większy, mniejszy lub równy, natomiast nie można, korzystając z niej, określić, o ile jeden obiekt różni się od drugiego, ponieważ nie występują tutaj jednostki miary. Możliwe jest obliczenie średnich pozycyjnych: mediany, kwartyli, percentyli [15, s. 81; 16, s. 35].

Pomiary porządkowe są stosunkowo prostym sposobem gromadzenia danych marketingowych w przekroju różnych krajów. W przypadku respondentów o niskim poziomie wykształcenia ustalenie przez nich rankingu obiektów lub wyrażenie preferencji w kategoriach rankingu porządkowego może być ułatwione dzięki wykorzystaniu, poza słowem pisanym, innego typu bodźców, np. obrazków, fotografii.

Jednak praktyczne wykorzystanie skali porządkowej w międzynarodowych badaniach marketingowych jest ograniczone ze względu na często występujący brak porównywalności zestawu obiektów porządkowanych w różnych krajach. Zgromadzone dane porządkowe mogą być wykorzystywane do celów komparatywnych tylko wówczas, gdy identyczny zestaw obiektów istnieje w różnych krajach. Uporządkowania (rankingi) różnych zbiorów obiektów nie mogą bowiem tworzyć wspólnej skali. Spełnienie tego założenia w przekroju wielu krajów napotyka na ogromne trudności [7, s. 258].

Skala przedziałowa jest skalą metryczną, pozwala na bezpośrednie porównywanie obiektów przez ustalenie odległości między nimi ze względu na analizowaną cechę. W tej skali ustalone są jednostki miary, dopuszczalne jest dodawanie i odejmowanie oraz określanie średniej arytmetycznej.

Udzielanie odpowiedzi na skali przedziałowej jest trudniejsze niż w przypadku skali nominalnej i porządkowej, wymaga wyższego poziomu wykształcenia i pewnego doświadczenia respondentów. Ponadto w badaniach międzynarodowych skale składające się z wielu kategorii (np. pięciu lub siedmiu) zazwyczaj są traktowane jako skale przedziałowe, mimo że uzyskane dane nie mają odpowiednich właści-

wości. Brak jest jednostek miary i wykonywanie wielu operacji matematycznych jest niedopuszczalne. Tego typu badania są realizowane w praktyce, jednak porównywalność danych w przekroju różnych krajów budzi wówczas sporo zastrzeżeń [7, s. 259].

Skala ilorazowa umożliwia pomiar na poziomie proporcjonalności obiektów ze względu na daną cechę. Występuje tutaj tzw. zero bezwzględne, dlatego każdą wartość można przedstawić jako wielokrotność innej wartości. Dopuszczalne są wszystkie operacje arytmetyczne.

Ten rodzaj skali jest rzadko używany do gromadzenia danych od indywidualnych respondentów, ze względu na duże trudności w formułowaniu odpowiedzi, a także zapewnieniu ich porównywalności w międzynarodowych badaniach marketingowych.

3. Metody oceny przydatności skal pomiarowych w międzynarodowych badaniach marketingowych

Dane marketingowe otrzymane w wyniku zastosowania opracowanych skal pomiaru są obciążone w mniejszym lub większym stopniu błędem. Wykorzystanie ich w badaniach marketingowych powinno być poprzedzone badaniem wiarygodności. Badacz bowiem musi mieć zaufanie do uzyskanych wyników pomiarów, by móc na ich podstawie dokonywać poprawnych analiz i wnioskowania.

W badaniach marketingowych rynku krajowego wyróżnia się zazwyczaj dwa aspekty wiarygodności pomiaru: rzetelność i trafność [11; 18]. W badaniach międzynarodowych, cechujących się większym stopniem trudności i kompleksowości, wynikającymi z bardziej zróżnicowanych uwarunkowań ekonomicznych, prawnych, kulturowych itp., w jakich dokonuje się gromadzenia danych, ocena wiarygodności pomiaru powinna być bardziej rozbudowana.

Można wyodrębnić następujące kryteria oceny przydatności skal pomiaru w przekroju międzynarodowym [7; 14]:

1. Równoważność instrumentu pomiaru.
2. Rzetelność skali.
3. Trafność skali.
4. Możliwość uogólnień.

Międzynarodowe badania marketingowe mają charakter porównawczy. Konieczne jest zatem stwierdzenie, czy zaobserwowane podobieństwa i różnice występujące między poszczególnymi krajami są rzeczywiste, czy wynikają jedynie z nierównoważności pomiaru. Kwestia równoważności pomiaru jest w badaniach międzynarodowych nadrzędna, ponieważ jej brak może skutkować zarówno nierzetelnością, jak i nietrafnością pomiaru [14].

Badanie równoważności polega na sprawdzeniu międzynarodowej porównywalności pomiarów, stanowiącej podstawowe uwarunkowanie poprawności wnioskowania. Na ogólną ocenę równoważności instrumentu pomiaru w przekroju wielu krajów składają się oceny cząstkowe [7]:

- równoważności badanej cechy,
- równoważności jednostek miary,
- równoważności pozycji skali.

Równoważność cechy oznacza, że w przekroju badanych krajów cecha taka występuje i ma ekwiwalentne znaczenie. Równoważność jednostek miary dotyczy jedynie skal przedziałowej i ilorazowej. Różne kraje na świecie odnoszą się zwyczajowo do odmiennych jednostek miary, dotyczy to m.in. pomiaru wagi, odległości, wartości wyrażonych w jednostkach pieniężnych itp. Badacz powinien zadbać o jednolitość stosowanych jednostek miary.

Elementami każdej skali są pojedyncze pozycje (kwestie, stwierdzenia), do których powinien odnieść się respondent. Ustalone pozycje skali łącznie mierzą daną cechę. W badaniach marketingowych rozróżnia się skale proste (jednopozycyjne, jednowymiarowe) i powszechniej stosowane skale złożone (wielopozycyjne, wielowymiarowe). Równoważność pozycji skali oznacza, że są one porównywalne w przekroju międzynarodowym, tzn. istnieją i funkcjonują w ten sam sposób w każdym z badanych krajów. Pozycje skali obciążone błędem systematycznym (funkcjonujące odmiennie w badanych krajach) powinny zostać usunięte.

W razie wykrycia nierównoważności instrumentu pomiaru badacz może [7]:

- wykluczyć porównania,
- zredukować nierównoważność,
- zinterpretować nierównoważność,
- zignorować nierównoważność.

Stosując pierwsze podejście, badacz może zdecydować, że międzynarodowe badania porównawcze nie mogą zostać zrealizowane z powodu nierównoważności pomiaru.

Podejście drugie jest najbardziej typowe w praktyce badań międzynarodowych. Stosowane jest w przypadku wykrycia nierównoważności pozycji skali i polega na usunięciu pozycji obciążonych błędem. Jeżeli jednak wiele pozycji skali cechuje się nierównoważnością, nasuwa to wątpliwość, czy istotnie zredukowana skala nadal jest reprezentatywna dla badanej cechy, a tym samym, czy może być stosowana w badaniach międzynarodowych.

Kolejną strategią badacza jest akceptacja nierównoważności jako nieuniknionej konsekwencji prowadzenia badań w więcej niż jednym kraju. Należy wówczas zinterpretować brak równoważności i wyjaśnić jej przyczyny.

Ostatnie podejście polegające na zignorowaniu nierównoważności pomiaru i przyjęciu założenia o wzajemnym znoszeniu się nierównoważności jest najbardziej ryzykowną, kontrowersyjną oraz budzącą wątpliwości natury teoretycznej strategią badacza.

Rzetelność skali oznacza konsekwencję pomiaru danej cechy. Odnosi się do stopnia, w jakim skala daje zgodne wyniki przy dokonaniu powtórnego pomiaru. Badanie stopnia rzetelności skali pomiaru polega na ocenie [7; 14; 18]:

- zgodności (stabilności) w czasie,
- zgodności niezależnych ocen,
- zgodności wewnętrznej (spójności) skali.

Im bardziej zgodne są pomiary, tym wyżej oceniana jest ich rzetelność.

Stabilność pomiarów w czasie jest szacowana metodą powtarzania pomiaru (*test-retest*) [7; 9; 14; 18]. Polega ona na co najmniej dwukrotnym pomiarze cechy w tej samej próbie w różnych momentach czasu i ustaleniu korelacji między otrzymanymi wynikami. Odstęp czasu między kolejnymi pomiarami powinien wynosić 2-5 tygodni. Przyjmuje się, że wysoka korelacja otrzymanych wyników potwierdza stabilność pomiaru i rzetelność stosowanej skali. Metoda ta obciążona jest jednak następującymi wadami [7; 14]:

– otrzymany wynik w dużym stopniu jest uwarunkowany okresem, jaki upłynął między pomiarami, zbyt krótki okres może odzwierciedlać jedynie konsekwencję respondentów, pamiętających poprzednią wypowiedź, i powodować przeniesienie jej z pierwszego do drugiego testu, okres zbyt długi może skutkować rzeczywistą zmianą wypowiedzi pod wpływem okoliczności zewnętrznych;

– pierwszy pomiar może uczulić respondenta na temat poddany badaniom i doprowadzić do zmiany postawy;

– niektóre cechy mogą być mierzone tylko raz, co uniemożliwia stosowanie tej metody.

Badanie zgodności niezależnych ocen przeprowadza się, stosując tzw. metodę pomiarów równoległych. Obejmuje ona dwukrotne badanie tej samej grupy osób dwoma równoważnymi formami skali opracowanymi przez ekspertów, a następnie ustalenie korelacji między nimi. Powinno to zostać poprzedzone oceną równoważności instrumentów pomiaru.

W międzynarodowych badaniach marketingowych zarówno *test-retest*, jak i metoda pomiarów równoległych są stosowane relatywnie rzadko, ponieważ już jednokrotne przeprowadzenie pomiarów w przekroju wielu krajów jest bardzo kosztowne i czasochłonne.

Wewnętrzna zgodność bądź spójność pozycji skali oznacza ich wzajemnie skorelowanie. Ponieważ pozycje skali odzwierciedlają poszczególne aspekty danej cechy, powinny pozostawać w istotnej współzależności. Do podstawowych metod oceny spójności należą [7; 9; 18]:

- test połówkowy (*split-half*),
- współczynnik α -Cronbacha,
- współczynnik *KR-20* Kudera-Richardsona.

Metoda połówkowa polega na podziale pozycji skali na połowy i zastosowaniu procedury pomiaru dla dwóch zestawów odpowiedzi. Zazwyczaj do jednego

zestawu przydziela się parzyste pozycje skali, a do drugiego nieparzyste. Następnie oblicza się miarę zgodności między otrzymanymi wynikami, która informuje o rzetelności połówek skali. Aby oszacować rzetelność całej skali, można zastosować współczynnik korelacji Spearmana-Browna opisany poniższym wzorem [9, s. 494]:

$$r_{SB} = \frac{2r_{hh}}{1 + r_{hh}}, \quad (1)$$

gdzie r_{SB} oznacza miarę rzetelności całej skali, r_{hh} – miarę rzetelności połowy skali.

Współczynnik r_{SB} określa, jaka byłaby rzetelność skali, gdyby każda z połówek była dwa razy dłuższa.

W badaniach międzynarodowych można traktować dwa kraje jako połówki, tzn. uwzględnić parzyste pytania z jednego kraju i nieparzyste z drugiego, następnie proces ten powtórzyć dla nieparzystych pytań z pierwszego kraju i parzystych z drugiego. Otrzymane wyniki porównuje się, obliczając korelację między nimi. Te wyniki są następnie porównane z otrzymanymi metodą połówkową dla parzystych i nieparzystych pozycji skali wewnątrz każdego kraju. Im wyższe są korelacje między nimi, tym wyżej oceniana jest rzetelność pomiarów otrzymanych w każdym kraju [7].

Metoda połówkowa może być stosowana w przypadku dostatecznie długiej i szczegółowej skali, wynoszącej ok. 20 pozycji. Ponadto test połówkowy jest obciążony pewną wadą, mianowicie ocena wewnętrznej zgodności pozycji skali zależy od sposobu podziału skali na połowy.

Następne wymienione metody – współczynnik α -Cronbacha oraz współczynnik $KR-20$ Kudera-Richardsona – są pozbawione tej wady, ponieważ bazują na średniej zgodności wszystkich możliwych podziałów skali na połowy. Pierwsza z wymienionych miar jest właściwa dla pomiarów dokonywanych z wykorzystaniem skal porządkowych, przedziałowych i stosunkowych, następna – dla skali dychotomicznej (zmienne zero-jedynkowe), stanowiącej szczególny przypadek skali nominalnej. Współczynnik α -Cronbacha i $KR-20$ Kudera-Richardsona przedstawiają poniższe formuły [9, s. 496]:

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k s_i^2}{s_x^2} \right), \quad (2)$$

gdzie $i = 1, 2, \dots, k$ (numer pozycji skali), s_i^2 oznacza wariancję i -tej pozycji skali, s_x^2 – całkowitą wariancję skali.

$$KR - 20 = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^k p_i q_i}{s_x^2} \right), \quad (3)$$

gdzie p_i oznacza proporcję odpowiedzi jednego typu (1 lub 0), q_i – proporcję odpowiedzi drugiego typu.

Współczynniki te przyjmują wartości z przedziału $\langle 0, 1 \rangle$. Skalę można ocenić jako rzetelną, jeżeli ich wartości przekraczają 0,6. Wewnętrzną spójność skali można poprawić, wzbogacając ją o kolejne pozycje lub dokonując ich redukcji przez usunięcie pozycji najslabiej skorelowanych z całą skalą.

W międzynarodowych badaniach marketingowych wymienione współczynniki wewnętrznej spójności skali są powszechnie stosowane.

Trafność pomiaru jest kolejnym kryterium oceny skali. Definiowana jest jako dokładność, z jaką skala mierzy określoną cechę. Określa stopień, w jakim różnice występujące między wynikami pomiarów odzwierciedlają rzeczywiste zróżnicowanie obiektów, a nie systematyczne lub przypadkowe błędy pomiaru. Ocenie należy poddać następujące rodzaje trafności [7; 14]:

- trafność treściową,
- trafność kryterium,
- trafność teoretyczną.

Trafność treściowa obejmuje subiektywną ocenę reprezentatywności treści skali dla pomiaru określonej cechy. Ze względu na subiektywny charakter, sama w sobie nie stanowi wystarczającej miary trafności.

Bardziej sformalizowaną ocenę skali można otrzymać, badając trafność kryterium, określającą związek między wynikami uzyskanymi za pomocą danej skali a wybranymi zmiennymi, stanowiącymi tzw. kryterium zewnętrzne. Funkcję kryterium mogą pełnić zmienne demograficzne, psychologiczne, dotyczące postaw lub zachowań, a nawet wyniki pomiarów otrzymanych z wykorzystaniem innych skal. Ze względu na okres, w którym stosowane jest kryterium zewnętrzne, wyodrębnia się dwie formy trafności: trafność równoległą (diagnostyczną) i trafność prognostyczną. Z trafnością równoległą mamy do czynienia wówczas, gdy gromadzenie danych z wykorzystaniem ocenianej skali, dotyczących innych zmiennych, pełniących funkcję kryteriów zewnętrznych odbywa się w tym samym czasie. Natomiast dokonując oceny predyktywnej trafności skali, badacz gromadzi dane dotyczące kryterium zewnętrznego dopiero po upływie pewnego czasu. Jeżeli np. jest wysoka korelacja między postawami respondentów wobec określonego produktu a wielkością ich późniejszej sprzedaży, to skalę cechuje wysoka trafność prognostyczna [7; 18].

Trafność teoretyczna określa, czy skala w rzeczywistości mierzy cechę, którą powinna mierzyć. Ten rodzaj trafności jest najistotniejszy a zarazem najtrudniejszy do ustalenia. Na trafność teoretyczną składają się [7]:

- trafność zbieżna,
- trafność dyskryminacyjna,
- trafność nomologiczna.

Trafność zbieżna oznacza wysoką dodatnią korelację wyników uzyskanych za pomocą badanej skali i innych skal mierzących tę samą cechę.

Trafność dyskryminacyjna skali występuje w przypadku niskiej korelacji z wynikami pomiarów uzyskanymi za pomocą skal mierzących odmienne cechy. Przykładem trafności dyskryminacyjnej znanym z praktyki badań marketingowych jest relatywnie niska korelacja wyników otrzymanych za pomocą CETSCALE – skali mierzącej etnocentryzm konsumencki – i pomiarów takich cech, jak patriotyzm, konserwatyzm polityczno-ekonomiczny i dogmatyzm [7]. Trafność nomologiczna to silna dodatnia korelacja wyników pomiaru cechy z pomiarami cech pokrewnych. Badaniu takiemu poddano również skalę CETSCALE, badając jej korelację z takimi cechami pokrewnymi, jak: postawy, intencje zakupów i posiadanie produktów pochodzenia zagranicznego. Konsumenci cechujący się etnocentryzmem wykazali negatywne postawy wobec tego typu produktów, co potwierdziło trafność nomologiczną tej skali [7].

Między rzetelnością i trafnością zachodzą określone relacje, które można wyjaśnić, przyjmując, że każdy pomiar składa się z kilku elementów, co przedstawia poniższa formuła [14, s. 196]:

$$X_0 = X_t + X_s + X_r, \quad (4)$$

gdzie X_0 oznacza zaobserwowana wartość cechy, X_t – rzeczywistą wartość cechy, X_s – systematyczny błąd pomiaru, X_r – przypadkowy błąd pomiaru.

Pomiar jest trafny, jeżeli odzwierciedla rzeczywisty wynik, a zarówno błąd systematyczny jak i przypadkowy równają się zero. Jeżeli skala jest nierzetelna, to występuje błąd przypadkowy i wówczas nie może być trafna. Jeżeli natomiast skala jest rzetelna, to niekoniecznie musi być trafna, ponieważ wyniki pomiarów mogą być obciążone błędem systematycznym, np. wynikającym z nierównoważności pomiarów. Jak z tego wynika, rzetelność jest warunkiem koniecznym, ale nie dostatecznym trafności.

Kolejnym kryterium przydatności skali stosowanej w międzynarodowych badaniach marketingowych jest dopuszczalny stopień jej uogólnień, tzn. łatwość jej stosowania i interpretacji w różnych okolicznościach i warunkach badawczych. Jest to zdeterminowane przede wszystkim możliwością stosowania różnorodnych metod gromadzenia informacji i uwzględnienia zróżnicowanej grupy respondentów oraz warunków jej interpretacji [14].

Wnioskowanie i podejmowanie decyzji na podstawie pomiarów dokonywanych w ramach międzynarodowych badań marketingowych powinno być poprzedzone oceną ich wiarygodności. Jedynie wykorzystanie równoważnych, rzetelnych i trafnych skal pomiaru cechy warunkuje poprawność badania marketingowego.

4. Alternatywne podejścia do opracowania wielowymiarowych skal pomiaru w międzynarodowych badaniach marketingowych

Opracowanie wielowymiarowych skal pomiaru wskaźników umożliwiających analizy porównawcze stanowi poważne wyzwanie dla międzynarodowych badań marketingowych. Można wyodrębnić następujące podejścia do tego zagadnienia [7]:

1. Wykorzystanie istniejącej skali wielowymiarowej opracowanej w pojedynczym kraju (*single context scale*):

- standaryzacja skali (dominacja podejścia *etic*),
- dostosowanie skali do specyfiki badanych krajów (dominacja podejścia *emic*),
- kombinacja elementów *etic-emic*.

2. Opracowanie nowej skali zdecentralizowanej, wielokontekstowej (*multiple context scale*).

Podejścia oparte na istniejącej skali są najpowszechniej stosowane w praktyce. Ich podstawą jest przekonanie, że badane wskaźniki (np. postawy) są uniwersalne, a odpowiednio zmodyfikowana skala może je zmierzyć.

Podejścia te polegają na opracowaniu skali pomiaru w jednym kraju, przetłumaczeniu jej, ewentualnej adaptacji i skierowaniu do innych krajów objętych badaniami międzynarodowymi. Wymaga to spełnienia założeń, że [7]:

- a) wskaźnik występuje w innych krajach i jest istotny,
- b) może być mierzony z wykorzystaniem tego samego instrumentu w innych krajach lub kulturach.

Należy również sprawdzić, czy oryginalna skala opracowana w jednym kraju ma te same właściwości, kiedy jest stosowana w innym. W tym celu określa się, czy:

- konsumenci w innych krajach odpowiadają w zgodny sposób (rzetelność skali),
- analizowany wskaźnik jest odróżnialny od innych w pozostałych krajach (trafność dyskryminacyjna),
- skala mierzy ten wskaźnik w innych krajach (trafność nomologiczna).

Podejście to skupia się na badaniu stopnia dopasowania skali stworzonej w jednym kraju do innych krajów, przez ocenę poziomu jej rzetelności i trafności. Pozytywna ocena skali opracowanej w kraju macierzystym a zastosowanej w innych wpływa na decyzję o jej użyteczności. Poprzedza ją dopasowanie do odmiennych kultur przejawiające się zazwyczaj koniecznością zmniejszenia liczby pozycji skali wielowymiarowej, przez eliminację pozycji uznanych za błędne. Zredukowa-

nie skali do jej najistotniejszych elementów może uwydatnić jej moc i efektywność predyktywną, co bywa postrzegane jako zaleta takiego podejścia. Dostosowanie skali przez eliminowanie pozycji może poprawić jej rzetelność, ale ten sposób nie jest w stanie przekształcić jej tak, by pasowała do kultury innego kraju. Podejście to jest ukierunkowane na poszukiwanie i uwypuklanie podobieństw (dominacja *etic*). Jego stosowanie generuje powstawanie obciążeń wewnętrznych skali (*pseudo-etic bias*), polegających na wymuszaniu pewnej standaryzacji odpowiedzi respondentów pochodzących z różnych krajów.

Dopasowanie skali oryginalnej do innych kultur może polegać również na zwiększaniu liczby pozycji skali w celu identyfikacji wymiarów specyficznych uwarunkowanych kulturowo, co oznacza dominację podejścia *emic*.

Kombinacja podejść *etic-emic* daje badaczowi możliwość wyboru wariantu pośredniego oraz określenia proporcji zachodzących między elementami uniwersalnymi i unikatowymi.

Implementacja podejść bazujących na istniejącej skali wielowymiarowej opracowanej w pojedynczym kraju budzi wątpliwość natury teoretycznej, czy skala zbudowana od podstaw w innym kraju byłaby identyczna w stosunku do testowanej. Występuje tutaj bowiem dominacja kraju macierzystego, co może generować obciążenie skali błędem odmiennej kultury.

Jest to zarazem najczęściej stosowane podejście ze względu na powszechną dostępność skal, które zostały opracowane i zweryfikowane w pojedynczym kraju, często w USA.

Budowa skali wielokontekstowej (*multiple context*), zwanej również zdecentralizowaną, to odmienne podejście polegające na opracowaniu nowej skali dla danego wskaźnika w każdym badanym kraju. Zapewnia to pomiar skupiony na dostosowaniu się do uwarunkowań każdego kraju i unikanie pułapek standaryzacji (*etic*). Skale pomiaru opracowywane są niezależnie i w konsekwencji każda z nich umożliwia wgląd w unikatowe komponenty postawy konsumentów. Występuje tutaj przewaga podejścia *emic* uwypuklającego występujące zróżnicowania międzynarodowe. Na procedurę opracowania wielowymiarowej skali zdecentralizowanej składają się następujące kroki [7]:

1) wypracowanie wstępnej definicji wskaźnika poddawanego pomiarowi, przy współudziale badaczy pochodzących z różnych krajów,

2) generowanie wymiarów (zmiennych) i pozycji skali odrębnie w każdym kraju,

3) określenie wymiarów uniwersalnych i specyficznych kulturowo (jeżeli wymiary wspólne nie występują, nie należy prowadzić międzynarodowych badań porównawczych z wykorzystaniem tego wskaźnika),

4) analiza i porównanie skal zdecentralizowanych (odrębnych dla każdego z badanych krajów) przez:

– porównywanie krajów parami (na zasadzie każdy z każdym),

– ustalenie kraju docelowego i porównywanie wszystkich pozostałych do niego, następnie zmiana rotacyjna kraju docelowego (jeżeli wskaźnik jest uniwersalny, zmiana kraju docelowego powinna dać te same rezultaty),

– porównanie łącznej puli wymiarów skali w każdym kraju, by wyeliminować błędne zmienne, następnie analizę wewnątrz kraju dla identyfikacji wymiarów specyficznych.

Podejście to nie pozwala na dominację kultury jednego kraju. Jest czasochłonne, ale bardziej uzasadnione teoretycznie.

5. Wnioski

Opracowanie wiarygodnej skali pomiaru, którą można byłoby z powodzeniem stosować w wielu krajach, zróżnicowanych pod względem kulturowym i ekonomicznym, jest ogromnie trudne. Konieczne bowiem jest zarówno zapewnienie porównywalności, jak i uwzględnienie specyfiki poszczególnych krajów. Niniejsze opracowanie wypukla zróżnicowane znaczenie znanych skal pomiaru w międzynarodowych badaniach marketingowych, konieczność rozszerzonej weryfikacji ich przydatności oraz podejścia do budowy wielowymiarowych skal pomiaru.

W międzynarodowych badaniach marketingowych najpowszechniejsze zastosowanie ze względu na większą prostotę mają skale: nominalna, porządkowa i przedziałowa. Stosowanie skali porządkowej bywa jednak niemożliwe, gdy porządkowane obiekty są nieporównywalne w przekroju krajów.

Skala ilorazowa ze względu na znacznie wyższy stopień trudności i ograniczoną porównywalność jest w praktyce niezmiernie rzadko stosowana w analizach międzynarodowych do gromadzenia danych od indywidualnych respondentów.

W badaniach rynków zagranicznych ogólna ocena wiarygodności pomiaru wymaga nie tylko weryfikacji rzetelności i trafności (stosowanej w badaniach krajowych), ale również równoważności instrumentu pomiaru i możliwości jego uogólnień. Dopiero pozytywna ocena wszystkich wymienionych właściwości skali uprawnia do jej stosowania w badaniach porównawczych realizowanych w przekroju krajów.

Problemem specyficznym dla badań międzynarodowych jest sposób opracowania wielowymiarowej skali pomiaru. Zazwyczaj wykorzystuje się pewną skalę oryginalną skonstruowaną w jednym kraju, a następnie dostosowuje się ją do warunkowań występujących w innych krajach. Korzystniejszym podejściem jest opracowanie nowej skali wielokontekstowej, która nie jest (w przeciwieństwie do poprzedniej) obciążona błędem dominacji jednej kultury. Jest to sposób znacznie bardziej złożony i czasochłonny, dlatego rzadko stosowany w praktyce, mimo że bardziej uzasadniony teoretycznie.

Następną kwestią, którą należałoby rozważyć w dalszych badaniach, jest analiza i ocena wybranych metod wielowymiarowej analizy statystycznej przydatnych w procesie opracowywania międzynarodowej skali pomiaru postaw.

Literatura

- [1] Ackoff R.L., *Decyzje optymalne w badaniach stosowanych*, PWN, Warszawa 1969.
- [2] *Badania przestrzenne rynku i konsumpcji. Przewodnik metodyczny*, red. S. Mynarski, PWN, Warszawa 1992.
- [3] Bazarnik J. i in., *Badania marketingowe. Metody i oprogramowanie komputerowe*, FOGRA, Kraków 1992.
- [4] Bąk A., *Dekompozycyjne metody pomiaru preferencji w badaniach marketingowych*, AE, Wrocław 2004.
- [5] Chojnicki Z., Czyż T., *Metody taksonomii numerycznej w regionalizacji geograficznej*, PWN, Warszawa 1973.
- [6] Churchill G.A., *Badania marketingowe. Podstawy metodologiczne*, PWN, Warszawa 2002.
- [7] Craig C.S., Douglas S.P., *International Marketing Research*, John Wiley & Sons, Chichester 2000.
- [8] *Ekonometria przestrzenna*, red. A. Zeliaś, PWE, Warszawa 1991.
- [9] Ferguson G.A., Takane Y., *Analiza statystyczna w psychologii i statystyce*, PWN, Warszawa 2002.
- [10] Galas Z., Nykowski I., Żółkiewski Z., *Programowanie wielokryterialne*, PWE, Warszawa 1987.
- [11] Kaczmarczyk S., *Badania marketingowe. Metody i techniki*, PWE, Warszawa 2002.
- [12] Kaczmarczyk S., *Skalowanie i pomiar w badaniach marketingowych*, [w:] *Zastosowanie metod wielowymiarowych w badaniach segmentacji i selektywności rynku*, red. S. Mynarski, AE, Kraków 1999.
- [13] Kowal J., *Metody statystyczne w badaniach sondażowych rynku*, PWN, Warszawa 1998.
- [14] Kumar V., *International Marketing Research*, Prentice Hall, Englewood Cliffs 2000.
- [15] Mynarski S., *Praktyczne metody analizy danych rynkowych marketingowych*, Zakamycze, Kraków 2000.
- [16] Pociecha J., *Statystyczne metody segmentacji rynku*, AE, Kraków 1986.
- [17] *Pomiar statystyczny*, red. W. Ostasiewicz, AE, Wrocław 2003.
- [18] Sagan A., *Badania marketingowe. Podstawowe kierunki*, AE, Kraków 1998.
- [19] Walesiak M., *Metody analizy danych marketingowych*, PWN, Warszawa 1996.

MEASUREMENT SCALES IN INTERNATIONAL MARKETING RESEARCH

Summary

The objective of the hereby study is the analysis of measurement scales types with regard to opportunities of their utilization in international research, as well as ways for the assessment of their usefulness. The following issues were discussed: equivalence of measurements run in many countries, correctness of measurement techniques applied for data collection and referring to it reliability and validity of data. Alternative approaches to applying multi-item scales were presented.