

Magdalena Jaciow

Uniwersytet Ekonomiczny w Katowicach
e-mail: magdalena.jaciow@ue.katowice.pl

EMPIRYCZNA OCENA EKWIWALENCJI WYNIKÓW BADAŃ – UWAGI METODYCZNE I PRÓBA IMPLEMENTACJI

EMPIRICAL EVALUATION OF EQUIVALENCE OF RESEARCH RESULTS – METHODOLOGICAL REMARKS AND AN ATTEMPT OF IMPLEMENTATION

DOI: 10.15611/pn.2018.525.16

JEL Classification: D2, M31

Streszczenie: Najważniejszym kryterium oceny jakości pomiaru w badaniach marketingowych jest trafność. W badaniach o charakterze międzynarodowym (międzykulturowym) kategoria trafności obejmuje również kwestię ekwiwalencji wskaźników w różnych kontekstach społeczno-kulturowych. Zachowanie ekwiwalencji na etapie projektowania badania (płaszczyzna interpretacyjna ekwiwalencji) nie gwarantuje uzyskania porównywalnych wyników. Ekwiwalencję należy sprawdzić na płaszczyźnie pomiaru (empirycznej) za pomocą metod analizy statystycznej. Celem artykułu jest określenie poziomu ekwiwalencji modelu pomiarowego postaw konsumentów wobec zakupów online w siedmiu krajach Europy. Postawiono hipotezę, która zakłada istnienie najniższego poziomu ekwiwalencji dla wszystkich badanych krajów. Na podstawie wyników analiz statystycznych hipotezę odrzucono. Najniższy poziom ekwiwalencji uzyskano dla sześciu z siedmiu krajów, a ekwiwalencję częściową jedynie dla dwóch krajów.

Słowa kluczowe: ekwiwalencja, międzynarodowe badania międzynarodowe, pomiar.

Summary: The most important criterion for measuring the quality of measurement in marketing research is relevance. In international (intercultural) research, the relevance category also includes the question of the equivalence of indicators in various socio-cultural contexts. Maintaining equivalence at the design stage of the study (the inter-pretatory level of equivalence) does not guarantee obtaining comparable results. Equivalence should be checked on the measurement plane (empirical) using statistical analysis methods. The aim of the article was to determine the level of equivalence of the measurement model of consumer attitudes towards online shopping in seven European countries. A hypothesis was put forward that assumed the existence of the lowest level of equivalence for all the studied countries. Based on the results of statistical analyzes, the hypothesis was rejected. The lowest level of equivalence was obtained for six out of seven countries, and partial equivalence only for two of them.

Keywords: equivalence, international marketing research, measurement.

1. Wstęp

Fundamentalną część nauk społecznych stanowią badania porównawcze. Poszukiwanie podobieństw i różnic między rynkami, krajami, kulturami i społecznościami pozwala odkrywać i poznawać prawidłowości rynków i społeczeństw. Świadczy o tym wiele prac naukowych, w tym m.in.: [Miller 1960; Tyree i in. 1979; Oyen (ed.) 1990; Mills i in. 2006; Niessen i in. (eds.) 2013; Ragin 2014; Treiman 2013]. Korzyści płynące z badań porównawczych dostrzegane są także w naukach medycznych [Röhmel i Mansmann 1999], edukacyjnych [Umiejętności 2013] czy politycznych [Peters 1998].

O ile w badaniach najważniejszym kryterium oceny jakości pomiaru jest trafność (czyli ocena, w jakim stopniu wskaźnik mierzy to, co ma mierzyć), o tyle w badaniach o charakterze międzynarodowym (międzykulturowym) kategoria trafności obejmuje również kwestię ekwiwalencji wskaźników w różnych kontekstach społeczno-kulturowych. Ekwiwalencja (porównywalność) badanych wskaźników w każdym badaniu porównawczym jest kwestią fundamentalną i trudną do osiągnięcia. Nie należy jej uznawać za pewnik, lecz zawsze traktować jako kwestię otwartą na weryfikację empiryczną. Zachowanie ekwiwalencji na etapie projektowania badania (płaszczyzna interpretacyjna ekwiwalencji) nie gwarantuje uzyskania porównywalnych wyników. Ekwiwalencję należy sprawdzić na płaszczyźnie pomiaru (empirycznej) za pomocą metod analizy statystycznej [Jaciow 2017].

W artykule przedstawione zostaną etapy oceny poziomu ekwiwalencji pomiaru ze wskazaniem statystycznych metod wykorzystywanych w tym celu. Ocenie empirycznej poddane zostaną wyniki badań międzynarodowych zrealizowanych w Katedrze Rynku i Konsumpcji Uniwersytetu Ekonomicznego w Katowicach (projekt badań pt. „Handel i usługi w Europie – diagnoza i perspektywy rozwoju”). Celem artykułu jest określenie poziomu ekwiwalencji modelu pomiarowego postaw konsumentów wobec zakupów online. Postawiono hipotezę badawczą, która zakłada istnienie najniższego poziomu ekwiwalencji dla wszystkich siedmiu badanych krajów. Założono również, że kolejne – wyższe poziomy ekwiwalencji uzyskane zostaną dla co najmniej czterech z siedmiu krajów.

2. Metodyka oceny ekwiwalencji wyników badań

W literaturze przedmiotu znaleźć można wiele sposobów testowania ekwiwalencji za pomocą statystycznej analizy wyników [van de Vijver i Leung 1997; Johnson 1998; van Deth 1998; Berry 2000]. Najważniejsze zestawiono w tab. 1.

Zachowanie ekwiwalencji pomiaru jest szczególnie ważne w badaniach prowadzonych z wykorzystaniem skal złożonych, w których jednostki pomiaru mają charakter umowny i uwzględniają uwarunkowania kulturowe badanego rynku. Badania przy użyciu takich skal będą ekwiwalentne na płaszczyźnie pomiaru w warunkach zachowania podobieństwa macierzy kowariancji dla badanych kultur, odpowiednio-

ści modeli pomiaru, porównywalności jednostek pomiaru, podobieństwa skali pomiarowej oraz jednorodności wpływu czynników swoistych [Sagan 2005].

Skale złożone przyjmują najczęściej postać skal Likerta. Oceny trafności, rzetelności i ekwiwalencji pomiaru za pomocą skali Likerta dokonuje się w nurcie klasycznej teorii testu za pomocą analizy czynnikowej [Vandenberg, Lance 2000]. Proces oceny ekwiwalencji pomiaru wiąże się z przeprowadzeniem serii testów sprawdzających hipotezy dotyczące występowania różnic międzykulturowych. Testy przeprowadza się w odpowiedniej kolejności, gdyż złe dopasowanie modelu bardziej ogólnego sprawia, że nieuzasadnione staje się testowanie kolejnego poziomu ekwiwalencji (stopniowalność poziomu ekwiwalencji) – rys. 1.

Tabela 1. Empiryczna ocena ekwiwalencji danych – typy, stopnie i sposoby

Płaszczyzna ekwiwalencji		Typ zgodności	Stopień ekwiwalencji	Sposób oceny
Płaszczyzna pomiaru	konstruktu	identyczność konstruktu	najniższy	ocena trafności i rzetelności pomiaru
	konfiguracyjna	„co do struktury”	podstawowy	analiza czynnikowa (CFA), skalowanie wielowymiarowe
	metryczna	„co do metryki” (zgodność ładunków czynnikowych)	ekwiwalencja częściowa	wielogrupowa analiza czynnikowa (MGCFA)
	kalibracji	skalarna (zgodność stałych czynnikowych)	ekwiwalencja częściowa	wielopoziomowa potwierdzająca analiza czynnikowa (MCFA)
zgodność błędów losowych		ekwiwalencja pełna		

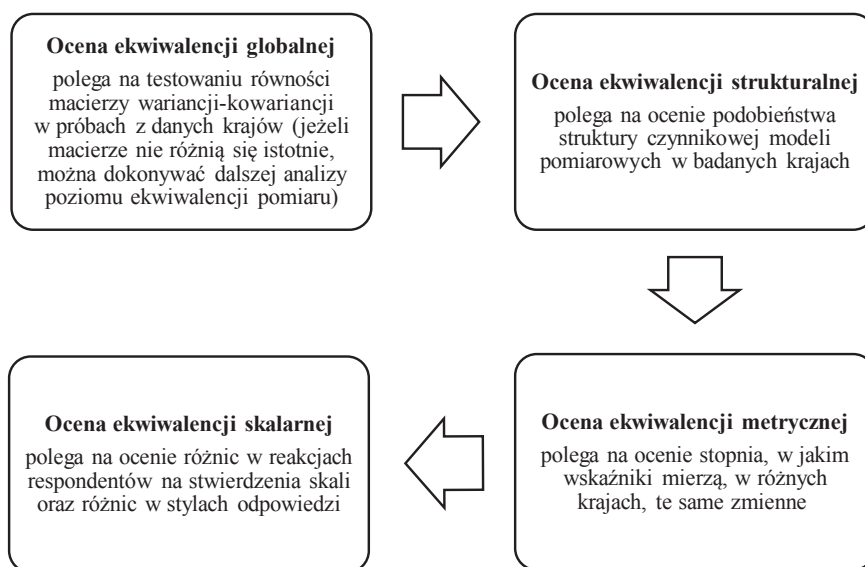
Źródło: opracowanie własne.

Dobre dopasowanie modeli do danych świadczy o istnieniu zgodności strukturalnej i umożliwia porównywanie między sobą wskaźników w przekroju krajów. Stopień dopasowania jest najczęściej testowany za pomocą statystyki χ^2 , wskaźnika CFI Bentlera, indeksu TLI Tuckera-Lewisa lub współczynnika RMSEA Steigera-Lindta. Jeżeli ten rodzaj zgodności nie jest spełniony (dopasowanie modelu jest słabe), wówczas wskaźniki są specyficzne dla poszczególnych kultur (brak ekwiwalencji).

Realizacja każdego kolejnego etapu oceny ekwiwalencji jest możliwa przy spełnieniu warunków dla etapu poprzedniego. Jeżeli np. nie istnieje strukturalna ekwiwalencja zmiennych ukrytych w badanych kulturach, to również nie można testować hipotezy o istnieniu ekwiwalencji metrycznej i skalarnej.

Na brak ekwiwalencji wpływa nie tylko zróżnicowanie kulturowe respondentów z badanych krajów (choć jest to główny czynnik braku porównywalności danych). Jak wskazuje Sagan [2014], do podstawowych źródeł braku ekwiwalencji pomia-

ru należą: hierarchiczne czynniki instytucjonalne, wywołujące współzależność obserwacji i występowanie silnych korelacji wewnątrzklasowych (tzw. zagnieżdżenie obserwacji), przynależność respondentów do grup kulturowych, a także: cechy społeczno-demograficzne badanych, czynniki sytuacyjne i kontekstowe pomiaru oraz styl reakcji na pytania.



Rys. 1. Etapy oceny poziomu ekwiwalencji pomiaru na skali Likerta

Źródło: na podstawie: [Sagan 2005].

3. Ocena poziomu ekwiwalencji badań własnych

Próba oceny poziomu ekwiwalencji na potrzeby niniejszego artykułu zostanie dokonana na przykładzie wyników badań międzynarodowych zrealizowanych w 2014 roku w Katedrze Rynku i Konsumpcji Uniwersytetu Ekonomicznego w Katowicach. Celem badań była identyfikacja opinii na temat funkcjonowania placówek handlowych w siedmiu krajach Europy (Finlandia, Francja, Niemcy, Polska, Rumunia, Słowacja, Węgry). Badania przeprowadzono wśród celowo dobranych konsumentów. W badaniach udział wzięło 2055 konsumentów¹. Badania zrealizowane zostały techniką ankiety rozdawanej. Narzędzie badawcze (kwestionariusz ankiety) składało się z 32 pytań. Pomiaru dokonano na skalach porządkowych, w tym na skali Likerta. Kwestionariusz przetłumaczono na języki narodowe, a następnie przesłano

¹ Charakterystykę próby badawczej i wyniki badań znaleźć można w: [Jaciow i in. 2015; Kucharśka i in. 2015].

do konsultacji koordynatorom w poszczególnych krajach. Badania przeprowadzono symultanicznie w kwietniu i maju 2014 roku.

3.1. Charakterystyka skali pomiarowej

Ze względu na szeroki zakres przedmiotowy badań opinii mieszkańców wybranych krajów Europy na temat funkcjonowania placówek handlu detalicznego i usług do analizy (na potrzeby artykułu) wybrano 8 zmiennych tworzących model pomiarowy postaw wobec zakupów online (oznaczonych: z1 – „Konsumenci mają zaufanie do zakupów w Internecie”, z2 – „Konsumenci obawiają się o bezpieczeństwo danych osobowych przy zakupach w Internecie”, z3 – „Konsumenci obawiają się oszustw związanych z płatnościami w Internecie”, z4 – „Zakupy produktów i korzystanie z usług w Internecie jest skomplikowane”, z5 – „Zakupy w Internecie ułatwiają dostęp do wielu produktów/usług”, z6 – „Zakupy w Internecie pozwalają zaoszczędzić pieniądze”, z7 – „Zakupy w Internecie pozwalają zaoszczędzić czas”, z8 – „W przyszłości konsumenci będą chętniej kupować usługi w Internecie niż tradycyjnie w placówkach usługowych”). Pomiaru postaw dokonano na pięciostopniowej skali Likerta (z wariantami odpowiedzi: całkowicie się nie zgadzam; raczej się nie zgadzam; ani się nie zgadzam, ani się zgadzam; raczej się zgadzam; całkowicie się zgadzam).

3.2. Ocena ekwiwalencji globalnej konstruktów

W ocenie ekwiwalencji wyników badań własnych wykorzystano klasyczną analizę czynnikową, która analizuje ukrytą strukturę przyczynową, będącą podstawą związków między cechami.

Aby dokonać oceny ekwiwalencji globalnej, sprawdzono rzetelność skali (za pomocą współczynnika α Cronbacha) oraz przydatność wybranych zmiennych (za pomocą statystyki Kaisera-Meyera-Olkina oraz testu sferyczności Bartletta). Wartości statystyk dla każdego kraju zestawiono w tab. 2.

Tabela 2. Wartości statystyki K-M-O dla badanych krajów

Wyszczególnienie	α Cronbacha	Wartość K-M-O	Test sferyczności Bartletta*		
			przybliżone χ^2	df	istotność
Finlandia	0,737	0,766	246,235	28	0,000
Francja	0,641	0,682	357,507	28	0,000
Niemcy	0,792	0,771	353,956	28	0,000
Polska	0,819	0,849	609,027	28	0,000
Rumunia	0,885	0,903	903,640	28	0,000
Słowacja	0,813	0,861	610,841	28	0,000
Węgry	0,744	0,803	224,149	28	0,000

* W oparciu o korelacje.

Źródło: opracowanie własne.

Uzyskane wyniki wskazują, iż założenia ekwiwalencji globalnej nie spełnia Francja (najniższy współczynnik α Cronbacha i za niska wartość K-M-O). Analiza kolejnego etapu oceny ekwiwalencji jest możliwa przy spełnieniu warunków dla etapu poprzedniego. W związku z tym Francja została wyłączona z dalszych analiz.

3.3. Ocena ekwiwalencji strukturalnej

Do sprawdzenia ekwiwalencji strukturalnej wykorzystano eksploracyjną analizę czynnikową. Analizę wykonano dla każdego kraju z osobna, co pozwoliło ocenić, czy otrzymana konfiguracja czynników jest dla badanych krajów taka sama.

W wyniku zastosowania powyższej procedury otrzymano kilka czynników głównych. Do ich wyboru zastosowano kryterium wartości własnej Keisera. W celu łatwiejszej interpretacji czynników zastosowano rotację czynników metodą varimax znormalizowaną. W tab. 3 zestawiono liczbę wyodrębnionych czynników wraz z ich wartościami własnymi i procentem wariancji.

Tabela 3. Liczba i wartości własne wyodrębnionych czynników

Wyszczególnienie	Liczba wyodrębnionych czynników	Wartość własna	Wyjaśnienia wariancji
Finlandia	2	2,175	35,385
		1,025	16,671
Niemcy	2	2,924	40,263
		1,209	16,641
Polska	2	3,959	44,577
		1,241	13,974
Rumunia	1	4,527	55,839
Słowacja	2	4,334	43,971
		1,462	14,836
Węgry	3	3,417	36,454
		1,512	16,129
		1,194	12,737

Źródło: opracowanie własne.

Uzyskane wyniki wskazują, iż założenia ekwiwalencji strukturalnej nie spełniają dwa kraje: Rumunia (w której wyodrębniono 1 czynnik) i Węgry (z wyodrębnionymi trzema czynnikami). W kolejnym etapie analizy uwzględniono cztery kraje.

3.4. Ocena ekwiwalencji metrycznej

W celu sprawdzenia ekwiwalencji metrycznej przeanalizowano wartości ładunków czynnikowych wyodrębnionych w poprzednim etapie analizy. Wartości te są odpo-

wiednikami korelacji pomiędzy zmiennymi a czynnikami. W tabeli 4 zestawiono wartości ładunków czynnikowych dla zmiennych w każdym kraju.

Tabela 4. Wartości ładunków czynnikowych metodą klasycznej analizy czynnikowej

Zmienne \ Czynniki	Finlandia		Niemcy		Polska		Słowacja	
	cz1	cz2	cz1	cz2	cz1	cz2	cz1	cz2
z1	0,418	-0,129	0,321	-0,087	0,328	-0,097	0,301	-0,001
z2	0,258	-0,004	0,333	-0,125	0,311	-0,095	0,277	-0,093
z3	0,244	0,004	0,311	-0,092	0,251	0,023	0,229	0,021
z4	-0,043	-0,027	0,140	0,097	0,211	-0,079	0,203	-0,034
z5	0,234	-0,121	0,145	0,102	0,152	0,069	0,258	-0,124
z6	0,183	-0,041	0,220	0,016	0,218	-0,099	0,150	-0,055
z7	-0,234	0,815	-0,244	0,662	0,073	0,284	0,112	0,143
z8	0,158	0,405	-0,060	0,483	-0,346	0,881	-0,289	0,982

Źródło: opracowanie własne.

Uzyskane wyniki wskazują, iż założenia ekwiwalencji metrycznej nie zostały spełnione w grupie wszystkich czterech analizowanych krajów. Gdyby brać pod uwagę pary krajów, to zachowana jest ekwiwalencja metryczna między Polską i Słowacją.

4. Zakończenie

Z problemem ekwiwalencji wyników badań spotykają się zespoły badaczy realizujących badania o zasięgu międzynarodowym. Wiele jest podejść do problemu ekwiwalencji. W literaturze przedmiotu wyróżnia się dwa główne podejścia do oceny poziomu ekwiwalencji badań rynku: jakościowe (eksperckie) oraz ilościowe (empiryczne). Zachowanie ekwiwalencji badań konieczne jest zarówno na płaszczyźnie interpretacyjnej (na poziomie definiowania obiektu badań, konstruowania narzędzi pomiarowych, projektowania przebiegu badań terenowych), jak i na płaszczyźnie proceduralnej (na poziomie statystycznej zgodności otrzymanych wyników dotyczących badanego konstruktów, jego struktury, jednostek i błędów pomiarowych). Na przykładzie międzynarodowego projektu badawczego dotyczącego funkcjonowania handlu detalicznego i usług w Europie dokonano analizy ekwiwalencji wybranych wyników na płaszczyźnie proceduralnej (empirycznej). Ocenie poddano model pomiarowy postaw konsumentów wobec zakupów online, składający się z ośmiu zmiennych mierzonych na skali Likerta. Postawiona hipoteza badawcza, która zakładała istnienie najniższego poziomu ekwiwalencji dla wszystkich badanych krajów, została odrzucona. Odrzucono także hipotezę zakładającą, że kolejne – wyższe

poziomy ekwiwalencji uzyskane zostaną dla co najmniej czterech z siedmiu krajów. Najniższy poziom ekwiwalencji osiągnięto dla sześciu z siedmiu krajów. Ekwiwalencję na poziomie podstawowym (strukturalną) spełniały cztery z siedmiu krajów, a częściową ekwiwalencję (na poziomie metrycznym) jedynie dwa kraje. W przypadku tej pary krajów zasadne byłoby sprawdzenie kolejnego poziomu ekwiwalencji (na poziomie kalibracji), jednak testowanie dobroci dopasowania modeli za pomocą confirmacyjnej analizy czynnikowej wykracza poza ramy niniejszego artykułu. Można jednak stwierdzić, iż uzyskany poziom ekwiwalencji metrycznej dla Polski i Słowacji uprawnia badaczy do porównań postaw konsumentów wobec zakupów online w tych krajach.

Literatura

- Berry J.W., 2000, *Introduction to methodology*, [w:] Triandis H.C., Berry J.W. (eds.), *Handbook of Cross-Cultural Psychology*, vol. 2, Allyn and Bacon, Boston.
- Gray R., Owen D., Sopher M.J., 1998, *Setting up a control system for your organization*, *Nonprofit World*, vol. 16, no. 3, s. 65-76.
- Jaciow M., 2017, *Ekwiwalencja w międzynarodowych badaniach marketingowych – identyfikacja problemu*, *Handel Wewnętrzny*, 6(371), s. 112-121.
- Jaciow M., Kolny B., Maciejewski G., Mikołajczyk B., Wolny R., 2015, *Services in Europe – Diagnosis and Development Perspectives*, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice.
- Johnson T.P., 1998, *Approaches to equivalence in cross-cultural and cross-national survey research*, *ZUMA-Nachrichten Spezial*, Nr 3, s. 1-40.
- Kucharska B., Kucia M., Maciejewski G., Malinowska M., Stolecka-Makowska A., 2015, *The Retail Trade in Europe – Diagnosis and Future Perspectives*, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice.
- Miller S.M., 1960, *Comparative social mobility*, *Current Sociology*, 9(1), s. 1-61.
- Mills M., Van de Bunt G.G., De Bruijn J., 2006, *Comparative research: Persistent problems and promising solutions*, *International Sociology*, 21(5), s. 619-631.
- Niessen M., Peschar J., Kourilsky C. (eds.), 2013, *International Comparative Research: Social Structures and Public Institutions in Eastern and Western Europe*, Pergamon Press, Oxford – New York.
- Oyen E. (ed.), 1990, *Comparative Methodology: Theory and Practice in International Social Research*, Sage Publications, London.
- Peters B.G., 1998, *Comparative Politics: Theory and Methods*, NYU Press, New York.
- Ragin C.C., 2014, *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*, University of California Press, Oakland, California.
- Röhmel J., Mansmann U., 1999, *Exact tests of equivalence and efficacy with a non-zero lower bound for comparative studies by I.S.F. Chan*, *Statistics in Medicine*, no. 17, s. 1403-1413.
- Sagan A., 2005, *Ocena ekwiwalencji skal pomiarowych w badaniach międzykulturowych*, *Zeszyty Naukowe Akademii Ekonomicznej w Krakowie*, nr 659, s. 59-73.
- Sagan A., 2014, *Rzetelność i funkcjonowanie pozycji skali a ukryta niejednorodność populacji*, *Studia Ekonomiczne*, nr 195, s. 151-163.
- Treiman D.J., 2013, *Occupational Prestige in Comparative Perspective*, Elsevier, New York – San Francisco – London.

- Tyree A., Semyonov M., Hodge R.W., 1979, *Gaps and glissandos: inequality, economic development, and social mobility in 24 countries*, American Sociological Review, vol. 44, no. 3, s. 410-424.
- Umiejętności Polaków – wyniki Międzynarodowego Badania Kompetencji Osób Dorosłych (PIAAC), 2013, IBE.
- Van de Vijver F.J., Leung K., 1997, *Methods and Data Analysis for Cross-Cultural Research*, vol. 1, Sage Publications, Thousand Oaks – London – New Delhi.
- Van Deth J.W., 1998, *Equivalence in Comparative Political Research. Comparative Politics. The Problem of Equivalence*, Routledge, London – New York.
- Vandenberg R.J., Lance C.E., 2000, *A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research*, Organizational Research Methods, no. 3(1), s. 4-70.