

# Problemy metodologiczne w badaniach międzykulturowych i kulturowych

Anna Kwiatkowska

*Instytut Psychologii, Polska Akademia Nauk*

Zdobywanie wiedzy o świecie, o człowieku, o jego funkcjonowaniu psychologicznym nie jest procesem neutralnym ideologicznie. Proces badawczy jest w swojej istocie produktem kulturowym i każdy z jego etapów wyraża w jakiś sposób idee, wartości i reguły zachowania właściwe dla danej kultury. Psycholodzy podejmujący badania międzykulturowe czasami zbyt optymistycznie zakładają, że wiedza metodologiczna, zdobyta w ramach wykształcenia w zakresie psychologii głównego nurtu, jest na tyle uniwersalna, iż wystarczy wiernie przetłumaczyć narzędzia na lokalny język lub zatrudnić przedstawiciela lokalnej społeczności w charakterze pomocnika badacza, by zdobyć znaczące, wartościowe dane na interesujący temat. O tym, że takie działania nie są wystarczające, powinien przekonać potencjalnego badacza ten artykuł, poświęcony metodologicznym aspektom prowadzenia badań naukowych w psychologii międzykulturowej i kulturowej. Zastanowimy się, na czym polega ich specyfika, z jakimi problemami może zetknąć się badacz i jak może je rozwiązywać. W pierwszej części artykułu zostanie zaprezentowana taksonomia badań międzykulturowych, a także paradygmaty badawcze funkcjonujące w tym obszarze. W drugiej części przyjrzymy się niektórym pułapkom metodologicznym, w jakie może wpaść psycholog międzykulturowy, głównie odnoszące się do problemu ekwiwalencji w badaniach i błędów tendencyjności metody.

Słowa kluczowe: *badania międzykulturowe, taksonomia metod, paradygmaty badawcze, równoważność metody, błędy tendencyjności*

## WSTĘP

Lawinowy wręcz przyrost badań międzykulturowych w psychologii w ciągu ostatnich kilkunastu lat jest faktem. Świadczy o tym chociażby ten numer *Psychologii Społecznej*, na który złożyły się prace polskich psychologów, czasem we współpracy z kolegami z innych krajów, czy niedawno wydane dzieło Pawła Boskiego (2009). Nie jest to powrót do klasycznego pytania gnębiącego przez dziesiątki lat psychologów: natura czy kultura, ponieważ to pytanie dawno już zostało rozstrzygnięte: i natura, i kultura. Obecnie nikt nie wątpi w istotną rolę kultury w kształtowaniu naszej psychiki, ale ciągle nie wiemy, jak dużo kulturze zawdzięczamy i jakie są mechanizmy wzajemnych oddziaływań. Stawiamy więc nowe pytania i ponawiamy stare, jeżeli nie jesteśmy zadowoleni z odpowiedzi. Umiejętność formułowania ważnych pytań nie

gwarantuje, że otrzymana odpowiedź będzie trafnie opisywać rzeczywistość. Każdy badacz dobrze wie, że proces badawczy musi spełniać wiele warunków, by można było zaakceptować jego wyniki. W tym artykule zostaną omówione tylko niektóre z warunków koniecznych, ważnych dla prowadzenia badań w psychologii międzykulturowej i kulturowej. Przegląd problemów, z jakimi można się zetknąć, podejmując badania uwzględniające kulturę, zostanie poprzedzony zaprezentowaniem typologii oraz faz rozwoju badań międzykulturowych i kulturowych.

Zacznijmy jednak od wyjaśnień w kwestii stosowania terminów „międzykulturowe/a” i „kulturowe/a” (badania/psychologia). Występowanie tych terminów obok siebie (jak w tytule artykułu) może sugerować, że mamy do czynienia z równorzędnymi, aczkolwiek odrębnymi paradygmatami badawczymi bądź subdyscyplinami psychologii. Jest to dość popularne stanowisko, mające związek, po pierwsze, z klasycznym już rozróżnieniem orientacji teoretycznych na absolutystyczną, relatywistyczną i uniwersalistyczną (Berry, Poortinga, Segall, Dasen, 2004). A po

---

Anna Kwiatkowska, Instytut Psychologii, Polska Akademia Nauk,  
ul. Jaracza 1, 00-378 Warszawa,  
e-mail: anna.kwiatkowska@psych.pan.pl

drugie, z formowaniem się różnych tradycji w prowadzeniu badań na styku kultury i psychologii, o czym będzie mowa w dalszej części tego artykułu.

W podejściu absolutystycznym przyjmuje się, że ludzkie zachowania są zasadniczo takie same na całym świecie, zatem czynniki kulturowe nie są istotne, a badania porównujące różne kultury są całkowicie zbędne. Jest to podejście zdecydowanie akulturowe, w przeciwieństwie do orientacji relatywistycznej, w której zakłada się, że zachowanie człowieka jest kulturowo zdeterminowane i może być wyjaśniane tylko poprzez odwołanie się do znaczeń podzielanych przez członków danej kultury. Porównania międzykulturowe są więc utrudnione, a czasem wręcz nie są możliwe. Zdaniem Johna Berry'ego i Ype Poortingi (2006), podejście relatywistyczne dało asumpt do powstania psychologii indygenicznej (psychologii uprawianej z pozycji unikalnej, niepowtarzalnej kultury) i w dużym stopniu – psychologii kulturowej (por. Shweder, Sullivan, 1993). Psychologia kulturowa ma na celu identyfikację specyficznych dla danej kultury uwarunkowań funkcjonowania człowieka. Zdaniem Richarda Shwedera, psychologia kulturowa to studiowanie, w jaki sposób kultura i psychika tworzą siebie nawzajem (*the way culture and psyche make each other up*; Shweder, Sullivan, 1993, s. 498).

Perspektywa uniwersalistyczna natomiast jest czymś w rodzaju pomostu między stanowiskiem absolutystycznym i relatywistycznym. Przyjmuje się w niej, że istnieją cechy wspólne dla wszystkich przedstawicieli rodzaju ludzkiego, które z kolei mogą manifestować się w sposób specyficzny dla danej kultury. Tak więc kontekst kulturowy jest bardzo istotny w wyjaśnianiu zachowania człowieka, lecz badania prowadzone w obrębie jednej tylko kultury nie pozwolą osiągnąć pełnego wglądu w naturę czynników kulturowych. Aby zidentyfikować to, co jest wspólne, uniwersalne, i to, co jest różne, specyficzne dla danej kultury, należy kultury porównywać. Można więc powiedzieć, że perspektywa uniwersalistyczna daje podstawy do uprawiania psychologii międzykulturowej (Berry i in., 2004).

Wydaje się jednak, że utrzymywanie podziału między psychologią międzykulturową i psychologią kulturową nie ma w tej chwili większego znaczenia. W badaniach prowadzonych na styku psychologii i kultury oba podejścia się przeplatają; nawet relatywistyczna psychologia indygeniczna, by odnaleźć własną specyfikę, nie cofa się przed porównaniami z innymi kulturami (por. Kim, Yang, Hwang, 2006). Van de Vijver i Leung (2001) przekonują, że badania prowadzone w ramach obu paradygmatów nawzajem się uzupełniają zarówno w kwestii uzasadniania istnienia uniwersaliów i specyfiki kulturowej, jak

również pod względem stosowanych metod (bardziej rygorystycznych w paradygmacie międzykulturowym i łączenie metod jakościowych i ilościowych w paradygmacie kulturowym). Niewykluczone że w niedalekiej przyszłości psychologia kulturowa podporządkuje sobie psychologię międzykulturową, traktując ją raczej jako podejście badawcze niż odrębną subdyscyplinę psychologii. Pojawianie się takich podręczników, jak *Handbook of cultural psychology* Kitayamy i Cohena (2007) czy *Cultural psychology* Heinego (2008), może świadczyć o rosnącej wśród psychologów tendencji do identyfikowania się z psychologią kulturową. Choć autorka tego artykułu byłaby skłonna używać głównie określenia „psychologia kulturowa” jako terminu bardziej ogólnego, to z racji prominentnego miejsca, jakie zajmują badania stricte międzykulturowe w historii badań na styku kultury i psychologii, stosowane tu będą oba terminy na zasadzie równorzędności<sup>1</sup>.

#### TYPY BADAŃ W PSYCHOLOGII MIĘDZYKULTUROWEJ I KULTUROWEJ

Fons van de Vijver i Kwok Leung (2001) wyróżniają dwa rodzaje badań nad różnicami i podobieństwami kulturowymi: badania skoncentrowane na strukturze relacji między zmiennymi (*structure-oriented studies*) oraz na różnicach w zakresie poziomu (natężenia) zmiennych (*level-oriented studies*). Struktura relacji może być opisywana w kategoriach prostych związków korelacyjnych lub zależności bardziej skomplikowanych, np. w postaci modeli ścieżkowych, jak również w terminach zależności przyczynowo-skutkowych weryfikowanych eksperymentalnie. Przykładem takiego podejścia zorientowanego na strukturę są badania Hui i Triandisa (1983), w których sprawdzano, czy umiejscowienie kontroli koreluje z tymi samymi czynnikami w różnych krajach, także badania Katarzyny Lubiewskiej (2014) nad wpływem wartości kolektywistycznych i indywidualistycznych na zachowania matki i dziecka w Polsce i w Niemczech czy badania Anny Kwiatkowskiej, Joanny Roszak, Renaty Sikory, Bena Kuo, Konstantego Karpinskiego, Tatiany Gushchiny i Greta Gober (w tym tomie) nad uwarunkowaniem strategii radzenia sobie ze stresem w kilku krajach europejskich. Badania drugiego rodzaju mają na celu sprawdzenie, czy kultury różnią się pod względem poziomu jakiejś zmiennej. Można tu wymienić badania Gidi Rubinsteina (1996) nad analizą różnic pod względem nasilenia autorytaryzmu u Palestyńczyków i Żydów, mieszkańców

<sup>1</sup> Czołowy metodolog w psychologii między/kulturowej w swoim artykule stosuje zapis: (*cross-)*cultural psychology (van de Vijver, Leung, 2001).

Tabela 1  
*Taksonomia badań w psychologii międzykulturowej i kulturowej według van de Vijvera i Leunga (2001)*

Zmienne kontekstowe	Cel: eksploracja	Cel: testowanie hipotez
Brak	(a) Badania różnic psychologicznych ( <i>psychological differences studies</i> )	(b) Badania uogólniające ( <i>generalizability studies</i> )
Są	(c) Badania nad trafnością zewnętrzną ( <i>external validation studies</i> )	(d) Badania kierowane teorią kontekstualną ( <i>contextual theory-driven studies</i> )

Izraela, czy badania Kuby Krysia i Karoliny Hansen (2014) nad różnicami kulturowymi w postrzeganiu osób uśmiechniętych.

Opracowując taksonomię badań, van de Vijver i Leung (2001) posłużyli się dwoma wymiarami, co pozwoliło uzyskać cztery typy badań (tabela 1). Należy podkreślić, że w obrębie każdego z wyróżnionych typów są możliwe badania zorientowane zarówno na strukturę, jak i na poziom, a ponadto mogą to być badania prowadzone zarówno w paradygmacie jakościowym, jak i ilościowym. Pierwszy wymiar odnosi się do znanego w naukach społecznych podziału na badania eksploracyjne i badania testujące hipotezy. Badania eksploracyjne prowadzone są bez przyjmowania wstępnych założeń co do spodziewanych wyników, natomiast badania drugiego typu są nastawione na sprawdzanie hipotez sformułowanych na podstawie teorii lub wyników z wcześniejszych badań. Drugi wymiar jest specyficzny dla badań międzykulturowych i kulturowych, mianowicie dotyczy pytania, czy badania są prowadzone z uwzględnieniem zmiennych kontekstowych, które mogą wpływać na obserwowane różnice i podobieństwa kulturowe. Do zmiennych kontekstowych możemy zaliczyć charakterystyki osób badanych (np. status ekonomiczno-społeczny, wiek, wykształcenie, płeć, właściwości psychologiczne nienależące do puli badanych zmiennych zależnych), również zmienne charakteryzujące badane społeczności (np. poziom zamożności danego kraju, produkt krajowy brutto, system edukacji, służba zdrowia).

### Badania nad różnicami psychologicznymi

Popularne podejście w badaniach międzykulturowych dotyczy różnic psychologicznych. Są to badania eksploracyjne, bez zmiennych kontekstowych (*psychological differences studies*). W wielu wypadkach narzędzia

opracowane w jednej kulturze stosowane są do pomiarów tego samego pojęcia w innych kontekstach kulturowych, by np. zdobyć pewien wgląd w jego strukturę w innej kulturze (np. czy neurotyzm znaczy to samo we wszystkich kulturach) lub po prostu porównać ze sobą odpowiedzi uzyskane w różnych grupach kulturowych (np. czy istnieją różnice w poziomie neurotyzmu w różnych kulturach). W badaniu Krysia i Hansen (2014) stawiano pytanie, czy osoby uśmiechające się są postrzegane bardziej lub mniej korzystnie w różnych krajach. Zaletą tego podejścia jest brak założeń na temat różnic kulturowych, a więc pewna otwartość umysłowa, niedogodnością zaś jest pokusa dokonywania zbyt swobodnych interpretacji.

### Badania uogólniające

Do badań uogólniających (*generalizability studies*) można zaliczyć badania Shaloma Schwartza dotyczące wartości (Schwartz, 1992), prace McCrae'a i Costy na temat uniwersalności pięcioczynnikowej struktury osobowości (McCrae, Terracciano, 2005), badania Eysencka dotyczące uniwersalności trójczynnikowej struktury osobowości (Eysenck, Eysenck, 1983), także badania Rene Möttusa i współpracowników (2012a, 2012b) nad sposobem rozumienia cechy *sumienność* w 21 krajach. Tego typu badania mają mocne podstawy teoretyczne, umożliwiające sformułowanie hipotez odnośnie do podobieństw i różnic międzykulturowych, jednakże nie prowadzi się w nich pomiaru zmiennych kontekstowych. Paradygmat ten jest więc wybierany wówczas, gdy psychologowie stawiają pytanie, w jakim stopniu badane zjawiska mogą być zgeneralizowane na różne kultury. Słabością tych badań jest nieobecność zmiennych kontekstowych, których potencjalny wpływ pozostaje nieustalony.

### Badania walidacyjne

Kolejny typ badań to badania walidacyjne (*external validation studies*), tj. nad trafnością zewnętrzną (ekologiczną) badanych konstruktów. Van de Vijver i Leung (2000) podają dwa przykłady badań prowadzonych w tym paradygmacie. Pierwszy przykład to badanie Williamsa, Satterwhite'a i Saiza (1998; za: van de Vijver, Leung, 2000), w którym proszono badanych w 20 krajach o ocenę psychologicznej ważności 300 cech z *Listy przymiotnikowej* (ACL), a następnie uzyskane średnie na poziomie krajów korelowano z różnymi parametrami rozwoju ekonomicznego i wskaźnikami demograficznymi. Drugi przykład to badanie Georgasa, van de Vijvera i Berry'ego (2004), w którym obliczano korelacje między pewnymi zmiennymi psychologicznymi a religią (liczba wiernych danego wyznania) i bogactwem kraju. Badania Magdaleny Żemojtel-Piotrowskiej i innych (2014),

prowadzone w 27 krajach, nad postawami roszczeniowymi i ich związkiem z rozwojem socjoekonomicznym kraju również są przykładem tego typu badań. Są to ważne badania dla rozwoju psychologii międzykulturowej, jednak słabością ich może się okazać arbitralny dobór kowariantów, czasem mało istotnych. Zwróćmy uwagę, że jedyna różnica między badaniami walidacyjnymi a badaniami uogólniającymi to obecność zmiennych kontekstowych.

### Badania kierowane teorią kontekstualną

Ostatni typ badań wyróżniony przez van de Vijvera i Leunga to badania kierowane teorią kontekstualną (*contextual theory-driven studies*). Mają one te zalety, że są silnie osadzone w teorii oraz biorą pod uwagę zmienne kontekstowe. Autorzy podają tu przykład badania Johna Berry'ego (1976) nad stylami poznawczymi. Berry założył, że orientacja przestrzenna i inne ważne aspekty postrzegania wzrokowego są ważniejsze dla przetrwania w kulturze łowców i zbieraczy niż w kulturze rolniczej. W związku z tym członkowie tej pierwszej grupy będą bardziej niezależni od pola niż rolnicy. Oznacza to, że łowcy będą mieć tendencję do używania wewnętrznego układu odniesienia, np. sygnałów ze strony własnego ciała, podczas gdy rolnicy będą przejawiać tendencję do używania zewnętrznego układu odniesienia. Wyniki badania potwierdziły hipotezę. Inny przykład badań należących do tej kategorii to metaanaliza różnych modeli wyjaśniających międzykulturowe różnice w poziomie wykonania testów procesów poznawczych, przeprowadzona przez van de Vijvera (1997). Autor potwierdził hipotezę, iż różnice w PKB i wielkości wydatków państwa na edukację pozostają w relacji do różnic w zakresie poziomu wykonania testów. Ten typ badań służy analizie relacji między czynnikami kulturowymi a zachowaniem, jednak jego słabością może być niedostrzeżenie alternatywnych interpretacji.

Jakiego typu badania przeważają w psychologii międzykulturowej? Trudno na to pytanie odpowiedzieć, w zasadzie w literaturze można znaleźć przykłady na każde z tych badań<sup>2</sup>. Być może na różnych etapach rozwoju dziedziny dominował określony typ, zgodnie z panującym w danym okresie paradygmatem badawczym. Prześledźmy zatem kolejne okresy.

## FAZY ROZWOJU BADAŃ MIĘDZYKULTUROWYCH I KULTUROWYCH

Badania prowadzone w ramach psychologii międzykulturowej ewoluowały przez parę dziesiątków lat: od opisywania różnic międzykulturowych, aż po identyfikację znaczących wymiarów kulturowej różnorodności; wymiary te były używane do tworzenia modeli teoretycznych, pozwalających z kolei przewidywać i wyjaśniać udokumentowane różnice. David Matsumoto i Seung Hee Yoo (2006) opisują cztery fazy ewoluowania badań międzykulturowych: (1) badania porównawcze, (2) identyfikacja wymiarów kulturowych, (3) badania kulturowe, (4) badania nad zależnościami, w tym badania „rozpakowujące” kulturę.

### Faza I – porównawcze badania międzykulturowe (*cross-cultural comparisons*)

Badania dokumentujące istnienie różnic między grupami kulturowymi stanowią trzon psychologii międzykulturowej. Prowadzone są jako quasi-eksperymenty, w których grupa kulturowa jest zmienną niezależną, a różne zmienne psychologiczne są zmiennymi zależnymi. Najczęściej grupy kulturowe to grupy narodowe, czyli kraje, aczkolwiek badane są także grupy etniczne, językowe i rasowe. Porównania międzykulturowe zostały zapoczątkowane ponad 100 lat temu, gdy Rivers w 1905 pokazał, że mieszkańcy Indii i Nowej Gwinei ulegają częściej złudzeniom optycznym niż Anglicy (za: Matsumoto, Yoo, 2006). Badania identyfikujące oraz dokumentujące podobieństwa i różnice międzykulturowe stanowią ogromny wkład w psychologię. Testują granice wiedzy wygenerowanej w badaniach monokulturowych, które dominowały przez dziesiątki lat w Ameryce i w innych krajach Zachodu. Są one użyteczne w tworzeniu wiedzy na temat uniwersalności i specyfiki kulturowej procesów psychologicznych. Badania te sprowokowały również bardziej krytyczne myślenie na temat możliwości zastosowań psychologicznego procesu badawczego w różnych kulturach.

Ograniczeniem badań prowadzonych w tej fazie było to, iż nie pozwalały one na empirycznie usprawiedliwioną interpretację, że to właśnie kultura jest źródłem różnic między grupami. Jeśli badacze odkrywali różnice, to na ogół wyciągali wniosek, że stało się tak za sprawą kultury. Jednakże grupy mogą się różnić między sobą na wiele różnych sposobów, a kultura może być tylko jednym z nich. A jeśli nawet różnice między grupami mają swe źródło w kulturze, nie wiadomo, jakie zmienne kulturowe są odpowiedzialne za te różnice i jaki jest mechanizm ich powstawania. Praktyka przypisywania różnic między

<sup>2</sup> Van de Vijver podaje przykłady badań w ramach omawianej typologii w artykule zamieszczonym w zbiorze zatytułowanym *Online readings in psychology and culture*; <http://dx.doi.org/10.9707/2307-0919.1017>.

grupami czynnikom kulturowym bez udowodnienia empirycznego to oczywisty błąd interpretacyjny, który można nazwać za Campbellem (1961, za: Matsumoto, Yoo, 2006) błędem atrybucji kulturowej (*cultural attribution fallacy*). Błąd ten polega na tym, że sądzi się, iż coś „kulturowego” wywołało różnice między porównywanymi grupami, podczas gdy w rzeczywistości nie ma empirycznych przesłanek dla takiego wniosku.

Ograniczenie to jest częściowo rezultatem sposobu, w jaki kultury są dobierane – jako kraje, grupy etniczne, grupy rasowe, a także częściowo dlatego, że porównywana jest zbyt mała liczba grup, np. tylko dwie lub zaledwie kilka. Grupy dobierane do badania niekoniecznie są grupami kulturowymi. Oczywiście wraca problem, czym jest kultura. Właściwie nie ma jednej zaakceptowanej przez wszystkich definicji kultury – ani w antropologii, ani w socjologii, ani w psychologii, dlatego wszelkie decyzje dotyczące porównywanych grup powinny być starannie przemyślane. Nie umniejszając zasług badaczy dokumentujących różnice międzykulturowe na tym etapie rozwoju dyscypliny, można jednak sformułować zarzut, iż większość badań koncentrowała się na porównaniach między Ameryką Północną a Azją Wschodnią. A przecież nie są to jedyne kultury na świecie. Steven Heine i Ara Norenzayan (2006) nawołują, by badacze w większym stopniu interesowali się innymi kontynentami, np. Ameryką Południową i Afryką, a zwłaszcza „małymi” kulturami.

Można wskazać na jeszcze jedną słabość badań prowadzonych w tej fazie. Wynika ona z faktu, że eksploruje się różnice w zakresie zjawisk charakterystycznych dla Zachodu. Sprawdza się, jak dalece konstrukty utworzone w kulturze Zachodu, takie jak samoocena, błędy atrybucji, można zaobserwować w kulturze innej niż zachodnia. Naturalnie, są to ważne pytania, jednak warto zwrócić uwagę na to, że w niezachodnich kulturach obecne są pewne charakterystyczne zjawiska psychologiczne, które mogą umknąć uwadze badaczy, albowiem w kulturach zachodnich są czymś marginalnym lub wręcz nierozpoznawalnym. Wiele takich fenomenów opisano w psychopatologii kulturowej, np. choroby koro, lataj czy hikikomori. Znane są także emocje, które zajmują centralne miejsce w życiu emocjonalnym ludzi Wschodu, np. *amae* w Japonii (por. Kwiatkowska, 2012b), a które nie mogą być włączone do żadnej znanej zachodniej taksonomii. Kolejnym przykładem silnego niezachodniego konstruktów kulturowego jest „zachowanie twarzy”, które jest dalece mniej ważne na Zachodzie. Jak piszą Heine i Norenzayan (2006), badanie takich zjawisk jest równie ważne i może przyczynić się do zwiększenia puli wiedzy o naturze ludzkiej.

## Faza II – identyfikacja znaczących wymiarów kulturowej różnorodności (*identifying meaningful dimensions of cultural variability*)

W procesie zbierania danych o różnicach międzykulturowych stało się oczywiste, że możliwość interpretacji kulturowych jest bardzo ograniczona. Powstała więc konieczność utworzenia konstruktów teoretycznych, które byłyby na tyle ogólne, że mogłyby gromadzić bardziej szczegółowe przejawy różnorodności wewnątrz danej kultury, a jednocześnie na tyle precyzyjne, by dobrze odróżniały od siebie poszczególne kultury. Należało więc zidentyfikować znaczące dymensje kulturowego zróżnicowania, przydatne w opisie subiektywnych elementów kultury. Jako pierwszy odpowiedział na to wyzwanie Geert Hofstede, który przeprowadził badania w 40 krajach na pracownikach IBM, a wkrótce potem dodał jeszcze 13 krajów (Hofstede, 2000), w ostatniej zaś pracy przedstawił dane zebrane w 73 krajach (Hofstede, 2001). Na podstawie tych badań udowodniono istnienie czterech wymiarów kulturowych: indywidualizm–kolektywizm, dystans władzy, unikanie niepewności, męskość–kobiecość oraz ostatnio dodany wymiar: orientacja długo- vs. krótkoterminowa (Hofstede, 2001).

Identyfikacja wymiarów kulturowych, ich kwantyfikacja i ulokowanie krajów na tych skalach były to znaczące kroki w rozwoju psychologii międzykulturowej. Za pomocą tak zdefiniowanych konstruktów badacze mogli przewidywać i wyjaśniać różnice międzykulturowe. Z pięciu dymensji, rozpoznanych przez Hofstede, najbardziej popularna okazała się dymensja kolektywizm–indywidualizm. Szczególnie dobrze znane są prace Harry’ego Triandisa, który używał tej dymensji do wyjaśniania wielu podobieństw i różnic międzykulturowych (Triandis, 1994, 1996). Wykazywano teoretyczne zależności między indywidualizmem a różnicami kulturowymi w zakresie wielu procesów psychicznych, w funkcjonowaniu jednostek i grup.

Chociaż zidentyfikowanie dymensji kulturowych było widocznym postępem w badaniach w psychologii międzykulturowej, nie rozwiązało to problemu błędu kulturowej atrybucji. Jak piszą Matsumoto i Yoo (2006), nazwy krajów po prostu zostały zastąpione nazwami dymensji. Na przykład założenie, że kultura jest albo kolektywistyczna albo indywidualistyczna oraz że członkowie kultury indywidualistycznej cenią wartości indywidualistyczne, a członkowie kultury kolektywistycznej cenią wartości kolektywistyczne, pozostaje tylko założeniem. W dalszym ciągu nie wiadomo, dlaczego tak jest. Przełamanie tego impasu nastąpiło wraz z przedstawieniem koncepcji Ja-niezależnego i Ja-współzależnego przez Hazel Markus i Shinobu Kitayamę (1993). Panuje powszechne

przekonanie, że był to milowy krok w kierunku przejścia do kolejnego stadium badań w psychologii międzykulturowej, a mianowicie do fazy badań kulturowych.

### Faza III – badania kulturowe (*cultural studies*)

Badania kulturowe można scharakteryzować jako bogate opisy złożonych modeli kulturowych, pozwalających wyjaśnić różnice kulturowe. Prace Markus i Kitayamy miały ogromne znaczenie dla rozwoju badań kulturowych, ponieważ zidentyfikowały ważny potencjalny mediator różnic kulturowych – konstrukt Ja (*self-construal*). Kultury indywidualistyczne sprzyjały kreowaniu Ja-niezależnego, które miało określony wpływ na procesy psychiczne i zachowanie, kultury kolektywistyczne zaś sprzyjały rozwojowi Ja-współzależnego, które z kolei miało odmienne konsekwencje dla procesów psychicznych i zachowania.

Ważny nurt badań kulturowych obserwujemy w obszarze procesów emocjonalnych. Na przykład Batja Mesquita (2001; Mesquita, Karasawa, 2002) opisała, jak odmienne formy Ja uwrażliwiają na różne sprawy. I tak, Ja-niezależne każe jednostce koncentrować się na sprawach ważnych dla niej samej, a emocje pozostają w relacji do stanów wewnętrznych, subiektywnych. Natomiast Ja-współzależne sprawia, że dla jednostki zarówno ona jako istota społeczna, jak i grupa są równie ważne, a zatem emocje odzwierciedlają określone aspekty relacji interpersonalnych.

W obszarze procesów poznawczych ważne odkrycia zostały dokonane dzięki pracom Richarda Nisbetta i jego kolegów (Masuda, Nisbett, 2001; Nisbett, 2009; Peng, Nisbett, 2000). Opisano wiele przykładów różnic w percepcji i myśleniu Amerykanów oraz Azjatów. Według Nisbetta i współpracowników, różnice te powstały jako odpowiedź na rozmaite warunki ekologiczne, kształtujące różne praktyki społeczne. W Ameryce wspierano aktywność podmiotu (*agency*), w Azji Wschodniej zaś – harmonię między ludźmi. Różnice te z kolei wpływają na to, jak ludzie kategoryzują obiekty, jak postrzegają innych, jak się komunikują, w jaki sposób czynią atrybucje.

To tylko kilka przykładów badań kulturowych. Powstaje ich wiele i wychodzą daleko poza dokumentowanie różnic międzykulturowych, co było typowe dla fazy pierwszej. Dostarczają fascynujących opisów praktyk kulturowych i potencjalnych mechanizmów odnośnie do tego, jak różnice kulturowe mogą powstawać (mechanizmy intrapsychiczne, interpersonalne, sytuacyjne, ideologiczne). Porównuje się w nich relacje między zmiennymi w różnych kulturach, co pozwala wyciągać wnioski na temat różnego sposobu funkcjonowania zmiennych w odmiennych kontekstach kulturowych. Prace te powstają

z uwzględnieniem kontekstu historycznego, często zawierają również obserwacje z dnia codziennego.

### Faza IV – badania nad zależnościami (*linkage studies*)

Jest to faza będąca kontynuacją badań kulturowych. Charakteryzuje się badaniem zależności między obserwowalnymi różnicami w zakresie średnich czy korelacji między zmiennymi a specyficznymi źródłami kulturowymi, które hipotetycznie są odpowiedzialne za te różnice. Na czym polega to podejście, Matsumoto i Yoo (2006) wyjaśniają na przykładzie badań Iwaty i Higuchiego (2000). Badacze ci porównywali poziom lęku (za pomocą skali STAI) między Japończykami a Amerykanami i stwierdzili u Japończyków niższy poziom afektu pozytywnego i wyższy poziom lęku niż u Amerykanów. Autorzy interpretują te różnice, odwołując się do takich „faktów”, jak kolektywizm Japonii, presja na podtrzymywanie harmonii w grupach, dyscyplina moralna w postaci hamowania afektu pozytywnego, praktyki socjalizacyjne nastawione na uczenie dzieci unikania asertywności etc. Jednak, według Matsumoty i Yoo (2006), w tym badaniu nie udowodniono empirycznie żadnego związku między wymienionymi czynnikami a różnicami między grupami. Co powinni zrobić autorzy? Udokumentować empirycznie, że (a) Japonia jest krajem kolektywistycznym, że (b) subiektywny dobrostan jednostki jest podporządkowany dobrostanowi grupy, że (c) podtrzymywanie harmonii jest najważniejszą wartością w Japonii etc. A następnie udowodnić, że te zmienne pozostają w relacji do różnic w poziomie wyników w STAI.

Chociaż ten paradygmat badawczy, zdaniem niektórych badaczy (np. Heine, Norenzayan, 2006), jest jeszcze w powijakach, możemy znaleźć już całkiem sporo badań prowadzonych w taki sposób. Jednak kilka kwestii teoretycznych i metodologicznych powinno zostać rozstrzygniętych, by rzeczywiście ten typ badań przynosił oczekiwane rezultaty.

O jakich kwestiach tu mowa? Matsumoto i Yoo (2006) twierdzą, że przede wszystkim należy się zastanowić, co jest kulturą, a co nią nie jest. A dokładniej mówiąc, chodzi o to, by nie mieszać czynników kulturowych z czynnikami ekologicznymi i indywidualnymi. Ponadto ważne jest to, jakie aspekty kultury subiektywnej – poza najbardziej popularną dymensją kolektywizm–indywidualizm – powinny być badane.

Jeżeli chodzi o pierwszy problem, to należy przyznać, że potencjalnym źródłem różnic obserwowanych między krajami mogą być czynniki występujące na poziomie ekologicznym, takie jak poziom zamożności kraju, gęstość zaludnienia, klimat (Georgas i in., 2004; Kwiatkowska,

2009), oraz na poziomie indywidualnym, związane z cechami demograficznymi prób poddanych badaniom, i osobowość. Stwierdzono, że istnieją różnice między krajami pod względem natężenia cech osobowości definiowanych w ramach teorii Wielkiej Piątki (Allik, McCrae, 2004): w poszczególnych krajach wykazywane są różnice w poziomie zaagregowanych wskaźników neurotyzmu, ekstrawersji, sumienności, otwartości na doświadczenia i ugodowości. Można więc popełnić błąd, przypisując stwierdzone różnice między danymi krajami (np. Włochami a Norwegią) różnicom kulturowym, podczas gdy będą to różnice osobowościowe między osobami badanymi (np. między ekstrawertywnymi Włochami a introwertywnymi Norwegami). Chociaż czynniki na poziomie ekologicznym i jednostkowym nie mogą być łatwo oddzielone od kultury, badacze powinni przeprowadzić pełne demograficzne oszacowanie swoich próbek, zbadać możliwe relacje między zmiennymi demograficznymi a badanymi zmiennymi psychologicznymi, by być świadomym ich potencjalnego wpływu na różnice kulturowe.

Jaki aspekt kultury subiektywnej powinien być badany? W ciągu ostatnich 20 lat teorię i badania zdominował jeden konstrukt: indywidualizm–kolektywizm. Oczywiście nie bez powodu. Ale są też inne konstrukty, które mogą być nie mniej ważne, a nawet być może bardziej użyteczne w rozumieniu zachowania człowieka. Przypomnijmy: Hofstede (2001) wyróżnił cztery inne dymensje oprócz kolektywizmu–indywidualizmu; badania Shaloma Schwartza nad wartościami wprowadziły kolejny zestaw siedmiu dymensji kulturowych (Schwartz, 1994); badania House'a i współpracowników (House, Hanges, Javidan, Dorfman, Gupta, 2004) nad zachowaniami w organizacjach biznesowych i nad przywództwem w organizacjach – dziewięć dymensji; badania Ingleharta (1997) dotyczące związku wartości z demokratyzacją społeczeństw – dwie dymensje, Leunga i in. (2002) nad społecznymi aksjomatami – pięć aksjomatów; badania Smitha, Dugana i Trompenaarsa (1996) nad wartościami pracowników organizacji – to kolejne dymensje. Problem związany z koncentrowaniem się na jednej dymensji ulega zaostreniu, gdy porównujemy tylko dwa kraje. Matsumoto i Yoo (2006) podają przykład Stanów Zjednoczonych i Japonii, które różnią się od siebie na wielu wymiarach, nie tylko na wymiarze kolektywizmu–indywidualizmu, w dodatku lokują się na różnych wymiarach w różnej odległości. W badaniach Hofstedego (2001) USA zajmuje pierwszą pozycję jako kraj najbardziej indywidualistyczny, Japonia zaś pozycję 27. na 70 krajów, czyli środkową, a mimo to kraje te są prezentowane jako zajmujące ekstremalne pozycje na tej skali. Na skali unikania niepewności Japonia jest ulokowana na pozycji 8., natomiast USA – mają rangę

59. Niewykluczone że wyniki badań porównujących te kraje mogłyby być lepiej wyjaśnione przy użyciu właśnie tego konstruktów (patrz: Iwata, Higuchi, 2000).

Badania prowadzone w ramach tego paradygmatu powinny dać w rezultacie informację o związkach, jakie mogą istnieć między różnicami międzygrupowymi a składnikami kultury. Zdaniem Matsumoto i Yoo (2006), szczególnie obiecujące wydają się tzw. badania rozpakowujące kulturę (*unpacking studies*) oraz badania eksperymentalne z użyciem próbek sytuacji (*situation sampling*) i torowania kulturowego (*cultural priming*). Kilka słów o tych badaniach.

**Badania rozpakowujące kulturę.** Badania te są rozszerzeniem badań mających na celu porównania międzykulturowe, lecz z dodatkowym pomiarem jednej lub wielu zmiennych, co do których zakłada się, że są odpowiedzialne za różnice między kulturami. Rozpakowywanie polega na tym, że kultura jest tu rozumiana jako ogromny zbiór zmiennych, często luźno ze sobą połączonych. W ten sposób niespecyficzna zmienna „kultura” jest zastąpiona przez bardziej specyficzne zmienne, zwane zmiennymi kontekstowymi (*context variables*). Zmienna kontekstowa powinna być zmierzona na poziomie każdego uczestnika badania ze wszystkich badanych kultur. W dalszej kolejności badacz powinien oszacować stopień, w jakim ta zmienna jest odpowiedzialna za obserwowane różnice, głównie poprzez analizę mediacyjną czy analizę kowariancji ANCOVA. Zmienne kontekstowe mogą być rozmaite: indywidualizm–kolektywizm, a w terminologii Triandisa – idiocentryzm i allocentryzm, unikanie niepewności, dystans władzy, egalitaryzm płciowy, orientacja długo- vs. krótkoterminowa, konstrukty Ja, postawy, wartości i przekonania, aksjomaty społeczne, by wyliczyć tylko najpopularniejsze i najlepiej opracowane teoretycznie (patrz: House i in., 2004; Hui, 1988; Leung i in., 2002; Schwartz, 1994; Singelis, 1994; Zimbardo, Boyd, 1999). Ważnymi zmiennymi kontekstowymi są zmienne behawioralne, które odnoszą się do praktyk kulturowych, np. metody wychowawcze, natura relacji interpersonalnych, ekspozycja na ikony kulturowe, doświadczenia edukacyjne, gotowość do przemieszczania się. Wiele z tych zmiennych korzysta z narzędzi opracowanych w naszej kulturze (np. Boski, 2009; Ciecuch, 2013; Pilarska, 2011).

**Podejście eksperymentalne: próbki sytuacji i torowanie kulturowe.** Z oczywistych względów nie można manipulować zmienną „kultura” i przypisywać ją osobom uczestniczącym w badaniu. Można jednak wykorzystać fakt, że kultura nie oddziałuje na ludzi w sposób abstrakcyjny, lecz poprzez konkretne sytuacje. Jak pisze

Heine (2008), kultura regularnie dostarcza pewnych typów sytuacji, na które zaczynamy reagować nawykowo, co wykształca w nas określony sposób myślenia o sobie i o świecie. Jeżeli porównamy reakcje ludzi należących do innej kultury na sytuacje specyficzne dla naszej kultury, otrzymamy informację, w jaki sposób kultura kształtuje nasze funkcjonowanie. Można to wyjaśnić na przykładzie badań przeprowadzonych przez Shinobu Kitayamę, Hazel Markus, Hisayę Matsumoto i Vinea Norasakkunkita (1997). Metoda dobierania próbek sytuacyjnych obejmuje dwa etapy. W pierwszym kroku są gromadzone opisy sytuacji, w których zdarza się coś specyficznego dla danej kultury, co następnie jest wykorzystywane w kolejnym badaniu. W badaniu Kitayamy i współpracowników (1997) Japończycy i Amerykanie opisywali sytuacje, w których doświadczali podwyższenia bądź obniżenia samooceny. Później druga grupa badanych – Japończyków i Amerykanów – otrzymała listę wygenerowanych sytuacji i opisywała swoje odczucia, wyobrażając sobie siebie w podanych sytuacjach. Należy podkreślić, że zarówno Japończycy, jak i Amerykanie reagowali na sytuacje charakterystyczne dla obu kultur: amerykańskiej i japońskiej. Metoda ta pozwala na przeprowadzenie kilku ważnych analiz. Badacz może zadać pytanie, jakie sytuacje zdarzają się częściej w jednej kulturze bądź w drugiej i analizować cechy sytuacji, wygenerowanych w obu kulturach. W cytowanym badaniu okazało się, że Japończycy wymienili więcej sytuacji, w których dochodziło do obniżenia samooceny, niż Amerykanie. Nasuwa się przypuszczenie, że spadek samooceny jest częstszym doświadczeniem kulturowym w Japonii niż w Stanach Zjednoczonych. Badacz może także sprawdzić, czy istnieją różnice kulturowe w reakcjach na te sytuacje. Jeśli bez względu na typ sytuacji badani z jednej kultury reagują w stały sposób, możemy wyciągnąć wniosek, że są to kulturowo wyuczone zachowania. W badaniu Kitayamy i innych (1997) Japończycy reagowali słabszym podwyższeniem samooceny niż Amerykanie w sytuacjach umożliwiających wzrost samooceny (np. pochwała), natomiast reagowali silniejszym spadkiem samooceny niż Amerykanie w sytuacjach obniżających samoocenę (np. nagana). Zdaniem badaczy świadczy to o tym, że Japończycy są wrażliwsi na sytuacje, które stwarzają okazję do samokrytycyzmu, Amerykanie zaś do sytuacji, w których można podnieść poczucie własnej wartości.

**Torowanie kulturowe.** Eksperymenty z torowaniem kulturowym (*cultural priming*) są jeszcze jednym sposobem na manipulowanie kulturą, a ściślej mówiąc, pewnymi jej elementami. Torowanie, inaczej: aktywizacja określonych pojęć czy idei, sprawia, że pojęcia te są łatwiej

dostępne jednostce i mogą wpływać na jej funkcjonowanie. Badania z torowaniem (poprzedzaniem) polegają na manipulowaniu nastawieniem umysłowym uczestników (*mindset*) i pomiarze zmian w zachowaniu. Przykładem może być badanie Trafimowa, Triandisa i Goto (1991), w którym badacze za pomocą odpowiedniej instrukcji wywoływali u Amerykanów i Chińczyków nastawienie na prywatność albo na kolektywność. Instrukcja do generowania prywatnego nastawienia miała następującą postać: „Przez następne dwie minuty nie musisz nic pisać. Pomyśl, co cię wyróżnia w twojej rodzinie, czym różnisz się od twoich przyjaciół”. Natomiast instrukcja do wzbudzania kolektywnego nastawienia brzmiała tak: „Pomyśl, co masz wspólnego z członkami twojej rodziny i z przyjaciółmi. Jakie są ich oczekiwania wobec ciebie?”. Po czym uczestnicy badania wykonali zadanie polegające na kończeniu zdań zaczynających się od: „Ja jestem...”. Odpowiedzi były kodowane zgodnie z tym, czy były zorientowane na Ja, czy na grupę. Okazało się, że Amerykanie udzielili więcej wypowiedzi o charakterze indywidualnym niż Chińczycy; zaobserwowano także silny efekt poprzedzania: od osób otrzymujących instrukcję „prywatną” uzyskano więcej odpowiedzi indywidualnych, od osób otrzymujących instrukcję kolektywną – więcej odpowiedzi kolektywnych, niezależnie od przynależności kulturowej.

#### PROCES BADAWCZY: PUŁAPKI METODOLOGICZNE

##### Uwagi ogólne

Fons van de Vijver i Kwok Leung (2000) opisują zestaw błędnych założeń, które często towarzyszą badaczom podejmującym badania międzykulturowe. Są to swoiste „pre-koncepcje”, działające jak schematy poznawcze, wpływające na cały proces badawczy. Na przykład, zaobserwowane różnice są bezkrytycznie interpretowane jako mające podłoże kulturowe, co nie musi być prawdą (tzw. błąd atrybucji kulturowej; Matsumoto, Yoo, 2006). Z drugiej strony badacze traktują obserwowane różnice w sferze poznawczej jako artefakty, a nie jako różnice międzykulturowe, związane z różnymi praktykami socjalizacyjnymi. Ponadto badacze nie przywiązują wystarczającej wagi do problemu równoważności pomiaru. I ostatni problem, na jaki wskazują autorzy, to nadmierna generalizacja. W badaniach międzykulturowych próbki bywają stosunkowo małe, czasem dobierane są dla wygody, przypadkowo, a nie ze względów merytorycznych. Narzędzia nie zawsze spełniają wymogi trafności i rzetelności, wiele badań trudno zreplikować. Generalizowanie wyników nie zawsze jest uprawnione.

Van de Vijver i Leung (2008) opisują też „paradoks interpretacji”, który polega na tym, że dane dotyczące



dużych i łatwych do zaobserwowania różnic międzykulturowych są trudniejsze do interpretacji. Van de Vijver i Leung paradoks ten wyjaśniają w następujący sposób: największe różnice międzykulturowe najczęściej obserwuje się między populacjami o dużym dystansie kulturowym względem siebie. Na przykład, jeśli test matrycy Ravena wykonują osoby umiejące czytać i pisać oraz analfabeci, to wynik jest łatwy do przewidzenia, ale czy jest równie łatwy do zinterpretowania? Oczywiście można powiedzieć, że analfabeci uzyskali gorszy wynik od osób czytających i piszących, ponieważ różni ich od siebie środowisko kulturowe. Ale co to oznacza? Wielość i różnorodność różnic jest przeogromna: decydują czynniki związane z wykształceniem, z socjalizacją, z codziennymi doświadczeniami, z ekspozycją na działania mediów itd. Istnieje więc problem zidentyfikowania prawdziwej przyczyny, przy czym wraz z dystansem kulturowym wzrasta liczba rywalizujących ze sobą wyjaśnień. Odwrotna sytuacja pojawia się, gdy badamy kultury o niewielkim dystansie. Różnice są trudne do znalezienia, choć są łatwiejsze do interpretacji. Na przykład, jeśli występują różnice w poziomie wykonania testu Ravena między dziećmi w tym samym wieku z dwóch krajów europejskich o mniej więcej podobnym poziomie rozwoju ekonomicznego i o podobnym systemie społeczno-politycznym, to prawdopodobnym czynnikiem odpowiedzialnym za te różnice będzie inny system kształcenia.

Można wskazać kilka źródeł słabości prowadzonych badań. Van de Vijver i Leung ubolewają, że psycholodzy (choć nie tylko oni) organizują swoje badania zgodnie z panującymi w danym czasie paradygmatami. Na przykład w psychologii międzykulturowej panuje tendencja, by różnice międzykulturowe badać i wyjaśniać, często bezkrytycznie, w paradygmacie kolektywizmu-indywidualizmu. Badacze posługują się tymi konstrukcjami zazwyczaj jako wyjaśnieniem czynionym *post factum*, nie przywiązując wagi do konieczności pomiaru tych zmiennych w badaniu. Ponadto badania porównawcze są zazwyczaj prowadzone między krajami Azji Południowo-Wschodniej i USA, które różnią się od siebie na wiele sposobów, nie tylko ze względu na kolektywizm i indywidualizm. Autorzy przypuszczają, że dymensja ta jest tak często używana w badaniach, iż może dramatycznie spaść jej atrakcyjność, a przecież może znakomicie wyjaśniać niektóre zjawiska, np. wzrost ekonomiczny.

Inną słabością jest przywiązanie zbyt wielkiej wagi do wartości  $p$  w badaniach, tj. do poziomu istotności. Badania uważa się za wartościowe i godne opublikowania, gdy  $p$  jest mniejsze od 0,05. Nie jest to najlepsza miara różnic między grupami kulturowymi, które wszak różnią się pod wieloma względami, a nie tylko – jak

w eksperymentach – pod względem jednej zmiennej. Powinno się podawać przynajmniej wielkość efektu, który lepiej odzwierciedla dystans między grupami.

Van de Vijver i Leung (2000) są zdania, że jakość narzędzi do pomiaru zmiennych indywidualnych przewyższa w znacznym stopniu jakość narzędzi używanych do pomiaru zmiennych kontekstowych. Taka sytuacja może nie dziwić w psychologii głównego nurtu, lecz w psychologii międzykulturowej, której naturalnym obiektem badania powinna być interakcja między jednostką a środowiskiem, badacze powinni zwłaszcza ten problem wziąć pod rozwagę.

Van de Vijver i Leung (2001) wyróżniają trzy typy narzędzi ze względu na ich umocowanie kulturowe. Można mówić o narzędziach monocentrycznych (*monocentred*), tj. takich, które powstały w jednej kulturze i mają dobre własności parametryczne oszacowane w badaniach prowadzonych w tej kulturze, np. kwestionariusz do pomiaru Wielkiej Piątki. Powstaje pytanie, czy te własności zostaną utrzymane po przetłumaczeniu narzędzia, gdy chcemy je zastosować w badaniach innej grupy kulturowej. Należy to rozstrzygnąć empirycznie. Stosowanie tego typu narzędzi może generować także inny problem, mianowicie procesy i zjawiska psychiczne, które nie są zbyt wyraziste w kraju, w którym powstało to narzędzie, mogą zostać pominięte.

Narzędzia decentryczne (*decentred*) powstały dzięki zbiorowej pracy badaczy z kilku kultur, by uzyskać narzędzia wolne od wpływów kulturowych (inne określenia: *culture-free*, *culture-fair*, *culture-reduced*). Założenie to spotkało się z krytyką, ponieważ trudno sobie wyobrazić, że pomiar psychologiczny może nie pozostawać pod wpływem czynników kulturowych. Niemniej jednak przekonanie, że są testy psychologiczne odpowiednie do prowadzenia badań międzykulturowych, ponieważ dobrze mierzą cechy ponadkulturowe, jak i specyficznie kulturowe, jest ciągle żywe (van de Vijver, 2002). Przykładem może być badanie wartości Schwartza (van de Vijver, Leung, 2001).

Trzeci typ instrumentów opisywany przez van de Vijvera i Leunga (2001) to narzędzia policentryczne, znane także pod nazwą międzyindygeniczne (*cross-indigenous*). Badacze należący do poszczególnych kultur opracowują narzędzia dla swoich kultur, powstaje w ten sposób zestaw narzędzi specyficznych kulturowo, który następnie jest stosowany w pełnym zestawie w każdej kulturze.

### **Problem równoważności (ekwiwalencji)**

Najważniejsze problemy metodologiczne w prowadzeniu badań międzykulturowych są opisywane jako problem równoważności, inaczej ekwiwalencji badania

(*equivalence*) i problem tendencyjności pomiaru (*bias*). Jeżeli badania międzykulturowe mają przynieść znaczące wyniki, musimy mieć pewność, że uczestnicy badania w taki sam sposób rozumieją sytuację badania, pytania kwestionariuszowe, występujące pojęcia. Innymi słowy, staramy się zapewnić równoważność metodologiczną. Jednak czasem nasze procedury badawcze nie mogą być zastosowane w tej samej postaci w każdej kulturze, dlatego czasem musimy je zmienić, zaadaptować do miejscowych warunków (Triandis, 1994). Taki kompromis czasem jest nieunikniony, zwłaszcza gdy badamy ekstremalnie różne kultury.

Pojęcie ekwiwalencji wiąże się z pytaniem, czy wyniki uzyskane w zakresie jednej pozycji skali czy całego narzędzia pomiarowego w jednej grupie kulturowej mogą być porównywane z wynikami uzyskanymi w innych grupach kulturowych (van de Vijver, 2002). Dysponowanie metodami pomiaru, charakteryzującymi się pełną ekwiwalencją, np. w dwóch różnych językach, jest jednym z najważniejszych wyzwań w prowadzeniu badań międzykulturowych. Ekwiwalencja powinna dotyczyć jak największej liczby aspektów procesu badawczego (najlepiej: wszystkich), tj. doboru prób badawczych, języka, procedur zbierania danych, ram teoretycznych (teorii), odpowiednich hipotez. Zjawisko ekwiwalencji pozostaje w ścisłej relacji z kwestią tendencyjności pomiaru. Jeśli zastosowana skala jest tendencyjna, wyniki są nieporównywalne. Na przykład 10 punktów na skali depresji powinno oznaczać ten sam poziom depresji we wszystkich badanych grupach kulturowych. Jeśli osoby charakteryzujące się tym samym poziomem depresji, np. oszacowanym w badaniach klinicznych, uzyskują różne wyniki na skali, przetłumaczonej na inny język, to skale te nie są równoważne, są tendencyjne.

Pozycja testu lub cały test jest tendencyjny, gdy nie mierzy tych samych własności (cech, zdolności, postaw) w poszczególnych grupach kulturowych. Tendencyjność podważa trafność narzędzia w różnych kontekstach kulturowych. Obrazuje to przykład podany przez Fonsa van de Vijvera i Ronalda Hambletona (1996). W angielskiej oryginalnej wersji pytanie testu brzmiało: „Where is a bird with webbed feet most likely to live?” (Gdzie najczęściej żyje ptak, który ma stopy połączone błonami?), natomiast w przekładzie na szwedzki *webbed feet* zostały zastąpione przez *swimming feet* (pływające stopy, stopy do pływania). Przyznać należy, że pytanie w wersji szwedzkiej dostarcza dość wyraźnej wskazówki, jakie środowiska taki ptak może zamieszkiwać.

W badaniach międzykulturowych problem ekwiwalencji i towarzyszących jej błędów tendencyjności rozpatruje się jako (a) ekwiwalencję funkcjonalną, (b) ekwiwalencję

strukturalną, (c) ekwiwalencją metryczną i (d) ekwiwalencją pełną (Fischer, Miu-Chi Lun, 2008).

**Ekwiwalencja funkcjonalna.** Ekwiwalencja funkcjonalna (*functional equivalence*), inaczej ekwiwalencja konstruktów lub pojęcia (*construct equivalence*) jest to równoważność pod względem konceptualnym konstruktów teoretycznych. Chodzi o to, by być pewnym, że konstrukt, który chcemy poddać badaniom, istnieje także w innych kulturach. Zanim podejmiemy decyzję dotyczące operacjonalizacji i narzędzi pomiaru, należy przeprowadzić gruntowne analizy teoretyczne i badania empiryczne (w tym jakościowe) w każdej grupie kulturowej. Podstawowe pytanie brzmi: czy ten sam konstrukt jest odpowiedzialny za różnice w zachowaniach osób obserwowanych w różnych grupach, tj. między grupami (Fischer, Miu-Chi Lun, 2008).

Problem ekwiwalencji funkcjonalnej może być przedstawiony na przykładzie międzykulturowych kontrowersji wokół modelu Wielkiej Piątki. Według McCrae'a i Costy (1997) istnieją poważne dowody świadczące o uniwersalności tego modelu w stosunku do Amerykanów, Niemców, Portugalczyków, Izraelczyków, Chińczyków i Japończyków. Jednak Cheung ze współpracownikami (1996) stwierdzili, że model ten wyklucza ważne aspekty psychologicznego funkcjonowania Chińczyków związane z czynnikami interpersonalnymi (harmonia w grupie, zachowanie twarzy). W swoim badaniu wykazali, że testowany model może być uniwersalny, ale tylko do pewnego stopnia, ponieważ nie obejmuje wszystkich możliwych konfiguracji cech. Przykładem badania ukierunkowanego na ustalenie precyzyjnego znaczenia jednej z cech Wielkiej Piątki – sumiennosci – jest projekt René Mõttusa i współpracowników (2012a) realizowany w 20 krajach, w tym w Polsce.

Inny przykład to różna konceptualizacja inteligencji. W kulturze zachodniej pojmowanie inteligencji opiera się głównie na osiągnięciach w zakresie zadań poznawczych, a w innych kulturach pojęcie inteligencji zawiera także, lub przede wszystkim, zachowania społeczne (Sternberg, 2007).

Z problemem braku równoważności konceptualnej możemy się spotkać w zupełnie nieoczekiwanych momentach, np. gdy zbieramy dane demograficzne, takie jak choćby wiek respondentów. Na przykład w kulturze !Kung, grup etnicznych związanych z pustynią Kalahari, nie istnieje pojęcie wieku chronologicznego. Zamiast liczenia wieku „absolutnego” w latach, ludzie !Kung stosują specyficzny ranking oparty na relacjach starszeństwa między członkami danej grupy. Na oznaczenie tak ustalonego wieku istnieje bogaty zbiór adekwatnych do danej

relacji określić, z uwzględnieniem płci i pokrewieństwa (Howell, 2009). Wyobraźmy sobie, że mamy w naszym języku specjalne słowa na określenie takich relacji, jak najstarszy syn drugiej żony dziadka czy młodsza córka starszej siostry ojca.

Z brakiem ekwiwalencji funkcjonalnej wiąże się tendencyjność nazywana kulturową specyfiką zmiennej teoretycznej (*cultural specificity of the theoretical variable*), co oznacza, że konstrukt istnieje tylko w pewnym specyficznym kontekście kulturowym, jak dysleksja czy dysgrafia, które występują tylko w społeczeństwie ludzi umiających czytać i pisać (Fischer, Miu-Chi Lun, 2008).

#### **Ekwiwalencja strukturalna (*structural equivalence*).**

Występuje wtedy, gdy w badanych grupach kulturowych na podstawie tych samych zmiennych obserwowalnych można wnioskować o interesujących nas zmiennych teoretycznych. Na przykład stawiamy pytanie, czy na podstawie wyników osiągniętych w tym samym zestawie zadań składających się na test inteligencji, zastosowanym w różnych grupach kulturowych, można wnioskować o poziomie inteligencji. Ten typ równoważności wymaga, by zmienne obserwowalne były dla zmiennych teoretycznych zarówno reprezentatywne, jak i relewantne (van de Vijver i Leung, 1997). W trakcie badania uczestnicy reagują na pewien niewielki zbiór bodźców (np. odpowiadają na pytania z testu inteligencji), który ma reprezentować jakąś zmienną teoretyczną (np. poziom inteligencji). Zakłada się, że uzyskany wynik odzwierciedla jakąś stałą cechę osoby badanej, która determinuje jej zachowanie także w innych sytuacjach, nie tylko w trakcie badania. Tak więc powinno się rozważyć, czy rzeczywiście ów mały zbiór bodźców dotyczy badanej zmiennej (jest relewantny dla inteligencji) i czy dobrze reprezentuje cały zakres tej zmiennej (czy jakieś specyficzne cechy inteligencji nie zostały zignorowane). Ronald Fischer i Vivian Miu-Chi Lun (2008) podają przykład niewłaściwego pomiaru motywacji wewnętrznej za pomocą narzędzia, w którym operowano skomplikowanym językiem, zawierającym rzadko spotykane słowa. W związku z tym wynik w takim badaniu mógł odzwierciedlać nie tylko poziom motywacji, lecz także zdolności werbalne.

Uzyskanie ekwiwalencji strukturalnej wymaga także, by stosowane narzędzia miały tę samą wewnętrzną strukturę we wszystkich interesujących badacza grupach kulturowych. Oznacza to, że w każdej grupie kulturowej te same pozycje ładują te same czynniki, np. w każdej grupie kulturowej stwierdzenie „Dbam o swoje zdrowie” należy do czynnika Ja-niezależne, co jak się wydaje nie zawsze jest prawdą (por. Kwiatkowska, 2012a). Problem ekwiwalencji strukturalnej może być rozwiązany za pomocą

konfirmacyjnej i eksploracyjnej analizy czynnikowej lub skalowania wielowymiarowego (van de Vijver, Leung, 1997). Przykładem rzetelnego dochodzenia do ekwiwalencji strukturalnej mogą być prace Joanny Roszak nad polską adaptacją amerykańskiej skali do pomiaru płci kulturowej (*Hoffman Gender Scale*; Roszak, 2010).

Ekwiwalencja strukturalna może być zagrożona z powodu niewystarczającej bądź zaburzonej reprezentatywności konstruktów, np. zmienne obserwowalne nie wyczerpują wszystkich domen, w których może się ujawniać zmienna teoretyczna. Jeśli w jakiejś kulturze inteligencja ma szerszy zakres pojęciowy (np. zawiera także zachowania społeczne), wtedy badanie wyłącznie kompetencji poznawczych jest niewystarczające.

#### **Ekwiwalencja metryczna (*metric equivalence*).**

Jest to kolejny poziom równoważności (Fischer, Miu-Chi Lun, 2008), a w terminologii Van de Vijvera i Leunga (1997) to ekwiwalencja jednostki pomiaru (*measurement unit equivalence*). Odpowiadamy tu na pytanie, czy możliwe jest porównywanie między grupami kulturowymi wzorów odpowiedzi. Inaczej mówiąc, czy zmiana w obrębie zmiennej latentnej prowadzi do takich samych zmian w wynikach uzyskanych w badanych grupach kulturowych. Ekwiwalencję metryczną graficznie można zobrazować za pomocą dwóch linii regresji, które mają ten sam przebieg w dwóch różnych grupach. Fischer i Miu-Chi Lun (2008) podają przykład pomiaru temperatury (zmienna latentna) wyrażonego w skali Kelvina i w skali Celsjusza (wskazania na termometrze – zmienna obserwowalna). Oczywiście nie można bezpośrednio porównywać wskazań na dwóch termometrach (w stopniach Kelvina i Celsjusza). Ale możliwe jest uchwycenie zmian we wskazaniach układających się według tego samego wzoru: zarówno termometr Kelvina, jak i termometr Celsjusza pokaże relatywnie tę samą zmianę między temperaturami dnia i nocy. Jeśli ekwiwalencja metryczna występuje w dwóch grupach kulturowych, uprawnione jest np. wyciągnięcie wniosku, iż w jednej grupie wyniki zadań werbalnych w teście inteligencji (np. 23 punkty) są wyższe o jedno odchylenie standardowe w porównaniu z wynikami w rozumowaniu abstrakcyjnym, natomiast w drugiej grupie (np. 18 punktów) są wyższe o pół odchylenia standardowego. Jednak nie powinno się porównywać bezpośrednio uzyskanych wyników (np. 23 i 18 punktów) i wnioskować na podstawie różnicy między nimi (np. pięć punktów) o wielkości rzeczywistej różnicy między grupami w zakresie zdolności werbalnych, jak piszą Fischer i Miu-Chi Lun (2008).

Zagrożeniem dla równoważności metrycznej mogą być różne formy tendencyjności pomiaru: tendencyjność

metody, próby, procedury. Sprawdzenie, czy równoważność metryczna jest zadowalająca, może być przeprowadzone za pomocą konfirmacyjnej analizy czynnikowej, również za pomocą dwuczynnikowej analizy wariancji (kultura  $\times$  pozycja testu; dokładniejsze informacje – van de Vijver, Leung, 1997a).

**Pełna ekwiwalencja wyników/wskaźników (*full score equivalence*).** Jest to ostatni, najwyższy poziom ekwiwalencji według Fischera i Miu-Chi Lun (2008). Jeśli faktycznie istnieje pełna równoważność wyników, możemy porównywać uzyskane wartości w sposób bezpośredni, zakładając na przykład, że osoby z różnych grup kulturowych, które uzyskały ten sam wynik w skali werbalnej testu inteligencji (23 punkty), charakteryzują się tym samym poziomem zdolności werbalnych (zmienna latentna), niezależnie od kontekstu kulturowego, a wszelkie różnice w liczbie punktów odzwierciedlają prawdziwe, realnie istniejące różnice w poziomie zdolności werbalnych. Tak jak poprzednio, istnieją odpowiednie procedury statystyczne umożliwiające ocenę, czy osiągnięto pełną ekwiwalencję wyników (np. konfirmacyjna analiza czynnikowa).

Można uznać, że na każdym etapie procesu badawczego psycholog międzykulturowy może napotkać przeszkody zagrażające ekwiwalencji badania. Omówimy najważniejsze z nich pod wspólnym hasłem „błędy tendencyjności metody”.

### **Błędy tendencyjności metody**

Pojęcie tendencyjności metody (*method bias*) odnosi się do tendencyjności próby (*sample bias*), tendencyjności narzędzia (*instrument bias*) i tendencyjności procedury (*administration bias*).

**Tendencyjność próby.** Dobieranie kultur do badania nie jest sprawą prostą. Jeśli potraktować kultury jako obiekty badania, to mamy kilka sposobów na uzyskanie odpowiedniej próby. Pierwszy sposób to randomizacja, tzn. ze wszystkich istniejących dobieramy kultury losowo (*random sampling*). Drugi sposób, najczęściej spotykany, to dobieranie prób ze względów praktycznych, bo tak nam pasuje, bo tak jest wygodniej (*convenience sampling*), np. z powodu ich dostępności, albo dlatego, że koszty związane z badaniami w danym kraju są niższe niż gdzie indziej. Trzeci sposób to dobór celowy (*systematic sampling*), przy czym ów cel jest wskazany przez teorię i pytania badawcze z niej wynikające, jak w wypadku badania, którego celem jest udokumentowanie różnic między kulturami męskimi a kobiecymi (Heine, 2008). Gdy badacz stawia pytanie, np. o to, jak kolektywizm

kształtuje sposób postrzegania relacji interpersonalnych, należy wybrać kultury, które różnią się od siebie pod względem poziomu kolektywizmu.

Chociaż wydaje się rzeczą oczywistą, że w doborze kultur do badania należy kierować się przede wszystkim pytaniem badawczym, to powszechność praktyki polegającej na dobieraniu prób ze względów praktycznych sprawia, iż badane są stale te same kraje. Można sądzić, że w badaniach międzykulturowych mamy nadreprezentację krajów takich jak USA, kraje Europy Zachodniej, Korea czy Japonia. Składa się na to kilka czynników, np. osobiste kontakty naukowców, bliskość położenia geograficznego czy to, że w niektórych krajach badania naukowe są łatwiejsze do przeprowadzenia. W rezultacie najwięcej badań międzykulturowych prowadzi się w krajach uprzemysłowionych (tj. w krajach bogatych – WEIRD<sup>3</sup>), takich jak Ameryka Północna i kraje Azji Wschodniej, w dodatku na najłatwiej dostępnych próbach studenckich.

Kolejny krok to wybór konkretnej próby do badania w każdej z wybranych kultur. Badacze się starają, by próbki badawcze były reprezentatywne dla kultur, będących przedmiotem badania, oraz by osoby badane nie różniły się pod względem zmiennych demograficznych, innych niż kulturowe. Jak wiadomo, uczestnikami większości badań międzykulturowych są studenci wyższych uczelni. Decydując się na badanie prób studenckich, musimy być świadomi, że łatwość, z jaką możemy przeprowadzać badania w tych grupach, pociąga za sobą znaczące koszty. Po pierwsze, powstaje pytanie, czy wyniki uzyskane na podstawie badań studentów mogą być generalizowane na całą populację. A zatem czy można mówić o różnicach kulturowych na podstawie stwierdzonych różnic między badanymi studentami w różnych krajach? Czyniąc takie generalizacje, należy zachować szczególną ostrożność. A przede wszystkim sięgać po próby bardziej zróżnicowane pod względem wieku, wykształcenia, warstwy społecznej itp.

Drugie pytanie, jakie można zadać, dotyczy mocy badania przeprowadzonego na studentach. Czy badanie zaprojektowane w ten sposób ma wystarczającą moc, by wykryć różnice międzykulturowe? Jeżeli porównujemy dwie kultury bardzo podobne do siebie, a za takie możemy uznać studentów dysponujących bardzo podobnymi doświadczeniami niezależnie od kraju zamieszkania, to nasze hipotezy dotyczące różnic międzykulturowych, chociaż teoretycznie uzasadnione, mogą nie zostać potwierdzone. Zdaniem Heinego (2008), porównania między studentami krajów rozwiniętych mają charakter konserwatywny, co

<sup>3</sup> WEIRD – akronim słów: *Western Educated Industrialized Rich Democratic*.

oznacza, że jeśli odkryjemy różnice między tak podobnymi grupami, to tym bardziej efekt kultury (przynajmniej na takim samym poziomie) będzie miał miejsce między grupami bardziej różniącymi się od siebie.

Zakłada się, że grupy studenckie są wystarczająco ekwiwalentne pod względem zmiennych demograficznych, jednak w wielu wypadkach tak nie jest. Chociażby z tego powodu, że wymagania, jakie musi spełniać ktoś, kto chce być przyjęty na uniwersytet, są odmienne w różnych krajach. W niektórych z nich studia uniwersyteckie to luksus, na który mogą sobie pozwolić tylko klasy uprzywilejowane. W innych wykształcenie wyższe jest szeroko dostępne. Tak więc obserwowane różnice między grupami narodowymi mogą nie być pochodną różnic między kulturami, lecz np. różnic między klasami społecznymi. Ale nawet między próbami złożonymi ze studentów mogą wystąpić znaczące różnice, np. wiekowe. W Stanach Zjednoczonych studenci są relatywnie młodzi – najczęściej są to osoby w wieku od 18 do 22 lat. W Polsce na uczelniach prywatnych często studiuje ludzie trzydziestoletni i starsi, zwłaszcza na studiach niestacjonarnych. Inną cechą różnicującą próbki studenckie jest sytuacja życiowa studentów. W USA studenci często już pracują i raczej nie mieszkają z rodzicami. W Japonii i Korei Południowej studenci są na całkowitym utrzymaniu rodziców. We Włoszech z rodzicami mieszkają głównie dorośli synowie, córki zaś wcześniej opuszczają dom. Zignorowanie wymienionych tu różnic między próbami studenckimi może okazać się czynnikiem podważającym wartość badania międzykulturowego.

Oprócz łatwości dostępu do studentów wielką zaletą takich prób jest fakt, że sytuacja badania jest dla nich czymś znanym. Jest to bardzo ważne, ponieważ odpowiadanie na pytania kwestionariuszowe to umiejętność wyuczona kulturowo. W krajach uprzemysłowionych ludzie mają szansę zdobycia pewnej wiedzy dotyczącej badań sondażowych. Ludzie na ogół są zorientowani, na czym polega udział w sondażu. Kiedy ankieter pyta o czyjąś opinię, to nie znaczy, że podejmuje z tą osobą dyskusję, że będzie ona nagrodzona lub ukarana za zgadzanie się lub niezgadzanie się z ankieterem. Jeśli w badaniu biorą udział ludzie mający i niemający takich doświadczeń, a tak czasem bywa, nie można mówić o ekwiwalencji badania.

Patricia Greenfield (1997) opisuje różnice, jakie istnieją między ludźmi z wykształceniem formalnym i bez takiego wykształcenia. Populacje edukowane w szkołach mają doświadczenie z abstrakcyjną rzeczywistością, z symbolicznym przedstawianiem zdarzeń. Większość wydarzeń, z jakimi stykają się osoby kształcące się w szkołach, jest opisywanych w książkach w postaci obrazków czy opisów werbalnych, abstrakcyjnych symboli. Są to zupełnie

różne doświadczenia od tych, które są udziałem dzieci niechodzących do szkoły. Ich kontakt z obiektami czy zdarzeniami jest bezpośredni, a nie przez symboliczną reprezentację. Problem polega na tym, że większość metod stosowanych przez psychologów międzykulturowych zawiera symboliczne przedstawienia obiektu, a więc osoby posiadające wykształcenie formalne są lepiej przygotowane do radzenia sobie w sytuacji badania niż osoby, które nie mają za sobą takich doświadczeń.

Kolejny problem wiąże się z faktem, że osoby bez wykształcenia formalnego pochodzą z kultur, w których apersonalny charakter kontaktów przyjęty w trakcie przeprowadzania badania psychologicznego może być trudny do zaakceptowania. W wielu kulturach kontakty z obcymi są czymś szalenie rzadkim, więc badacz psycholog – ten obcy – powinien poświęcić sporo czasu, żeby nawiązać właściwy kontakt z osobą badaną. Oprócz tego w kulturach kolektywistycznych normy kulturowe nakazują, by nie mieć własnych opinii, a przynajmniej ich nie wyrażać. W zasadzie członek grupy powinien wyrażać opinie grupy, a nie swoje. W badaniu dzieci pojawia się problem dodatkowy, gdy zgodnie z panującymi obyczajami dziecko nie powinno mówić w obecności dorosłych, tylko słuchać, a więc odpytywanie dzieci może być w konflikcie z lokalnymi normami kulturowymi. Jaką więc będzie miało wartość badanie oparte na odpowiedziach na pytania zawarte w kwestionariuszu, starannie opracowanym według zachodnich zasad?

Kolejny problem może być konsekwencją faktu, iż skale używane w psychologii głównego nurtu najczęściej składają się z kilku czy więcej podobnie brzmiących stwierdzeń dotyczących danego konstruktów. Badani mogą być niemile zaskoczeni, że muszą odpowiadać wielokrotnie na to samo pytanie, co wprowadza dyskomfort do sytuacji badania.

Trzeba przyznać, że metody badania społeczeństw, które dzieli duży dystans kulturowy (czego wskaźnikiem mogą być m.in. rodzaj i poziom wykształcenia badanych) podlegają dużo większym ograniczeniom niż metody służące do badania kultur podobnych.

W związku z tym, że obecnie upowszechnia się pozyskiwanie osób badanych przez internet, kilka uwag na temat równoważności prób zdobywanych tą drogą. Przede wszystkim należy sprawdzić, jaka jest dostępność internetu w danym kraju i jakie grupy ludności z niego korzystają. Z polskich badań wynika, że internauci nie stanowią reprezentatywnej próby społeczeństwa. Są młodszy, lepiej wykształceni i mieszkają w większych miejscowościach niż statystyczny Polak (Batorski, Olcoń-Kubicka, 2006). Ponadto internauci częściej angażują się w życie społeczne, mają określone poglądy polityczne, osiągają

wyższe dochody i są bardziej zadowoleni z życia (Wenzel, Feliksiak, Toczyski, 2009). Należy zadać pytanie, jak jest w innych krajach.

**Tendencyjność lingwistyczna.** Podstawowy problem pojawia się wtedy, gdy badacz i uczestnicy używają innych języków. Rozwiązaniem czasem stosowanym jest używanie narzędzi w języku oryginalnym, np. w angielskim, i badanie osób, które są biegłe w obu językach, tj. w angielskim i w ojczystym. Unikamy wówczas kosztów związanych z tłumaczeniem, ale problemy mogą być innego rodzaju. Po pierwsze, nie wiadomo, czy uczestnicy znają tak dobrze język angielski, jak potencjalni tłumacze, z usług których zrezygnowaliśmy. Po drugie, badacz powinien zastanowić się, czy osoby dwujęzyczne są reprezentatywne dla kultury. A trzecia wątpliwość wiąże się z dobrze znanym faktem, że język, jakiego używamy i jakiego używają osoby dwujęzyczne, ma ogromny wpływ na to, jak myślimy i jak się zachowujemy. Chińczyk wypełniający test po angielsku może myśleć i czuć inaczej niż gdyby wypełniał test w języku mandaryńskim. Pouczające są wyniki klasycznego już badania Nan Sussman i Howarda Rosenfelda (1982). Testowano hipotezę o wpływie posługiwania się językiem ojczystym lub angielskim na zachowywany dystans w czasie konwersacji. Kiedy badani (Wenezuelczycy, Japończycy, Amerykanie) rozmawiali w swoim ojczystym języku, zaznaczyły się następujące tendencje: Wenezuelczycy siadali blisko siebie, Japończycy – najdalej, natomiast Amerykanie zajmowali dystans mniejszy niż Japończycy, a większy niż Wenezuelczycy. Gdy jednak wszyscy konwersowali w języku angielskim, uczestnicy badania zajmowali mniej więcej podobny dystans względem swego rozmówcy, niezależnie od kultury pochodzenia.

Tłumaczenie na inny język nie jest wyłącznie problemem technicznym. Powszechnie stosowanym sposobem zapewniającym wierność wersji tłumaczonej oryginałowi jest podwójne tłumaczenie (*back translation*; Brislin, 2000). Najpierw skala jest tłumaczona z angielskiego na polski, następnie przez innego tłumacza z polskiego na angielski. Jeśli ta ostatnia wersja nie odbiega od oryginału, tłumaczenie uznaje się za poprawne. Jeśli obie wersje są różne, badacz powinien przedyskutować sprawy sporne z obydwoma tłumaczami. Istnieje jednak wiele terminów psychologicznych, które nie mają swoich słownikowych odpowiedników, jak np. angielskie *self-esteem* nie istnieje w słowniku chińskim, czy brak w europejskich językach odpowiednika japońskiego słowa *amae*.

**Błędy odpowiedzi (*response bias*).** Drugi problem w badaniach międzykulturowych, zwłaszcza prowadzo-

nych metodą kwestionariuszową, wynika z możliwości popełniania błędów odpowiedzi (*response bias*). Jest to systematyczna tendencja do odpowiadania w pewien określony sposób. Może być kilka rodzajów takiej tendencyjności. Jedną z nich jest skłonność do odpowiadania w sposób społecznie pożądanym (*socially desirable responding*), która może być w pewnych kulturach silniejsza. Ta tendencja może występować w dwóch postaciach: jako skłonność do samooszukiwania się (*self-deceptive enhancement*), czyli stwarzanie obrazu pozytywnej osoby w swoich własnych oczach, oraz jako skłonność do wywierania wrażenia na innych (*impression management*). Lalwani, Shavitt i Johnson (2006) wykazali, że ta pierwsza tendencja jest silniejsza u studentów amerykańskich pochodzenia europejskiego niż u studentów amerykańskich pochodzenia koreańskiego, podczas gdy ta druga – była silniejsza w drugiej grupie badanych.

Inne rodzaje stronniczości w odpowiadaniu to wybieranie odpowiedzi w pewien stały sposób, niezależnie od treści. A więc może to być tendencja do zgadzania się (*acquiescence bias*), czyli raczej do zgadzania się z pytaniami niż do niezgadzania się; tendencja do udzielania skrajnych odpowiedzi (*extreme response bias*), czyli do używania krańców skal bez względu na ich treść; tendencja do udzielania przeciętnych odpowiedzi (*moderacy effect*); efekt grupy odniesienia (*reference group effect*) i efekt deprywacji (*deprivation effect*; Heine, 2008).

Opisane błędy mogą być uznane za artefakty metodologiczne, które powinny być kontrolowane, by uzyskać „prawdziwe” odpowiedzi. Należy jednak spojrzeć na nie również jako na nieodłączne elementy kultury, które wpływają na proces zbierania danych. Johnson, Kulesa, Cho i Shavitt (2004) stwierdzili, że tendencja do udzielania skrajnych odpowiedzi była częstsza w kulturach, w których wysoko wartościowano męskość, władzę i status. Afroamerykanie i Amerykanie pochodzenia hiszpańskiego wykazują tendencję do udzielania bardziej skrajnych odpowiedzi niż Amerykanie pochodzenia europejskiego (Hui, Triandis, 1989). Azjaci z kolei są bardziej umiarkowani w udzielaniu odpowiedzi niż Amerykanie pochodzenia europejskiego (Chen, Lee, Stevenson, 1995). Udzielanie skrajnych lub umiarkowanych odpowiedzi da się kontrolować w ten sposób, że można skalę zmniejszyć do dwóch pozycji: „tak” i „nie”, lub usunąć środkową odpowiedź ze skali, tyle że w ten sposób skala przestaje reagować na subtelniejsze odcienie odpowiedzi, co jest dużym kosztem ponoszonym przez badacza.

Tendencja do zgadzania się jest rzadsza w kulturach indywidualistycznych (Johnson, Kulesa, Cho, Shavitt, 2004), ale może mieć także związek z holistycznym stylem poznawczym, charakterystycznym dla Azji Wschod-

niej (Nisbett, 2009). Jedną z konsekwencji holistycznego sposobu percepcji świata jest założenie, że istnieje więcej prawd niż jedna. Na przykład, zarówno stwierdzenie „Jestem introwertywna”, jak i stwierdzenie „Jestem ekstrawertywna”, są prawdziwe, jeśli przyjmiemy perspektywę holistyczną. W pewnych sytuacjach osoba może być ekstrawertywna, w innych introwertywna. Tak więc ludzie należący do takich kultur mogą mieć tendencję do zgadzania się z obydwojema twierdzeniami, chociaż z punktu widzenia osoby z zachodniej kultury twierdzenia te są sprzeczne (Choi, Choi, 2002). Najprostszym sposobem radzenia sobie z tą tendencją jest konstruowanie narzędzi, gdzie oceny badanego konstruktów na niektórych skalach są odwrócone. Na przykład gdy mierzymy samoocenę, w zestawie powinny się znaleźć takie stwierdzenia, jak: „Często myślę o sobie dobrze” oraz takie jak: „Często czuję się bezwartościowy”.

**Inne zjawiska zakłócające pomiar. Efekt grupy odniesienia (reference-group effect).** Efekt ten występuje wtedy, gdy respondent czyni porównania do swojej grupy odniesienia, zamiast kierować się własnymi wartościami, standardami itd. Zjawisko porównywania się do innych ludzi, zwłaszcza bliskich, podobnych do siebie, gdy dokonujemy ewaluacji jakiejś cechy u siebie, jest dobrze znane w psychologii od czasów Festingera (1954). W badaniach międzykulturowych zjawisko to przysparza kłopotów, ponieważ porównywanie się do różnych grup odniesienia może oznaczać używanie innych standardów do porównań. Szczególnie jest to ważne, gdy w badaniach używamy skal typu Likerta, tzn. badani przekazują swoje subiektywne oceny w danej sprawie. Heine, Lehman, Peng i Greenholtz (2002), opisując problemy ze skalami typu Likerta, posługują się następującym przykładem. Załóżmy, że Amerykanie są rzeczywiście bardziej niezależni niż Japończycy. Gdy przeciętny Japończyk wybierze odpowiedź „Zdecydowanie się zgadzam” w reakcji na jakieś stwierdzenie dotyczące niezależności, np. „Lubię być unikalny i różny od innych”, to przecież nie będzie się odnosił do tego, jak unikalność i bycie różnym pojmują Amerykanie, tylko do tego, jacy w jego mniemaniu są pod tym względem inni Japończycy. Jeśli taką samą odpowiedź wybierze Amerykanin, to wcale nie znaczy, że jest tak samo „unikalny i różny” jak Japończyk, albowiem jego grupą odniesienia są inni – bardziej niezależni niż Japończycy – Amerykanie. Tak więc można nie zaobserwować spodziewanych różnic międzykulturowych, ponieważ znaczenie danego stwierdzenia wynika z zastosowanego standardu – innego dla Japończyka, innego dla Amerykanina.

Jedną z technik umożliwiających zmniejszenie efektu grupy odniesienia, jak radzi Heine (2008), jest konstruowanie skal odnoszących się do zjawisk na bardziej konkretnym poziomie, np. zamiast użyć stwierdzenia zawierającego abstrakcyjną cechę „Jestem pomocny”, można zastosować formę z bardziej konkretnym, jednoznacznie rozumianym opisem sytuacji: „Kiedy przyjaciel poprosi mnie o pomoc w odrobieniu zadania domowego, pomagam mu, nawet za cenę zmiany moich planów”. Można także uczynić odpowiedź bardziej konkretną: zamiast „zgadzam się” czy „zdecydowanie się zgadzam” – „przynajmniej raz dziennie”, „przez 10% czasu” itd. Pozostaje pytanie otwarte, o czym już Steven Heine nie wspomina, czy w ten sposób nie dochodzi do zmiany badanego konstruktów. „Być pomocnym” jest cechą ogólną, zapewniającą stabilność zachowania w różnych sytuacjach, natomiast „pomagać przyjacielowi w odrabianiu zadań” może być zachowaniem reprezentatywnym dla cechy „bycie pomocnym”, a może być tylko jakąś osobliwością w życiu człowieka, który z wyjątkiem tego jednego przypadku raczej nikomu nie pomaga. Co więcej, konkretniejszy sposób formułowania pozycji na skali (także odpowiedzi) jest bliższy koncepcji Ja-współzależnego niż Ja-niezależnego; taka forma testu może być więc trafniejsza w kulturze wschodniej niż w zachodniej.

Można uniknąć efektu grupy odniesienia przez stosowanie wskaźników behawioralnych, np. Levine i Norenzayan (1999) badali tempo życia (*pace of life*), wykorzystując kilka obiektywnych miar zachowania. Była to szybkość pokonywania pewnej odległości przez mieszkańców wybranych miast, dokładność zegarów publicznych oraz szybkość załatwiania klienta na poczcie.

**Efekt deprywacji.** Shalom Schwartz w swoich badaniach prowadzonych w 38 krajach odkrył, że w odpowiedzi na pytanie, jak bardzo cieszysz się życiem, jak bardzo cenisz przyjemności, na trzecim miejscu znalazły się Niemcy Wschodnie (była NRD). Włosi zaś, o których wiemy, że ich styl życia jest pełen przyjemności, zajęli przedostatnią pozycję od końca (Schwartz, 1994). W innym badaniu Amerykanie cenili „pokorę” wyżej niż Chińczycy, natomiast Chińczycy cenili wyżej niż Amerykanie „kierowanie się własnymi celami” (Peng, Nisbett, Wong, 1997). Takie wyniki wydają się niedorzeczne, ale spróbujmy je zrozumieć, czyniąc rozróżnienie między tym, co ludzie rzeczywiście mają, a tym, czego nie mają i pożądamy (Heine, 2008). Jak można odpowiedzieć na pytanie: „Jak bardzo cenisz bezpieczeństwo osobiste?”. Bezpieczeństwo osobiste jest dobrem pożądanym przez wszystkich ludzi na całym świecie, jednak zaczynamy przywiązywać więcej wagi do tego stanu, jeśli nasze

bezpieczeństwo jest zagrożone. W badaniach międzykulturowych można więc otrzymać wynik wskazujący na to, że tam, gdzie bezpieczeństwo ludzi jest długotrwale zagrożone, tam ludzie cenią je bardziej niż w krajach, gdzie ludzie czują się bezpiecznie. Należy więc być bardzo ostrożnym w wyciąganiu wniosków na temat kultury na podstawie subiektywnych ocen wartości. Jedynym sposobem na sprawdzenie trafności tego typu badań jest porównanie ich z innymi miarami cenięcia wartości.

**Błąd ekologiczny (*ecological fallacy*)** (Oyserman, Uskul, 2008) lub **paradoks Simpsona** (Fiedler, 2000). Poleganie na jednostce jako źródle informacji i jednocześnie generalizowanie informacji na społeczeństwo może prowadzić do problemów opisywanych jako błąd ekologiczny, czyli wykorzystanie wiedzy zdobytej na jednym poziomie analizy do wyciągania wniosków na temat danego zjawiska na innym poziomie analizy. Paradoks Simpsona polega na tym, że czasem obserwujemy wyraziste zależności na różnych poziomach analizy, jednak gdy dokonujemy dekompozycji, zależności mogą być słabsze albo o przeciwnym kierunku lub w ogóle się nie pojawić. Inaczej mówiąc, korelacja między dwiema zmiennymi może zniknąć lub się odwrócić, gdy podzielimy próbę według trzeciej zmiennej (Fiedler, 2000). Na przykład można stwierdzić w populacji generalnej pozytywną korelację między rozmiarem obuwia a wysokością wynagrodzenia za pracę. Korelacja znika, gdy badamy ten związek osobno w grupie kobiet i w grupie mężczyzn. Wyjaśnienie jest proste: mężczyźni zarówno noszą większe obuwie, jak i mają wyższe pensje. Oyserman i Uskul (2008) podają przykład związku między opiekowaniem się krewnymi w domu a poziomem indywidualizmu. Badaczki uważają, że chociaż związek ten jest silny, może być tu zaangażowany trzeci czynnik, a mianowicie stopień zamożności osobistej lub zamożności społeczeństwa. Chociaż poziom indywidualizmu i stopień zamożności, mierzone na poziomie kraju, mogą być ze sobą skorelowane, to zamożność może być mediatorem między indywidualizmem a chęcią zapewnienia swoim bliskim domowej, a nie instytucjonalnej opieki. Na przykład od zamożności osobistej i zamożności społeczeństwa zależy to, czy w ogóle istnieją lub jakie są dostępne alternatywne w stosunku do opieki domowej miejsca opieki nad osobami starymi lub chorymi. W związku z tym mieszkańcy krajów o niskim indywidualizmie (a więc krajów biedniejszych) mogą wyrażać chęć opiekowania się swoimi bliskimi nie dlatego, że się czują silniej z nimi związani lub zobligowani do opieki czy definiują siebie bardziej w kategoriach społecznych zobowiązań niż w kategoriach osobistej szczęśliwości, ale dlatego, że nie mają

realnej alternatywy. Podobnie mieszkańcy krajów o wysokim indywidualizmie mogą mniej chętnie opiekować się bliskimi w domu nie dlatego, że czują się mniej związani z nimi, ale dlatego, że wymogi rynku pracy są takie, iż żaden z członków rodziny nie może zostać w domu, by się opiekować starymi rodzicami. Inny przykład: wyobraźmy sobie, że jakieś społeczeństwo charakteryzuje się wysokim procentem matek pozostających w domu i niewielkim procentem matek pracujących, co sprawia wrażenie, że obywatele tego kraju cenią wysoko tradycyjne role rodzinne i macierzyństwo. Jednak poszczególne matki mogą mieć niewiele do powiedzenia w kwestii pozostania w domu lub pójścia do pracy, ponieważ nie ma zbyt dobrze rozwiniętej opieki instytucjonalnej nad dzieckiem, dzień pracy w szkole jest krótki, a rynek oferuje tradycyjne, mało przyjazne dla matek miejsca pracy. A zatem obywatele tego kraju nie wybrali tradycyjnego modelu rodziny, tylko po prostu muszą z nim żyć – piszą Oyserman i Uskul (2008).

Psycholodzy kulturowi ten problem już podnosili, np. Schwartz (1994), który zilustrował niestosowność wnioskowania o cechach jednostek na podstawie cech grupy, posługując się przykładem z pracy Zito (1975; za Schwartz, 1994). Jeśli ława przysięgłych (*hung jury*) nie może dojść do sformułowania konkretnego wniosku o winie, można powiedzieć, że jako grupa jest niezdolna do podjęcia decyzji, jako grupa nie ma opinii. Ale błędem byłoby wnioskowanie, że poszczególni członkowie grupy także są niezdolni do podjęcia decyzji, że nie mają własnych opinii. Wręcz przeciwnie, ława przysięgłych jest niezdecydowana właśnie dlatego, że jej członkowie mają silne własne opinie i nie mogą dojść do porozumienia w efekcie powodują, że ława przysięgłych nie ma opinii.

## ZAKOŃCZENIE

Problemy, z jakimi może się zetknąć psycholog międzykulturowy czy kulturowy, nie powinny zniechęcać do prowadzenia badań w tym obszarze. Przede wszystkim istnieją już szczegółowe rozwiązania, umożliwiające ograniczenie czy nawet pozbycie się błędów wynikających z tendencyjności, a zapewniające równowagę teoretyczną i metodologiczną badania, jak również specjalne procedury statystyczne, które pozwalają na poprawienie jakości metod, np. *item response theory* (patrz: Davidov, Schmidt, Billiet, 2011; Harkness i in., 2010; van de Vijver, Leung, 2000; van de Vijver, van Hemert, Poortinga, 2008). Ale nawet najlepiej opracowane i najbardziej szczegółowe kompendia wiedzy metodologicznej nie zwolnią nas z krytycznej postawy wobec własnych badań i z uświadomienia sobie, że każda metoda ma jakieś niedostatki, a żadne pojedyncze badanie nie może



dać zadowalającej odpowiedzi na postawione pytanie. Dlatego warto rozważyć prowadzenie badań za pomocą wielu metod, tak jak uczynili to np. Richard Nisbett i Dov Cohen (1996, 1999; Cohen, Vandello, Puente, Rantilla, 1999), którzy opracowali interesujące pod względem metodologicznym badanie kultury honoru. Jest to w zasadzie studium przypadku, jak to określa Heine (2008), przeprowadzone z wykorzystaniem danych archiwalnych dotyczących Południa Stanów Zjednoczonych, danych pochodzących z badań sondażowych, z badań eksperymentalnych laboratoryjnych na studentach pochodzących z Południa i Północy, także z eksperymentów naturalnych. To wszystko umożliwiło uchwycenie złożoności psychologiczno-socjologicznej kultury honoru w sposób bogatszy i głębszy niż mogło dać nawet najlepiej zaplanowane pojedyncze badanie.

Łączenie różnych metod badawczych występuje w podejściu zwanym triangulacją. Jest to termin o wielu znaczeniach. Prawdopodobnie wywodzi się z „wielorakiej operacjonalizacji” (*multiple operationalism*) zaproponowanej przez Campbella i Fiske (1995), polegającej na dokonywaniu wielu pomiarów tej samej cechy za pomocą różnych metod, co zapewnia większe prawdopodobieństwo, iż obserwowana wariancja jest związana z mierzoną cechą, a nie z pomiarem. Na przykład to samo zjawisko jest opisywane przez osoby udzielające wywiadu niezależnie od siebie (Denzin, Lincoln, 1998).

Triangulacja oznacza także zbieżność (konwergencję) danych uzyskanych przez poszczególnych badaczy (np. prowadzących badania terenowe), jak również konwergencję między teoriami. Ogólnie rzecz biorąc, w procesie badawczym należy wykorzystywać różne źródła i posługiwać się różnymi metodami uzupełniającymi się nawzajem, o specyficznych dla siebie zaletach i wadach (Denzin, Lincoln, 1998).

Dla van de Vijvera i Chasiotis (2010) triangulacja to kluczowe pojęcie w badaniach łączących metody jakościowe i ilościowe. Jeżeli dane pochodzące z różnych źródeł wskazują na ten sam kierunek przebiegu procesu, tj. wspierają hipotezę, taka zbieżność znacząco wzmacnia siłę konkluzji danego badania. Druga możliwość pojawia się, gdy owe źródła dostarczają informacji dotyczących różnych aspektów badanego zjawiska, są względem siebie komplementarne, tj. ani nie potwierdzają hipotezy, ani jej nie zaprzeczają. Z punktu widzenia metodologii poszerzają zakres wysnuwanych wniosków, ale nie czynią ich trafniejszymi. I trzecia możliwość wyłaniająca się z zastosowania triangulacji to sytuacja, gdy uzyskane informacje nawzajem sobie przeczą. Niekompatybilność danych może wskazywać na to, że przynajmniej jedna metoda nie jest właściwa, a zatem formułowanie konkluzji

nie jest uprawnione (van de Vijver, Chasiotis, 2010). Dov Cohen (2007), będąc zdeklarowanym zwolennikiem stosowania wielu metod w badaniu, proponuje, by w obliczu rozbieżności wyników (*divergent*) zadać sobie dwa pytania. Pierwsze, czy jest to może artefakt metody, a drugie pytanie, czy rozbieżność ta nie odzwierciedla aby jakiegoś znaczącego zjawiska kulturowego. Być może badany problem jest znacznie bardziej skomplikowany niż badacz pierwotnie zakładał i zaobserwowana niespójność to wskazówka, w jakim kierunku podążać, by odkryć coś naprawdę ważnego w kulturze (por. także Mazurowska i Boski, 2014).

Badanie zjawisk występujących na styku jednostki i kultury nastęrcza wiele problemów metodologicznych, ale z drugiej strony przyzwolenie na testowanie różnych metod może stać się ekscytującą przygodą intelektualną, dopóki – jak pisze Cohen (2007) – psychologia kulturowa nie stała się jeszcze dogmatyczną dyscypliną, narzucającą „jedynie słuszny” sposób postępowania.

#### LITERATURA CYTOWANA

- Allik, J., McCrae, R. R. (2004). Towards a geography of personality traits: Patterns of profiles across 36 countries. *Journal of Cross-Cultural Psychology*, 35, 13–28.
- Batorski, D., Olcoń-Kubicka, M. (2006). Prowadzenie badań przez internet – podstawowe zagadnienia metodologiczne. *Studia Socjologiczne*, 3 (182), 99–132.
- Berry, J. W. (1976). *Human ecology and cognitive style. Comparative studies in cultural and psychological adaptation*. Beverly Hills: Sage.
- Berry, J. W., Poortinga, Y. W. (2006). Cross-cultural theory and methodology. W: J. Georgas, J. W. Berry, F. J. van de Vijver, C. Kagitcibasi, Y. W. Poortinga (red.), *Families across cultures. A 30-nation psychological study* (s. 51–71). Cambridge: Cambridge University Press.
- Berry, J. W., Poortinga, Y. H., Segall, M. H., Dasen, P. R. (2004). *Cross-cultural psychology: Research and applications* (wyd. 2). Cambridge: Cambridge University Press.
- Boski, P. (2009). *Kulturowe ramy zachowań społecznych*. Warszawa: Wydawnictwo Naukowe PWN i Academica.
- Brislin, R. W. (2000). *Understanding culture's influence on behavior* (wyd. 2). New York: Wadsworth.
- Campbell, D. T., Fiske, D. W. (1995). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Chen, C., Lee, S.-Y., Stevenson, H. W. (1995). Response style and cross-cultural comparisons of rating scales among East Asians and North American students. *Psychological Science*, 6, 170–175.
- Cheung, F. M., Leung, K., Fan, R. M., Song, W. Z., Zhang, J. X., Chang, J. P. (1996). Development of the Chinese Personality Assessment Inventory. *Journal of Cross-Cultural Psychology*, 27, 181–199.

- Choi, I., Choi, Y. (2002). Culture and self-concept flexibility. *Personality and Social Psychology Bulletin*, 28 (11), 1508–1517.
- Cieciuch, J. (2013). Pomiar wartości w zmodyfikowanym modelu Shaloma Schwartza. *Psychologia Społeczna*, 1 (24), 22–41.
- Cohen, D. (2007). Methods in cultural psychology. W: S. Kitayama, D. Cohen (red.), *Handbook of cultural psychology* (s. 196–236). New York–London: Guilford Press.
- Cohen, D., Nisbett, R. E., Bowdle, B. F., Schwarz, N. (1996). Insult, aggression and the Southern culture of honor: An „experimental ethnography”. *Journal of Personality and Social Psychology*, 70, 945–960.
- Cohen, D., Vandello, J., Puente, S., Rantilla, A. (1999). „When you call me that, smile!” How norms for politeness, interaction styles, and aggression work together in Southern culture. *Social Psychology Quarterly*, 62, 257–275.
- Davidov, E., Schmidt, P., Billiet, J. (red.). (2011). *Cross-cultural analysis. Methods and applications*. New York–London: Routledge.
- Denzin, N. K., Lincoln, Y. S. (red.) (1998). *Collecting and interpreting qualitative materials*. Thousand Oakes, CA: Sage.
- Eysenck, H. J., Eysenck, S. B. (1983). Recent advances in the cross-cultural study of personality. W: J. N. Butcher, C. D. Spielberger (red.), *Advances in personality assessment* (t. 2, s. 41–69). Hillsdale, NJ: Erlbaum.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140.
- Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgement biases. *Psychological Review*, 107, 659–676.
- Fischer, R., Miu-Chi Lun, V. (2008). Measuring cognition and motivation across cultural groups. W: R. M. Sorrentino, S. Yamaguchi (red.), *Handbook of motivation and cognition across cultures* (s. 565–589). Amsterdam: Academic Press Elsevier.
- Georgas, J., van de Vijver, F. J., Berry, J. W. (2004). The eco-cultural framework, ecosocial indices, and psychological variables in cross-cultural research. *Journal of Cross-Cultural Psychology*, 35, 74–86.
- Greenfield, P. M. (1997). You can't take it with you. Why ability assessments don't cross cultures. *American Psychologist*, 52, 1115–1124.
- Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L., Mohler, P. P. i inni (red.). (2010). *Survey methods in multinational, multiregional, and multicultural contexts*. Hoboken, NJ: Wiley.
- Heine, S. J. (2008). *Cultural psychology*. New York–London: W. W. Norton.
- Heine, S. J., Norenzayan, A. (2006). Toward a psychological science for a cultural species. *Perspectives on Psychological Science*, 1, 251–269.
- Heine, S. J., Lehman, D. R., Peng, K., Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales? The reference-group effect. *Journal of Personality and Social Psychology*, 82, 903–918.
- Hofstede, G. (2000). *Kultury i organizacje. Zaprogramowanie umysłu* (przeł. M. Durska). Warszawa: Polskie Wydawnictwo Ekonomiczne.
- Hofstede, G. H. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (wyd. 2). Thousand Oakes, CA: Sage.
- House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., Gupta, V. (red.) (2004). *Culture, leadership, and organizations. The GLOBE Study of 62 societies*. Thousand Oakes, CA: Sage.
- Howell, N. (2009). *Demography of the Dobe !Kung* (wyd. 2). New Jersey: Aldine Transaction Rutgers.
- Hui, C. H. (1988). Measurement of individualism–collectivism. *Journal of Research in Personality*, 22, 17–36.
- Hui, C. H., Triandis, H. C. (1983). Multistrategy approach to cross-cultural research: The case of locus of control. *Journal of Cross-Cultural Psychology*, 14, 65–83.
- Hui, C. H., Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology*, 20, 296–309.
- Inglehart, R. (1997). *The silent revolution: Changing values and political styles among Western publics*. Princeton: Princeton University Press.
- Iwata, N., Higuchi, H. R. (2000). Responses of Japanese and American university students to the STAI items that assess the presence of absence of anxiety. *Journal of Personality Assessment*, 74, 48–62.
- Johnson, T., Kulesa, P., Cho, Y. I., Shavitt, S. (2005). The relation between cultures and response styles: Evidence from 19 countries. *Journal of Cross-Cultural Psychology*, 36, 264–277.
- Kim, U., Yang, K.-S., Hwang, K.-K. (red.) (2006). *Indigenous and cultural psychology. Understanding people in context*. New York: Springer.
- Kitayama, S., Cohen, D. (red.) (2007). *Handbook of cultural psychology*. New York–London: Guilford Press.
- Kitayama, S., Markus, H. R., Matsumoto, H., Norasakkunkit, V. (1997). Individual and collective processes in the construction of the Self: Self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology*, 72, 1245–1267.
- Kryś, K., Hansen, K. (2014). Tylko głupcy uśmiechają się do obcych? Analiza różnic kulturowych w postrzeganiu inteligencji i szczerości osób uśmiechniętych. *Psychologia Społeczna*, 9 (2), 163–176.
- Kwiatkowska, A. (2009). Zróżnicowanie kulturowe świata a rozwój gospodarczy – czy mapa bogactwa i biedy pokrywa się z kulturową mapą świata. *Zeszyty Naukowe Wyższej Szkoły Zarządzania i Prawa im. Heleny Chodkowskiej*, 33, 5–32.
- Kwiatkowska, A. (2012a). Tożsamość a kultura. W: A. Szuster, D. Maison, D. Karwowska (red.), *W stronę podmiotowości. O emocjonalności, tożsamości, dobrych uczynkach i pożytkach płynących z porannego wstawania* (s. 145–180). Sopot: Smak Słowa.

- Kwiatkowska, A. (2012b). Uniwersalność i specyfika kulturowa relacji przywiązania. *Psychologia Wychowawcza, XLIV* (LVIII), 26–43.
- Kwiatkowska, A., Roszak, J., Sikora, R., Kuo, B. C. H., Karpinski, K., Gushchina, T. i in. (2014). Kultura a strategie radzenia sobie ze stresem. Badania międzykulturowe. *Psychologia Społeczna, 1* (28), 67–90.
- Lalwani, A. K., Shavitt, S., Johnson, T. (2006). What is the relation between cultural orientation and socially desirable responding? *Journal of Personality and Social Psychology, 90*, 165–178.
- Leung, K., Bond, M. H., de Carrasquel, S. R., Munoz, C., Hernandez, M., Murakami, F. i in. (2002). Social axioms. The search for universal dimensions of general beliefs about how the world functions. *Journal of Cross-Cultural Psychology, 33*, 286–302.
- Levine, R., Norenzayan, A. (1999). The pace of life in 31 countries. *Journal of Cross-Cultural Psychology, 30*, 178–205.
- Lubiewska, K. (2014). Wpływy polskiego kolektywizmu i niemieckiego indywidualizmu na zachowania rodzicielskie matki i przywiązanie jej nastoletniego dziecka. *Psychologia Społeczna, 9* (2), 198–217.
- Markus, H., Kitayama, S. (1993). Kultura i Ja: Implikacje dla procesów poznawczych, emocji i motywacji. *Nowiny Psychologiczne, 3*, 5–20.
- Masuda, T., Nisbett, R. (2001). Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology, 81*, 922–934.
- Matsumoto, D., Yoo, S. H. (2006). Toward a new generation of cross-cultural research. *Perspectives on Psychological Sciences, 1*, 234–250.
- Mazurowska, K., Boski, P. (2014). Czynniki istotne w procesie rozwiązywania konfliktu w międzynarodowym środowisku pracy. Triangulacja w interpretacji danych. *Psychologia Społeczna, 9* (1), 38–53.
- McCrae, R. R., Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist, 52*, 509–516.
- McCrae, R. R., Terracciano, A. E. (2005). Universal features of personality traits from the observer's perspective: Data from 50 cultures. *Journal of Personality and Social Psychology, 88*, 547–561.
- Mesquita, B. (2001). Emotions in collectivist and individualist contexts. *Journal of Personality and Social Psychology, 80*, 68–74.
- Mesquita, B., Karasawa, M. (2002). Different emotional lives. *Cognition & Emotion, 16*, 127–141.
- Möttus, R., Allik, J., Realo, A., Pullmann, H., Rossier, J., Zecca, G. i in. (2012). Comparability of Self-Reported Conscientiousness Across 21 Countries. *European Journal of Personality, 26*, 303–317.
- Möttus, R., Allik, J., Realo, A., Rossier, J., Zecca, G., Ah-Kion, J. i in. (2012). The Effect of Response Style on Self-Reported Conscientiousness across 20 Countries. *Personality and Social Psychology Bulletin, 38*, 1423–1436.
- Nisbett, R. E. (2009). *Geografia myślenia. Dlaczego ludzie Wschodu i Zachodu myślą inaczej*. Sopot: Smak Słowa.
- Nisbett, R. E., Cohen, D. (1996). *Culture of honor: The psychology of violence in the South*. Boulder, CO: Westview Press.
- Nisbett, R. E., Cohen, D. (1999). Men, honor and murder. *Scientific American Presents, 10*, 16–19.
- Norenzayan, A., Heine, S. (2005). Psychological universals: What are they and how can we know. *Psychological Bulletin, 131*, 763–784.
- Oyserman, D., Uskul, A. K. (2008). Individualism and collectivism: Societal-level processes with implications for individual-level and society level outcomes. W: F. J. van de Vijver, D. A. van Hemert, Y. H. Poortinga (red.), *Multilevel analysis of individuals and cultures* (s. 145–173). New York–London: Lawrence Erlbaum.
- Peng, K., Nisbett, R. (2000). Dialectical responses to questions about dialectical thinking. *American Psychologist, 55*, 1067–1068.
- Peng, K., Nisbett, R. E., Wong, N. Y. (1997). Validity problems comparing values across cultures and possible solutions. *Psychological Methods, 2*, 329–344.
- Pilarska, A. (2011). Polska adaptacja *Skali konstruktorów Ja*. *Studia Psychologiczne, 49*, 21–34.
- Roszak, J. (2010). Culture as a factor influencing expression of one's own identity: Measuring gender identity. W: A. Chybicka, S. Safdar, A. Kwiatkowska (red.), *Culture and gender: An intimate relation* (s. 117–138). Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- Różycka, J., Żemojtel-Piotrowska, M., Truong, T. K. H. (2013). Wartości osobiste i kulturowe w ujęciu Shaloma Schwartza, w kulturze polskiej i wietnamskiej. *Psychologia Społeczna, 4* (27), 396–408.
- Rubinstein, G. (1996). Two people in one land: A validation study of Altemeyer's Right-Wing Authoritarian scale in the Palestinian and Jewish societies in Israel. *Journal of Cross-Cultural Psychology, 27*, 216–230.
- Schwartz, S. H. (1992). The universal content and structure of values: Theoretical advances and empirical tests in 20 countries. W: M. P. Zanna (red.), *Advances in experimental social psychology* (t. 25, s. 1–65). New York: Academic Press.
- Schwartz, S. H. (1994). Beyond individualism/collectivism: New dimensions of values. W: U. Kim, H. C. Traandis, C. Kagitcibasi, S. C. Choi, G. Yoon (red.), *Individualism and collectivism: Theory, method, and applications* (s. 85–119). Newbury Park, CA: Sage.
- Shweder, R. A., Sullivan, M. A. (1993). Cultural psychology: Who needs it? *Annual Review of Psychology, 44*, 497–523.
- Singelis, T. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin, 20*, 580–591.
- Smith, P. B., Dugan, S., Trompenaars, F. (1996). National culture and the values of organizational employees: A dimensional analysis across 43 nations. *Journal of Cross-Cultural Psychology, 27*, 231–264.
- Sorrentino, R. M., Yamaguchi, S. (red.) (2008). *Handbook of motivation and cognition across cultures*. Amsterdam: Academic Press Elsevier.

- Sternberg, R. J. (2007). Intelligence and culture. W: S. Kitayama, D. Cohen (red.), *Handbook of cultural psychology* (s. 547–568). New York–London: Guilford Press.
- Sussman, N. M., Rosenfeld, H. M. (1982). Influence of culture, language and sex on conversational distance. *Journal of Personality and Social Psychology*, 42, 66–74.
- Trafimow, D., Triandis, H. C., Goto, S. G. (1991). Some tests of the distinction between the private self and the collective self. *Journal of Personality and Social Psychology*, 60, 649–655.
- Triandis, H. C. (1994). *Culture and social behaviour*. New York: McGraw-Hill.
- Triandis, H. C. (1996). The psychological measurement of cultural syndromes. *American Psychologist*, 51, 407–415.
- van de Vijver, F. J. (1997). Meta-analysis of cross-cultural comparisons of cognitive test performance. *Journal of Cross-Cultural Psychology*, 28, 678–709.
- van de Vijver, F. J. (2002). Cross-cultural assessment: Value for money? *Applied Psychology: An International Review*, 51, 545–566.
- van de Vijver, F. J., Chasiotis, A. (2010). Making methods meet: Mixed designs in cross-cultural research. W: J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. P. Mohler i in. (red.), *Survey methods in multinational, multiregional, and multicultural contexts* (s. 455–473). Hoboken, NJ: Wiley.
- van de Vijver, F., Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1, 89–99.
- van de Vijver, F. J., van Hemert, D., Poortinga, Y. H. (red.) (2008). *Multilevel analysis of individuals and cultures*. New York–London: Lawrence Erlbaum Associates.
- van de Vijver, F. J., Leung, K. (1997a). *Methods and data analysis for cross-cultural research*. Newbury Park, CA: Sage.
- van de Vijver, F. J., Leung, K. (1997b). Methods and data analysis of comparative research. W: J. W. Berry, Y. H. Poortinga, J. Pandey (red.), *Handbook of cross-cultural psychology* (wyd. 2, t. 1, s. 257–300). Boston: Allyn & Bacon.
- van de Vijver, F. J., Leung, K. (2000). Methodological issues in psychological research on culture. *Journal of Cross-Cultural Psychology*, 31, 33–51.
- van de Vijver, F. J., Leung, K. (2001). Personality in cultural context: Methodological issues. *Journal of Personality*, 69, 1007–1031.
- Wenzel, M., Feliksiak, M., Toczyski, P. (red.) (2009). *Portret internauty. Raport*. Warszawa: CBOS, Gazeta.Pl.
- Zimbardo, P. G., Boyd, J. N. (1999). Putting time in perspective: A valid, reliable individual-differences metric. *Journal of Personality and Social Psychology*, 77, 1271–1288.
- Żemojtel-Piotrowska, M. A., Argentero, P., Baltatescu, S., Bardhwaj, G., Bukowski, M., Calogero, R. i in. (2014). Kulturowe wymiary postaw roszczeniowych. *Psychologia Społeczna*, 9 (2), 218–238.

## Methodological issues in cultural and cross-cultural research

Anna Kwiatkowska

*Institute of Psychology, Polish Academy of Science*

### ABSTRACT

Acquirement of knowledge about human beings, their psychological functioning, is not an ideological neutral process. Research process is a cultural product, and each of its stages reflects ideas, values, behavior rules specific for the given culture. Quite often, psychologists assume, rather too optimistically, that their expertise in main-stream methodology is sufficient to conduct interesting and valid research in cultural and cross-cultural psychology. The purpose of this paper is to make researchers sensitive to specific problems that appear in studying individuals in cultural contexts, and to provide an overview of issues that can influence the process of developing research project, collecting and interpreting data. In the first part of the paper the taxonomy of methods and of the most popular research paradigms was presented, while the second part focused on problems of equivalence and of methodological biases in cross-cultural domains.

Keywords: *cross-cultural research, equivalence, bias*

Złożono: 18.03.2013

Złożono poprawiony tekst: 13.06.2013

Zaakceptowano do druku: 24.09.2013