Forum Oświatowe Vol 38, No 2(74) (2025) ISSN 0867-0323
https://doi.org/10.34862/fo.2025.2.5

Monika Maciejewska
Uniwersytet Warmińsko-Mazurski w Olsztynie
ORCID: https://orcid.org/0000-0003-2945-0884

Kulturowa adaptacja narzędzi badawczych w pedagogice – argumenty za upowszechnieniem tej praktyki

Abstrakt: Artykuł podejmuje kwestie możliwości wykorzystywania kulturowej adaptacji narzędzi badawczych w pedagogice. Procedura ta opiera się na przystosowaniu oryginalnej wersji narzędzia do innej specyfiki kulturowej i może być dokonywana przy przenoszeniu narzędzia stworzonego, funkcjonującego w jednym kraju do innego kraju. Adaptacja nie eliminuje potrzeby tworzenia nowych narzędzi, ale oferuje racjonalną drogę wykorzystania istniejących, teoretycznie zakotwiczonych i empirycznie zweryfikowanych instrumentów tam, gdzie badaczom zależy na porównywalności międzykulturowej oraz oparciu się na już zgromadzonej bazie dowodów. W opracowaniu zaprezentowano dyskusje toczące się nad standardami adaptacji, przedstawiono kolejne etapy tego procesu oraz zaprezentowano dwa przykłady kulturowej adaptacji narzędzi w obszarze pedagogiki: Kwestionariusza Źródeł Poczucia Skuteczności Nauczyciela (STEQ) oraz Kwestionariusza Praktyki Refleksyjnej (RPQ). W konkluzji wskazano, że kulturowa adaptacja narzędzi badawczych powinna być traktowana jako odrębny projekt badawczy, a nie czynność techniczna. W takim rozumieniu może stanowić istotny element prowadzenia rzetelnych badań porównawczych, umiędzynarodowienia pedagogiki oraz budowania bazy standaryzowanych narzędzi badawczych.

Słowa kluczowe: kulturowa adaptacja narzędzia badawczego, narzędzie badawcze

Kontakt:	Monika Maciejewska monika.maciejewska@uwm.edu.pl
Jak cytować:	Maciejewska, M. (2025). Kulturowa adaptacja narzędzi badawczych w pedagogice – argumenty za upowszechnieniem tej praktyki. *Forum Oświatowe*, 38(2), 75–90. https://doi.org/10.34862/fo.2024.2.5
How to cite:	Maciejewska, M. (2025). Kulturowa adaptacja narzędzi badawczych w pedagogice – argumenty za upowszechnieniem tej praktyki. *Forum Oświatowe*, 38(2), 75–90. https://doi.org/10.34862/fo.2024.2.5

Wprowadzenie

Kulturowa adaptacja narzędzi badawczych, czyli przystosowanie oryginalnej wersji narzędzia do innej specyfiki kulturowej, dokonywane przy przenoszeniu narzędzia stworzonego i funkcjonującego w jednym kraju do innego kraju (Matczak, 2011), nie jest popularna w badaniach pedagogicznych. Widać to szczególnie jeśli porównamy zainteresowanie tą formą tworzenia narzędzi wśród psychologów czy przedstawicieli nauk medycznych. Przyczyn takiego stanu rzeczy jest zapewne wiele. W prezentowanym opracowaniu chcę zwrócić uwagę na dwie kwestie z tym związane: wymagania metodologiczne tworzenia standaryzowanych narzędzi badawczych oraz niski poziom umiędzynarodowienia badań pedagogicznych.

Problematykę związaną z tworzeniem narzędzi badawczych w pedagogice trafnie zdiagnozował Krzysztof Rubacha (2013) pokazując, że w badaniach nad edukacją dominuje praktyka doraźnego tworzenia autorskich narzędzi, które w większości nie przechodzą pełnej procedury standaryzacji ani empirycznej weryfikacji. Skutkuje to dużym rozproszeniem wskaźników pomiarowych, brakiem możliwości prowadzenia metaanaliz oraz wątpliwą wiarygodnością wielu wyników. Lektura polskich raportów z badań edukacyjnych pozwala zaobserwować, że większość badaczy rozpoczyna projekt od budowania własnego narzędzia. Warto jednak podkreślić, że takie działanie jest często koniecznością, ponieważ w ramach polskiej pedagogiki (inaczej niż ma to miejsce w psychologii) nie utworzono zorganizowanej bazy i dystrybucji empirycznie zweryfikowanych narzędzi badawczych. K. Rubacha podkreśla jednocześnie, że opracowanie dobrego narzędzia wymaga czasu, zespołu, kompetencji psychometrycznych oraz odpowiednio dużej próby badawczej do analiz. Dobrze przygotowane narzędzie standaryzowane przechodzi pełny cykl opracowania: jasną konceptualizację konstruktu, dobór pozycji zgodny z teorią, pilotaż, analizę mocy dyskryminacyjnej pozycji, szacowanie rzetelności, badanie struktury czynnikowej i normalizację na odpowiednio dobranych próbach. Dopiero po zakończeniu tego złożonego procesu mamy empiryczne podstawy, by twierdzić, iż narzędzie rzeczywiście mierzy to, co mierzyć powinno i robi to w sposób powtarzalny oraz precyzyjny (Rubacha, 2013). Adaptacja kulturowa narzędzia, choć również nie jest procesem prostym, pozwala skrócić proces budowania narzędzia i obniżyć koszty jego

przygotowania oraz dodatkowo korzystać z wzorców, które już zostały sprawdzone w badaniach.

Niewielkie zainteresowanie pedagogów adaptacją kulturową narzędzi badawczych może się również wiązać ze stosunkowo niewielką aktywnością polskich badaczy w międzynarodowym dyskursie o edukacji. Po części jest to zapewne spuścizna powojennej historii rozwoju polskiej pedagogiki, z dominacją indoktrynacji pedagogiki radzieckiej, prowadzącą do zamierania wewnętrznego i zewnętrznego dyskursu. Natomiast jednym z głównych celów kulturowej adaptacji narzędzi badawczych jest uzyskanie międzykulturowej/międzynarodowej porównywalności wyników. Standaryzowane narzędzia są jednym z warunków widoczności międzynarodowej. Dysponując instrumentami o dobrze udokumentowanych właściwościach psychometrycznych badacz może włączać się w badania porównawcze, metaanalizy, projekty międzynarodowe oraz dyskurs różnych dyscyplin. Prowadzona w ostatnich latach polityka umiędzynarodowiania polskich badań, oraz postępujące w tym kierunku strategie rozwoju polskich uczelni również zachęcają (lub też zmuszają) polskich badaczy praktyki edukacyjnej do zwrócenia uwagi na efektywność procedur kulturowej adaptacji narzędzi.

W tym kontekście kulturowa adaptacja narzędzi badawczych jawi się jako strategia, która z jednej strony pozwala podnieść standard metodologiczny badań pedagogicznych przy relatywnie mniejszych nakładach czasowych i finansowych niż konstrukcja autorskiego narzędzia badawczego od podstaw, z drugiej zaś otwiera możliwość włączania wyników badań w szerszy, ponadnarodowy obieg naukowy. Wymaga ona jednak przyjęcia określonej perspektywy metodologicznej, w której narzędzie traktowane jest jako element szerszego projektu standaryzacji, a nie jednorazowy „formularz do zebrania danych", oraz uznania, że porównywalność międzynarodowa wyników jest istotną wartością także dla pedagogiki.

W dalszej części artykułu zostaną więc omówione podstawowe założenia adaptacji kulturowej narzędzi, ze szczególnym uwzględnieniem procedur tłumaczenia, wymogów trafności i rzetelności w badaniach społecznych oraz logiki kolejnych etapów procedury adaptacyjnej. Następnie zostaną zaprezentowane dwa przykłady adaptacji narzędzi badawczych przeprowadzonych przez pedagogów, które posłużą jako ilustracja praktycznych możliwości wykorzystania adaptacji kulturowej w badaniach nad edukacją.

Kulturowa adaptacja i jej podstawowe procedury

Termin kulturowa adaptacja (*cross-cultural adaptation*) odnosi się zarówno do takiego rodzaju przekładu, w którym zachowana jest wierność tłumaczenia opisu oryginalnego instrumentu badawczego, jak i jego adaptacji do warunków kulturowych społeczności, w której będzie to narzędzie używane (Beaton et al., 1998; Brzeziński, 2004). Jest to procedura wykorzystywana z powodzeniem przede wszystkim na polu psychologii, nauk medycznych oraz badań nad zarządzaniem. Pozwala nie tylko na wykorzystanie dobrze opracowanego, standaryzowanego narzędzia badawczego,

ale także odniesienia się do doświadczeń i wyników badań związanych z jego zastosowaniem przez innych badaczy.

W Polsce analizy związane z procesem adaptacji kulturowej podejmowane są przede wszystkim przez psychologów. Badacze dokonujący adaptacji różnego rodzaju kwestionariuszy czy testów powołują się najczęściej na opracowania procedur metodologicznych autorstwa Radosława Ł. Drwala (1995), Elżbiety Hornowskiej i Władysława J. Paluchowskiego (2011)[1].

Radosław Ł. Drwal porządkując problematykę procesu adaptacji wskazał na możliwość stosowania różnych procedur, które prowadzą do odmiennych rezultatów. Jego zdaniem prace nad kulturową adaptacją lokują się na kontinuum od maksymalnie wiernego, dosłownego tłumaczenia, po zbudowanie od podstaw nowego narzędzia, jedynie inspirowanego pierwowzorem. Przytaczane przez niego odmiany procedur to: transkrypcja, translacja, trawestacja, parafraza i rekonstrukcja (Drwal, 1995, s. 15). Najbliżej oryginału znajduje się transkrypcja, rozumiana jako maksymalnie wierne tłumaczenie przy zachowaniu tej samej formy narzędzia i dosłownym oddaniu treści, nawet wtedy, gdy nie są one zgodne z polskimi doświadczeniami kulturowymi. Taka procedura może nasilać różnice kulturowe i zwykle skutkuje niższą rzetelnością i trafnością niż wersja oryginalna, co ogranicza jej użyteczność w badaniach praktycznych. Z kolei procedura translacji oznacza tłumaczenie nastawione na wierność, ale dopuszczające modyfikacje tam, gdzie przekład dosłowny jest niemożliwy. Kolejne z podejść, to trawestacja, która idzie dalej w kierunku przekształcania narzędzia i zakłada bardziej znaczące modyfikacje treści. Mogą one wynikać z intuicji badacza albo opierać się na danych empirycznych, np. poprzez modyfikowanie poszczególnych pozycji narzędzia do momentu osiągnięcia przyjętego kryterium równoważności psychometrycznej. Procedura parafrazy polega natomiast na opracowaniu nowego narzędzia, które nawiązuje do koncepcji teoretycznej i wybranych rozwiązań pierwowzoru. Takie narzędzia nie służą do bezpośrednich porównań międzykulturowych, ale mogą być użyteczne w pomiarze analogicznej zmiennej w warunkach lokalnych. Najdalej posuniętą procedurą jest rekonstrukcja, czyli stworzenie polskiego narzędzia inspirowanego oryginałem od nowa. (Drwal, 1995).

Dyskusję nad procesami adaptacji kulturowej narzędzi (testów i procedur pomiaru) wielokrotnie podejmowali także E. Hornowska i W. Paluchowski. Autorzy wychodzą od tezy, że przenoszenie narzędzi między krajami nie jest głównie problemem języka, lecz różnic kontekstu kulturowego (m.in. norm, instytucji, systemu edukacji, tradycji), które wpływają na interpretację treści, znaczenie konstruktów i zachowania testowe. Ich zdaniem adaptacja ma więc prowadzić do równoważności wersji (nie tylko poprawnego przekładu) i obejmuje takie etapy działania, jak uzasadnienie decyzji o adaptacji, tłumaczenie i dostosowanie materiału bodźcowego, badanie równoważności, a następnie walidację i oszacowanie parametrów psychometrycznych wersji finalnej. Autorzy zwracają także uwagę na pięć mitów (Hornowska i Paluchowski 2006) będących konsekwencją niezrozumienia procedury adaptacji kultu-

1 Do bardzo ważnych tekstów analizujących i opisujących proces adaptacji kulturowej zaliczają się również opracowania J.M. Brzezińskiego (2019) oraz J.M. Brzezińskiego i E. Hornowskiej (1998).

rowej, w wyniku których powstają narzędzia zewnętrznie przypominające oryginał, ale nietrafne w nowej, docelowej kulturze. Do mitów deformujących praktykę adaptacji należy m.in. przekonanie, że jest ona „zawsze" lepsza niż budowa nowego narzędzia. W rzeczywistości sens adaptacji pojawia się zwłaszcza wtedy, gdy badaczom zależy na wykorzystaniu istniejącej bazy empirycznej albo na porównaniach międzynarodowych. W innych sytuacjach trafniejsze może być konstruowanie narzędzia lokalnego. Po drugie, autorzy odrzucają pogląd, że do dobrej adaptacji wystarczy biegła znajomość dwóch języków. Ich zdaniem kompetencje tłumacza powinny obejmować również znajomość kultury docelowej, podstawową wiedzę o mierzonym obszarze oraz rozumienie ograniczeń konstrukcyjnych testu (np. długości pozycji). Po trzecie, podkreślają, że nawet „dobre" tłumaczenie nie gwarantuje trafności pomiarów w nowym kontekście kulturowym. Trzeba zatem wykazać, że narzędzie mierzy ten sam konstrukt, inaczej rośnie ryzyko stronniczości i błędnej interpretacji uzyskanych wyników. Po czwarte, polemizują z założeniem o uniwersalności konstruktów: część pojęć (np. to, co w danej kulturze uznaje się za „inteligencję") może mieć różną treść i wagę, więc adaptacja musi być poprzedzona analizą teoretyczną i empiryczną ekwiwalencji. Po piąte, autorzy pokazują, że badacze podejmujący się adaptacji narzędzia badawczego nie są w stanie przewidzieć wszystkich pułapek wynikających z jego tłumaczenia, dlatego konieczne są badania sprawdzające, m.in. walidacyjne.

Zatem niezwykle istotne jest aby wykazać równoważność oryginału i adaptacji względem szeregu kryteriów, które pozwalają określić, że narzędzie nadal mierzy to, co mierzyło pierwotnie. W polskich opracowaniach wskazuje się na osiem takich kryteriów. Jest to: równoważność teoretyczna (koncepcyjna) odnosząca się do tożsamości badanej cechy; równoważność społeczna, związana ze znaczeniami przypisywanymi badanym zachowaniom w obu kulturach; równoważność funkcjonalna, odnosząca się do tożsamego przeznaczenia narzędzia i badania przez nie tej samej zmiennej; równoważność fasadowa, która dotyczy identyczności formy narzędzia (instrukcji, formy pozycji testowych, procedury badania); równoważność tłumaczenia odnosząca się do wierności przekładu i stopnia nasycenia kulturą poszczególnych pozycji narzędzia; równoważność rekonstrukcji, odnosząca się do identyczności procedury badania danym narzędziem; równoważność psychometryczna, związana z uzyskaniem identyczności wskaźników statystycznych i psychometrycznych oryginału i wersji adaptowanej; równoważność pomiarowa, która odnosi się do porównywalności wyników badania (Drwal, 1995; Strelau i Doliński, 2008).

Podejście takie prezentuje również Jerzy M. Brzeziński (2019), który wskazuje, że istotne w procesie oceny adaptacji kulturowej danego narzędzia jest odniesienie się do siedmiu podstawowych kontekstów: kontekstu teorii psychologicznej (lub pedagogicznej), kontekstu psychometrycznego, kontekstu celu, kontekstu doświadczenia kulturowego osoby badanej, kontekstu świadomości społecznej, kontekstu kwalifikacji zawodowych użytkownika narzędzia (badacza) i kontekstu użycia. Zdaniem autora każdy z zaznaczonych przez niego momentów/kontekstów może sprzyjać popełnieniu błędów, których uważny badacz podejmujący wyzwanie adaptacji kulturowej powinien unikać (Brzeziński, 2019, s. 416–423).

Etapy procesu kulturowej adaptacji narzędzi badawczych

Zachowanie przedstawionych powyżej standardów poprawności kulturowej adaptacji narzędzi badawczych, które mają zapewnić nie tylko poprawność tłumaczenia, ale przede wszystkim równoważność znaczeniową i funkcjonalną narzędzia w różnych kontekstach kulturowych, wymaga zastosowania wieloetapowej procedury, łączącej perspektywę językoznawczą, psychometryczną i społeczno-kulturową. Strukturę i działanie takiego procesu opisał międzynarodowy zespół badaczy: Dorcas E. Beaton, Clarie Bombardier, Francis Guillemin i Marcos B. Ferraz (2000). Jest to schemat, na który najczęściej powołują się badacze dokonujący adaptacji różnych kwestionariuszy (Yasir Arafat et al., 2016; Cruchinho et al., 2024).

Wytyczne opisują pięć podstawowych etapów oraz dodatkowy etap oceny końcowej. Pierwszym krokiem jest tłumaczenie w przód, w którym co najmniej dwóch niezależnych tłumaczy przekłada narzędzie z języka źródłowego na docelowy. Chodzi nie tyle o znalezienie „jedynego słusznego" wariantu, ile o uzyskanie różnych propozycji oddania sensu pozycji, instrukcji i odpowiedzi. Wskazuje się, że tłumacze powinni znać język i kulturę zarówno oryginału, jak i adaptacji, a przynajmniej jeden z nich także terminologię dziedzinową.

W etapie drugim dokonywana jest synteza tłumaczeń, w ramach której badacz wraz z tłumaczami porównuje wersje, omawia rozbieżności, negocjuje brzmienie poszczególnych pozycji i tworzy roboczą, uzgodnioną wersję kwestionariusza w języku docelowym. Ten etap ma charakter konceptualny i wymaga odwołania się do teorii, na której opiera się narzędzie, oraz do wiedzy o kontekście społeczno-kulturowym populacji, w której będzie stosowane.

Trzeci etap to tłumaczenie zwrotne. Uzgodniona wersja w języku docelowym jest przekładana z powrotem na język oryginału przez nowych tłumaczy, którzy nie znają tekstu wyjściowego ani szczegółów konstruktu. Procedura ta jest traktowana jako forma kontroli jakości, ponieważ pozwala wykryć przesunięcia znaczenia, uproszczenia, „dopowiedzenia" czy pominięcia, które mogły pojawić się we wcześniejszym tłumaczeniu. Porównanie *back translation* z oryginałem nie służy uzyskaniu identycznego sformułowania, ale ocenie, czy sens pozycji pozostał zachowany.

Centralne znaczenie w procesie ma etap czwarty związany z pracą zespołu ekspertów (*expert committee review*). W jego skład powinni wchodzić tłumacze, specjaliści dziedzinowi (np. z zakresu zdrowia, pedagogiki, psychologii), metodyk lub psychometra oraz osoba o wysokich kompetencjach językowych, a także opcjonalnie przedstawiciel grupy, dla której narzędzie jest przeznaczone. Zespół ekspertów analizuje równolegle wersję oryginalną, tłumaczenia, wersję syntetyczną i tłumaczenia zwrotne, oceniając ekwiwalencję na czterech poziomach: semantycznym (znaczenie słów i zdań), idiomatycznym (zwroty potoczne, frazeologizmy), empirycznym (czy opisywane sytuacje są rozpoznawalne i typowe w danej kulturze) oraz konceptualnym (czy pojęcia funkcjonują społecznie w podobny sposób). Na podstawie tej analizy formułowana jest wersja przedfinalna narzędzia w języku docelowym oraz

dokumentowane są wszystkie podjęte decyzje, wraz z ich uzasadnieniem teoretycznym i językowym.

W prezentowanej procedurze adaptacji podkreślana jest także konieczność empirycznego sprawdzenia przygotowanej wersji poprzez badanie pilotażowe w populacji docelowej (etap piąty). W niewielkiej, ale zróżnicowanej próbie badani wypełniają kwestionariusz, a następnie wyjaśniają, jak rozumieją poszczególne pytania, jakie znaczenia im przypisują i dlaczego wybierają określone odpowiedzi. Tego typu badanie najczęściej przeprowadzane jest w formie wywiadów kognitywnych. Ten etap pozwala wychwycić niejasne, wieloznaczne lub kulturowo nieadekwatne sformułowania, a także zidentyfikować pozycje, które w danym kontekście społecznym budzą opór, niezrozumienie lub niepokój. Na podstawie wyników tego badania wprowadza się ostatnie korekty, przygotowując wersję narzędzia przeznaczoną do właściwych badań psychometrycznych.

Ostatnim, niejako dodatkowym etapem procesu adaptacji jest wymóg rzetelnej dokumentacji i raportowania całego procesu. Opis adaptacji kulturowej powinien obejmować szczegóły dotyczące doboru tłumaczy, przebiegu poszczególnych etapów, kryteriów podejmowania decyzji przez zespół ekspertów, charakterystyki próby w badaniach pilotażowych oraz wprowadzonych zmian w stosunku do oryginału. Taka dokumentacja jest niezbędnym warunkiem intersubiektywnej kontroli naukowej, czyli pozwala innym badaczom ocenić jakość adaptacji, odtworzyć procedurę w kolejnych badaniach oraz interpretować wyniki z uwzględnieniem ewentualnych ograniczeń narzędzia. W efekcie proces adaptacji kulturowej, zgodnie z propozycją Beaton i jej współpracowników, nie jest traktowany jako techniczna operacja tłumaczenia, lecz jako pełnoprawny projekt badawczy, którego celem jest zapewnienie, że w różnych kulturach mierzone jest to samo zjawisko społeczne w sposób trafny, rzetelny i porównywalny.

Kolejnym krokiem, dopełniającym kulturową adaptację, jest określenie właściwości psychometrycznych ostatecznej wersji nowopowstałego narzędzia badawczego. W literaturze przedmiotu zaznacza się dyskusja na temat tego, które właściwości psychometryczne powinny być brane pod uwagę jako istotne przy ocenie narzędzi pomiarowych, a także w jaki sposób ta ocena może być wykonana. Określone standardy w tym zakresie zostały zaproponowane m.in. przez American Psychological Association (APA) oraz Scientific Advisory Committee of the Medical Outcomes Trust (SAC-MOS) (Aaronson et al., 2002). W naukach medycznych dyskusja ta doprowadziła do powołania tzw. inicjatywy COSMIN (*Consensus-based Standards for the selection of health status Measurement Instruments*), w którą zaangażowani byli specjaliści z rożnych dziedzin (w tym w zakresie psychologii i szeroko rozumianej statystyki). W jej ramach określono kluczowe właściwości psychometryczne istotne dla każdego narzędzia badawczego: rzetelność, trafność, czułość oraz interpretowalność (Mokkin et al., 2010).

Rzetelność jest rozumiana jako wskaźnik dokładności i powtarzalności pomiaru. Narzędzie jest rzetelne wtedy, gdy wielokrotne pomiary, przy braku rzeczywistej zmiany badanej cechy, dają zbliżone wyniki. Wysoka rzetelność oznacza niski błąd

pomiaru. Wyróżnia się kilka sposobów jej oceny: rzetelność test-retest, polegającą na powtórnym zastosowaniu tego samego narzędzia po krótkim czasie; zgodność między oceniającymi, kiedy porównuje się wyniki uzyskane równolegle przez co najmniej dwóch badaczy; stałość pomiarów jednego oceniającego w różnych terminach; spójność wewnętrzną, czyli stopień, w jakim poszczególne pozycje narzędzia badają ten sam aspekt konstruktu (najczęściej szacowaną współczynnikiem alfa Cronbacha), oraz analizę błędu pomiarowego. Ten ostatni określa, jak bardzo mogą różnić się wyniki przy powtarzanych pomiarach w tych samych warunkach i stanowi podstawę do wyznaczenia minimalnej wykrywalnej zmiany, czyli najmniejszej różnicy między dwoma pomiarami, którą można uznać za statystycznie istotną, a nie będącą jedynie fluktuacją błędu. (Hornowska, 2005; Mokkin et al., 2010).

Trafność opisuje, na ile narzędzie rzeczywiście mierzy ten konstrukt, który miało mierzyć, a więc dotyczy wartości diagnostycznej i teoretycznej instrumentu. W ujęciu COSMIN wyróżnia się trzy główne obszary. Trafność treściowa odnosi się do tego, czy zakres pozycji i struktura narzędzia adekwatnie pokrywają wszystkie elementy założonego konstruktu. Jej ocena wiąże się z analizą procesu konstruowania narzędzia i jego struktury wewnętrznej. Trafność teoretyczna (konstruktu) jest uznawana za najważniejszy parametr i dotyczy zgodności wyników pomiaru z założeniami teorii, czyli odpowiedniego odzwierciedlenia wymiarowości konstruktu, powiązań między wskaźnikami wewnątrz narzędzia oraz związków z innymi zmiennymi zewnętrznymi. Może być oceniana poprzez analizę strukturalną (np. analizy czynnikowe), testowanie hipotez teoretycznych oraz analizę międzykulturową, która sprawdza, czy wyniki uzyskiwane wersją przetłumaczoną i zaadaptowaną kulturowo odzwierciedlają wyniki narzędzia oryginalnego. Trafność kryterialna natomiast opiera się na korelacji badanego narzędzia z zewnętrznym kryterium, to znaczy z innym instrumentem o uznanej trafności („złoty standard") lub wskaźnikami przyszłego funkcjonowania. W jej ramach wyróżnia się trafność diagnostyczną, gdy porównuje się wyniki dwóch narzędzi stosowanych w tym samym czasie, oraz trafność prognostyczną, gdy bada się zdolność narzędzia do przewidywania przyszłych zjawisk, np. stanu zdrowia (Brzeziński, 2004; Hornowska, 2019).

Czułość na zmiany dotyczy zdolności narzędzia do wykrywania rzeczywistych zmian w stanie badanego obszaru w określonym czasie. Wyróżnia się czułość wewnętrzną, odnoszącą się do wykrywania zmian pod wpływem określonej interwencji na podstawie porównania wyników przed i po jej zastosowaniu, oraz czułość zewnętrzną, która opisuje zdolność narzędzia do rejestrowania istotnych zmian w odniesieniu do danych referencyjnych pochodzących z innych badań z użyciem tego samego instrumentu. Parametr ten ma szczególne znaczenie w badaniach ewaluacyjnych i interwencyjnych, w których istotne jest nie tylko „złapanie" poziomu cechy, ale także zarejestrowanie jej dynamiki (Mokkin et al., 2018).

Interpretowalność odnosi się do jakości i czytelności wyników uzyskiwanych za pomocą narzędzia, a więc do tego, w jakim stopniu liczby mogą zostać przełożone na sensowne, znaczące informacje. Chodzi o możliwość przypisania wartości ilościowych określonym cechom jakościowym oraz zrozumienia zmian wyników w czasie.

Interpretowalność można wzmacniać poprzez analizy wyników z wielu badań, np. w przeglądach systematycznych, które pokazują jak rozkładają się wyniki w różnych podgrupach i w populacji ogółem (Mokkin et al., 2018).

Przykłady kulturowej adaptacji narzędzi badawczych w pedagogice

Pierwszym przykładem narzędzia badawczego poddanego adaptacji kulturowej na język polski oraz ocenie właściwości psychometrycznych jest Kwestionariusz Źródeł Poczucia Skuteczności Nauczyciela (*Sources of Teacher Efficacy Questionnaire*, STEQ) opracowany przez grupę chińskich badaczy C.K.W. Hoi, M. Zhou, T. Teo i Y. Nie (2017). Polska wersja narzędzia została przygotowana przez Elżbietę Strutyńską-Laskus, Magdalenę Boczkowską, Macieja Karwowskiego i opisana w artykule „Kwestionariusz Źródeł Poczucia Skuteczności Nauczyciela: Charakterystyka psychometryczna polskiej adaptacji" opublikowanym w czasopiśmie Studia z Teorii Wychowania (2023).

Jak argumentują autorzy tego opracowania narzędzie STEQ zostało wybrane do adaptacji kulturowej dlatego, że w odróżnieniu od wielu stosowanych w pedagogice skal mierzących ogólne poczucie skuteczności nauczycieli lub ich przekonania w wybranych domenach (np. zarządzanie klasą, wspieranie uczniów o specjalnych potrzebach), koncentruje się ono na tym, skąd to poczucie się bierze, a więc na mechanizmach jego kształtowania. STEQ jest bowiem zakorzeniony w teorii społeczno-poznawczej Alberta Bandury i bezpośrednio odwołuje się do klasycznego modelu czterech źródeł poczucia własnej skuteczności: doświadczeń mistrzostwa, doświadczeń zastępczych, perswazji społecznej oraz doświadczeń fizjologicznych i emocjonalnych. W tym rozumieniu poczucie skuteczności jest jednym z zasobów człowieka związanym z subiektywnym przekonaniem na temat własnej zdolności do efektywnego zakończenia podejmowanych działań i wyrasta z teorii sprawstwa (*human agency*). W przypadku nauczycieli poczucie własnej skuteczności odnosi się do ich przekonania o posiadaniu wiedzy i zdolności do wpływania na osiągniecia uczniów i ich motywacje do uczenia się (Strutyńska-Laskus et al., 2023).

Z metodologicznego punktu widzenia takie odniesienie teoretyczne oznacza, że narzędzie nie tylko pozwala odnotować poziom samoskuteczności, ale umożliwia rekonstrukcję procesów leżących u jej podstaw. Uzyskanie profilu źródeł poczucia skuteczności jest niezwykle cenne zarówno w badaniach wyjaśniających, projektowaniu oddziaływań ukierunkowanych na rozwój zawodowy nauczycieli, jak i w kształtowaniu środowiska pracy sprzyjającego rozwojowi tego typu kompetencji. Kwestionariusz stosowany w polskich warunkach może zatem nie tylko dostarczać danych do analiz statystycznych, ale także być wykorzystywany jako podstawa projektowania programów kształcenia i doskonalenia nauczycieli, ewaluacji ich efektów oraz monitorowania zmian w czasie.

Proces adaptacji kulturowej Kwestionariusza Źródeł Poczucia Skuteczności Nauczyciela przeprowadzono w oparciu o procedury stosowane w adaptacji narzędzi

psychologicznych (Hornowska i Paluchowski, 2011). Uwzględniając przepisy prawa autorskiego uzyskano zgodę autorów skali na wykorzystanie jej w polskich badaniach. Pierwszy etap prac związany z tłumaczeniem rozpoczął się od tłumaczenia wprost angielskojęzycznej wersji skali z uwzględnieniem jej wszystkich elementów (instrukcji, twierdzeń, odpowiedzi i nazw podskal). Zrealizowało je dwóch niezależnych tłumaczy języka angielskiego. Następnie wyznaczeni sędziowie dokonali przeglądu i porównania tłumaczeń, na podstawie którego uzgodnili ujednoliconą wersję roboczą narzędzia. Wersja ta została poddana tłumaczeniu zwrotnemu. Po porównaniu wersji roboczej z tłumaczeniem zwrotnym uzgodniono ujednoliconą wersję docelową. Kolejnym etapem działania było przeprowadzenie badań za pomocą baterii kwestionariuszy, w tym także skali poczucia skuteczności nauczycieli TSES (Strutyńska-Laskus et al., 2023) wykorzystywanej do oceny trafności zbieżnej, które zostało zrealizowane przez przeszkolonych ankieterów metodą ołówek-papier.

Polska wersja kwestionariusza STEQ, analogicznie do wersji oryginalnej, zawiera 26 pozycji pogrupowanych w cztery podskale odpowiadające źródłom samoskuteczności wg Bandury. Odpowiedzi udziela się na siedmiostopniowej skali Likerta (od „zdecydowanie się nie zgadzam" do „ zdecydowanie się zgadzam") rozszerzonej w stosunku do klasycznej skali o dwa stopnie.

Właściwości psychometryczne badano na próbie 946 nauczycieli o zróżnicowanym wieku, stażu pracy, stopniu awansu zawodowego oraz typie i lokalizacji szkoły, co czyni próbę relatywnie dobrą reprezentacją populacji zawodowej w polskim systemie edukacji.

Badanie właściwości psychometrycznych rozpoczęto od analizy konfirmacyjnej z użyciem estymatora WLSMV, którego celem było ustalenie struktury wewnętrznej skali oraz jej trafności teoretycznej sprawdzającej, czy narzędzie faktycznie mierzy źródła poczucia skuteczności wskazane przez A. Bandurę. Na podstawie uzyskanych wyników stwierdzono bardzo dobre dopasowanie modelu do danych. Badane na kolejnym etapie związki korelacyjne między podskalami okazały się istotne statystycznie i wysokie, co zostało zinterpretowane jako dowód na ich współzależność, ale pozwalającą na uznanie, że mierzą odrębne aspekty źródeł samoskuteczności.

Kolejnym krokiem było określenie rzetelności skali z wykorzystaniem współczynnika alfa Cronbacha oraz konfirmacyjnej analizy czynnikowej (H), które wykazały wysoką spójność wewnętrzną podskal. Wykonano także test stabilności czasowej na próbie 68 nauczycieli, który po dwóch tygodniach wykazał istotne, choć zróżnicowane współczynniki, co autorzy zinterpretowali jako satysfakcjonującą stabilność konstruktu przy jednoczesnej wrażliwości na zmiany kontekstowe. Ustalono również trafność teoretyczną, zbieżną i kryterialną dla kwestionariusza, analizując korelacje wszystkich podskal STEQ ze Skalą Poczucia Skuteczności Nauczycieli (TSES). Zastosowanie regresji wielorakiej pozwoliło określić względną wagę poszczególnych źródeł w przewidywaniu ogólnej skuteczności nauczycieli. W tym wypadku pojawiła się rozbieżność między badaniami polskimi a badaniami nad oryginalną wersją STEQ. W badaniach podstawowych w roli najsilniejszego predyktora występowała perswazja społeczna natomiast w warunkach polskich było to doświadczanie mi-

strzostwa. Różnice te zostały zinterpretowane w kategoriach kontekstu kulturowego i specyfiki warunków pracy nauczycieli w Polsce.

Dodatkowe analizy socjodemograficzne pokazały, że płeć nie różnicuje źródeł poczucia skuteczności, natomiast wiek i staż pracy wiążą się z wyższymi wynikami w zakresie doświadczania mistrzostwa, perswazji społecznej i doświadczeń fizjologiczno-emocjonalnych, co może odzwierciedlać kumulację doświadczeń zawodowych w przypadku polskich nauczycieli.

W podsumowaniu autorzy wskazują, że polska adaptacja STEQ stanowi trafne i rzetelne narzędzie do ilościowego pomiaru czterech źródeł poczucia skuteczności nauczycieli i wypełnia istotną lukę w krajowych badaniach pedagogicznych. Z perspektywy praktyki edukacyjnej wiedza o strukturze i sile działania poszczególnych źródeł samoskuteczności może być wykorzystywana zarówno przez dyrektorów szkół, jak i instytucje kształcące nauczycieli do projektowania interwencji wzmacniających obszary szczególnie istotne dla budowania wysokiego poczucia skuteczności. Autorzy podkreślają, że rozwijanie środowiska pracy sprzyjającego doświadczaniu mistrzostwa, konstruktywnej perswazji społecznej oraz pozytywnym doświadczeniom emocjonalnym z pracy może przekładać się na większe zaangażowanie nauczycieli, wyższą satysfakcję zawodową i mniejsze ryzyko wypalenia, a sam STEQ może stać się ważnym narzędziem diagnostycznym w dalszych badaniach nad funkcjonowaniem zawodowym nauczycieli.

Kolejny przykład procesu adaptacji kulturowej ciekawego narzędzia badawczego, zrealizowany w obrębie pedagogiki, to przystosowanie Kwestionariusza Praktyki Refleksyjnej (*Reflective Practice Questionnaire*, RPQ) w wersji dla nauczycieli i pedagogów. Podstawowa wersja kwestionariusza została opracowana w Stanach Zjednoczonych przez Lynn Priddis i Shanę L. Rogers (2017). Polską wersję narzędzia przygotowała Zuzanna Sury, a jej analiza została zaprezentowana w Przeglądzie Badań Edukacyjnych (2025). Głównym celem badania było opracowanie polskiej wersji narzędzia umożliwiającego ilościowy pomiar „zdolności refleksyjnej" nauczycieli rozumianej w duchu Johna Deweya i Donalda Schöna jako zdolność do krytycznego namysłu nad własnym działaniem oraz do modyfikowania praktyki zawodowej w oparciu o doświadczenie. Jak argumentuje autorka polskiej wersji, RPQ jest obecnie jedynym kwestionariuszem, którego rzetelność została zweryfikowana i który umożliwia eksplorację praktyki refleksyjnej zarówno wśród studentów, jak i profesjonalistów, a jednocześnie jest na tyle elastyczny, że można go stosować w odniesieniu do różnych grup zawodowych (m.in. pracowników socjalnych, psychologów i innych specjalistów świadczących usługi społeczne).

Procedura adaptacji została przeprowadzona zgodnie z wytycznymi dotyczącymi procesu międzykulturowej adaptacji narzędzi samooceny (Beaton et al., 2000). Obejmowała tłumaczenie w przód dokonane przez dwóch filologów znających koncepcję praktyki refleksyjnej, syntezę tych wersji tłumaczeń, a następnie niezależne tłumaczenie zwrotne wykonane przez kolejnych dwóch tłumaczy. W etapie końcowym pięcioosobowy panel ekspertów (pedagodzy specjalizujący się w refleksyjności, psycholog, pracownik socjalny, specjalista nauk o zdrowiu) dokonał oceny wy-

ników tłumaczeń zgodnie z kryteriami ekwiwalencji semantycznej, idiomatycznej, empirycznej i pojęciowej. W miejscach, gdzie współczynnik zgodności W Kendalla wskazywał na rozbieżności ocen, poszczególne pozycje modyfikowano, kładąc nacisk na dopasowanie sformułowań do polskiego kontekstu kulturowego, a nie na dokładność tłumaczenia. Wersję roboczą kwestionariusza, która została skonstruowana na podstawie tych prac, poddano pilotażowi na próbie 63 osób dwujęzycznych – studentów anglistyki z doświadczeniem dydaktycznym.

Właściwa walidacja psychometryczna została zrealizowana na próbie 403 nauczycieli różnych typów szkół (przedszkola, szkoły podstawowe, szkoły ponadpodstawowe) z Polski, o zróżnicowanym wieku i stażu zawodowym. Dobór próby miał charakter kolejny (*consecutive*), wynikający z ograniczeń organizacyjnych i niskiej gotowości do udziału w badaniu, co autorka uznała za kompromis między wymogami metodologicznymi a realiami terenowymi. Kwestionariusz RPQ obejmuje 40 pozycji ocenianych na sześciostopniowej skali Likerta i składa się z kluczowej skali „Zdolności refleksyjnej" (*Reflective Capacity*, RC) – 16 pozycji, z czterema komponentami: refleksja w działaniu (RiA), refleksja nad działaniem (RoA), refleksja z innymi (RO), samoocena (SA). Ta skala RPQ jest uważana za kluczową dla eksploracji praktyki refleksyjnej i jest przedmiotem analizy czynnikowej w prowadzonych badaniach. Oprócz skali RC, Kwestionariusz Praktyki Refleksyjnej obejmuje sześć czteroelementowych skal, które służą do eksploracji cech istotnych dla praktyki refleksyjnej: pragnienie doskonalenia, ogólna pewność siebie, pewność siebie w komunikacji, niepewność, stres w kontakcie z uczniami oraz satysfakcja zawodowa. Jak podkreśla autorka dziesięciu podskal stanowiących pełen kwestionariusz Praktyki Refleksyjnej tworzy zwięzłe narzędzie samooceny, a jednocześnie wszystkie skale mogą być stosowane oddzielnie w różnych kontekstach badawczych (Sury, 2025).

Analizy statystyczne objęły ocenę stabilności czasowej, rzetelności wewnętrznej, struktury czynnikowej oraz związków z zewnętrznymi miarami. Na ich podstawie po niewielkich korektach sformułowań uznano równoważność wersji polskiej i angielskiej za zadowalającą. Współczynniki alfa Cronbacha i omega McDonalda wykorzystane do badania rzetelności wewnętrznej wskazały dobrą spójność wewnętrzną narzędzia w polskiej wersji. Struktura korelacji między podskalami okazała się zbliżona do wcześniejszych badań z użyciem oryginalnej wersji RPQ. Podskale refleksji silnie korelowały ze sobą, a pragnienie doskonalenia i samoocena były wyraźnie powiązane z ogólną zdolnością refleksyjną.

W analizie trafności teoretycznej została przeprowadzona konfirmacyjna analiza czynnikowa dla skali *Reflective Capacity.* Zadowalające, ale dalekie od dobrych wskaźniki uzyskano dla modelu czteroczynnikowego oraz modelu drugiego rzędu, w którym cztery komponenty refleksji (w działaniu, nad działaniem, z innymi, samoocena) są podporządkowane ogólnemu czynnikowi zdolności refleksyjnej. Skłoniło to autorkę do wniosku, że struktura RPQ wymaga dalszych badań, również w innych grupach zawodowych. Zwrócono uwagę, że podskala „refleksja z innymi" (RO) najsłabiej wiąże się z czynnikiem ogólnym, co koresponduje z nowszymi badaniami

zagranicznymi i może wynikać z ograniczonego doświadczenia nauczycieli w formalnych formach refleksji grupowej, takich jak superwizja.

Trafność kryterialną i zbieżną oceniano poprzez korelacje podskal ze skalami innych kwestionariuszy: skróconej wersji Kwestionariusza Ruminacji-Refleksji oraz Mini COPE Kwestionariusza Radzenia Sobie ze Stresem.

W dyskusji podkreślono, że polska adaptacja RPQ jest narzędziem o dobrej rzetelności i akceptowalnej trafności, przydatnym zarówno w badaniach naukowych nad praktyką refleksyjną nauczycieli, jak i w ewaluacji programów kształcenia oraz doskonalenia zawodowego. Jednocześnie autorka zwraca uwagę na istotne ograniczenia: niewystarczająco losowy dobór próby, nadreprezentację kobiet, stosunkowo niewielką liczebność próby pilotażowej, a także samoopisowy charakter narzędzia i trudności związane z operacjonalizacją samego konstruktu refleksyjności. W konkluzji wskazuje, że mimo niejednoznaczności struktury czynnikowej RPQ, kwestionariusz może pełnić ważną funkcję w badaniach i praktyce edukacyjnej, zwłaszcza jako narzędzie inicjujące refleksję nauczycieli nad własną praktyką. Zauważa także potrzebę dalszych badań w tym zakresie, które umożliwiłyby doprecyzowanie modelu teoretycznego i zakresu zastosowań kwestionariusza.

Wnioski końcowe

Przedstawione rozważania wskazują na fakt, że adaptacja kulturowa standaryzowanych narzędzi badawczych ma potencjał do tego, aby stać się jedną z kluczowych praktyk metodologicznych we współczesnej pedagogice. Z jednej strony odpowiada ona na zdiagnozowane przez K. Rubachę problemy: doraźności konstrukcji narzędzi, rozproszenia wskaźników i braku kumulatywności wiedzy. Z drugiej strony wpisuje się w proces umiędzynarodowienia badań, umożliwiając włączanie wyników polskich projektów w szerszy, porównawczy obieg naukowy. Adaptacja nie eliminuje potrzeby tworzenia nowych narzędzi, ale oferuje racjonalną drogę wykorzystania istniejących, teoretycznie zakotwiczonych i empirycznie zweryfikowanych instrumentów tam, gdzie badaczom zależy na porównywalności międzykulturowej oraz oparciu się na już zgromadzonej bazie dowodów.

Analiza standardów wypracowanych w innych naukach pokazuje, że adaptacja kulturowa może (i powinna) być prowadzona według jasno określonych, rygorystycznych procedur. Warto zatem rozpatrzeć możliwość, aby adaptacje i walidacje narzędzi stanowiły odrębne projekty badawcze, w ramach których systematycznie analizuje się rzetelność, różne wymiary trafności, czułość na zmiany oraz interpretowalność wyników. Takie podejście wyraźnie kontrastuje z częstą w pedagogice praktyką traktowania przekładu narzędzia jako krótkiego, technicznego „wstępu" do właściwych badań. Przeniesienie standardów znanych z medycyny i psychometrii (np. wytycznych COSMIN, schematów adaptacyjnych w duchu Beaton et al.) do pedagogiki pozwoliło by na podniesienie statusu metodologicznego badań nad edukacją do poziomu oczekiwanego we współczesnych naukach społecznych.

Przytoczone w niniejszym opracowaniu przykłady polskich adaptacji kwestionariuszy STEQ i RPQ pokazują, że wykorzystanie tych standardów w badaniach pedagogicznych jest nie tylko możliwe, ale także poznawczo i praktycznie owocne. W obu przypadkach adaptacja kulturowa została zrealizowana zgodnie z wieloetapową procedurą (tłumaczenia, praca zespołu ekspertów, pilotaż, walidacja psychometryczna), a uzyskane narzędzia okazały się rzetelne i użyteczne zarówno dla badań wyjaśniających, jak i dla praktyki edukacyjnej. Jednocześnie wyniki analiz (np. różnice w sile poszczególnych źródeł samoskuteczności nauczycieli w STEQ czy niejednoznaczna struktura czynnikowa RPQ) ujawniają, że adaptacja nie polega na prostym „przeniesieniu" testu, lecz prowadzi do ważnych odkryć na temat specyfiki lokalnego kontekstu kulturowego i instytucjonalnego dla badanego konstruktu.

W szerszej perspektywie oznacza to, że adaptacja kulturowa narzędzi może pełnić w pedagogice podwójną rolę. Z jednej strony dostarczać wiarygodnych, standaryzowanych instrumentów pomiaru, z drugiej natomiast stawać się narzędziem krytycznej refleksji nad tym, jak uniwersalne konstrukty teoretyczne „wpisują się" w lokalne realia systemu edukacji. Takie działania mogłyby zbliżyć badania pedagogiczne do międzynarodowych standardów metodologicznych, a jednocześnie pozwalały by na zachowanie wrażliwość na kontekst społeczny i kulturowy, w którym działają nauczyciele, uczniowie i instytucje edukacyjne.

Konsekwencją takiego podejścia do kulturowej adaptacji narzędzi badawczych byłyby oczywiście zmiany w praktyce badawczej. Z metodologicznego punktu widzenia konieczne jest, aby projekty adaptacyjne były traktowane jako pełnoprawne zadania badawcze (z odpowiednim czasem, budżetem i miejscem w dorobku naukowym), a nie jako uboczny etap realizacji „właściwych" badań. Z instytucjonalnego punktu widzenia potrzebne jest w tym wypadku tworzenie zespołów międzydyscyplinarnych (łączących pedagogów, psychometrów, statystyków, językoznawców), rozwój krajowych baz standaryzowanych narzędzi oraz włączanie problematyki kulturowej adaptacji narzędzi badawczych do kształcenia metodologicznego studentów i doktorantów.

Podsumowując, argumenty przedstawione w artykule, które zostały wzmocnione doświadczeniami psychologii i nauk medycznych oraz przykładami wybranych, udanych adaptacji w pedagogice, prowadzą do wniosku, że upowszechnienie kulturowej adaptacji narzędzi badawczych w znaczący sposób może poprawić rzetelność badań nad edukacją oraz zwiększyć uczestnictwo polskich badaczy w międzynarodowym dyskursie.

Bibliografia

Aaronson, N., Alonso, J., Burnam, A., Lohr, K. N., Patrick, D. L., Perrin, E., Stein, R. E. (2002). Assessing health status and quality-of-life instruments: attributes and review criteria. *Quality of life research*, *11*(3), 193–205. https://doi.org/10.1023/a:1015291021312

Beaton, D. E., Bombardier, C., Guillemin, F., Ferraz, M. B. (2000). Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Brzeziński, J. (2019). *Metodologia badań psychologicznych*. Wydawnictwo Naukowe PWN.

Brzeziński, J. (red.). (2004). *Metodologia badań psychologicznych. Wybór tekstów*. Wydawnictwo Naukowe PWN.

Brzeziński, J., Hornowska, E. (red.). (1998). *Skala Inteligencji Wechslera WAIS-R*. Wydawnictwo Naukowe PWN.

Ciechanowicz, A. (red.). (1990). *Kulturowa adaptacja testów*. Pracownia Testów Psychologicznych PTP.

Cruchinho, P., López-Franco, M. D., Capelas, M. L., Almeida, S., Bennett, P.M., Miranda da Silva, M., Teixeira, G., Nunes, E., Lucas, P., Gaspar. F., Handovers-4SafeCare. (2024). Translation, Cross-Cultural Adaptation, and Validation of Measurement Instruments: A Practical Guideline for Novice Researchers. *Journal of Multidisciplinary Healthcare*, *17*, 2701–2728. https://doi.org/10.2147/JMDH.S419714

Drwal, R. Ł. (1995). *Adaptacja kwestionariuszy osobowości. Wybrane zagadnienia i techniki*. Wydawnictwo Naukowe PWN.

Hoi, C. K. W., Zhou, M., Teo, T., Nie, Y. (2017). Measuring effi cacy sources: Development and validation of the Sources of Teacher Effi cacy Questionnaire (STEQ) for Chinese teachers. *Psychology in the Schools*, *54*(7), 756–769. https://doi.org/10.1002/pits.22025

Hornowska, E. (2005). *Testy Psychologiczne*. Wydawnictwo Naukowe Scholar.

Hornowska, E. (2019). *Testy psychologiczne. Teoria i praktyka*. Wydawnictwo Naukowe Scholar.

Hornowska, E., Paluchowski, W. J. (2006). Wokół mitów dotyczących adaptacji kulturowej narzędzi diagnostycznych. *Prace Naukowe Akademii Ekonomicznej we Wrocławiu*, (1130), 468–476. https://dbc.wroc.pl/dlibra/publication/170618

Hornowska, E., Paluchowski, W. J. (2011). Kulturowa adaptacja testów psychologicznych. W: J. Brzeziński (red.), *Metodologia badań społecznych. Wybór tekstów* (s. 177–242). Wydawnictwo Zysk i s-ka.

Matczak, A. (2011). Kulturowa adaptacja testów psychologicznych. W: T. Szustrowa (red.), *Testy w biznesie. Standardy i praktyka* (s. -). Pracownia Testów Psychologicznych PTP.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*(7), 737–745. https://doi.org/10.1016/j.jclinepi.2010.02.006

Mokkink, L. B., Prinsen, C. A., Patrick, D. L., Alonso, J. Bouter, L. M. de Vet H. C., Terwee, C. B. (2018). COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs). VU University Medical Center. https://

cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf
Priddis, L., Rogers, S. L. (2017). Development of the Reflective Practice Questionnaire: Preliminary Findings. *Reflective Practice*, *19*(1), 1–16. https://doi.org/10.1080/14623943.2017.1379384
Rubacha, K. (2013). Standardy metodologiczne w zbieraniu i analizie danych. Pracownia Narzędzi Badawczych Komitetu Nauk Pedagogicznych PAN. *Przegląd Badań Edukacyjnych*, 1(16), 35–41. https://doi.org/10.12775/PBE.2013.003
Strelau, J., Doliński, D. (red.). (2008). *Psychologia. Podręcznik akademicki*. Gdańskie Wydawnictwo Psychologiczne.
Strutyńska-Laskus, E., Boczkowska, M., Karwowski, M. (2023). Kwestionariusz Źródeł Poczucia Skuteczności Nauczyciela: Charakterystyka psychometryczna polskiej adaptacji. *Studia z Teorii Wychowania*, 4(45), 261–275.
Sury, Z. (2025). Polish Adaptation and Psychometric Evaluation of the Reflective Practice Questionnaire: Teachers and Pedagogues Version. *Przegląd Badań Edukacyjnych*, *49*(1), 201–230. https://doi.org/10.12775/PBE.2025.012
Yasir Arafat, S. M., Chowdhury, H. R., Shalahuddin Qusar, M. M. A., Hafez, M. A. (2016). Cross-cultural adaptation and psychometric validation of research instruments: A methodological review. *Journal of Behavioral Health*, *5*(3), 129–136.

Cross-Cultural Adaptation of Research Instruments in Pedagogy: Arguments for the Wider Adoption of This Practice

Abstract: The article addresses the possibilities of using cross-cultural adaptations of research instruments in the field of pedagogy. This procedure is based on adjusting an original version of an instrument to a different cultural context, and may be applied when transferring a tool developed and employed in one country to another national setting. Adaptation does not eliminate the need to construct new instruments, but it offers a rational strategy for utilising existing, theoretically grounded, and empirically validated measures in situations where researchers seek cross-cultural comparability and intend to build on an already accumulated body of evidence. The paper presents ongoing debates on standards of adaptation, outlines the subsequent stages of the adaptation process, and discusses two examples of cross-cultural adaptation of instruments in pedagogy: the Sources of Teacher Efficacy Questionnaire (STEQ) and the Reflective Practice Questionnaire (RPQ). In conclusion, it is argued that cross-cultural adaptation of research instruments should be treated as an independent research project, rather than a merely technical operation. Understood in this way, it can become a crucial component of conducting robust comparative studies, advancing the internationalisation of educational research, and developing a repository of standardised research instruments.

Keywords: cross-cultural adaptation of a research instrument, research instrument