



Sławomir Pasikowski

Uniwersytet Łódzki, Wydział Nauk o Wychowaniu, e-mail: pasiks@wp.pl

## **Zagrożenia pomiaru cech niejednorodnych i cech rozmytych w badaniach nad edukacją**

DOI: <http://dx.doi.org/10.12775/PBE.2017.012>

### **The Threats of Heterogeneous and Fuzzy Variables Measurement in Research in Education**

#### **Abstract**

The aim of the article is to present the main threats to the validity of measuring the fuzzy variables and the variables characterized by a dynamic structure of the opposition. These threats are mainly due to inadequate modeling of this type of variables, routine choice of rating scales in the process of measurement design and the use of numerical methods unsuited to a verbal format of data in which these variables could be expressed. In the article the schemes of heterogeneous and fuzzy variables modeling were presented. These schemes are based on models developed in the theory of attitudes and on the system of fuzzy logic.

**Key words:** education, measurement, validity, heterogeneous variables, fuzzy variables.

Pomiar w pedagogice nie musi wiązać się z operacjonalizacją cech zmierzającą w jej końcowej fazie do nadania danym empirycznym postaci liczb. Argumentów dostarcza nie tylko reprezentacyjna koncepcja pomiaru, która znosi konieczność tzw. addytywności wielkości (w uproszczeniu: łączenia/sumowania wartości lub wariantów danej zmiennej) i za której prekursorów uchodzą S.S. Stevens (1946) oraz P. Suppes i J.L. Zinnes (1963). Uzasadnienie przynosi też obraz kreślony na poziomie definicji pomiaru. Utożsamianie pomiaru z liczbo-

wym formatem danych samo może być zresztą traktowana jako jedno z podstawowych zagrożeń zniekształcających odbiór projektów badawczych i prezentowanych w nich rezultatów, bo identyfikującym pomiar wyłącznie z operacjami arytmetycznymi. Tymczasem teoria zbiorów rozmytych, do której odwołuję się niżej, dostarcza podstawy pod dokonywanie pomiaru na danych werbalnych, czyli subiektywnych i niejednoznacznych, z nastawieniem na zachowanie tych własności jako cennych, a nie bezwzględnie koniecznych do eliminowania.

### **Trafność jako podstawowa własność pomiaru**

Celem pomiaru jest zdanie sprawy z własności mierzonego zjawiska. Najogólniej rzecz ujmując, polega to na dostarczeniu opisu tych własności za pośrednictwem zastosowanego instrumentarium i w języku charakterystycznym dla tego instrumentarium. Jeśli więc o zagrożeniach pomiaru mowa, to chodzi nie tyle o wąsko rozumiany akt pomiaru, co wspomniany cel, któremu on służy. Własności zjawisk interesujących pedagogikę z rzadka tylko są bezpośrednio obserwowalne. Bywa raczej tak, że i o samym zjawisku orzeka się pośrednio w oparciu o symptomy uznane za takie na podstawie przyjętej teorii, utrwalonych przekonań, lub poprzednich obserwacji pozwalających łączyć antecedynty i następstwa. Kwestia opisu niewidocznych lub wręcz latentnych zjawisk i ich własności czyni tym wyraźniejsze zagadnienie wiarygodności i adekwatności pomiaru, które wyrażają się przede wszystkim w pojęciu trafności. Pojęcie to najkrócej można wyrazić jako adekwatność narzędzi i rezultatów pomiaru względem mierzonej dziedziny zachowania, na temat której formułowane będą wnioski dotyczące bieżącej oraz przyszłej sytuacji (por. Guilford, 2005; Messick, 2005)<sup>1</sup>. To właśnie trafność bywa ofiarą błędów, mniej lub bardziej

---

<sup>1</sup> Inną ważną właściwością, przez którą wyraża się wiarygodność pomiaru, jest rzetelność. Rzetelność jako własność narzędzia pozwala określać dokładność rezultatu pomiarowego na zasadzie powtarzania pomiaru. Rzetelność jest zatem tym większa, im bardziej podobne do siebie są rezultaty kolejnych pomiarów przy użyciu tego samego narzędzia i w tych samych (lub przynajmniej znacznie zbliżonych) warunkach, w jakich pomiar przebiega. Do jej oceny wykorzystuje się różne metody traktujące poszczególne akty pomiaru jako odrębne, ale równoległe. Ocena rzetelności może więc polegać na: ponownym użyciu tego samego narzędzia (stabilność bezwzględna i względna, czyli stałość narzędzia w czasie), porównywaniu wyników dwóch części wydzielonych z tego samego narzędzia (rzetelność połówkowa) lub badaniu związku pomiędzy wynikami dostarczonymi przez poszczególne pozycje (itemy) instrumentu (zgodność wewnętrzna). Obok tych sposobów ustalania rzetelności jako osobny wyróżniana jest tzw. metoda sędziów kompetentnych, polegająca na testowaniu zgodności pomiędzy ocenami przydzielanymi narzędziu przez niezależne osoby (Jankowski, Zajenkowski, 2009). Ocenie podlega odpowiedniość pomiar względem

świadomych zaniedbań albo czynionych z premedytacją nadużyć podczas realizacji czynności związanych z pomiarem. Obniżanie trafności pomiaru skutkuje zniekształceniami obrazu zjawiska albo wręcz jego mylną identyfikacją. Podkreślenia wymaga, że kwestia trafności dotyczy tak samo badań realizowanych w strategii ilościowej, jak i jakościowej. W obu bowiem przypadkach dąży się do zbudowania modelu zjawiska, choć przy użyciu odmiennego medium symbolicznego: liczb albo słów. Istnieją w tej sprawie liczne argumenty pokazujące, że najczęściej dyskusja toczy się wokół różnic terminologicznych przy jednoczesnym utrzymywaniu zgodności co do idei wiarygodności rezultatów badawczych oraz adekwatności działań podejmowanych przez badaczy (Guba, 1981; Morse, Barret, Mayan, Olson, Spiers, 2002).

Trafność pomiaru w badaniach pedagogicznych stanowi wypadkową wielu warunków, wśród których podstawowym jest dopasowanie koncepcji modelującej jakąś nieobserwowalną wprost cechę. Zwykle cechy będące przedmiotem zainteresowania takich badań okazują się na tyle złożone, że typowe działania badawcze oparte na formułowaniu pytań lub zadań wymagają budowania całych ich zestawów do testowania każdego z aspektów lub wymiarów przewidywanych przez model danej cechy, i to jedynie z nadzieją na przybliżenie. Nie można więc zapominać, że model ten jest idealizacją i ostatecznie poddany musi zostać empirycznej weryfikacji zanim instrument pomiaru będzie mógł być uznany za nadający się do użycia. Brak dostatecznej dbałości o model cechy oraz niedostateczna weryfikacja narzędzi pomiaru skutkuje obniżaniem trafności pomiaru. Nie wdając się w pogłębione analizy złożonego zagadnia trafności<sup>2</sup>, niezbędne tu nam będzie jedynie rozróżnienie jej trzech podstawowych rodzajów (Guion, 2005; Hornowska, 2009): trafność teoretyczna (związek narzędzia pomiaru z konstruktem teoretycznym), trafność treściowa (zakres, w jakim treść pozycji narzędzia reprezentuje uniwersum treści), trafność kryterialna (związek narzędzia z zewnętrznym wobec niego kryterium; nazywana w związku z tym trafnością empiryczną (Nowakowska, 1975, s.47)). Osobno uwzględnić warto tzw. trafność fasadową (*face validity*), niepozostającą w bezpośrednim związku z trzema wymienionymi, ale rzutującą na uzyskiwane rezultaty pomiaru. Trafność fasadowa odnosi się do ogólnej orientacji badanych

---

zjawiska, które jest przez dane narzędzie mierzone. Uważam, że tę metodę można ustawić na początku szeregu, jaki tworzą poprzednie, uzyskując tym samym „kontinuum» oceny rzetelności”, począwszy właśnie od metody sędziów kompetentnych, w której za odrębne pomiary (realizujące się w metaujęciu sędziów) uznawane są oceny poszczególnych sędziów.

<sup>2</sup> Samuel Messick, pisząc o pomiarze psychometrycznym przy użyciu testów, wyróżnia 15 rodzajów trafności (Messick, 2005).

osób w zakresie tego, co mierzy dany instrument, a polega na wzbudzaniu ich wrażenia, wiary czy też odczucia sensu i odpowiedniości narzędzia względem tego, do czego ma być przeznaczone (por. Guilford, 2005, s. 93–94; Anastasi, Urbina, 1999, s. 164). Analityczne podejście do zagadnienia trafności generuje obraz tej własności dużo bardziej różnorodny i złożony, a niekiedy nastrożający trudności w utrzymywaniu rozdziału pomiędzy poszczególnymi rodzajami trafności. Nadmienię więc tylko, że istnieją też holistyczne ujęcia zabezpieczające przed gubieniem z oczu sensu trafności. Takim podejściem jest stanowisko Guion (2005), który z całą mocą podkreśla, że trafność jest jedna, a wszelkie rozróżnienia nie mają charakteru rodzajowego, lecz aspektualny. Jego unitarystyczna koncepcja trafności nie tyle przeciwstawia się ujęciu trynitarnemu, co raczej je wchłania. W związku z tym problematyka pomiaru cech niejednorodnych i rozmytych prezentowana tu będzie w kontekście zagadnienia trafności jako unitarnej własności pomiaru.

## **Modelowanie cech niejednorodnych i cech rozmytych**

Aby jednak zrozumiała stała się idea adekwatności modelu względem modelowanej cechy, jasności nabrać musi samo pojęcie cechy. Otóż w badaniach empirycznych pomiarowi poddawane są w istocie wybrane właściwości obiektów i zjawisk, o których wyborze decyduje cel badań. Te właśnie właściwości to cechy, przez które wyraża się specyfika danego obiektu, czyniąca go w jakimś zakresie podobnym, a w jakimś różnym od innych obiektów. W metodologii badań empirycznych cecha jest pojęciem absolutnie podstawowym. Przysługują mu też inne podziały pozwalające wyodrębnić różne postaci cech. Na przykład, wśród podstawowych rozróżnień znajduje się podział ze względu na zachodzenie występowania wariantów lub wartości, jakie dana cecha może przybierać (stałe i zmienne), a także podział ze względu na liczbę tych wariantów lub wartości (dychotomiczne i politomiczne). W przypadku cech ilościowych za ważny uznaje się też podział ze względu na to, czy pomiędzy sąsiednimi wartościami cechy możliwe jest ustalenie wartości pośredniej, a pomiędzy nią i sąsiednią kolejną pośrednią (dyskretne i ciągłe). Największe jednak znaczenie w praktyce ma podział ze względu na format danych, w jakich cecha się wyraża (jakościowy, ilościowy) oraz podział ze względu na poziom pomiaru cechy (głównie: nominalny, porządkowy, interwałowy, ilorazowy)<sup>3</sup>. Oba te podziały do pewnego

---

<sup>3</sup> Podział ze względu na poziom pomiaru (skalę pomiaru rozumianą jako system matematyczny) jest z kolei podstawą współczesnej teorii pomiaru w jej ujęciu reprezentacyjnym. Sta-

stopnia zachodzą na siebie, tzn. cechy wyrażające się w formacie werbalnym są mierzalne na poziomie nominalnym, najwyższej porządkowym. Cechy ilościowe, czyli wyrażane w formacie liczbowym, na dwóch pozostałych.

Nasuwa się w tym momencie pytanie o charakterystykę przywołanych w tytule cech niejednorodnych i rozmytych, która pozwoliłaby wyobrazić sobie ich miejsce w podanych podziałach.

CECHY NIEJEDNORODNE to takie, które wykazują biegunową strukturę, opartą na opozycyjnych aspektach. W zbiorze wartości, które przyjmują, można więc wyodrębnić osobne podzbiory pozostające w relacji przeciwieństwa. Przykładami takich cech są postawy (pozytywne – negatywne), motywacja („do” – „od”), gotowość do działań o przeciwnym charakterze (np. inkluzyjne – ekskluzyjne), przeciwstawne dążenia w obrębie tych samych celów lub też sprzeczne elementy tożsamości. Pomiar polegający bezwzględnie na granicach pomiędzy wartościami zmiennych i klasami tych wartości fałszywie ujednoznacznia obraz takich cech i ogranicza możliwość uchwycenia charakteryzującego je napięcia wewnętrznego czy też naprężenia.

Z kolei CECHY ROZMYTE, nazywane też zmiennymi lingwistycznymi, charakteryzują się niejednoznacznością i brakiem wyraźnej granicy pomiędzy wartościami występującymi w konkretnych obserwacjach empirycznych, zależnością od kontekstu oraz subiektywizmem w opisie, ocenie i ekspresji<sup>4</sup>. Są al-

---

nowi kluczowe zaplecze budowania modeli pomiaru cech oraz regulowania precyzji pomiaru. W kolejnych zdaniach podaję ich charakterystykę, choć w zdecydowanie uproszczonej postaci. Nominalny poziom pomiaru zapewnia jedynie możliwość dokonywania klasyfikacji ze względu na posiadanie (lub nie) danego wariantu cechy. Porządkowy oprócz możliwości poprzedniego gwarantuje też rangowanie wariantów cechy, a wraz z nimi obiektów, którym są one przypisane. Interwałowy poziom pomiaru daje dodatkowo możliwość ustalania stałej jednostki pomiaru, co jest podstawą dokonywania operacji dodawania i odejmowania na wartościach, jakie przyjmuje zmienna w zbiorze obserwacji. Ilorazowy poziom pomiaru umożliwia ustalenie punktu, w którym zmienna przyjmuje zerowe natężenie. Dzięki tej własności wartości przyjmowane przez zmienną można poddawać mnożeniu, co stwarza warunki do formułowania twierdzeń na przykład takich, że mając 10 jednostek (np. złotych), ma się dwa razy więcej niż pięć jednostek. Tego rodzaju twierdzenia nie dają się sensownie formułować wobec wyników testowych pomiarów inteligencji, wiedzy lub postaw.

<sup>4</sup> Pojęcie tego rodzaju cechy związane jest z systemem teoretycznym logiki rozmytej Lotfi Zadeha. System obejmuje wąsko rozumianą logikę, teorię zbiorów, teorię relacji oraz teorię poznania (Zadeh, 2009). U jego podstaw tkwi spostrzeżenie przez Zadeha (1973) różnicy pomiędzy systemami mechanicznymi i systemami humanistycznymi. Dużo większa złożoność tych ostatnich powoduje poważne ograniczenia trafności opisów i analiz wykorzystujących instrumenty oparte na klasycznej logice dwuwartościowej i inspirowanych nią metodach numerycznych. To, co charakteryzuje zjawiska związane z systemami humanistycznymi, to nieprecyzyjność, niepewność, wieloznaczność i subiektywizm. Stąd w przypadku ich badania pożądana jest metodologia

ternatywą dla zmiennych numerycznych. Z cechami możliwymi do traktowania jako rozmyte ma się do czynienia choćby wtedy, gdy wskazaniu tego samego wariantu cechy przez różne osoby odpowiada odmienna reprezentacja umysłowa u każdej z tych osób. Cechy rozmyte wyrażają się w pojęciach języka naturalnego (Zadeh, 1973) pod postacią rzeczowników, czasowników, przymiotników, przysłówków, ale też całych fraz i sentencji. Do prostszych przykładów należy: „wiek” (młody, średni, stary), „uczy się” (bardzo łatwo, łatwo, przeciętnie, trudno, bardzo trudno), „szybkość rozwiązywania problemów” (szybko, przeciętnie, wolno) oraz „dom” (element bardziej lub mniej „domowy”). Poszczególne wartości zmiennej lingwistycznej są etykietami podzbiorów rozmytych. Im większą liczbę wartości przyjmuje zmienna, tym rozleglejsza jest kolekcja podzbiorów rozmytych mogących ją reprezentować. Pojęcie cechy lingwistycznej koresponduje więc wyraźnie z pojęciem cechy jakościowej.

Ustalenie poziomu pomiaru w przypadku cech niejednorodnych i cech rozmytych wymaga wcześniejszego odwołania się do teorii mierzonego zjawiska oraz do teorii samego pomiaru. Cechy rozmyte, potraktowane jako jakościowe, nie muszą bowiem być bezwarunkowo rozpatrywane na nominalnym poziomie pomiaru. Podobnie cechy niejednorodne nie wymagają traktowania jako bezwzględne całości konieczne do oddania na jakiejś pojedynczej skali szacunkowej. W obu przypadkach podstawowe znaczenie ma konceptualizacja tych cech oraz dobór do nich adekwatnego modelu pomiaru. Cechy rozmyte i cechy niejednorodne daje się bowiem rozpatrywać jako te, które przysługują obiektom w różnym stopniu (albo przynajmniej jako przypisywane w różnym stopniu przez obserwatora) oraz jako możliwe do wyrażania przy pomocy dymensji. Pozostawanie w ich przypadku przy klasycznych formułach grozi obniżeniem trafności późniejszego pomiaru, co wyrażać się może na przykład niezdolnością obejmowania obserwacją niektórych symptomów cechy lub wartości, które może ona przyjmować. Próby osiągnięcia precyzji na drodze transformacji numerycznej, która zwykle w takich przypadkach sprowadza się do zamiany słów na cyfry, sztucznie zawyża poziom pomiaru i daje złudne wrażenie docierania do danych głębiej ukrytych, a w rzeczywistości pozostaje niedopasowaną formułą odwzorowania obserwowanych zjawisk.

---

o adekwatnej elastyczności, czyli tolerancji na nieprecyzyjność i dopuszczanie stopniowości prawdziwości twierdzeń. Zadeh (1973) uzasadniał to tym, że podstawowe własności ludzkiego myślenia zdecydowanie bardziej związane są z kategoriami rozmytych zbiorów niż z liczbami. Oznacza to zatem, że przynależność do zbiorów i relacje między nimi nie są ujmowane przez umysł jako kategoriyczne, dyskretne czy zerojedynkowe, lecz podlegają stopniowaniu, które może przyjmować dowolne wartości z zadanego kontinuum.

## **Zagrożenia realizacji pomiaru cech niejednorodnych i cech rozmytych**

Pomiar w klasycznym ujęciu polega na przypisywaniu symboli (najczęściej liczbowych) poszczególnym wartościom zmiennej według określonych wcześniej reguł, wśród których najważniejsza jest ta mówiąca o oddawaniu relacji empirycznych zachodzących między obiektami przez relacje pomiędzy symbolami (np. Nowakowska, 1975, s. 205). Systemu opisu tych reguł dostarcza teoria zbiorów. Aby pomiar cechował się trafnością, konieczne jest wcześniejsze stworzenie REPREZENTACJI niewidocznych własności danych obiektów empirycznych i uczynienie tego w sposób zapewniający pożądany stopień JEDNOZNACZNOŚCI tej reprezentacji (Walenta, 1971). Najogólniej rzecz biorąc, pomiar opiera się na ustaleniu odpowiedniości między zbiorem obiektów empirycznych oraz jego modelem, czyli zbiorem symboli (w teorii skalowania zwanym systemem relacyjnym) wykorzystywanym do odwzorowania tych pierwszych. Tę odpowiedniość opisuje funkcja przyporządkowująca, która pozwala ustalać adekwatność pomiędzy obiektami obu zbiorów, ale też odzwierciedlać relacje na elementach zbioru symboli, jakie zachodzą pomiędzy elementami zbioru empirycznego<sup>5</sup>. Wartości zmiennej (czyli stopień, w jakim dana cecha charakteryzuje dany obiekt) traktowane są tu jako rozłączne klasy, tzn. przypisanie jakiejś obserwacji danej wartości wyklucza zaliczenie jej do klasy innej wartości. Każda obserwacja przyjmująca dowolną wartość podlega więc przyporządkowaniu do określonej klasy według zasady zerojedynkowej. Tę specyfikę widać szczególnie wyraźnie przy pomiarach na poziomie

---

<sup>5</sup> Ta odpowiedniość może być jednoznaczna lub nie. Jednoznaczność przyporządkowania określonych symboli danym obiektom empirycznym stanowi kluczowe zagadnienie dla trafności pomiaru. Im ta jednoznaczność jest wyższa, tym pomiar bardziej precyzyjny (Brzeziński, 1999, s. 205; Choynowski, 1971; Walenta, 1971), odwzorowanie jednego systemu przez drugi dokładniejsze, a co wyrażane jest zwykle za pomocą pojęcia skali/poziomu pomiaru. Należy jednak pamiętać, że precyzja, z jaką pomiar jest dokonywany, nie musi być zmaksymalizowana, aby trafnie orzekać o interesującym badacza zjawisku. Wszystko zależy od celu i dokładności, z jaką badacz zamierza później formułować wnioski. Przyporządkowania symboli obiektom empirycznym mogą podlegać przekształceniom (np. matematycznym), ale dowolność tych przekształceń jest znacznie ograniczona przy funkcjach coraz bardziej jednoznacznych. W teorii skalowania funkcja (odwzorowanie) o najwyższym stopniu jednoznaczności nazywana jest bijekcją. W tym kontekście mówi się o odwzorowaniu homomorficznym i izomorficznym. To pierwsze cechuje się jednoznacznością, tzn. każdemu elementowi systemu empirycznego przyporządkowany zostaje tylko jeden symbol. Z kolei te drugie charakteryzują się wzajemną jednoznacznością, czyli jednocześnie każdemu symbolowi odpowiada tylko jeden element systemu empirycznego (Choynowski, 1971; Walenta, 1971).

nominalnym i porządkowym bazujących na klasyfikacji obiektów, takich jak np. opinie odpowiedzi osób badanych. Jeśli jednak obserwowana cecha ma ciągły charakter (np. wysokość) lub, co ważniejsze, nie poddaje się modelowaniu za pomocą idei rozłącznych zbiorów wartości, to klasyczna formuła pomiaru spowoduje utratę tych informacji, które znajdują się pomiędzy sąsiadującymi klasami. Nawet sumowanie wartości, jak to się niekiedy odbywa przy pomiarach z użyciem szacunkowych skal numerycznych (np. Likerta), nie znosi tego zagrożenia. Udaje się je minimalizować dopiero przez zastosowanie modelu kontinuum w operacjonalizowaniu zmiennych. Dzieje się tak dlatego, że przynależność danego obiektu do klasy, jaką tworzy wariant zmiennej jakościowej (lingwistycznej), daje się stopniować w modelu kontinuum. W przypadku tego rodzaju zmiennych sam model kontinuum jednak nie wystarcza. Wymaga on wsparcia ideą rozmytej operacjonalizacji cech (Noworol, 2013), czyli takiej, która dopuszcza płynność granic między klasami i możliwość przynależenie do kilku klas naraz<sup>6</sup>.

Zatem modelowanie (konceptualizacja) zjawisk wieloznacznych, niejednorodnych i złożonych stwarza znacznie mniej ograniczeń dla trafności, jeśli w miejsce zerowego klasyfikowania lub szacunkowych skal dyskretnych wprowadza się pojęcie wymiaru i kontinuum. Warto tu jednak dodać, że i ujęcia bipolarne leżące u podstaw jednowymiarowych biegunowych skalach odcinkowych lub punktowych (jak w przypadku dyferencjału semantycznego) mogą niewystarczająco oddawać własności mierzonych cech. Tym bardziej więc jako nieskuteczne przedstawiają się otwarte formaty odpowiedzi (Schuman, 2013). Ale także nawet najbardziej wymyślne zestawy kafeterii nie są

---

<sup>6</sup> Zwiększanie trafności pomiaru zmiennych lingwistycznych odbywa się zwykle przez transformację ich wariantów do postaci wyrażeń rozmytych (liczb rozmytych). Taka transformacja może odbywać się: metodą ekspertów (*expert scaling methods*), wybraną metodą klasyfikacyjną, jak np. sieci neuronowe (*machine-learning methods*) lub w drodze indywidualnej oceny respondentów (*pollster methods*) (Jefmański, 2011a). W przypadku ekspertów i indywidualnych ocen osób badanych polega to na określeniu zakresu wartości liczbowych możliwego do reprezentowania przez dane wyrażenie lingwistyczne. Może to dotyczyć dopasowywania wartości liczbowych (np. wyrażających liczbę centymetrów) do słowa „wysoki”, lub wartości liczbowych do słowa oznaczającego ogólniejszą kategorię (np. „szkoła”). W tym ostatnim przypadku chodziłoby o przypisywanie wartości liczbowych obiektom zależnie od stopnia, w jakim reprezentują tę daną kategorię. Pierwszy przykład dotyczy operacji na danych z porządkowego poziomu pomiaru, a drugi z poziomu nominalnego. W obu transformacja polega na zastosowaniu kontinuum wartości. Transformacja wyrażeń lingwistycznych pozwala też na odkrycie charakteru funkcji przynależności elementów do zbioru rozmytego. Ta może bowiem przyjmować różne postaci, wśród których najpowszechniejsze to trójkątna i trapezoidalna (por. Zadeh, 2009, s. 3990). Może też przybierać wartości tak lingwistyczne, jak i liczbowe (Noworol, 2013, s. 32–33).



w stanie osiągnąć tej efektywności, którą mają ciągłe skale szacunkowe. W ich przypadku formuła udzielania odpowiedzi może oprócz odcinkowej mieć także postać pozycjonowania natężenia cechy przez podawanie cyfrowej wartości z zadanego ciągu liczb (por. Jefmański, 2011). Zastrzec tu jednak należy, że również przy pomiarach z użyciem skal ciągłych głównym zagrożeniem trafności odwzorowania cech rozmytych jest zbyt szybkie i pochopne przechodzenie do zamiany wartości lingwistycznych na wartości liczbowe, co skutkuje zauważonym wcześniej sztucznym zawyżaniem poziomu pomiaru. Tego rodzaju transformacje są bardzo powszechne, na co wskazuje choćby popularność skali Likerta w badaniach sondażowych i pomiarach testowych.

Wśród przykładów ugruntowanych podejść konceptualizujących zjawisko niejednorodności z wykorzystaniem idei wymiarów podać można teorię postaw jako wymiarów ewaluatywnych (Böhner, Wänke, 2002; Cacioppo, Gardner, Berntson, 1997; Dormandy, Hankins, Martean, 2006), model postaw dualnych (Wilson, Lindsey, Schooler, 2000), jak też teorię identyfikacji społecznej i autokategoryzacji (Aron, Aron, Smollan, 1992; Schubert, Otten, 2002). Z kolei wśród konkretnych rozwiązań pomiarowych wykorzystujących ideę kontinuum warto przywołać skale piktogramowe, choć pomiar przy ich użyciu nadal dokonuje się na poziomie porządkowym. Stąd też korzystniejszy w przypadku kontinualnego modelowania cechy okazuje się pomiar ciągły, polegający na zastosowaniu wizualnej skali analogowej (Funke, Reips, 2012). Tego rodzaju skale szacunkowe mają postać odcinków, na których osoby badane zaznaczają miejsca w odległości odzwierciedlającej natężenia danej cechy, o którą są pytane. Skale wizualne są też intuicyjnie zrozumiałe, a podstawowe ich ograniczenie polega na przestrzeganiu określonej długości odcinka (por. Brzezińska, Brzeziński, 2006; Reips, Funke, 2008). Surowe wartości muszą być następnie przeliczone przy użyciu odpowiednich formuł matematycznych, które to w trybie pomiaru pośredniego dają dopiero wgląd w charakterystykę mierzonego zjawiska (np. Scott, 1969).

Warto pamiętać, że zagrożenie trafności kryje się też w instrukcjach opisujących sposób wykonywania zadań. Słuszne jest kierowanie się zasadą tworzenia poleceń z nastawieniem na adresata wymagającego maksymalnej jasności opisu. Poza tym każdy adresat wymaga ukierunkowania w zakresie tego, jaki obszar jego/jej umysłowości, w jakim zakresie i z uwzględnieniem jakich aspektów wymagać będzie zaangażowania podczas wykonywania zadań pomiarowych. Bez tego jego/jej odpowiedzi formułowane mogą być z pominięciem tych elementów, które dla celu badań są kluczowe, lub z uwzględnieniem tych, które wgląd w specyfikę zjawiska mogą utrudniać. Traci na tym tzw. traf-

ność fasadowa, ale zasadniczo i tak zagrożony jest w szczególności teoretyczny aspekt trafności. To dlatego niezbędne jest informowanie w instrukcjach o celu i przedmiocie badań. Wyjątkiem są sytuacje, gdy świadomość celu badań lub znaczenia poszczególnych czynności zniekształca wyniki pomiaru. Stosuje się wówczas działania regulowane przez zasady tzw. decepcji, polegającej na kontrolowanym opóźnieniu dostępu badanych osób do wiedzy o projekcie, w którym uczestniczą. Ale i decepcja prowadzona być musi zgodnie z zasadami etyki badań (Brzeziński, 1999).

Mimo zabezpieczenia pomiaru na etapie budowy narzędzia oraz wcześniej konceptualizacji i operacjonalizacji cech niejednorodnych i rozmytych zagrożenia trafności wystąpić mogą też podczas doboru rozwiązań analizy danych. Wartości zdjęte ze skal ciągłych, np. analogowych skal wizualnych, same jeszcze nie zapewnią warunków pozwalających odzwierciedlić charakter obserwowanego zjawiska. Tak jest na przykład wówczas, gdy dane z pomiaru postaw prowadzonego na oddzielnych wymiarach poddawane są integracji. Nawet jeśli na etapie konceptualizacji nie pojawiły się elementy ograniczające adekwatną integrację danych empirycznych, to nie jest wykluczone, że takie wystąpią w późniejszych etapach procesu badawczego. Metody analizy danych, w tym formuły przewidziane do ich integracji, również mogą ograniczać trafność rezultatów badawczych. Przykładów w tym zakresie dostarczają badania nad efektywnością formuł przeliczających dane pochodzące z pomiarów cech traktowanych jako konglomeraty sprzecznych tendencji (Breckler, 1994; Priester, Petty, 1996), ale też symulacje przeliczeń możliwe do samodzielnego przeprowadzenia z użyciem dostępnych formuł na tych samych danych. Osobne zagrożenia dotyczą metod analizy danych wyrastających z systemu logiki rozmytej. Chodzi tu zasadniczo o błędy związane z doбором funkcji i formuł (por. Zadeh, 2009; Noworol, 2013).

Zagrożenia w pomiarze cech niejednorodnych i rozmytych dotyczą więc generalnie zagrożeń trafności pomiaru, które występować mogą na każdym etapie konstrukcji i realizacji badania. Część z nich daje się przewidywać, inne jako niewiadome wymagają nieustannej czujności. W obu przypadkach kluczowym warunkiem ich kontrolowania jest metodologiczna świadomość badaczy.

## Bibliografia

- Anastasi A., Urbina S. (1999), *Testy psychologiczne*, Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego, Warszawa.

- Aron A., Aron E.N., Smollan D. (1992), *Inclusion of Other in the Self Scale and the Structure of Interpersonal Closeness*, „Journal of Personality and Social Psychology”, No. 63(4), s. 596–612.
- Böhner G., Wänke M. (2004), *Postawy i zmiana postaw*, GWP, Gdańsk.
- Breckler S.J. (1994), *A Comparison of Numerical Indexes for Measuring Attitude Ambivalence*, „Educational and Psychological Measurement” No. 54(2), s. 350–365.
- Brzezińska A., Brzeziński J. (2006), *Skale szacunkowe*, w: J. Brzeziński (red.), *Metodologia badań psychologicznych. Wybór tekstów*, PWN, Warszawa, s. 232–306.
- Brzeziński J. (1999), *Metodologia badań psychologicznych*, PWN, Warszawa.
- Cacioppo J., Gardner W., Berntson G. (1997), *Beyond Bipolar Conceptualizations and Measures: The Case of Attitudes and Evaluative Space*, „Personality and Social Psychology Review”, No. 1(1), s. 3–25.
- Chojnowski M. (1971), *Pomiar w psychologii*, w: J. Koziński (red.), *Problemy psychologii matematycznej*, PWN, Warszawa, s. 15–42.
- Dormandy E., Hankins M., Marteau T.M. (2006), *Attitudes and Uptake of a Screening test: The Moderating Role of Ambivalence*, „Psychology and Health”, No. 21(4), s. 499–511.
- Funke F., Reips U. D. (2012), *Why Semantic Differentials in Web-Based Research Should be Made From Visual Analogue Scales and Not From 5-Point Scales*, „Field Methods”, No. 24(3), s. 310–327.
- Guba G.E. (1981), *Criteria for Assessing the Trustworthiness of Naturalistic*, „Educational Communication and Technology”, No. 29(2), s. 75–91.
- Guilford J.P. (2005), *Rzetelność i trafność pomiarów*, w: J. Brzeziński (red.), *Trafność i rzetelność testów psychologicznych*, tłum. E. Hornowska, M. Zakrzewska, GWP, Gdańsk, s. 62–106.
- Guion R.M. (2005), *O trynitarnej doktrynie trafności*, w: J. Brzeziński (red.), *Trafność i rzetelność testów psychologicznych*, tłum. E. Hornowska, M. Zakrzewska, GWP, Gdańsk, s. 272–287.
- Hornowska E. (2009), *Testy psychologiczne. Teoria i praktyka*, Scholar, Warszawa.
- Jankowski K., Zajenkowski M. (2009), *Metody szacowania rzetelności pomiaru testem*, w: K. Fronczyk (red.), *Psychometria, Podstawowe zagadnienia*, Wyższa Szkoła Finansów i Zarządzania w Warszawie, Warszawa, s. 84–110.
- Jefmański B. (2011a), *Nowe podejście w pomiarze opinii respondentów z zastosowaniem skal porządkowych i elementów teorii zbiorów rozmytych – charakterystyka wybranych aspektów metodologicznych*, w: K. Mazurek-Łopacińska, M. Sobocińska (red.),

- Badania marketingowe – metody, nowe podejścia i konteksty badawcze*, Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław, s. 184–191.
- Jefmański B. (2011b), *Pomiar i ocena jakości usług z zastosowaniem liczb rozmytych – aspekty metodologiczne i przykłady zastosowań*, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, nr 151, s. 288–300.
- Messick S. (2005), *Trafność testu a etyka oceny (diagnozy)*, w: J. Brzeziński (red.), *Trafność i rzetelność testów psychologicznych*, tłum. E. Hornowska, M. Zakrzewska, GWP, Gdańsk, s. 460–485.
- Morse J., Barret M., Mayan M., Olson K., Spiers J. (2002), *Verification Strategies for Establishing Reliability and Validity in Qualitative Research*, „International Journal of Qualitative Methods”, No. 1(2), s. 13–22.
- Nowakowska M. (1975), *Psychologia ilościowa z elementami naukometrii*, PWN, Warszawa.
- Noworol C. (2013), *Heurystyki kategorii rozmytych w koncepcji pomiaru psychologicznego*, PWN, Warszawa.
- Priester J.R., Petty R.E. (1996), *The Gradual Threshold Model of Ambivalence: Relating the Positive and Negative Bases of Attitudes to Subjective Ambivalence*, „Journal of Personality and Social Psychology”, No. 71(3), s. 431–449.
- Schubert T.W., Otten S. (2002), *Overlap of Self, Ingroup, and Outgroup: Pictorial Measures of Self-Categorization*, „Self and Identity”, No. 1, s. 353–376.
- Schuman H. (2013), *Metoda i znaczenie w badaniach sondażowych*, tłum. A. Grabowska, Oficyna Naukowa, Warszawa.
- Scott W.A. (1969), *Structure of Natural Cognition*, „Journal of Personality and Social Psychology”, 12(4), s. 261–278.
- Stevens S.S. (1946), *On the Theory of Scales of Measurement*, „Science”, No. 103(2684), s. 677–680.
- Suppes P., Zinnes J.L. (1963), *Basic Measurement Theory*, w: D. Luce, R.R. Bush, E. Galanter (red.), *Handbook of Mathematical Psychology* (Vol. 1), Willey, New York – Oxford, s. 1–76.
- Walenta K. (1971), *Podstawowe pojęcia teorii pomiaru*, w: J. Kozielski (red.), *Problemy psychologii matematycznej*, PWN, Warszawa, s. 43–64.
- Wilson T.D., Lindsey S., Schooler T.Y. (2000), *A Model of Dual Attitudes*, „Psychological Review”, No. 107(1), s. 101–126.
- Zadeh L.A. (1973), *Outline of a New Approach to the Analysis of Complex Systems and*

*Decision Processes*, „IEEE Transactions on Systems, Man, and Cybernetics”, No. 3(1), s. 28–44.

Zadeh L.A. (2009), *Fuzzy logic*, w: R.A. Meyers (red.), *Encyclopedia of Complexity*, Springer, New York, s. 3985–4009.