

Studia z Psychologii w KUL, tom 17
red. O. Gorbaniuk, B. Kostrubiec-Wojtachnio,
D. Musiał, M. Wiechetek, A. Błachnio, A. Przepiórka
ISBN 978-83-7702-473-7 Lublin, Wyd. KUL 2011, s. 213-245

Andrzej Januszewski¹

Instytut Psychologii

Katolicki Uniwersytet Lubelski Jana Pawła II

**MODELE RÓWNAŃ STRUKTURALNYCH
W METODOLOGII BADAŃ PSYCHOLOGICZNYCH.
PROBLEMATYKA PRZYCZYNOWOŚCI
W MODELACH STRUKTURALNYCH
I DOPUSZCZALNOŚĆ MODELI**

Treść artykułu koncentruje się na wybranych możliwościach modeli równań strukturalnych (SEM – ang. *Structural Equation Models*). Korzystając z nich, badacze reprezentujący nauki społeczne (psychologia, pedagogika, socjologia, ekonomia i in.) śmielej podejmują problem przyczynowości w badaniach, budując modele teoretyczne aspirujące do wyjaśniania zjawisk psychicznych i empirycznie je weryfikując. Stało się to możliwe dzięki temu, że w ostatnim dwudziestoleciu opracowano wiele zaawansowanych programów do modelowania strukturalnego. Pozwalają one na wielowymiarową i wielozmiennową precyzyjną analizę danych empirycznych odnoszących się do określonych aspektów rzeczywistości, z większymi niż dotąd możliwościami niż te, które oferowała klasyczna statystyka. Stosowanie SEM umożliwia także rozwój tych nauk, gdyż bardziej bezpośrednio i efektywniej niż dotąd skonfrontowane są „teoria” z „empirią”, co owocuje koniecznością optymalizowania modeli teoretycznych – z jednej strony, a z drugiej – narzędzi służących do diagnozy rzeczywistości psychologicznej. Ponadto, dzięki SEM, tradycyjne hipotezy badawcze zostają wzbogacone o aspekt przyczynowości.

¹ Adres do korespondencji: Katedra Psychologii Ogólnej, Instytut Psychologii, Katolicki Uniwersytet Lubelski Jana Pawła II, Al. Raclawickie 14, 20-950 Lublin; e-mail: andrzej.januszewski@kul.pl

Niniejszy artykuł, w formie ogólnej oraz bardziej zobrazowanej na rycinach, przedstawia Czytelnikowi kilka wybranych aspektów hipotetycznej przyczynowości, które dzięki SEM można uczynić przedmiotem testowania. Przedstawia także kilka podstawowych kryteriów (testy dobroci dopasowania), na podstawie których dokonuje się oceny dopasowania modelu teoretycznego do empirycznej macierzy kowariancji kontrolowanej układem równań.

Słowa kluczowe: modele równań strukturalnych (SEM – ang. *Structural Equation Models*), problem przyczynowości w SEM, wybrane kryteria dopuszczalności SEM

STRUCTURAL EQUATION MODELS IN METHODOLOGY OF PSYCHOLOGICAL RESEARCH. THE ISSUE OF CAUSALITY IN STRUCTURAL MODELS AND MODEL ADMISSIBILITY

The contents of the article focuses on selected possibilities of structural equation models (SEM). Using SEM, researchers representing social sciences (psychology, educational studies, sociology, economy etc.) are more entitled to undertake the issue of causality in their studies, they build theoretical models aspiring to explain mental phenomena and they verify them empirically. This became possible due to the fact that many advanced software programs for structural modeling were developed over the last two decades. They enable precise multidimensional and multivariable analysis of empirical data concerning particular aspects of reality with greater possibilities than those offered by classical statistics. Using SEM also facilitates development of these disciplines since “theory” can be confronted with “reality” more directly and effectively than in the past, which bring about the necessity of optimizing theoretical models, on the one hand, and on the other hand – the tools serving to diagnose psychological reality. Additionally, the traditional research hypotheses can be enriched by the aspect of causality. This article presents several selected aspects of hypothetical causality in the general form and more concretely – in the figures, which can become the subject of testing owing to SEM. It also presents some fundamental criteria (tests of goodness of fit), on which the evaluation is made of the fit between the theoretical model and the empirical covariance matrix controlled by the set of equations.

Keywords: Structural Equation Models (SEM), the problem of causality in SEM, selected criteria of admissibility in SEM

Wprowadzenie

Celem tego artykułu jest przybliżenie sposobów opracowania wyników badań przy użyciu współczesnych zaawansowanych narzędzi statystycznych do modelowania strukturalnego, takich jak: AMOS, CALIS, EQS, HLM, LISREL, Mplus, Mx, SIMPLIS, które pozwalają na wielowymiarową i wielozmiennową precyzyjną analizę danych empirycznych odnoszących się do określonych aspektów rzeczywistości, z większymi niż dotąd możliwościami opracowania ich z uwzględnieniem problemu przyczynowości. Warto również wspomnieć o możliwościach przeprowadzania imitacji planów eksperymentalnych, których celem jest zazwyczaj dokładniejsza eksploracja współzależności między analizowanymi zmiennymi, z dopuszczalnością testowania zależności przyczynowo-skutkowych. Dzięki możliwościom udostępnionym badaczowi przez ww. programy oraz szybszym i bardziej niezawodnym komputerom, można w dość krótkim czasie przeprowadzić imitację eksperymentów, nawet z wielomilionową liczbą kombinacji, dochodząc do wykazania dopuszczalnej przyczyny X_i dla skutku Y_i . Wszystko to oczywiście pod warunkiem, że zebrany materiał badawczy (macierz danych z rozkładami zmiennych) pozwoli na rozwiązanie układu równań i oszacowanie (estymację) występujących w równaniach wyrazów niewiadomych. Praktyczne korzyści ze stosowania programów do modelowania strukturalnego to przede wszystkim możliwości formalnego: (1) integrowania wiedzy w obszarach danych empirycznych, (2) dokładniejszego niż dotąd, badania zależności między konstruktami teoretycznymi a ich empirycznymi argumentami, (3) przygotowania do określania praw i prawidłowości uwzględniających złożoność przyczynową, (4) wykazania przyczynowej roli czynników konstytucjonalnych i nieświadomych, które wymykają się spod kontroli, szczególnie wtedy, gdy koncentracja badacza ukierunkowana jest na atrybucyjnym wyjaśnianiu i opisie zachowań, u podłoża których leży świadomy, intencjonalny wybór (co ma ponadto podstawowe znaczenie w diagnostyce indywidualnej), (5) przygotowania do różnicowania wiedzy dzięki lepszym możliwościom kontrolowania źródeł zmienności wyników, czego efektem są trafniejsze typologizacje osób lub efekty aglomerowania cech psychicznych.

Dodatkową zachętą dla badaczy, którzy sięgają po narzędzia do modelowania, jest to, że większość spośród nich, oprócz podstawowego interfejsu użytkownika, zazwyczaj wykorzystującego profesjonalny język programowania (wymagane są odpowiednie kompetencje użytkownika), posiada także przystępny interfejs graficzny. Dzięki niemu nawet początkujący, po przeszkoleniu, mogą zaryzykować przeprowadzenie obliczeń bazujących na rozwiązywaniu układów równań.

W niniejszym artykule chciałbym przedstawić, jak z perspektywy modeli równań strukturalnych (SEM) można: (1) podejmować problem przyczynowości w badaniach typowych dla nauk społecznych (głównie: psychologii, pedagogice, socjologii i in.), (2) budować modele z wykorzystaniem notacji, której użycie sprzyja poprawności formalnej i merytorycznej opracowań². Początkujący badacz będzie musiał zmierzyć się z zupełnie nową dla siebie sytuacją, nawet jeśli już korzystał z dość zaawansowanych narzędzi statystycznego opracowania wyników. Owo wyzwanie polegać będzie na swoim zmaganiu się z problemem zintegrowania niekiedy ogromnej liczby informacji uzyskanych z obliczeń oraz powiązania ich z licznymi warunkami, wymaganymi jeszcze przed przystąpieniem do estymacji modelu.

Nie-programistom zainteresowanym praktyczną stroną posługiwania się programami do SEM polecam wersje z graficznym interfejsem użytkownika. Przykładowo, do kolejnych edycji programu „AMOS Graphics” dołączany jest przystępnie napisany podręcznik (Arbuckle i Wothke, 1995, 1999, 2005; Arbuckle, 2007, i nast.) zawierający analizy przykładów zastosowań SEM, który z powodzeniem spełnia rolę poradnika dla użytkowników-samouków od początkowego do średniozaawansowanego poziomu umiejętności.

Przyczynowy model strukturalny

W tym artykule, z uwagi na postawiony cel, omówię wybraną spośród wielu metod należących do grupy tzw. modeli równań strukturalnych (SEM – ang. *Structural Equation Models*)³ konfirmacyjną analizę czynnikową (CFA – ang. *Confirmatory Factor Analysis*) oraz przyczynowe modelowanie strukturalne (CM – ang. *Causal Modeling*). W obu przypadkach skorzystam

² Po wyczerpujące informacje odsyłam do klasycznych opracowań, np. bogatego w treść polskiego podręcznika R. Konarskiego (2009), czy wcześniejszych anglojęzycznych (por. Bollen, 1989; por. Brown, 2006; i inne).

³ Do najpopularniejszych modeli równań strukturalnych zalicza się:

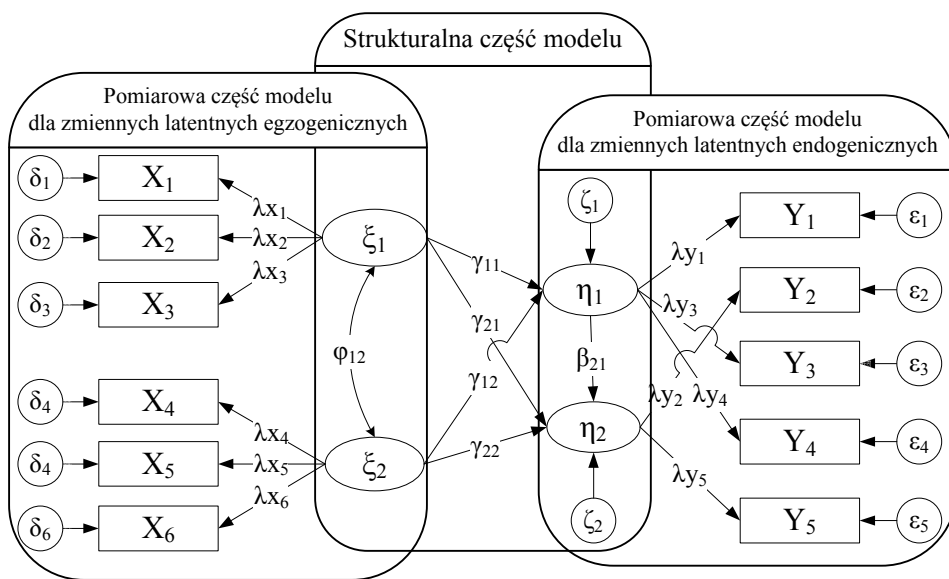
- konfirmacyjną analizę czynnikową (CFA – ang. *Confirmatory Factor Analysis*),
- wielowymiarową analizę regresji (MRA – ang. *Multiple Regression Analysis*),
- analizę ścieżek (PA – ang. *Path Analysis*),
- modelowanie przyczynowe (CM – ang. *Causal Modeling*),
- analizę struktur latentnych (LSA – ang. *Latent Structure Analysis*),
- analizę struktury kowariancji (CSA – ang. *Covariance Structure Analysis*) (Bollen, 1989; por. Zakrzewska, 2004, s. 442).

W Polsce pierwszą publikację nt. zastosowania analizy ścieżek (PA) do interpretacji przyczynowej przedstawili M. Gaul i A. Machowski (1987).

z programu AMOS (ang. *Analysis of Moments Structures*) (Arbuckle i Wothke, 1995/1999; Arbuckle, 2007), współpracującego z pakietem statystycznym SPSS (obecnie IBM SPSS Statistics).

Badaczy, którzy podejmują kluczową decyzję co do wyboru planu badawczego – eksperymentalny lub korelacyjny (por. Brzeziński, 2000, s. 355) – warto poinformować, że SEM umożliwia wykonanie każdego z nich, oczywiście po spełnieniu określonych warunków, ale za to z precyzją większą niż procedury należące do „tradycyjnej” statystyki. Trzeba jednak dodać, że wyniki SEM nie zawsze są równoznaczne z wynikami uzyskanymi przy użyciu klasycznych metod statystycznych. Ponadto, specyficzne funkcje w szeroko rozumianej analizie przyczynowości w SEM umożliwiają mediatory i/lub moderatory, których rola w ogólnym zarysie zostanie przedstawiona w tym artykule.

Przykładowo, w typowym paradygmacie badań korelacyjnych, odwołującym się do klasycznego użycia regresji wielokrotnej, lecz rozszerzonym o udział zmiennych latentnych, można posłużyć się następującym modelem (Rycina 1):



Rycina 1. SEM: regresja wielokrotna z nieobserwowanymi zmiennymi.
Ogólny model przyczynowy

Powyższy wariant zalicza się do podstawowych modeli przyczynowych, które w typowy sposób uwzględniają: (a) część pomiarową dla latentnych zmiennych egzogenicznych, (b) część pomiarową dla latentnych zmiennych endogenicznych oraz (c) wiążącą obie – część strukturalną modelu (Bollen, 1989; por. Ringle, 2004; por. Arbuckle, 2005, s. 89).

Warto podkreślić, że część strukturalna tego modelu umożliwia testowanie podstawowej hipotezy badawczej (na podstawie weryfikacji H_0 , o matematycznym braku różnicy), tzn. hipotezy o braku podstaw formalnych do odrzucenia proponowanego modelu teoretycznego, jeżeli tradycyjna miara, którą jest wynik otrzymany w teście χ^2 nie przekroczył wartości krytycznej rozkładu (χ^2 ; $p > 0,05$). W tej sytuacji macierz implikowana przez model teoretyczny (Σ) nie różni się zasadniczo od empirycznej macierzy kowariancji (S) kontrolowanej układem równań. Innymi słowy, wynik weryfikacji H_0 stanowi podstawę do przyjęcia bądź odrzucenia hipotezy badawczej: **o dopuszczalności przyczynowego wpływu rzeczywistości psychologicznej reprezentowanej przez latentne zmienne egzogeniczne na rzeczywistość reprezentowaną przez latentne zmienne endogeniczne**. Co ważne, konfirmacja⁴ nie oznacza końca dociekań, gdyż dopuszczalnym jest prawdopodobieństwo zbudowania kolejnego modelu teoretycznego, który byłby znacząco lepszy od poprzedniego. Problem, o którym tutaj mowa, jest jednym z zasadniczych zagadnień analizowanych przez badaczy. Jego rozwiązywanie polega na optymalizowaniu modeli teoretycznych – z jednej strony, z drugiej – narzędzi służących do diagnozy rzeczywistości psychologicznej. Chodzi o takie wprowadzanie korekt zarówno do modeli, jak również narzędzi, aby precyzyjniej analizować i interpretować wyniki w świetle teorii (por. Januszewski, 2009). Warto dodać, że możliwości optymalizacji teorii i narzędzia są traktowane przez użytkowników programów do SEM jako ich wyjątkowa zaleta w porównaniu z psychometrią bazującą na tradycyjnej statystyce.

⁴ Konfirmacja (łac. *confirmare* – wzmacniać, dowodzić) to szczególny przypadek uzasadniania twierdzeń, przeciwny ich odkrywaniu i formułowaniu. Jest to wynik, a także czynność potwierdzania hipotezy (i/lub tezy) poprzez pośrednie sprawdzanie empiryczne, np. danymi obserwacyjnymi. Jeśli wszystkie znane konsekwencje okażą się prawdziwe, to hipoteza jest w pewnym stopniu potwierdzona (por. Kamiński, 1989, s. 32). W konfirmacji hipotez lub teorii empirycznych występują zabiegi popierania bądź kwestionowania hipotez na podstawie obserwacji. Wyróżnia się trzy główne typy konfirmacji: (1) koroborację – będącą wynikiem rzetelnych prób nieefektywnej falsyfikacji dedukcyjnych konsekwencji hipotezy, (2) indukcyjną – gdzie hipoteza jest konfirmowana przez zgodność jej konsekwencji ze zdaniem obserwacyjnymi, (3) hemplowską – uwzględniającą aspekty indukcyjne (indukcja enumeracyjna) i dedukcyjne (rekonstrukcja formalna). Przy ocenie naukowej przydatności teorii konfirmacji należy uwzględnić elementy idealizacji hipotetyczno-dedukcyjnego modelu testowania hipotez implikowanego przez daną teorię (por. Hajduk, 2002, s. 580-584).

Notacje w modelach równań strukturalnych (SEM)

Specjalistyczna literatura przedmiotu, w tym oficjalne wystąpienia badaczy prezentujących wyniki badań opracowane przy użyciu SEM, koncentruje się na szczegółach, np. na wartościach współczynników ścieżek, ale nie dba o status określonej estymanty w odniesieniu do całości modelu lub jego istotnych substruktur. Stąd rodzą się trudności ze zidentyfikowaniem głównej prawidłowości i zinterpretowaniem jej w świetle teorii, której tezy i hipotezy przyjęto do zbudowania modelu teoretycznego. Dlatego też proponuję zapoznać się z obudową formalno-merytoryczną do SEM proponowaną przez klasyków zagadnienia, która pozwoli upewnić się w zakresie kryteriów poprawności i dopuszczalności zabiegów w budowanych, testowanych lub recenzowanych modelach. Owa zdobyta wiedza jest bardzo potrzebna na etapie konceptualizacji modeli, a także umożliwia rzetelniejszą analizę i interpretację testowanych modeli. Odwołując się do klasycznych założeń SEM, warto przedstawić kilka podstawowych (wzorcowych) modeli, które mogą być pomocne dla początkujących użytkowników programów do modelowania równań strukturalnych. Inną, równie ważną wskazówką dla zainteresowanych tą problematyką jest zachęta do korzystania z klasycznej notacji SEM. Moim zdaniem, jest ona ważna na każdym etapie budowania i testowania modeli, ponieważ dyscyplinuje postępowanie merytoryczne i formalne badacza, dając przejrzysty obraz postępowania naukowego, mimo że domaga się opanowania „małego alfabetu” znaków. System notacji najlepiej odnieść do przedstawionego już modelu na rycinie 1.

Notacje SEM: symbole – nazwy – skrótowe definicje (por. Bollen, 1989, s. 14):

- (1) notacje dla modelu latentnych (nieobserwacyjnych) zmiennych
zmienne:
 - ξ_i – (gr. *ksi*) – zmienne latentne egzogeniczne (czynniki)
 - η_i – (gr. *eta*) – zmienne latentne endogeniczne (czynniki)
 - ζ_i – (gr. *dzeta*) – w równaniach, błędy (zakłócenia) związane z określonymi zmiennymi latentnymi endogenicznymi
- współczynniki:
 - β_{ij} – (gr. *beta*) – współczynnik odnoszący się do wpływu η_i na η_j (efekt bezpośredni)
 - γ_{ij} – (gr. *gamma*) – współczynnik odnoszący się do wpływu ξ_i na η_j (najważniejszy efekt bezpośredni)
- kowariancje:
 - φ_{ij} – (gr. *phi*) – kowariancje (korelacje) między latentnymi egzogenicznymi zmiennymi ξ_i

ψ_{ij} – (gr. *psi*) – kowariancje (korelacje) między błędami (zakłóceniami) ζ_i związanymi z pomiarem zmiennych η_i

(2) notacje dla modelu pomiarowego (por. Bollen, 1989, s. 20)

zmienne:

Y_i – (gr. *ipsylon*) – wskaźniki obserwacyjne dla η_i

X_i – (gr. *chi*) – wskaźniki obserwacyjne dla ξ_i

ε_i – (gr. *epsilon*) – błąd związany z pomiarem określonego wskaźnika Y_i

δ_i – (gr. *delta*) – błąd związany z pomiarem określonego wskaźnika X_i

współczynniki:

λ_{y_i} – (gr. *lambda*) – współczynnik odnoszący się do wpływu η_i na Y_i

λ_{x_i} – (gr. *lambda*) – współczynnik odnoszący się do wpływu ξ_i na X_i

kowariancje:

θ_ε – (gr. *teta* ε) – kowariancje (korelacje) między błędami pomiaru związanymi z pomiarem zmiennych Y_i

θ_δ – (gr. *teta* δ) – kowariancje (korelacje) między błędami pomiaru związanymi z pomiarem zmiennych X_i

Podstawą oceny dopuszczalności modelu teoretycznego, w którym hipotetycznie przyjęto kierunki oraz istotną siłę zależności między zmiennymi, jest pozytywny wynik przeprowadzonej procedury weryfikującej hipotezę badawczą, o dopuszczalności modelu (podstawa odpowiedzi na główny i ogólny problem), oraz estymacja nieznanymi parametrów modelu wraz z oszacowaniem znaczenia tych parametrów. Programy do SEM wyliczają zarówno estymanty standaryzowane, jak również niestandaryzowane, które można poddać interpretacji zgodnie z celem postawionym sobie przez badacza. Estymantami, które najczęściej bierze się pod uwagę, są: wagi regresyjne (ang. *regression weights*), potocznie nazywane współczynnikami ścieżek, wariancje zmiennych egzogenicznych (ang. *variances of exogenous variables*), kowariancje lub korelacje między zmiennymi egzogenicznymi (ang. *covariances or correlations among exogenous variables*), średnie dla zmiennych egzogenicznych (ang. *means of exogenous variables*), wyrazy wolne służące predykcji zmiennych endogenicznych (ang. *intercepts for predicting endogenous variables*), kwadraty współczynników korelacji wielokrotnej (ang. *squared multiple correlations*). Wartości estymant niestandaryzowanych podawane są przez program w kontekście ich błędów standardowych, współczynników krytycznych (C.R.)⁵ oraz poziomów istotności.

⁵ Współczynnik krytyczny (C.R. – ang. *Critical Ratio*) jest równowartością statystyki „z” osiągającej istotność ($z < -1,96$ lub $z > 1,96$; $p \leq 0,05$). W przypadku estymant (np. wag regresyjnych, kowariancji, wariancji, różnic między parametrami) C.R. jest wartością otrzymaną z ilorazu E/SE (niestandaryzowana wartość estymanty podzielona przez jej błąd standardowy) (por. Arbuckle, 2005). Warto podkreślić, że przekroczenie wartości krytycznej dowodzi istotnego udziału konkretnej estymanty na rzecz całości modelu. Nieistotna wartość estymanty ($p > 0,05$) oznacza, że związany z nią rozkład wyników jest przypadkowy, tzn. nie

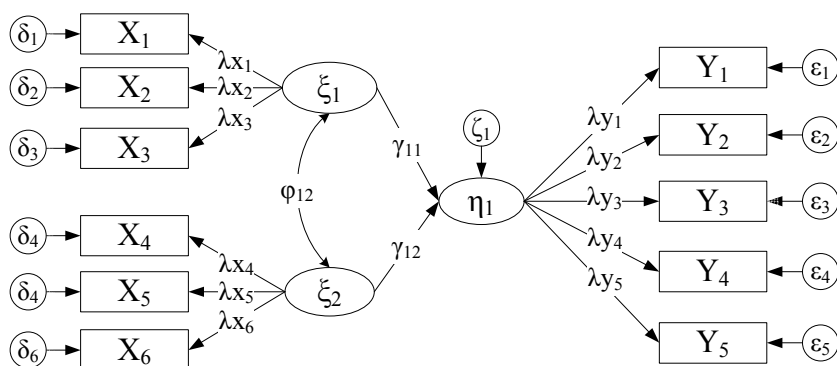
W źródłowej koncepcji modele strukturalne zmiennych ukrytych są połączeniem konfirmacyjnej analizy czynnikowej i analizy ścieżkowej. Testują liniowe efekty ukrytych zmiennych egzogenicznych (niezależnych, objaśniających) w zakresie innych zmiennych ukrytych endogenicznych (zależnych, objaśnianych). Każda z ukrytych zmiennych mierzona jest przyporządkowanym zestawem zmiennych empirycznych, czyli wskaźników (por. Judd, 1995/2002). Zobrazowano to na rycinie 1. Hipotetyczne wymienienie przez badacza określonych zmiennych egzogenicznych jako przyczynowych dla endogenicznych (z merytorycznym uzasadnieniem) nie znaczy, że są to jedyne źródła przyczynowości. Mogą istnieć także inne wpływy przyczynowe, które z różnych powodów merytorycznych i formalnych nie zostały w danym modelu uwzględnione. Oczywiście są również pytania: czym jest uwarunkowana bieżąca zmienna egzogeniczna? Czy można ją tłumaczyć inną zmienną egzogeniczną, która ma status źródłowy względem niej? Czy też określona zmienna pierwotnie uznana za egzogeniczną, faktycznie pełni rolę mediatora między inną zmienną egzogeniczną a endogeniczną? Pełni zatem rolę pośredniczącą a nie przyczynową (por. Rycina 4).

Innym argumentem przemawiającym za odwoływaniem się do notacji jest to, że badacz zapoznający się z programami komputerowymi wykorzystującymi założenia SEM, szybko podejmuje samodzielne próby budowania modelu, aby następnie, odwołując się do macierzy danych, przetestować go – ocenić jego poprawność (dopuszczalność). Ze względu na dość powszechne korzystanie przez badaczy z tzw. wersji graficznych programów, budowanie modeli przypomina konstrukcyjną, „wciągającą” zabawę. Jeśli wynik owych badań recenzuje osoba o gruntownej wiedzy merytorycznej z zakresu modelowania równań strukturalnych, zapewne będzie się kierować kryteriami statusu zmiennych, dopuszczalności lub niedopuszczalności wprowadzanych założeń. Dlatego uzasadnione jest stałe odwoływanie się do klasycznych dzieł takich autorów jak: P. M. Bentler, K. A. Bollen, K. G. Jöreskog i D. Sörbom oraz inni. Podręcznikowa wiedza o którymś z ww. programów może okazać się niewystarczająca, chociaż ich kolejne edycje są coraz bogatsze, jak np. *Amos 6.0, 7.0, 16.0, 17.0 User's Guide* (Arbuckle, 2006 i nast.), znajdujemy tam jednak głównie „podpowiedzi” nt. technicznej obsługi programu.

ma związku z rozkładem współzmienności (kowariancji) wyników w empirycznej macierzy danych; wartość istotna dowodzi, że rozkład ten nie jest przypadkowy.

Regresja z nieobserwowanymi zmiennymi (prosta i złożona)

Przyczynowe modele równań strukturalnych⁶, oparte na wynikach weryfikacji hipotez statystycznych (H_0), umożliwiają badaczowi zajęcie stanowiska wobec przyjęcych psychologicznych hipotez badawczych, np. wykazania wpływu struktury egzogenicznej na endogeniczną – ścieżki γ_i (gr. *gamma*). Ważne, by przed weryfikacją głównej hipotezy przygotować submodele, które umożliwią zweryfikowanie hipotez szczegółowych, np. o dopuszczalności określonego konstruktów teoretycznego (ξ_i , lub η_i) odzwierciedlającego dany mentalny konstrukt poznawczy. Od strony technicznej wykorzystywana jest w tym celu confirmacyjna analiza czynnikowa. Dla zobrazowania powyższego wywodu proponuję sięgnąć do klasycznych modeli SEM. Oto przykład prostego modelu strukturalnego (Rycina 2), w którym testowany jest wpływ dwóch zmiennych egzogenicznych (ξ_1 oraz ξ_2 – gr. *ksi*) o porównywalnym statusie na zmienną endogeniczną (η_1 – gr. *eta*).



Rycina 2. SEM (CM): prosta regresja wielokrotna z nieobserwowanymi zmiennymi

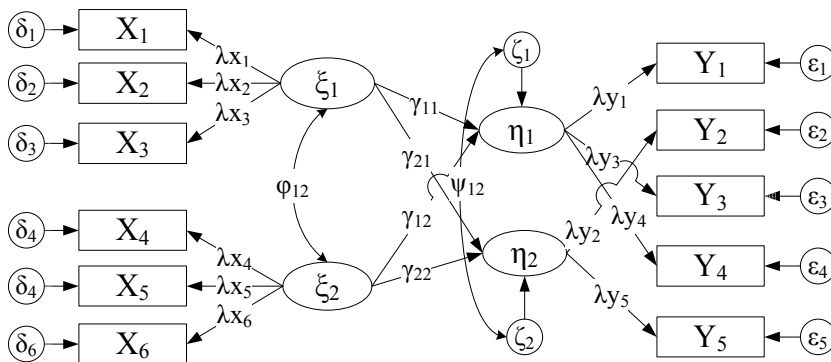
Wszystkie trzy wymienione zmienne to zmienne latentne, które można porównać do konstruktów osobistych w rozumieniu G. A. Kelly'ego (1955, 1970). Każdy z egzogenicznych konstruktów ma swoją argumentację empiryczną. Dla ξ_1 są to wskaźniki obserwacyjne $X_{1, 2, 3}$ (gr. *chi*), a dla ξ_2 – wskaźniki obserwacyjne $X_{4, 5, 6}$, stanowiące materiał badawczy uporządkowany w empirycznej macierzy danych (S). Podobnie, dla konstruktów endogenicznych η_1 są to wskaźniki obserwacyjne $Y_{1, 2, 3, 4, 5}$ (gr. *ipsylon*). Macierz jest podstawą do testowania, czyli wykazania dopuszczalności wykoncypowa-

⁶ Odwołuję się tutaj do modeli bazujących na sprawdzonych wzorcach analizy ścieżek i modeli przyczynowych (por. Bollen, 1989, s. 33 i s. 53; por. Ringle, 2004, s. 9).

nego modelu teoretycznego (Σ), jeśli w wyniku przeprowadzonej estymacji modelu wartość w teście χ^2 nie będzie znacząca (tzn. $p > 0,05$). Zanim jednak zostanie przetestowana dopuszczalność najważniejszej hipotezy o istotnym wpływie zmiennych egzogenicznych na endogeniczną, modele cząstkowe (submodele) powinny przejść pozytywnie weryfikację. Za taką strategią przemawia argumentacja merytoryczna i formalna, której istotne elementy przedstawiłem poniżej, omawiając konfirmacyjną analizę czynnikową oraz problem kierunku przyczynowości.

Wracając do typowego modelu kauzalnego (regresja z nieobserwowanymi zmiennymi), warto przypomnieć, jak brzmi interpretacja standaryzowanych estymant dla ścieżek: jeśli wartość zmiennej, od której wychodzi strzałka, wzrośnie o 1 odchylenie standardowe, to wartość zmiennej, do której strzałka dochodzi, wzrośnie o wyestymowaną wartość współczynnika ścieżki (standaryzowana waga regresyjna), jeśli znak tego współczynnika jest dodatni; wartość zmiennej skutkowej obniży się, jeśli znak współczynnika jest ujemny. Co więcej, wynik przeprowadzonej estymacji modelu dostarcza informacji o wartościach kwadratu współczynnika korelacji wielokrotnej, które pojawiają się przy zmiennych endogenicznych zarówno latentnych (η_i – gr. *eta*), jak i obserwacyjnych Y_i (gr. *ipsylon*) – wskaźników dla η_i oraz X_i (gr. *chi*) – wskaźników dla ξ_i (gr. *ksi*). Każda z tych informacji określa procent (jeśli wyliczoną wartość estymanty pomnoży się $\times 100$), w jakim dana zmienna zostaje wyjaśniona przez wpływy tych zmiennych, które do niej wysyłają strzałki. Warto dodać, że zmienne wysyłające strzałki są predyktorami dla zmiennych, które przyjmują strzałki. Przykładowo, jeśli wspomniana estymanta dla określonej zmiennej endogenicznej wynosi 0,28, oznacza to, że jej rozkład w 28 procentach można wytłumaczyć wpływem zmiennych, które są jej predyktorami. W takiej sytuacji dla badacza dopuszczalną jest ekstrapolowana informacja, że 72 procent zmienności tej zmiennej endogenicznej nie jest wyjaśnione przez predyktory uwzględnione w tym modelu. Analogicznie można interpretować bardziej rozbudowany model (por. Rycina 3).

Jeśli w modelu kauzalnym nie uwzględnia się roli mediatora dla zmiennej latentnej endogenicznej (η_1 lub η_2) i w ten sposób kontroluje się jedynie najważniejsze wpływy bezpośrednie (np. γ_{11} reprezentuje wpływ między ξ_1 a η_1), a rezygnuje się z kontroli wpływów bezpośrednich (β_{21} lub β_{12} – gr. *beta*) między zmiennymi endogenicznymi, wówczas na etapie specyfikacji modelu należy dopuścić korelację (ψ_{12} – gr. *psi*) między błędami pomiaru (ζ_1 oraz ζ_2 – gr. *dzeta*) zmiennych latentnych endogenicznych (η_1 lub η_2). Przyjmuje się, że źródło zmienności η_1 i η_2 może być wyjaśnione nie tylko wpływem zmiennych egzogenicznych (ξ_1 i ξ_2), ale również innymi zmiennymi egzogenicznymi, które nie zostały uwzględnione w modelu. Dopuszczenie korelacji (ψ_{12}) oznacza w domyśle istotną współzmiennność konstruktów η_1 i η_2 oraz współzmiennność między ich błędami (ζ_1 i ζ_2). Podobnie jak na poziomie obligatoryjnych założeń, należy dopuścić korelację (ϕ_{12} – gr. *phi*) między



Rycina 3. SEM (CM): złożona regresja wielokrotna z nieobserwowanymi zmiennymi (bez udziału mediatora)

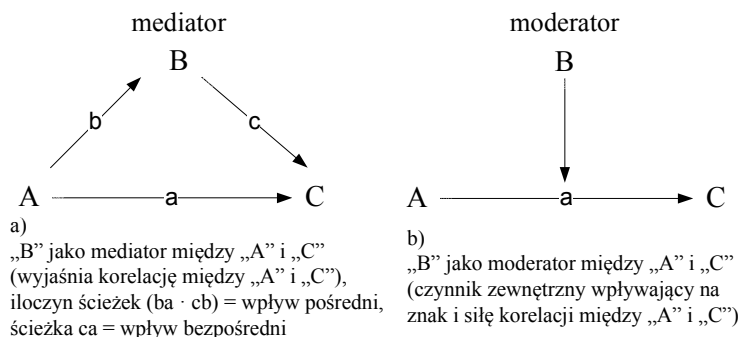
konstruktami ξ_1 i ξ_2 , uznając, że poszczególne konstrukty ξ_i odnoszą się do hipotetycznej, zwartej sfery psychologicznej, której dopuszczalność pozytywnie wcześniej zweryfikowano na podstawie oceny wskaźników dopasowania modelu do danych. Przyjmuje się, że źródło współzmienności (ϕ_{12}) leży poza rozpatrywanym modelem.

Rola mediatorów i moderatorów w przyczynowości pośredniej i bezpośredniej

Przykłady przedstawione na rycinie 4 obrazują sytuacje mediowania oraz moderowania przyczynowości przez czynnik B. Z hipotetycznym wpływem moderatora mamy do czynienia w sytuacji, w której tylko pod określonymi warunkami B, A jest znacząco powiązane z C. Natomiast z hipotetycznym wpływem B jako mediatora między A i C, a dokładniej hipotetycznymi wpływami A na B oraz B na C, mamy do czynienia wtedy, gdy spełnione są warunki: (1) A jako predyktor jest znacząco powiązane z C, (2) A jako predyktor jest znacząco powiązane z B, (3) B jest znacząco powiązane z C, lecz pod kontrolującym wpływem A, (4) wpływ A na C jest mniejszy, aniżeli pod kontrolującym wpływem B (za: Holmbeck, 1997).

Moderatory w SEM pełnią rolę czynnika eksperymentalnego, który pozwala badaczowi testować dopuszczalność hipotezy o związku przyczynowo-skutkowym. Badacz podejmuje ten problem, jeśli zastosuje odpowiedni schemat eksperymentalny (por. Judd, 1995/2002; por. Shaughnessy i in.,

1990/2002, por. Wojciszke, 2004), w którym kluczową rolę w wyjaśnianiu przyczyn różnic między strukturami mogą odegrać moderatory (moim zda-



Rycina 4. Funkcja mediatora i moderatora w wyjaśnianiu problemu przyczynowości w SEM

niem można je porównać do funkcji kanonu jedynej różnicy zdefiniowanego przez J. S. Milla). Moderatorami (przyczynowymi regulatorami zmian) mogą być zmienne nominalne, takie jak: przynależność do grupy płci, przedziału wiekowego, kontrolowane w badaniach zmienne socjodemograficzne, a nade wszystko specyficzne właściwości osobowościowe (zwłaszcza temperamentu⁷ i charakteru), które mogą być podstawą przynależności osób do określonego typu strukturalnego osobowości. Warto podkreślić, że testowanie roli moderatora, czyli np. wpływu specyfiki grupy, polega na przyjmowaniu (na różnych poziomach analiz) podobieństwa estymant tj.: wag pomiarowych (ang. *measurement weights*), pomiarowych wyrazów wolnych (ang. *measurement intercepts*), wag strukturalnych (ang. *structural weights*), strukturalnych wyrazów wolnych (ang. *structural intercepts*), średnich strukturalnych (ang. *structural means*), kowariancji strukturalnych (ang. *structural covariances*), reszt strukturalnych (ang. *structural residuals*), reszt pomiarowych (ang. *measurement residuals*) między porównywanymi grupami (por. przykład moderatora, w: A. Januszewski, 2009).

⁷ Argumentację merytoryczną dla moderującego wpływu temperamentu można znaleźć w pracy J. Strelaua (2002, s. 273). Przykładowo, temperament w rozwoju osobniczym występuje bądź kształtuje się najwcześniej, stąd rozpatrywanie jego związku ze stresem jest jak najbardziej zasadne. Cechy temperamentu w odniesieniu do radzenia sobie ze stresem mogą mieć status moderatora, ponieważ stanowią warunek poprzedzający występowanie innych zjawisk, na które wpływają. Jak się wydaje, Autor tej argumentacji będzie weryfikował (lub poddawał konfirmacji) układ hipotez, że cecha konstytucjonalna, która chronologicznie była wcześniej, może być przyczyną bardzo ważnej cechy lub zachowania, które – jako efekt uczenia – ukształtowały się później. W polskiej literaturze cennym źródłem informacji na temat roli mediatorów i moderatorów jest praca B. Wojciszke (2004), odwołująca się do klasycznego tekstu R. Barona i D. Kenny’ego (1986).

Z formalnego punktu widzenia możliwa liczba założeń o podobieństwie estymant jest zależna od poziomu rozbudowania modelu. Najmniej będzie ich w przypadku prostych modeli jednoczynnikowych, więcej – w wieloczynnikowych, jeszcze więcej – w hierarchicznych, dalej – kauzalnych z wyraźnie wyodrębnioną częścią strukturalną itd. Jednak z merytorycznego punktu widzenia (argumentacja teoretyczna) ważniejsze jest uzasadnienie założeń o podobieństwie lub niepodobieństwie określonych estymant charakteryzujących porównywane grupy osób. Przy założonym podobieństwie wszystkich estymant, uzyskana wartość ($p \leq 0,05$) pozwalająca oszacować istotność wyniku w najczęściej używanym (por. Arbuckle, 2007) teście podobieństwa (χ^2) jest podstawą do nadania moderatorowi statusu czynnika wyjaśniającego źródło różnicy estymant między porównywanymi grupami osób. Tak jak wspomniałem, czynnikiem moderującym może być jakaś zmienna nominalna, która chociaż leży poza modelem, jest jednak jedyną zmienną, której specyfika jest ewentualną przyczyną różnic analogicznych estymant w modelu między porównywanymi grupami osób.

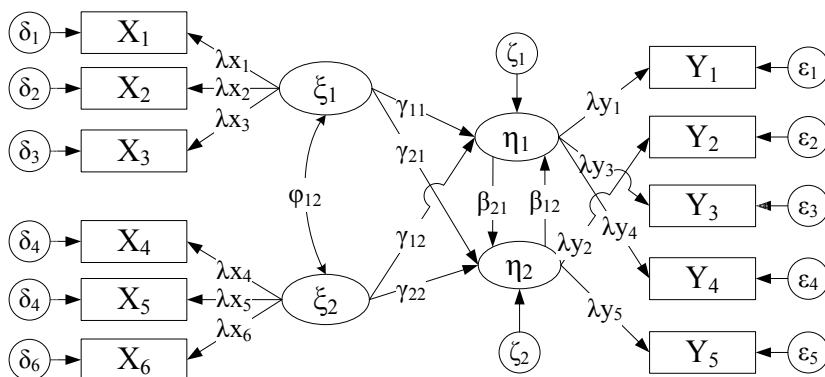
Z merytorycznego punktu widzenia przyczyna leży zawsze poza modelem; należy do sfery pozamatematycznej, pozaformalnej. Może nią być np. płeć, jeśli: (1) założy się podobieństwo wszystkich estymant w modelu, (2) spełni się warunki jednorodności wewnętrznej dla grup kobiet i mężczyzn, (3) wartości wszystkich innych zmiennych psychologicznych mają podobne rozkłady⁸, a uzyskany wynik w teście (χ^2) okaże się istotny ($p \leq 0,05$). Praktyka dowodzi jednak, że większość badaczy stawia sobie cele bardziej eksploracyjne niż arbitralnie rozstrzygające to, co jest „pierwszą przyczyną”. Dlatego „przyczynowość” najczęściej jest rozumiana jedynie jako „dopuszczalna przyczynowość”. Takie jej zdefiniowanie proponują też klasycy SEM.

Z kolei rola mediatora może być porównana do klasycznej zmiennej pośredniczącej, której status między zmienną niezależną a zależną polega na pośredniczeniu wzmacniającym, osłabiającym lub obojętnym. Dlatego też status przyczynowości dla mediatora jest niższy, gdyż pełni on rolę filtra a nie przyczyny głównej. Jego rolę można porównać do funkcji kanonu zmian towarzyszących zdefiniowanego przez J. S. Milla. Jednak z uwagi na to, że znacznie łatwiej jest włączyć mediator do modelu SEM, niż przeprowadzić żmudne hipotetyczne uzasadnienie merytoryczne i formalne dla wyizolowania moderatora, a następnie empirycznie wykazać jego rolę, wariant z mediatorem zdaje się być „bezpieczniejszy”. Między innymi te właśnie argumenty przyczyniły się do wzrostu popularności modeli SEM z jego użyciem.

⁸ Chodzi o te, które dodatkowo, oprócz przykładowo wymienionego czynnika płci, mogą wpłynąć na różnicę między strukturami estymant porównywanych grup, a których nie uwzględniono wewnątrz modelu, tzn. nie są przedmiotem estymacji, np. dane socjodemograficzne, poziom i struktura inteligencji lub inne.

Literatura przedmiotu podaje, że jest niewielka liczba analiz z użyciem moderatora w znaczeniu omówionym powyżej. Z pewnością wynika to z konieczności dokonania licznych szczegółowszych analiz oraz interpretacji, bez gwarancji, że próba wyizolowania „wyłącznej” przyczyny zakończy się powodzeniem. Z tego powodu badacze zajmujący się problematyką przyczynowości w naukach społecznych zdają się być rzecznikami stanowiska, że: (1) przyczyna jest specyficzną strukturą, w której istotną rolę odgrywają także interakcje między zmiennymi, (2) operowanie pojęciem „dopuszczalna przyczyna” zamiast arbitralne orzekanie o niej jest bezpieczniejsze oraz dokładniej odnosi się do natury zjawiska. Z drugiej strony, moderatorów wśród zmiennych psychologicznych warto szukać między „twardymi” zmiennymi, np. raczej wśród cech temperamentu niż charakteru, cech centralnych osobowości aniżeli powierzchownych itp. Rzetelniej opracowane wyniki badań dowodzą bowiem, że „twarde” zmienne psychologiczne, nawet na poziomie tradycyjnych metod statystycznych (np. ANOVA, MANOVA), uzyskują szczególny status wyjaśniający. Dla przykładu, operując statystyką η^2 (gr. *eta* kwadrat), otrzymujemy informację o tym, w jakim procencie zróżnicowanie wyników zmiennej zależnej jest wyjaśnione wpływem zmiennej niezależnej (por. Francuz i Mackiewicz, 2005, s. 382), np. specyfiką wynikającą z przynależności do określonej grupy lub typu osób. Doświadczeni badacze wiedzą, że jeśli zmienną niezależną będzie „twarda” zmienna (nominalna), np. dobrze wyodrębnione typy strukturalne temperamentu, to ów wyliczony procent jest z reguły znacząco wyższy, niż wtedy, gdyby były to kulturowo uwarunkowane typy struktur osobowości i wynikające z tego powiązania z innymi zmiennymi osobowościowymi.

Jeśli do modelu kauzalnego (regresji wielokrotnej z nieobserwowanymi zmiennymi) włączony zostanie mediator, to wariant rozwiązania pokazany na rycinie 3 będzie wyglądał następująco:



Rycina 5. SEM (CM): prosta regresja wielokrotna z nieobserwowanymi zmiennymi (z udziałem mediatora)

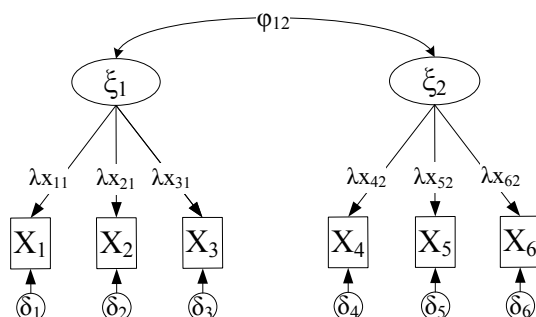
Rozwiązanie przedstawione na rycinie 5 pozwala analizować i testować istotność wpływów bezpośrednich, głównie γ_i (gr. *gamma*), ale także λ_i (gr. *lambda*), np. λx_i oraz λy_i), oraz najprostsze postacie wpływów pośrednich, np. jako iloczyn $\gamma_i \lambda y_i$. Szczególnie wartościowe dla powyższego przykładu jest uwzględnienie roli mediatora. Gdyby była nim zmienna η_1 , to jeden z ważniejszych efektów pośrednich – między ξ_1 a η_2 – byłby obliczany jako iloczyn $\gamma_{11}\beta_{21}$ i mógłby być konfrontowany z efektem bezpośrednim γ_{21} . Natomiast gdyby zmienna η_2 była mediatorem między zmienną egzogeniczną ξ_2 a endogeniczną η_1 , to jeden z ważniejszych efektów pośrednich byłby obliczany jako iloczyn $\gamma_{22}\beta_{12}$ i mógłby być konfrontowany z efektem bezpośrednim γ_{12} . Warto dodać, że testowanie wpływów β_{21} i β_{12} powinno się traktować jako alternatywne a nie jako jednocześnie.

Sprawa kontrolowania wpływów pośrednich jest dość złożona, ponieważ badacz może eksperymentalnie poddać analizie wiele modeli uwzględniających dopuszczalne ustawienia. Po drugie, jak podkreśla Bollen (1989), dopuszczenie pewnej zmiennej endogenicznej do roli mediatora między określoną zmienną egzogeniczną a inną zmienną endogeniczną, wymaga szczególnego uzasadnienia merytorycznego. Po trzecie, aby dopuścić w złożonym modelu strukturalnym udział współczynnika β_{ij} (gr. *beta*), celowe wydaje się wcześniejsze potwierdzenie hipotezy, że wspomniana zmienna endogeniczna pełni rolę mediatora. Dlatego też, mimo formalnej dopuszczalności powyższego modelu, musi on być wcześniej przedyskutowany i uzasadniony merytorycznie. Ważne jest również to, aby przed ostatecznym testowaniem hipotezy o przyczynowości w modelach strukturalnych, ustalić status teoretyczny i pomiarowy zmiennych latentnych egzogenicznych i endogenicznych, za każdym razem w odniesieniu do ich empirycznych argumentów (wskaźników). Z uwagi na to, że związane ze zmiennymi latentnymi błędy – ζ_i (gr. *dzeta*) w znaczym zakresie mają wspólne źródło zmienności, ich wpływ można ograniczyć, dopuszczając kowariancję – ψ_{ij} (gr. *psi*), wymaganą z formalnego punktu widzenia (por. Bollen, 1989).

Konfirmacja konstruktów poznawczych jako przyczyny zachowań

Przed przystąpieniem do budowy pełnego modelu kauzalnego warto uzyskać potwierdzenie dopuszczalności submodeli. Można przyjąć, iż każdy pojedynczy submodel jest próbą zobrazowania jednostkowego konstruktów osobistego w rozumieniu G. A. Kelly'ego (1955, 1970), a zebrana przez badacza empiryczna macierz kowariancji służy jako strukturalny zbiór argumentów konfirmujący (potwierdzający) hipotezę o istnieniu takiego ukrytego

konstruktu. W tym celu wykorzystuje się konfirmacyjną analizę czynnikową (CFA – ang. *Confirmatory Factor Analysis*). Chcę podkreślić, że nie jest to wymaganie, lecz zalecenie merytoryczno-formalne wyniesione z doświadczeń. Badacze zainspirowani dużymi możliwościami SEM, szczególnie ci, którzy budują wielowymiarowe i wielozmiennowe modele, doznają różnych niepowodzeń na etapie ich testowania. Najczęściej jest to negatywna weryfikacja H_0 o braku rozbieżności między obserwowaną macierzą kowariancji (S) a implikowaną przez model (Σ). Źródła owych porażek tkwią zarówno w formułach modeli teoretycznych, jak również w samej macierzy danych (m.in. skośne rozkłady zmiennych, dublujące się zmienne). Tak zwana trafność teoretyczna jest zasadniczym walorem SEM (w porównaniu z tradycyjnymi metodami statystycznymi), dlatego każdorazowe odrzucenie H_0 wiąże się z uznaniem modelu teoretycznego za nietrafny. Szczególnie dotkliwie odczuwa się negatywnie zweryfikowany model kauzalny z wyraźnie wyodrębnionymi częściami: egzogeniczną, strukturalną i endogeniczną, ponieważ wielki wysiłek badacza włożony w prace przygotowawcze okazał się daremny. Aby zapobiec takiej sytuacji, warto rozbudowane modele wpierw podzielić na fragmenty i przeprowadzić ich weryfikację. Takimi submodelami mogą być wyekstrahowane i wyraźnie zdefiniowane obszary badanej rzeczywistości psychologicznej (konstrukty osobiste) oraz odpowiednio dobrane wskaźniki empiryczne, których macierz kowariancji jest podstawą do przyjęcia konstruktu (uznania go za trafny) bądź odrzucenia (uznania za nietrafny). Submodele, które spełnią kryteria dopuszczalności, mają szansę odegrać szczególnie wartościową rolę w rozbudowanym modelu docelowym, w którego strukturalnej części odbywa się testowanie najważniejszej hipotezy (por. Rycina 1). Prosty, dwuczynnikowy model CFA można zobrazować następującym przykładem:

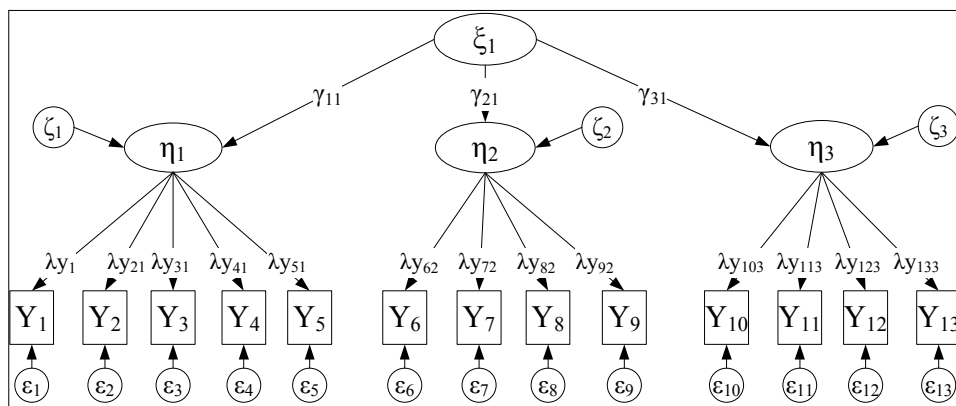


Rycina 6. SEM (CFA): dwuczynnikowy model konfirmacyjnej analizy czynnikowej

(za: Bollen, 1989, s. 260), w którym uwzględniono dwa egzogeniczne konstrukty (ξ_1 i ξ_2 – gr. *ksi*), mimo że suwerenne, to jednak o wspólnym źródle zmienności (ϕ_{12} – gr. *phi*). Empirycznym potwierdzeniem każdego z nich są

wskaźniki obserwacyjne – X_i , reprezentujące dany konstrukt ze stosowną siłą argumentacji (λx_i). Chodzi tu o wyestymowane współczynniki ścieżek (λx_i oraz λy_i) odnoszące się odpowiednio do wpływu ξ_i na X_i oraz η_i na Y_i . Chociaż w wyniku estymacji współczynniki te są zaopatrzone w informację o ich istotności (*p.i.*), to jednak przyjmuje się, że ich wartość standaryzowana nie mniejsza niż 0,3 (a nawet 0,4 przez bardziej wymagających badaczy) traktowana jest jako satysfakcjonująca. Wartość ta jest argumentem formalnej dopuszczalności wskaźnika dla konstruktów, wystarczająco wolnej od wpływu czynnika losowości.

Istnieją także struktury zmiennych latentnych o bardziej złożonej postaci. Do ich zidentyfikowania można posłużyć się przykładem hierarchicznej konfirmacyjnej analizy czynnikowej, omówionym przez K. Bollen (1989, s. 315):



Rycina 7. SEM (CFA): Konfirmacyjna analiza czynnikowa. Model hierarchiczny

W prezentowanym przykładzie (Rycina 7) status każdej ze zmiennych latentnych endogenicznych (η_1 , η_2 , η_3 – gr. *eta*) jest porównywalny, a ukrytą strukturą, w stosunku do nich nadrzędną, jest latentna zmienna egzogeniczna (ξ_1 – gr. *ksi*), która dla endogenicznych ma status źródłowej. Jej dopuszczalność jest argumentowana dwustopniowo. Na pierwszym poziomie są to empiryczne wskaźniki obserwacyjne (Y_i), na drugim – zmienne latentne (η_i).

Warto podkreślić, że badacz konstruujący na bazie CFA narzędzie diagnostyczne, np. test czy kwestionariusz, bazujący na założeniach określonej teorii, wypracowuje jednocześnie klucz do tego narzędzia (współczynniki ścieżek – λx_i lub λy_i – gr. *lambda*, zależnie od wariantu CFA – prosta lub hierarchiczna). Wyliczone wartości tych estymant można porównać z rolą, jaką pełnią ładunki czynnikowe w eksploracyjnej analizie czynnikowej. Natomiast spełnienie przez model CFA warunków dopuszczalności (por. kryteria w tab. 2) jest jednocześnie spełnieniem kryterium trafności teoretycznej narzędzia.

Wyniki ustalone na podstawie CFA można skonfrontować z tymi, które uzyskano przy zastosowaniu tradycyjnej, eksploracyjnej analizy czynnikowej (EFA). Każdorazowe porównanie zwykle dowodzi poprawy walorów trafności i rzetelności narzędzia diagnostycznego opracowywanego metodą CFA. Warto dodać, że bardzo rzadko czynniki wcześniej wyodrębnione metodą EFA wiernie odtwarzają się przy zastosowaniu CFA, mimo że często istnieje ogólne podobieństwo struktur ustalonych przez te dwie odmienne strategie.

Tabela 1. Różnice między eksploracyjną i konfirmacyjną analizą czynnikową

Eksploracyjna analiza czynnikowa (EFA)	Konfirmacyjna analiza czynnikowa (CFA)
jest metodą typu indukcyjnego, realizującą procedurę wnioskowania od szczegółu (obserwacyjnego) do ogółu (syntezy teoretycznej)	jest metodą typu dedukcyjnego, realizującą procedurę wnioskowania od ogółu (konstruktu teoretycznego) do szczegółu (obserwacji)
przeznaczona głównie do redukcji danych	przeznaczona głównie do weryfikacji konstruktu teoretycznego na podstawie danych empirycznych
dopuszcza się związki każdego z czynników ze wszystkimi pytaniami (wskaźnikami)	w wyniku wprowadzonych ograniczeń dopuszcza się niezerowe współczynniki regresji (ładunki czynnikowe) tylko między określonym konstruktom a wskaźnikami przypisanymi jemu na mocy teorii
wynik zwykle jest uproszczeniem pierwotnych zamierzeń teoretycznych i empirycznych badacza	wynik jest źródłem optymalizacji: (1) założeń teorii, (2) narzędzi pomiarowych służących weryfikacji teorii
błąd pomiaru jest informacją wskazującą na konieczność pogodzenia się badacza z ograniczonym walorem trafności narzędzia diagnostycznego	błąd pomiaru jest dla badacza źródłem informacji zwrotnej będącej podstawą do: (1) modyfikowania formuły pytania, jeśli wykazano, że nie było ono suwerenne i testowało równocześnie różne konstrukty, (2) wprowadzania nowych pytań, których brakowało do testowania jakiegoś aspektu teorii
Inne cechy specyficzne konfirmacyjnej analizy czynnikowej:	
(1) umożliwia sprawdzenie dopasowania zarówno całkowitego hipotetycznego modelu czynnikowego, jak również estymacji rozmaitych parametrów wewnątrz tego modelu (2) pozwala na porównanie różnych konkurencyjnych hipotetycznych modeli czynnikowych (3) umożliwia oszacowanie adekwatności danego modelu czynnikowego przy uwzględnieniu różnych grup osób współtworzących badaną populację; podczas analizy tego problemu możliwe jest testowanie roli moderatora (4) pozwala na obliczenie różnych wskaźników dopasowania, które pomagają w porównywaniu różnych modeli czynnikowych	

(por.: Bollen, 1989; Arbuckle i Wothke, 1995/1999; Arbuckle, 1995/2006; Górniak, 2000, 2005; Zakrzewska, 2004, s. 443).

W tej części artykułu chcę zwrócić uwagę na jeszcze jedną kwestię. Otóż badacz, chcąc zastosować SEM do opracowania wyników pochodzących z większego programu badawczego, w którym zastosowano liczne wieloskalowe metody, wcześniej opracowane za pomocą EFA, przed przystąpieniem do obliczeń powinien je uzgodnić z wymogami SEM. Następstwem tego dopracowania jest korekta w kluczach do skal, testów i – oczywiście – w definicjach zmiennych. Ten efekt także optymalizuje walory trafnościowe i rzetelnościowe całego przedsięwzięcia badawczego. Warto podkreślić, że pojawiające się „pokusy”, aby część skal i testów, opracowanych wcześniej przy użyciu tradycyjnego podejścia EFA, pozostawić, inną część – dopracować zgodnie ze standardami SEM, a następnie wszystkie te zmienne włączyć do rozbudowanego, np. kauzalnego modelu SEM, zwykle kończy się porażką. Zostaje to potwierdzone w wynikach testów dopasowania modelu teoretycznego do empirycznej macierzy kowariancji (por. tab. 2). Chociaż takie rozwiązanie jest dopuszczalne z formalnego punktu widzenia, to jednak aspekt merytoryczny (trafnościowy) nie jest tutaj rozpracowany z należytymi standardami. Negatywną konsekwencją tego jest duża zmienność wyników powodowana rolą czynnika losowości a nie siłą związku wskaźnika z konstruktem.

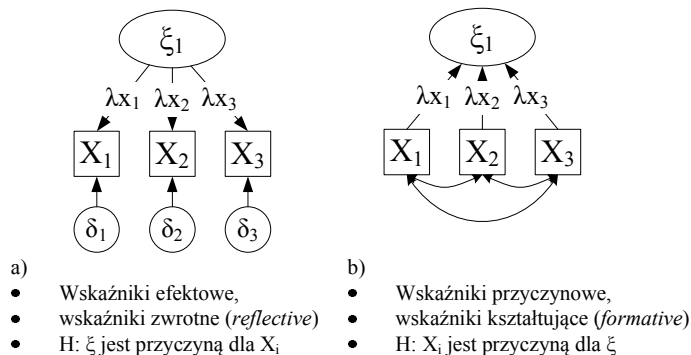
Mając na uwadze doświadczenia badacza stosującego eksploracyjną oraz confirmacyjną analizę czynnikową do opracowania danych psychologicznych, przydatne może się okazać ich porównanie, przynajmniej w aspekcie jakościowym.

Kierunek przyczynowości (wskaźniki efektowe czy przyczynowe?)

Problem kierunku przyczynowości w programach do SEM, z technicznego punktu widzenia, można podjąć alternatywnie: (1) testując hipotezę o tym, że latentna zmienna egzogeniczna (ξ_i) jest przyczyną zmienności dla wskaźnika obserwacyjnego (X_i) lub (2) X_i jest przyczyną dla ξ_i . Dokonując wyboru wariantu, badacz powinien jednak postępować zgodnie z kryterium merytorycznym a nie formalnym. K. Bollen podkreśla, że wyznaczanie kierunku przyczynowości między zmiennymi ukrytymi a wskaźnikami w modelu pomiarowym może stanowić problem. Zgodnie z sugestią H. Blalocka (1975, za: Bollen, 1989) należy rozróżnić wskaźniki kształtujące oraz zwrotne, co w odniesieniu do poniższej Ryciny 8 można przedstawić w sposób następujący:

ad a) W modelu ze „wskaźnikami zwrotnymi” (ang. *reflective indicators*) ścieżki wychodzą od zmiennych ukrytych (konstruktywów) do zmiennych

obserwowalnych (wskaźników) (Bollen, 1989; por. Brown, 2006, s. 351), co w nawiązaniu do klasycznej teorii testów (Lord i Novick, 1968) odzwierciedla bezpośrednią przyczynowość między ukrytym konstruktem a jego obserwacyjną miarą. Jak podkreśla K. Bollen, mamy tu do czynienia z tzw.



Rycina 8. Kierunek przyczynowości. Zmienna latentna jako przyczyny lub synteza danych obserwacyjnych

„wskaźnikami efektowymi” (ang. *effect indicators*). Można dodać, że w takim przypadku zmienna ukryta jest hipotetyczną przyczyną dla zmiennej obserwacyjnej, tzn. wartości zmiennych obserwacyjnych (X_i) są konsekwencją wpływu zmiennej ukrytej (ξ_i). W pracach z zakresu nauk społecznych większość planów badawczych opiera się na takim właśnie paradygmacie wykorzystywania wskaźników efektowych.

Dla zobrazowania statusu wskaźników efektowych proponuję posłużyć się ideą rozumowania dedukcyjnego, w którym wcześniej uporządkowany „ogół”, czyli zbiór prawdziwych przesłanek teoretycznych, pozwala poprawnie przejść do określonego „szczegółu”, np. wniosku empirycznego.

W terminologii SEM zewnętrznymi kryteriami oceny dopuszczalności, prawdziwości lub fałszywości „szczegółu” mogą być wyniki w odpowiednich testach dobroci dopasowania modeli teoretycznych do macierzy danych empirycznych (por. Tabela 2). Z psychologicznego punktu widzenia, badacz korzystający z reguł heurystycznych, w porównaniu z regułami algorytmicznymi, obiera rozumowanie hipotetyczno-dedukcyjne. Jak podkreśla Z. Chlewiński, funkcjonowanie heurystyczne, dla którego najistotniejsze są informacje oraz reprezentacje poznawcze, jest wyższym poziomem funkcjonowania układu nerwowego człowieka (1993, s. 826). Sądzę, że reprezentacje poznawcze można porównać do pojęć albo konstruktów osobistych (ang. *personal constructs*) w rozumieniu G. A. Kelly’ego (1955, 1970) oraz do konstruktów w rozumieniu modelowania strukturalnego, jako zmiennych latentnych egzogenicznych (ξ_i) lub endogenicznych (η_i) zgodnie z tezą głoszoną

przez L. Cronbacha i P. Meehla (1955, s. 283), że „konstrukt to pewna postulowana cecha ludzi, znajdująca odbicie w wykonaniu testu”. Z. Chlewiński (1993) dodaje, że człowiek nabywa uzdolnienia heurystyczne na poszczególnych etapach rozwoju intelektualnego, a następnie odkrywa ich wartość praktyczną w rozwiązywaniu problemów, m.in. dzięki odkrywaniu analogii między danym problemem a problemami ze znanym już rozwiązaniem. Zatem istotą rozumowania dedukcyjnego, heurystycznego zdaje się być formułowanie hipotezy wygenerowanej na podstawie konstruktu pojęciowego o prawdopodobnym jego następstwie, czyli zachowaniu. Opisany powyżej wariant kierunku przyczynowości jest szczególnie przydatny do opracowywania submodeli.

ad **b)** W modelach ze „wskaźnikami kształtującymi” (ang. *formative indicators*) kierunek przyczynowości wiedzie od obserwacyjnego wskaźnika do konstruktu. Dlatego złożona przyczynowość, o którą tutaj chodzi, to następstwo „wskaźników przyczynowych” (ang. *cause indicators*), których zmienne obserwacyjne są hipotetyczną przyczyną istnienia zmiennej ukrytej (Bollen, 1989; por. Brown, 2006, s. 351). W takim przypadku wartości zmiennych obserwacyjnych X_i mają argumentujący wpływ na istnienie zmiennej ukrytej ξ_i . Ta postać modelu – wskaźników przyczynowych, pomimo jej właściwego zastosowania w wielu przypadkach, nie cieszy się jednak dużym zainteresowaniem badaczy, być może dlatego, że bardziej przypomina eksploracyjne niż confirmacyjne postępowanie badawcze. Co więcej, status wskaźników przyczynowych można porównać do typowego dla idei rozumowania indukcyjnego wnioskowania „od szczegółu do ogółu”, które jest podstawą tworzenia uogólnień teoretycznych, chociaż jest ono zarazem redukcją informacji. Uogólniająca synteza dokonana na bazie empirii zwykle nie satysfakcjonuje wnikliwego badacza, który poszukuje nie tyle dopełnienia teoretycznego, ile wypracowania systemu teoretycznego umożliwiającego trafne przewidywanie faktów empirycznie właściwych dla danej dziedziny. Karol Popper (1992, s. 120) wręcz twierdzi, że indukcja jest nieporozumieniem i nie odgrywa żadnej roli w rozwoju wiedzy, a nawet prowadzi do jej regresu. Pytając dlaczego większość rozsądnych ludzi jest przekonana o poprawność indukcji, Popper odpowiada, że jest to problem psychologiczny, u którego podstaw leży mechanizm kojarzenia nakazujący wierzyć ludziom (nawyki, przyzwyczajenia), że to, co zdarzyło się w przeszłości, będzie zdarzać się w przyszłości (ib. s. 127).

Wnikliwy badacz jest bardziej zainteresowany budowaniem teorii tłumaczących rzeczywistość empiryczną niż odwrotnie; teorii, której system tez i hipotez pozwala trafniej analizować, interpretować i tworzyć pojęcia nt. zdarzeń psychologicznych (por. Januszewski, 1997).

Synteza teoretyczna, jaką człowiek tworzy na podstawie jednostkowej obserwacji, ma ograniczoną wartość. Z psychologicznego punktu widzenia badacz odkrywa owo ograniczenie, kiedy – przykładowo – dokonuje analizy

rozumowania transduktywnego. Zdaniem J. Piageta, występuje ono u dzieci w stadium myślenia przedoperacyjnego i polega na tym, że wynik rozumienia pojedynczego zdarzenia (sytuacji, przypadku) jest podstawą utworzenia pseudopojęcia lub „przed-pojęcia”, które jest ściśle związane z danym zdarzeniem. Należące do wspólnej kategorii kolejne zdarzenie jest podstawą utworzenia następnego „przed-pojęcia”, które – dodane do poprzedniego – tworzy zestaw wniosków szczegółowych. Rozumowanie transduktywne polega na przechodzeniu od przypadku do przypadku bez uwzględnienia zasady, która je łączy. Tworzone na ich podstawie pseudopojęcia mogą być odizolowane od siebie a nawet sprzeczne ze sobą. Dzięki usystematyzowanej edukacji szkolnej pseudopojęcia zmieniają się w bardziej uporządkowane, logiczne i zhierarchizowane, stając się pojęciami (por. Przetacznikowa i Spiołek, 1976, s. 396)⁹. Jest prawdopodobne, że badacz, który zaczyna posługiwać się SEM bazującym na wskaźnikach przyczynowych, zapewne sprawniej „przechodzi/zalicza” fazy rozwoju myślenia przedoperacyjnego niż dziecko, które jeszcze nigdy nie doświadczyło myślenia wykorzystującego operacje konkretne, a tym bardziej formalne. Podobnie badacz, posługujący się „klasyczną” eksploracyjną analizą czynnikową, sprawniej niż dziecko przechodzi przez kolejne stadia rozwoju myślenia, o ile porównuje znaczenia konotacyjne poszczególnych pytań, które zawarł w modelu. W ten sposób zmierza do zidentyfikowania jakiejś ukrytej zmiennej psychologicznej, aby ostatecznie sformułować jej wstępną definicję. Zazwyczaj po wykonanym zadaniu badacz albo odczuwa stysfakcję, bo udało mu się „coś” wyodrębnić czy odkryć, co można używać jako tezy w dalszych dociekaniach teoretycznych i empirycznych, albo przeżywa frustrację, gdyż wiele subtelności psychologicznych uwzględnionych w pojedynczych pytaniach nie ma istotnego znaczenia, ponieważ nie pozwalają na to wskaźniki matematyczne, które nie osiągnęły oczekiwanego minimum. Właśnie w tych sytuacjach badacz doświadcza przykrych konsekwencji wynikających z rozumowania indukcyjnego. Zamierzał budować wiedzę poprzez jej różnicowanie, tymczasem staje wobec konieczności zadowolenia się uogólnieniem redukującym bogatą rzeczywistość.

⁹ Tak jak to można zaobserwować w klasycznym eksperymencie piagetowskim, dotyczącym rozwoju poznawczego (okres przedoperacyjny, 2-7 lat), polegającym na przelewaniu cieczy z jednego naczynia do drugiego, a zarejestrowana obserwacja wyrażona jest interpretacją dziecka (... *więcej jest jej w tym naczyniu* ...), które wskazuje na naczynie węższe i wyższe. Punktem odniesienia jest sytuacja początkowa, kiedy w polu widzenia dziecka wlewano tę samą objętość cieczy do dwóch identycznych, transparentnych naczyń. Wobec braku ukształtowanego pojęcia stałości objętości, koncentrowania się na stanach a nie na przekształcaniach stanów, mogą być wypowiadane karykaturalne wnioski, korygowane zazwyczaj przez dziecko posługujące się już zasadą stałości objętości (okres operacji konkretnych, 7-11 lat). Podobnie w dziecięcej ocenie sytuacji ... *a kiedy będzie obiad* ...?, mimo spożytego obfitego posiłku jednodaniowego, lecz poprzedzonego niemal cyklicznym i rytualnym podawaniem obiadu w sekwencji 3-daniowej. Nieprzestrzeganie rytuału może być podstawą do nieuznania takiego posiłku za obiad.

Jednak i w takiej sytuacji można mówić o postępie w tworzeniu wiedzy, bowiem badacz daną zmienną psychologiczną: (1) zidentyfikował, (2) może poszukiwać jej miejsca wśród znanych wcześniej zmiennych psychologicznych, (3) może uczynić ją przedmiotem precyzyjniejszej eksploracji, zgodnie z przekonaniem klasyków SEM, że spełnienie przez model kryteriów jego dopuszczalności wcale nie znaczy, że jest on jedyny i najlepszy.

Wskaźniki efektowe a operacjonalizacja zmiennych

Zagadnienie kierunku przyczynowości koresponduje bezpośrednio z sygnalizowanym przez M. Maloney'a i M. Warda problemem trafności diagnozy psychologicznej i naturą pomiaru psychologicznego. Będąc rzecznikami psychologów zajmujących się diagnozą testową, podkreślają oni obiektywną trudność pomiaru psychologicznego cech, które nie mogą być mierzone bezpośrednio, tak jak czynią to na przykład fizycy, ponieważ są one konstrukcjami pojęciowymi bez namacalnego istnienia (1976, s. 131). Mimo braku bezpośredniego dostępu do konstruktów psychologicznych wymaga się, by podczas budowania narzędzia diagnostycznego i późniejszego posługiwania się nim, ustalić jego trafność. Ustosunkowując się do kwestii trafności testów psychologicznych: treściowej, kryterialnej i teoretycznej, Maloney i Ward podkreślają szczególną wartość tej ostatniej, gdyż „...odnosi się do stopnia, w jakim test mierzy określony konstrukt teoretyczny czy cechę. Podstawową metodą określania tego rodzaju trafności jest teoretyczna analiza natury badanej zmiennej i jej związków z innymi zmiennymi” (1976, s. 136). Autorzy podają jednak, że w obrębie trafności teoretycznej istnieje trafność pojęciowa, której teoria jest efektem końcowym a nie punktem wyjścia (1976, s. 140). Warto w tym momencie dodać, że klasycznie rozumiana trafność teoretyczna odnosi się do stopnia, w jakim test mierzy określony konstrukt teoretyczny (zmienną latentną, przyp. A.J.) bądź cechę (ib. s.136), innymi słowy – im większy stopień badanej zmiennej (wskaźnikowej, np. pytanie lub skala, przyp. A.J.) wiąże się ze zmiennymi (latentnymi, przyp. A.J.), tym większą można mieć pewność, że test spełnia kryterium trafności teoretycznej i mierzy to, co było celem diagnozy. Tym samym osiąga się cel sformułowany przez Amerykańskie Towarzystwo Psychologiczne (APA, 1954, za: Wiggins, 1973, s. 116), a mianowicie, że istotą trafności testów psychologicznych jest pomiar określonych bytów pojęciowych. A zatem mamy tu najprawdopodobniej do czynienia z sytuacją, w której konstrukt teoretyczny jest czymś pierwotnym, a jego związek ze zmiennymi obserwacyjnymi – czymś wtórnym, co odpowiada idei rozumowania dedukcyjnego, na której

bazują wspomniane wyżej modele ze „wskaźnikami efektowymi” (zwrotnymi). Przedmiotem testowania jest tutaj hipoteza, iż konstrukt (ξ_i) jest przyczyną następstw mierzonych wskaźnikami (X_i). Dodatkowo uwzględnia się tutaj udziały błędów pomiaru (δ_i oraz ϵ_i), których analiza umożliwia badaczowi optymalizację zarówno modelu teoretycznego, jak i wskaźników empirycznych (np. korekta pytań). W sytuacji modeli ze „wskaźnikami kształtującymi” testowana jest hipoteza, że indeks (zagregowany zbiór X_i) dowodzi istnienia konstruktów (ξ_i).

Niezależnie od wariantu kierunku przyczynowości w CFA (wskaźniki efektowe bądź kształtujące) badacz zawsze zdobywa wiedzę, którą można wykorzystać do zbudowania trafniejszego modelu teoretycznego (por. hemplowski typ konfirmacji). Zatem mimo naturalnych ograniczeń rozumowanie indukcyjne, charakterystyczne dla modeli ze wskaźnikami kształtującymi, nie jest bezwartościowe, wręcz przeciwnie – jest źródłem cennego doświadczenia empirycznego, np. takiego, które można wykorzystać na rzecz uporządkowanego rozumowania dedukcyjnego, wykorzystującego modele ze wskaźnikami efektowymi. Ta ostatnia sytuacja odnosi się do maksymy K. Poppera (1992) o prymacie „dobrej” teorii nad empirią.

Z merytorycznego i formalnego punktu widzenia, badacz, który według założeń SEM, przeprowadza procedurę specyfikacji modelu, zgodnie z postulatem L. Cronbacha i P. Meehla (1955), wykonuje czynności podstawowe i niezbędne dla konfirmacji modelu teoretycznego. Powinien zatem korzystać z wariantu CFA przeznaczonego dla wskaźników efektowych.

Jedną z najważniejszych czynności w tym zakresie jest operacjonalizacja terminów teoretycznych (zmiennych), która polega na nadawaniu sensu empirycznego terminom psychologicznym (por. Hornowska, 1989, 2000; por. Brzeziński, 1996, s. 206-215). Takie wiązanie terminów teoretycznych z obserwacjami jest podstawowym sposobem wyjaśniania ich znaczenia. Należy być również przygotowanym na to, aby zainteresowanej osobie wyjaśnić treść pojęcia poprzez dokładne pokazanie jego powiązań z obserwowalnymi zachowaniami. Taka „procedura wiązania terminów teoretycznych (odnoszących się do nieobserwowalnych właściwości zdarzeń i obiektów) z terminami obserwacyjnymi (oznaczającymi obserwowalne właściwości i relacje) nazywa się procedurą operacjonalizacji” (Roskam, 1989, za: Hornowska, 2000, s. 390). Badacz formułuje i testuje hipotezę o tym, że istnieje latentny konstrukt pojęciowy, który odnosi się do danego obszaru rzeczywistości psychologicznej. O istnieniu i jakości tego konstruktów zapewnia specyficzny zbiór zachowań, wyrażony np. sposobem odpowiedzi na próby testowe albo na pytania kwestionariusza. Przyjmuje się, że empiryczna macierz danych (osoby $\times$ zmienne wskaźnikowe) odzwierciedla ten zbiór i stanowi podstawę weryfikacji hipotezy o dopuszczalności modelu teoretycznego.

Koncepcja operacjonalizacji powstała w ramach idealizacyjnej teorii nauki i wymaga sprecyzowania następujących pojęć: czynnik, wielkość,

zmienna oraz relacja istotnościowego następstwa. **Czynnik** to pojęcie teoretyczne o charakterze **ontologicznym**, które odnosi się do opisywanej przez badacza rzeczywistości. Czynników może być wiele, a istniejące pomiędzy nimi powiązania można nazwać, np. wykazując istotnościowy wpływ jednego czynnika na drugi. **Wielkość** jest terminem teoretycznym o charakterze epistemologicznym, który pozwala opisać sposób postrzegania rzeczywistości przez badacza. Jest więc **obrazem czynnika** ustalany przez badacza na drodze teoretycznej rekonstrukcji i wynika z przyjętej teorii. Z uwagi na to, że wielkość ma charakter teoretyczny, opisuje zjawiska niedające się bezpośrednio zaobserwować. **Zmienna** jest także pojęciem epistemologicznym, odnoszącym się do sposobu ujmowania i opisywania przez badacza świata rzeczywistego. Problemy badawcze są formułowane w języku określonych pojęć. Aby przejść z poziomu pojęciowego na poziom empiryczny, pojęcia muszą być przekształcone w zmienne. Pod pojęciem „zmienna” rozumie się przyjęty przez badacza sposób przejawiania się wielkości na poziomie obserwacji. Ponadto, wpisująca się w te trzy kluczowe pojęcia teza o „zasadzie empiryzmu” dyscyplinuje badacza w jego dociekaniach przypominając, że dla każdego czynnika teoretycznego istnieje w jego przestrzeni wpływu jakiś jego obszar identyfikacji lub dające się zaobserwować jego konsekwencje (por. Hornowska, 2000, s. 392-394). E. Hornowska (2000, s. 394-400) wymienia następujące etapy procedury operacjonalizacji: (1) rekonstrukcję czynnika teoretycznego – polegającą na ogólnym, teoretycznym zdefiniowaniu danego pojęcia odnoszącego się do rzeczywistości; (2) konceptualizację wielkości – jej efektem jest rekonstrukcja modelu pojęciowego wraz z jego wymiarami; (3) dobór zmiennych operacjonalizujących wielkość – polegający na wyselekcjonowaniu wskaźników (danych obserwowalnych), które pozwolą na trafną identyfikację i testowanie zrekonstruowanego modelu pojęciowego.

W głoszonej wspólnie idei celowości wiązania teorii z empirią szczególny status naukowy zajmują wskaźniki efektowe. Zdaniem K. Poppera (1992), konsekwencją prymatu teorii nad empirią jest poszukiwanie abstrakcyjnych wyjaśnień dla zjawisk, które objęte są procesem badawczym. Tak więc istotą poznania naukowego jest weryfikowanie teorii, a dokładniej – konfirmowanie hipotetyczno-dedukcyjnego modelu, który obejmuje zarówno czynności, jak również wynik potwierdzania hipotezy; bądź ujmując ogólniej, uporządkowanego systemu tez i hipotez dotyczących określonego aspektu rzeczywistości, poprzez skonfrontowanie jej z danymi empirycznymi, które są rejestrowalne w obszarze rzeczywistości znajdującym się w zasięgu teorii.

Kryteria ocen dopasowania modeli

Po przeprowadzonej estymacji testowanego modelu teoretycznego badacz otrzymuje zbiór informacji szczegółowych o jakości dopasowania modelu do danych empirycznych. Są one wyrażone wynikami w testach dobroci dopasowania modelu (walory trafności i rzetelności modelu) w relacji do testującej go macierzy danych.

Tabela 2. Testy dobroci dopasowania stosowane powszechnie do oceny modeli równań strukturalnych (SEM – ang. *Structural Equation Models*) (za: Górniak, 2000)

	test	w. kryt.	interpretacja wyniku testowego
postacie testu χ^2	χ^2		Tradycyjna miara umożliwiająca testowanie H_0 o braku rozbieżności między obserwowaną macierzą kowariancji (S) a implikowaną przez model (Σ) (jeśli $p > 0,05$); tzn. reszty standaryzowane obu macierzy wynoszą 0. Oznacza to, że ograniczenia nałożone przez badacza w modelu teoretycznym są trafne. Użyteczność tego testu, zwłaszcza w przypadku dużych prób, jest ograniczona, gdyż zbyt łatwo prowadzi do odrzucenia prawdziwych modeli.
	p	$p > 0,05$	Cytowanie wyniku tego testu jest raczej konsekwencją pewnej tradycji, a nie racji merytorycznych odnoszących się do trafności i rzetelności modelu. Ten wymóg lepiej spełnia miara χ^2/df . Większa liczba stopni swobody (df) dla modelu oznacza, że jest on „elastyczniejszy”, bardziej uniwersalny, np. łatwiej odtworzy się na innych próbach. Uzyskuje się to, minimalizując liczbę przyjętych założeń w modelu
	χ^2/df	do 2 do 4~5	Maksymalna wartość dla dobrze dopasowanych modeli (mniejsza lub równa 2 – dla kryterium konserwatywnego; mniejsza lub równa 4, a nawet 5 – dla kryterium bardziej liberalnego)
testy RMSEA	RMSEA	$\leq 0,05$	<i>Root Mean Square Error of Approximation</i> (pierwiastek średniokwadratowego błędu przybliżenia); test Steigera-Linda mający wysoką rangę w opinii ekspertów. Jest to miara tego, jak źle dopasowany jest model, z uwzględnieniem jego parametrów wymagających oszacowania. Wynik im bliższy jest 0, tym lepiej model teoretyczny jest dopasowany do macierzy wyników. Wskaźnik nieprzekraczający wartości 0,05 oznacza dobre dopasowanie modelu do danych empirycznych.
	RMSEA LO	$\leq 0,05$	Podstawowa górna granica błędu dla modeli dobrze dopasowanych.

testy RMSEA	RMSEA HI	$\leq 0,08$	Aproksymowana górna granica błędu przyzwoitego dopasowania modelu to wartość 0,08 lub mniejsza, co oznacza dopuszczalny błąd aproksymacji; wartość 0,1 jest podstawą odrzucenia modelu
	PCLOSE	$\geq 0,05$	lub wyższy, zmierzający do 1; tym lepiej dla modelu; wynik 0,05 jest minimalną wartością dającą argument, iż nie ma podstaw do odrzucenia modelu; PCLOSE to tzw. test bliskości.
testy GFI	GFI	$> 0,9$	<i>Goodness of Fit Index</i> – wskaźnik dobroci (jakości) dopasowania analogiczny do współczynnika determinacji w regresji. Mierzy wielkość wariancji-kowariancji macierzy (S), która jest przewidywana przez odtworzoną macierz (Σ). Wartość powyżej 0,9 – oznacza akceptowalny model, 0,95 – satysfakcjonujący, a 1 – doskonałe dopasowanie modelu.
	AGFI	$> 0,9$	<i>Adjusted Goodness of Fit Index</i> – skorygowany wskaźnik dobroci dopasowania; miara analogiczna do skorygowanego współczynnika determinacji w modelach regresyjnych, w którym uwzględnia się liczbę stopni swobody względem liczby zmiennych. Im wyższa liczba <i>df</i> , tym lepszy jest model. Wartość powyżej 0,9 – oznacza akceptowalny model, od 0,95 – satysfakcjonujący, a 1 oznacza doskonałe jego dopasowanie.
	„N”- Hoeltera (0,05)	≥ 200	Wyliczona dla modelu wartość nie powinna być mniejsza niż 200. Jest to dolna granica akceptowalności modelu wyrażona testem „N”-Hoeltera. Jeśli jest mniej, świadczy to, że model może zostać odrzucony, w przypadku gdy liczebność próby spadnie poniżej 200 (dla $p < 0,05$). Jeśli wyliczona dla modelu wartość N jest większa, np. 254, oznacza to, że model może ulec zachwianiu, jeśli skonfrontujemy go z macierzą danych pochodzącą od więcej niż 254 osób.

Testy dopasowania, które najczęściej (niezależnie od „szkoły”) używane są do oceny dopuszczalności modeli teoretycznych przedstawiam w tabeli 2, przyjmując ich syntetyczne interpretacje (za: Górniak, 2000).¹⁰ Z powyższego zestawienia wynika, że największe znaczenie ma grupa testów RMSEA, dalej GFI, test „N”-Hoeltera, postacie testu *chi*² (którego najwcześniej używano do oceny dopuszczalności modeli). Zależnie od „szkoły”, badacze używają także – indeksów NFI, TLI, CFI.

Warto podkreślić fakt, że dorobek matematyków na tym obszarze jest imponujący. Przykładowo, w wersji AMOS-a nr 7 (Arbuckle, 2007) jest 38 testów, przyporządkowanych merytorycznie do 10 grup. Dyskusja nt. zalet

¹⁰ Zastrzegam, iż są to testy odpowiednie dla szerzej sygnalizowanych w tym artykule modeli CFA i CA. W przypadku innych wariantów wykorzystania SEM, sugeruję użycie dodatkowych testów, w czym każdorazowo pomocny okaże się manual (por. Arbuckle, 2005, 2006, 2007 i nast.).

i słabości testów jest bardzo bogata. Badacz, który śledzi jej przebieg, może kierować się w doborze testów m.in. kryterium częstości cytowania, merytoryczną argumentacją dotyczącą znaczenia wyniku w danym teście lub zgodnością wyników z różnych grup testów.

Na koniec, warto choćby zasygnalizować, że badacz posługujący się SEM musi dokonać kilku rozstrzygnięć natury formalnej wobec prezentacji opracowanego materiału badawczego. Musi zająć stanowisko w sprawach: prezentacji wyników rozwiązania modelu: graficznie na diagramie lub tabelarycznie, przedstawienia rozwiązań standaryzowanych lub niestandaryzowanych, bądź jednych i drugich, co w ostatnim wariancie czyni sprawozdanie dość obszernym. Sprawa prezentacji wyników standaryzowanych i/lub niestandaryzowanych jest wyzwaniem podobnym do klasycznego modelu wielokrotnej regresji liniowej: estymanty standaryzowane są bardziej przydatne do spełnienia funkcji eksplanacyjnej równania, estymanty niestandaryzowane – do funkcji predyktywnej. Ponadto, źródła optymalizacji modelu teoretycznego tkwią w szczegółowej analizie estymant niestandaryzowanych, które dla badacza są najcenniejszym źródłem informacji zwrotnej o newralgicznych fragmentach modelu i możliwościach dokonania w nim zmian (zabiegi re-specyfikacji), aby po powtórnej estymacji stał się modelem dopuszczalnym.

W przypadku estymant niestandaryzowanych badacz otrzymuje także oszacowania wartości błędów, np.: (ζ_i – gr. *dzeta*) – w równaniach, błędy (zakłócenia) związane z określonymi zmiennymi latentnymi endogenicznymi, (ψ_{ij} – gr. *psi*) – kowariancje (korelacje – w przypadku danych standaryzowanych) między błędami (zakłóceniami) ζ_i związanymi z pomiarem zmiennych η_i , (ε_i – gr. *epsilon*) – błąd związany z pomiarem określonego wskaźnika Y_i , (δ_i – gr. *delta*) – błąd związany z pomiarem określonego wskaźnika X_i , i wreszcie powiązane z ostatnimi – kowariancje (θ_ε – gr. *teta* _{ε}) – między błędami pomiaru związanymi z pomiarem zmiennych Y_i oraz (θ_δ – gr. *teta* _{δ}) – między błędami pomiaru związanymi z pomiarem zmiennych X_i .

Wiedza na temat oszacowanych błędów może być przez badacza twórczo wykorzystywana, np. do optymalizacji modeli teoretycznych, a nade wszystko – konstruowania testów psychologicznych jako trafniejszych niż dotąd narzędzi badawczych, konstruowanych na podstawie założeń tradycyjnej psychometrii.

Jeśli jednak zasadniczym celem opracowania jest prezentacja gotowych modeli, które na etapach zwykle żmudnych przygotowań spełniły kryteria identyfikowalności, poprawności i dopuszczalności, wówczas, jak sądzę, badacz może poprzestać na prezentacji estymant standaryzowanych, ponieważ na ich podstawie może dokonać opisu i analiz tego, co z teoretycznej (dedukcyjnej) perspektywy wydaje się najważniejsze, a mianowicie określić wewnętrzną zasadę organizacji konstruktów (lub konstruktów) jej uwarunkowania i następstwa. Łatwiej rozwiązać problemy związane z optymalizacją modeli, kiedy odwołujemy się do macierzy estymant niestandaryzowanych.

To ostatnie rozwiązanie warto polecić badaczom przygotowującym podręczniki do konstrukcji testów psychologicznych. Mimo, że w większości sprawozdań SEM przedstawia się standaryzowane estymanty (por. Brown, 2006, s. 147), to jednak należy być bardzo czujnym, ponieważ wyłączenie bazowanie na rozwiązaniach standaryzowanych może zakończyć się przyjęciem mylącego lub błędnego rozwiązania (por. Bollen, 1989).

Warto podkreślić praktyczny aspekt estymant standaryzowanych: ujęcie ich w przedziale (od -1 do +1), niejako sprowadzenie ich do wspólnego mianownika, ułatwia prowadzenie analiz, gdyż mogą być one bezpośrednio porównywane ze sobą. Jest to przydatne zwłaszcza wtedy, gdy rozkłady wyników surowych dla zmiennych wskaźnikowych bazowały na różnych rodzajach skal przedziałowych.

Szczególnym przedsięwzięciem wieńczącym całość analiz jest raport z badania wykorzystującego SEM. Schemat raportu z uwagi na jego wielość merytoryczną i formalną wymaga oddzielnego studium.

Literatura cytowana

- APA/American Psychological Association (1954). Technical recommendations for psychological tests and diagnostic techniques. *Psychological Bulletin*, 51(2), Part 2.
- Arbuckle, J. L., Wothke W. (1995/1999). *Amos 4.0. User's Guide*. Chicago, IL: Small/Waters Corporation.
- Arbuckle, J. L. (1995/2005). *Amos 6.0. User's Guide*. Spring House, PA: Amos Development Corporation.
- Arbuckle, J. L. (1995/2006). *Amos 7.0. User's Guide*. Spring House, PA: Amos Development Corporation.
- Arbuckle, J. L. (1995/2007). *Amos 16.0. User's Guide*. Spring House, PA: Amos Development Corporation.
- Baron, R. M., Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 65, 1173-1182.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, London: The Guilford Press.
- Brzeziński, J. (1996). *Metodologia badań psychologicznych*. Warszawa: Wydawnictwo Naukowe PWN.
- Brzeziński, J. (2000). Podstawowe modele badawcze. W: J. Brzeziński (red.) *Metodologia badań naukowych i diagnostycznych*. W: J. Strelau (red.), *Psychologia. Podręcznik akademicki* (t. 1, s. 355-387). Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- Chlewiński, Z. (1993). Hasło – heurystyka w psychologii. W: *Encyklopedia katolicka* (t. 6, s. 826). Lublin: Towarzystwo Naukowe KUL.
- Cronbach, L. J., Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.

- Francuz, P., Mackiewicz, R. (2005). *Liczby nie wiedzą skąd pochodzą. Przewodnik po metodologii i statystyce nie tylko dla psychologów*. Lublin: Wydawnictwo KUL.
- Gaul, M., Machowski, A. (1987). Elementy analizy ścieżek. W: J. Brzeziński (red.). *Wielozmiennowe modele statystyczne w badaniach psychologicznych* (s. 82-112). Warszawa: PWN.
- Górniak, J. (2000). *My i nasze pieniądze. Studium postaw wobec pieniądza*. Kraków: Aureus.
- Górniak, J. (2005). *Modelowanie strukturalne z wykorzystaniem programu AMOS*. Kraków: SPSS Polska Sp. z o. o. (skrypt kursowy do użytku wewnętrznego).
- Hajduk, Z. (2002). Hasło – Konfirmacja. W: B. Migut i in. (red.). *Encyklopedia Katolicka* (t.9, s. 582-584). Lublin: Towarzystwo Naukowe KUL.
- Holmbeck, G. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators. Examples from child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology*, 65(4), 599-610.
- Hornowska, E. (1989). *Operacjonalizacja wielkości psychologicznych. Założenia – struktura – konsekwencje*. Wrocław: Ossolineum.
- Hornowska, E. (2000). Operacjonalizacja terminów teoretycznych: czynnik, wielkość, zmienna. W: J. Brzeziński i E. Hornowska (red.), *Podstawowe metody badawcze – teoria i praktyka testowania*. W: J. Strelau (red.). *Psychologia. Podręcznik akademicki* (t. 1, s. 389-400). Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- Januszewski, A. (2009). *Modele równań strukturalnych w metodologii badań psychologicznych. Wybrane aspekty optymalizacji trafności i rzetelności modeli*. Lublin: Katedra Psychologii Ogólnej KUL (mps.).
- Judd, Ch. M. (1995/2002). Hasła: analiza ścieżkowa (*path analysis*), modelowanie równań strukturalnych (*structural equation modeling*), W: A.S.R. Monasted i M. Hewstone (red.), *Encyklopedia Blackwella. Psychologia społeczna*. Warszawa: Jacek Santorski & Wydawnictwo.
- Kamiński, S. (1989). Hasło – Falsyfikacja. W: P. Hempterek i in. (red.), *Encyklopedia Katolicka* (s. 32). Lublin: Towarzystwo Naukowe KUL.
- Kelly, G. A. (1955). *The psychology of personal constructs*. New York: Norton.
- Kelly, G. A. (1970). A brief introduction to personal construct theory. W: D. Banister (red.), *Perspectives in personal construct theory* (s. 1-30). London: Academic Press.
- Konarski, R. (2009). *Modele równań strukturalnych. Teoria i praktyka*. Warszawa: Wydawnictwo Naukowe PWN.
- Lord, F. M. & Novick, M. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley. [8]
- Maloney, M. P. i Ward, M. P. (1976/1995). Testy psychologiczne jako metoda zbierania danych. W: W. J. Paluchowski (red.), *Testowanie a ocena kliniczna. Wybrane problemy* (t. 6, s. 94-141). Warszawa: Polskie Towarzystwo Psychologiczne.
- Popper, K. (1992). *Wiedza obiektywna. Ewolucyjna teoria epistemologiczna*. Warszawa: PWN.
- Przetacznikowa, M. i Spionek, H. (1976). Wiek poniemowlęcy. W: M. Żebrowska (red.). *Psychologia rozwojowa dzieci i młodzieży* (s. 343-415). Warszawa: PWN.
- Ringle, Ch. M. (2004). *Messung von Kausalmodellen. Ein Methodenvergleich*. Hamburg: Karl-Werner Hansmann, Universität Hamburg.

- Shaughnessy, J. J., Zechmeister, E. B., Zechmeister, J. S. (1990/2002). *Metody badawcze w psychologii*. (tłum. z j. ang. M. Rucińska). Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- SPSS Inc. (2006). *SPSS 14.0 PL for Windows*. Chicago, IL: SPSS Inc.
- Strelau, J. (2002/2006). *Psychologia różnic indywidualnych*. Warszawa: Wydawnictwo Naukowe „Scholar”.
- Wiggins, J. S. (1973/1994). Techniki ustrukturuwane. W: W. J. Paluchowski (red.), *Kwestionariusze w diagnostyce psychologicznej i badaniach naukowych* (t. 5, s. 93-170). Warszawa: Polskie Towarzystwo Psychologiczne.
- Wojciszke, B. (2004). Systematyczne Modyfikowane Autoreplikacje: logika programu badań w psychologii. W: J. Brzeziński (red.), *Metodologia badań psychologicznych. Wybór tekstów* (s. 44-68). Warszawa: Wydawnictwo Naukowe PWN.
- Zakrzewska, M. (2004). Konfirmacyjna analiza czynnikowa w ujęciu pakietu statystycznego LISREL 8.51 (2001) Karla G. Jöreskoga i Daga Sörboma. W: J. Brzeziński (red.), *Metodologia badań psychologicznych. Wybór tekstów* (s. 442-478). Warszawa: Wydawnictwo Naukowe PWN.