

## ZAŁOŻENIA FORMALNE MODELI WERYFIKOWANYCH ZA POMOCĄ UKŁADÓW RÓWNAŃ STRUKTURALNYCH

### STRESZCZENIE

Celem artykułu jest opisanie założeń formalnych dotyczących wykorzystywania metody równań strukturalnych (*Structural Equation Model* – SEM). Układy równań strukturalnych służą do weryfikowania teorii i jej założeń.

W artykule zostało opisanych pięć założeń budowy każdego modelu pomiarowego, którego stworzenie poprzedza weryfikację założeń teoretycznych za pomocą układu równań strukturalnych. Głównym celem budowy modelu pomiarowego jest sprawdzenie budowy zmiennych latentnych, w tym: a) sprawdzenie rzetelności zmiennych latentnych, b) obliczenie wariancji zmiennych latentnych, c) sprawdzenie trafności różnicowej zmiennych latentnych, d) sprawdzenie wysokości ładunków czynnikowych ( $\lambda$ ) oraz e) sprawdzenie statystyk dopasowania modelu.

Kolejno omówione zostały założenia modelu weryfikowanego za pomocą jednopoziomowego układu równań strukturalnych SEM: a) sposoby wyłączenia ścieżek w modelu, b) sprawdzenie dopasowania modelu teoretycznego do danych empirycznych.

W celu przybliżenia omawianych treści zostały przedstawione grafy prezentujące model pomiarowy, a także model SEM.

Słowa kluczowe: model, struktura, teoria, elementy teoretyczne, pojęcia teoretyczne, układ równań

### FORMAL ASSUMPTIONS OF THE STRUCTURAL EQUATION MODELS

#### ABSTRACT

The purpose of this article is to describe the formal assumptions regarding the use of *Structural Equation Models* (SEM). *Structural Equation Models* are used to verify the theory and its assumptions.

Five assumptions concerning construction of the measurement model, which precedes the verification of the theoretical assumptions with the aid of the structural equation model, were described. These are: a) checking the reliability of the latent variables, b) calculation of the variance of latent variable, c) checking the discriminant validity

---

<sup>1</sup> Adres do korespondencji: elysium5678@gmail.com.

of the latent variables, d) checking the amount of factor loadings ( $\lambda$ ) and e) model fit statistics.

In the next step, the one-level SEM model assumptions were discussed: a) turning off the paths of the model, b) checking the fit of the theoretical model to empirical data. For imaging of the content, graphs showing the measurement model and model SEM were presented.

Keywords: model, structure, theory, theoretical elements, theoretical notions, structural equations

## WPROWADZENIE

Celem artykułu jest przybliżenie kluczowych założeń modeli weryfikowanych za pomocą układów równań strukturalnych. Przedstawiony w artykule materiał stanowi syntetyczną wiedzę na temat etapów budowy modeli oraz najważniejszych procedur związanych z ich weryfikacją. Zebrany materiał ma przybliżyć czytelnikowi wagę poszczególnych etapów weryfikacji modeli oraz ich implikacji (np. nie można mówić o weryfikacji modelu, jeżeli uprzednio nie skonstruowano modelu teoretycznego itp.), a także pokazać potencjalne błędy popełniane podczas weryfikacji modeli za pomocą układów równań strukturalnych.

Układ równań strukturalnych jest zaawansowaną, wielozmiennową metodą statystyczną służącą do weryfikacji teorii i jej założeń. Teoria to „uporządkowany zbiór, system praw nauki, wyjaśniający pewien typ zjawisk czy zdarzeń wedle jednolitej zasady” (Nowak, 2007, s. 397). Teoria powinna opisywać (deskrybować) oraz wyjaśniać proces lub strukturę jakiegoś zjawiska. Powinna być również przydatna do przewidywań. Nie powinna zatem pomijać warunków początkowych, uruchamiających prawa nią rządzące. Ta ostatnia uwaga jest szczególnie istotna dla strukturalno-formalnej zasady teorii, której zadaniem jest ujawnianie relacji między jej elementami.

Każda teoria powinna zostać zweryfikowana empirycznie. Jak zauważa Nowak: „Żadne założenia teoriiotwórcze nie mogą zastąpić empirii w orzekaniu o prawdziwości praw nauki” (2007, s. 401). Weryfikacja empiryczna hipotez danej teorii czyni zasadnym stawianie nowych hipotez i rozwój teorii, z drugiej strony pozwala na modyfikowanie założeń teoretycznych, gdy te okazują się błędne.

Rozwój teorii odbywa się zatem w dwóch etapach. W etapie pierwszym następuje opis praw rządzących daną teorią, tzn. wyłonienie pojęć i twierdzeń teoretycznych. Z kolei w etapie drugim teoria zostaje zweryfikowana. Weryfikacja teorii ma dla nauki doniosłe znaczenie. Znalezienie empirycznego potwierdzenia przedstawionych założeń teoretycznych utrzymuje w mocy ich słuszność. Pozwala to na kolejne rozszerzenie teorii i/lub wykorzystywanie jej do budowania nowych teorii. Stąd we współczesnej metodologii badań tak doniosłe znaczenie modeli strukturalnych.

Współczesne układy równań strukturalnych dzielą się na dwa rodzaje, ze względu na hierarchiczność weryfikowanej teorii. Wyróżnia się zatem jednopoziomowe układy równań strukturalnych (*one level structural equation models* – SEM) oraz wielopoziomowe, zwane również hierarchicznymi (*multi-level structural equation models* – MSEM; Heck, Thomas, Tabata, 2010; Heck, Thomas, 2009).

Jeżeli model teoretyczny ma charakter jednopoziomowy, wówczas jest weryfikowany za pomocą jednopoziomowego układu równań strukturalnych. Jeżeli ma charakter wielopoziomowy (hierarchiczny), wówczas do obliczenia wykorzystuje się wielopoziomowy układ równań strukturalnych.

Modele wielopoziomowe zakładają, że na zmienne latentne z niższego poziomu hierarchicznego wpływają zmienne latentne, które pochodzą z wyższego poziomu hierarchicznego i które moderują związki między zmiennymi z poziomu niższego. I tak np. temperament dziecka może być potraktowany jako zmienna z wyższego poziomu hierarchii, moderująca związek między dwiema zmiennymi z poziomu podrzędnego, np. szybkością uczenia się i satysfakcją z chodzenia do szkoły.

Konstrukcja modelu hierarchicznego powinna zostać zaprezentowana na poziomie opisywania weryfikowanej struktury teoretycznej. Do weryfikacji tej struktury należy zastosować wielopoziomowy układ równań strukturalnych (MSEM). Wykorzystuje on efekty międzygrupowe, czyli procedurę analizy wariancji, co umożliwia ujawnianie efektów różnic między grupami w zakresie związków zmiennych z poziomu niższego (Heck i in., 2010; Heck, Thomas, 2009; Rabe-Hesketh, Skrondal, Zheng, 2007; Tranmer, Elliot, 2007). Nie wszystkie jednak programy obliczające modele SEM obliczają modele MSEM (Garson, 2013; Raudenbush, Bryk, Cheong, Congdon, Toit, 2011).

Hierarchiczne i jednopoziomowe modele SEM można również łączyć z hierarchicznymi i jednopoziomowymi modelami pomiarowymi. Nie ma w tej kwestii żadnych ograniczeń, ponieważ w teoriach jednopoziomowych mogą istnieć elementy, które mają charakter wielopoziomowy, i odwrotnie. Zatem również model pomiarowy może przyjąć postać zarówno modelu jednopoziomowego, jak i hierarchicznego albo modelu drugiego poziomu (*second order*; Milfont, Duckitt, 2004). To, jaką przyjmuje on postać, również zależy od hierarchiczności, tym razem jednak jest ona przypisana właściwościom, charakterystykom – inaczej mówiąc: elementom teoretycznym – a nie samej teorii. W dalszej części artykułu opisany jest wyłącznie model jednopoziomowy.

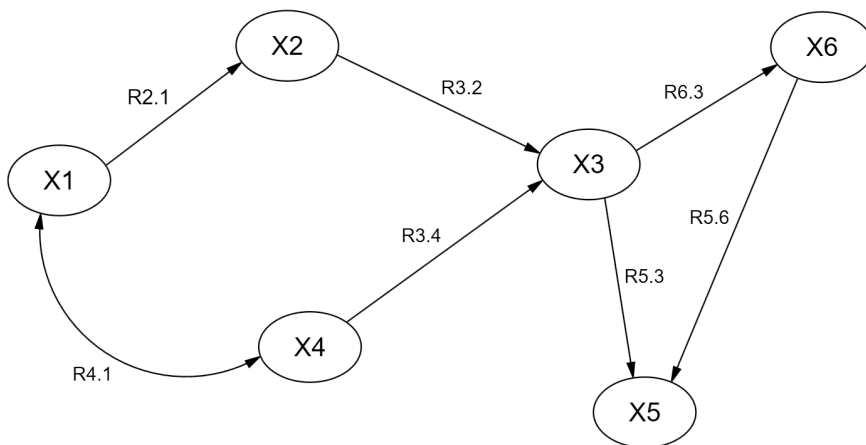
Ujęcie strukturalne wymaga prezentacji teorii w sposób odbiegający od tradycyjnego jej przedstawiania. To nowe ujęcie, zaproponowane przez Suppessa, jest nazywane *aksjomatyzacją teorii* (Suppes, 1972). Kluczowym elementem w takiej rekonstrukcji teorii staje się pojęcie *modelu* rozumianego jako interpretacja aksjomatów teorii. Jak podaje Jonkis: „Ujęcie strukturalne jest wyróżnione stosowanym w nim sposobem definiowania modeli teorii” (Jonkis, 1998, s. 14),

i dalej: „a modele (właściwe) teorii, rozumianej jako system dedukcyjny, uznaje się za takie interpretacje standardowe, które są realizacjami aksjomatów teorii” (1998, s. 14). Gdy mówimy więc o budowie modelu, odnosimy się do strukturalnej rekonstrukcji teorii.

### REKONSTRUKCJA TEORII W UJĘCIU STRUKTURALNYM

Rekonstrukcja teorii wymaga przeprowadzenia analizy materiału dotyczącego badanej teorii, w tym: a) zakresów podstawowych pojęć teorii (podziału pojęć na swoiste i zapożyczone), b) podstawowych praw teorii, c) zasad ograniczających działanie teorii. Kolejno ważne jest rozpoznanie struktury teorii i jej związków z innymi teoriami. W końcu przedstawienie teorii w języku rekonstrukcji, a więc zaprezentowanie struktury teorii w postaci modelu. Współcześnie przyjęło się prezentować modele teorii również za pomocą matematycznych grafów, które stanowią wizualną reprezentację weryfikowanego modelu teoretycznego. Podczas rekonstrukcji teorii ważne jest również (o ile to możliwe) przedstawienie innych możliwych modeli badanej teorii (tzw. modeli alternatywnych).

Model strukturalny, stanowiący rekonstrukcję weryfikowanej teorii, zwykle się nazywać modelem substancywnym (*substantive model*) bądź modelem procesu substancywnego (*model of the substantive process*). Podsumowując: model jest strukturalną rekonstrukcją teorii lub mówiąc inaczej: model reprezentuje założenia teorii (Gospodarek, 2009).



Rysunek 1. Zakres i charakterystyka struktury.

Rekonstruując teorię w ujęciu strukturalnym, nadaje się jej strukturę. Struktura to układ o postaci  $\langle X_j, R_j \rangle$ , gdzie  $X_j$  to ciąg elementów, z których składa się struktura, zwanych również *zakresem struktury*. Z kolei  $R_j$  to relacje

między elementami struktury, czyli *charakterystyka zakresu tej struktury*. Typy relacji między elementami danej struktury nazywane są aksjomatami. Metoda rekonstrukcji teorii, zwana *aksjomatyzacją teorii*, to zatem określanie typów relacji między jej elementami. Nie istnieją ograniczenia co do ilości elementów  $X_j$ , jak również  $R_j$ . Nie ma zatem ograniczeń co do zakresu struktury.

Rysunek 1 przedstawia przykładowe relacje strukturalne ( $R_j$ ) między elementami ( $X_j$ ).

Typów możliwych relacji jest kilka; mogą to być relacje: związku między elementami, relacje moderacji, mediacji i relacje wpływu.

Relacja R4.1 to relacja związku. Na grafach prezentujących strukturalne ujęcia teorii jest prezentowana za pomocą zakrzywionej strzałki z dwukierunkowym grotom.

Relacje R2.1, R3.2 oraz R3.4 to relacje wpływu. Na grafach są prezentowane za pomocą prostych strzałek z jednokierunkowymi grotami, symbolizującymi kierunek wpływu.

Relacje R2.1 i R3.2 to relacja mediacji. Zmienna  $X_2$  pośredniczy (mediuje) w związku zmiennych  $X_1$  i  $X_3$ .

Relacje R5.3, R6.3 oraz R5.6 to relacje z efektem moderacji. W tym wypadku element  $X_3$  moderuje relację R5.6, prowadząc do efektu synergii bądź buforowania (Nowak, 2007).

Podczas rekonstrukcji teorii podaje się zakres struktury, tzn. określa się wszystkie jej elementy: podaje zakresy pojęć, określa się, które pojęcia są swoiste dla danej teorii, a które zapożyczone, oraz relacje między nimi. Pokazuje się związki teorii z innymi teoriami. Przedstawia się zakres możliwych modeli badanej teorii. Kolejno prezentuje się model za pomocą matematycznego grafu ukazującego zakres struktury, tzn. opisuje wszystkie elementy i relacje między nimi. W ten sposób stawia się hipotezy szczegółowe, które dotyczą relacji między elementami w strukturze modelu. W omawianym tu przykładzie hipoteza pierwsza (H1) odnosiłaby się do relacji R4.1 i brzmiałaby: istnieje związek między elementami  $X_1$  i  $X_4$  itd.

Jak zostało powiedziane wcześniej, teoria nie powinna pomijać warunków początkowych uruchamiających jej prawa. Kwestią niezmiernie ważną, jak również często poruszaną w literaturze przedmiotu (Konarski, 2009) jest problem umieszczenia kluczowych dla zjawiska zmiennych podczas opisywania modelu. Badacz powinien w nim uwzględnić wszystkie zmienne kluczowe do wyjaśnienia analizowanego zjawiska. Ich opuszczenie może się przyczynić do zbudowania złego modelu, który prawdopodobnie zostanie odrzucony albo w niewielkim stopniu będzie objaśniał analizowany proces. Stąd ważne jest nie tylko dopasowanie modeli do danych, lecz także siła związku między zmiennymi oraz kierunek zależności (Hair, Black, Babin, Anderson, Tatham, 2006).

Oczywiście kwestia wprowadzania wszystkich zmiennych do modelu jest rzeczą problematyczną. Nigdy nie można być absolutnie pewnym, że uwzględniono

wszystkie zmienne w modelu substancywnym, można jedynie zbliżyć się do pewnego ideału – zawsze może się okazać, że istnieją jeszcze takie zmienne, które dotychczas nie zostały przewidziane na poziomie teoretycznym, a które jednak pośredniczą w analizowanym procesie.

Podczas rekonstrukcji modelu należy również unikać przesady w jego upraszczaniu. Może to doprowadzić do niewystarczającego, pobieżnego opisu zjawiska. Z drugiej strony, prostota modelu ma bardzo duże implikacje w modelowaniu. Modele bardzo złożone są trudne w interpretacji. Stara naukowa zasada brzytwy Ockhama (*Ockham's razor*) ma w przypadku budowy modeli substancywnych szczególne znaczenie. Należy dbać o to, aby budowane modele miały nie tylko jasną strukturę, lecz także przedstawiały sensowne, interpretowalne powiązania między zmiennymi w modelu. Modele, które przedstawiają prawie wszystkie możliwe powiązania między zmiennymi, są w rzeczywistości bardzo trudne do interpretacji. Wątpliwa jest również wartość teorii, która niczego nie wyjaśnia, tylko utrzymuje, że powiązane jest wszystko ze wszystkim. Budowa modelu nie polega więc na połączeniu między sobą wszystkich zmiennych siecią powiązań<sup>2</sup>, przeciwnie: polega na uwolnieniu związków między zmiennymi i pozostawieniu jedynie istotnych z punktu widzenia teorii.

Budowa modelu substancywnego wymaga od badacza wyłonienia zmiennych egzogenicznych oraz zmiennych endogenicznych. Zmienne egzogeniczne to główne zmienne, niezależne modelu, od nich rozpoczyna się proces opisywany przez model. Wskazanie zmiennej egzogenicznej/zmiennych egzogenicznych w modelach jest rzeczą niezmiernie ważną. W przypadku, gdy w modelu znajduje się więcej niż jedna zmienna egzogeniczna, założeniem formalnym modelu jest skorelowanie zmiennych egzogenicznych<sup>3</sup>. W modelu zaprezentowanym na rysunku 1 zmiennymi egzogenicznymi są zmienne X1 i X4.

Nie sposób przecenić wagi zmiennej egzogenicznej w modelu. Można ją porównać do wpływu takich zmiennych, jakimi są np. siła i kierunek wyrzucenia kamienia, na przebieg jego lotu oraz miejsce upadku. Określenie zmiennej egzogenicznej w modelu jest sprawą, do której należy przykładać specjalne znaczenie. Należy pamiętać, że m.in. dzięki wskazaniu zmiennych egzogenicznych w modelach SEM przypisuje się im status modeli, które nie tylko weryfikują teorię, lecz także odtwarzają procesy (Konarski, 2009).

Zmienne endogeniczne to zmienne, które zależą od zmiennych egzogenicznych oraz od siebie wzajemnie; zwane są również zmiennymi *wewnętrznymi*.

---

<sup>2</sup> Rozwiązanie takie nie ma sensu również z poziomu statystycznego. Modele, w których obliczane są wszystkie związki między zmiennymi, to modele nasycone – są one zawsze dopasowane do danych. Pozostawienie więc wszystkich związków między zmiennymi i stwierdzenie, że model jest dopasowany do danych, jest truizmem.

<sup>3</sup> Koreluje się zmienne, którym w przypadku zmiennych egzogenicznych nie jest przypisany błąd pomiaru.

### OPERACJONALIZACJA POJĘĆ TEORII

Poprawna operacjonalizacja pojęć teorii jest kluczowym elementem całego procesu weryfikacji. Operacjonalizacja wymaga podania zakresu pojęć, tzn. ich konotacji, w taki sposób, aby można było wyłonić zbiór wszystkich desygnatów nazwy, czyli wskazać jej denotację. Kolejno należy przeprowadzić procedury operacjonalizacji, tzn. stworzyć metody umożliwiające pomiar badanego zjawiska.

W naukach psychologicznych często właściwość (charakterystyka), którą badamy, ma charakter ukryty (latentny), inferujemy (wnioskujemy) o niej na podstawie obserwowalnych wskaźników. Wskaźnik musi zatem mieć dużą moc zawierania (tzn. zawierać w sobie to, co dotyczy *indicatum*, czyli badanego pojęcia), jak również dużą moc odrzucania (tzn. odrzucać poza swój zakres to, co nie dotyczy *indicatum* – badanego pojęcia; Nowak, 2007).

Wskaźnik, który miałby doskonałą moc zawierania (zawierałby w sobie wszystko, co dotyczy *indicatum*) oraz doskonałą moc odrzucania (odrzucałby poza swój zakres wszystko, co do *indicatum* nie należy), byłby *czystym markerem*. Znalezienie czystego markera wymaga jednak długoletnich badań, dlatego najczęściej do operacjonalizowania pojęć służy wiele wskaźników. W przypadku modeli rekonstruujących teorię, a weryfikowanych kolejno przez układy równań strukturalnych, liczba wskaźników staje się sprawą niebagatelną, gdyż wyznacza złożoność modelu, a tym samym kształtuje warunki proceduralne całych badań.

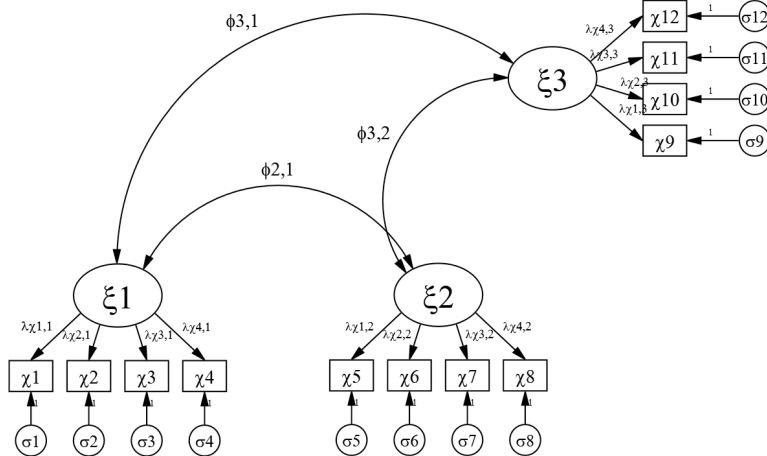
W modelach weryfikowanych przez układ równań strukturalnych zmienne latentne reprezentujące cechy ukryte (symbole  $\xi$  i  $\eta$ ) objaśniają wariancję zmiennych obserwowalnych (symbolizowanych za pomocą  $X$  i  $Y$ ). Zmienne obserwowalne to wskaźniki zmiennej ukrytej (latentnej). Na poziomie modelu pomiarowego<sup>4</sup>, między zmiennymi pojawia się nowa relacja, której nie ma w modelu rekonstruującym teorię. W modelu tym pojawiają się dwie relacje, które można by za Figurską nazwać relacjami dominacji. Są odwrotne wobec siebie: zmienna latentna jest nadrzędna wobec swoich zmiennych obserwowalnych, a zmienne obserwowalne są podrzędne w stosunku do zmiennej latentnej (Figurska, 1993).

Rysunek 2 przedstawia model pomiarowy z trzema zmiennymi latentnymi. Model ten jest odpowiednikiem modelu substancywnego, reprezentującego teorię wówczas, gdy założy się istnienie wszystkich możliwych relacji między jej elementami (zmiennymi). Budowa modelu pomiarowego służy określeniu, czy wszystkie zmienne reprezentujące właściwości psychiczne zostały poprawnie

---

<sup>4</sup> Model pomiarowy stanowi pierwszy etap budowy pełnego modelu weryfikowanego za pomocą układu równań strukturalnych. Jego celem jest sprawdzenie, czy zmienne latentne modelu zostały dobrze skonstruowane. Model ten różni się od modelu strukturalnego tym, że ma wszystkie możliwe powiązania między zmiennymi latentnymi, ma też uwolnione powiązania między błędami pomiarowymi.

zoperacjonalizowane. Jest to pierwszy etap weryfikowania modelu za pomocą równań strukturalnych.



Rysunek 2. Graf prezentujący jednopoziomowy model pomiarowy (*Measurement Model*), gdzie  $\lambda$  jest ładunkiem czynnikowym,  $\sigma$  jest wariancją błędów,  $\chi$  symbolizuje zmienne obserwowalne,  $\xi$  symbolizuje konstrukty latentne,  $\phi$  symbolizuje korelacje między zmiennymi latentnymi.

Jak widać, każda zmienna latentna (zaznaczona symbolem  $\xi$ ) wyjaśnia zmienność czterech zmiennych obserwowalnych (oznaczonych symbolem  $\chi$ ). Między wszystkimi zmiennymi latentnymi zachodzi relacja – jest to relacja związku, zaznaczona zakrzywną strzałką z dwukierunkowym grotem. Każdej zmiennej obserwowalnej przypisany jest błąd pomiaru, oznaczony symbolem  $\sigma$ .

Dla tak zoperacjonalizowanych trzech zmiennych latentnych mamy aż 12 zmiennych obserwowalnych i 12 błędów pomiaru. Model ten ma 51 stopni swobody. Ich liczba wynika z wzoru 1:

$$(1) \quad \begin{aligned} df &= [p(p+1)/2] - t \\ t &= 2p + k \end{aligned}$$

gdzie:  $p$  to liczba zmiennych obserwowalnych,  $t$  to liczba parametrów obliczanych w modelu,  $k$  to liczba powiązań między zmiennymi latentnymi. Zatem:

$$df = [12(13)/2] - [2 \cdot 12 + 3] = 51$$

Stopnie swobody w modelu pomiarowym to uwolnione związki między błędami pomiaru ( $\sigma$ ). Jak widać na rysunku 2, nie przeprowadzono między nimi relacji. Tych wolnych związków dla takiego modelu jak przedstawiony na rysunku 2 jest aż 51. Wraz ze wzrostem liczby zmiennych obserwowalnych przyrasta liczba stopni swobody, ponieważ zostaje uwolnionych coraz więcej



związków między błędami pomiaru. Tym samym, im większa liczba zmiennych obserwowalnych, tym więcej zostaje uwolnionych stopni swobody. Uwolnienie tych związków jest wymogiem formalnym modeli weryfikowanych za pomocą układu równań strukturalnych<sup>5</sup>.

Wraz z liczbą zmiennych obserwowalnych i latentnych w modelach weryfikowanych za pomocą układu równań strukturalnych wzrasta liczba koniecznych do oszacowania parametrów. Są to: związki między zmiennymi latentnymi, związki zmiennych obserwowalnych ze zmiennymi latentnymi, wariancje zmiennych obserwowalnych oraz błędy pomiaru. Tym samym wzrasta złożoność modelu. Złożone modele wymagają bardzo dużych prób. Stąd przemyślenie liczby wskaźników w celu zoperacjonalizowania zmiennej staje się rzeczą kluczową. Jak podaje Aranowska (2005), liczba ta nie powinna być mniejsza niż cztery, ponieważ w przeciwnym razie wątpliwym staje się trafny pomiar zmiennej. Z drugiej strony, duża liczba zmiennych obserwowalnych sprawi, że model będzie bardzo złożony, jego weryfikacja będzie wymagała bardzo dużej próby. Liczba zmiennych obserwowalnych w modelu przekłada się na jego złożoność i, jak widać, realnie kształtuje warunki proceduralne całych badań.

Podczas operacjonalizacji pojęć należy przemyśleć, ile wskaźników jest potrzebnych do trafnego, a z drugiej strony – w miarę prostego odwzorowania pojęć. Warto również się zastanowić nad skalą pomiarową dla wskaźników. Estymator *Maximum Likelihood*, przeznaczony do modeli SEM, działa na zbiorze danych uciąglonych. Format odpowiedzi, jakim jest np. pięciostopniowa skala Likerta, jest raczej dyskretny niż ciągły. Warto stosować formaty dłuższe, np. format od 1 do 7 lub od 1 do 10.

Należy również wspomnieć, że w modelach weryfikowanych za pomocą układów równań strukturalnych niektóre zmienne obserwowalne mogą występować jako samodzielne reprezentanty elementów teorii, np. wiek czy wzrost. Są to jednowymiarowe, ciągłe zmienne liczbowe, które podlegają bezpośredniej obserwacji w przyrodzie. Takie zmienne mogą stanowić samodzielne elementy struktury i występować w modelach wraz z innymi zmiennymi latentnymi. Model, w którym wszystkie zmienne są reprezentowane przez zmienne obserwowalne, zwany analizą ścieżkową (*path analysis*), a przez niektórych autorów włączony do klasy modeli strukturalnych (Konarski, 2009), ma jednak zupełnie inne założenia niż modele równań strukturalnych. Nie dotyczy go bowiem warunek o wielkości lambd (ładunków czynnikowych), ponieważ te w modelach analizy ścieżkowej nie występują.

---

<sup>5</sup> Taki sam model obliczany za pomocą analizy ścieżkowej, gdzie jedna zmienna obserwowalna służy do pomiaru całego elementu struktury (co najczęściej wynika z tego, że zmienna powstaje jako suma poszczególnych pozycji testowych), miałby zero stopni swobody. Wynika to z faktu, że w przypadku analizy ścieżkowej nie są obliczane związki między epsilonami (resztami regresji). Więcej w: Gaul, Machowski, 1987.

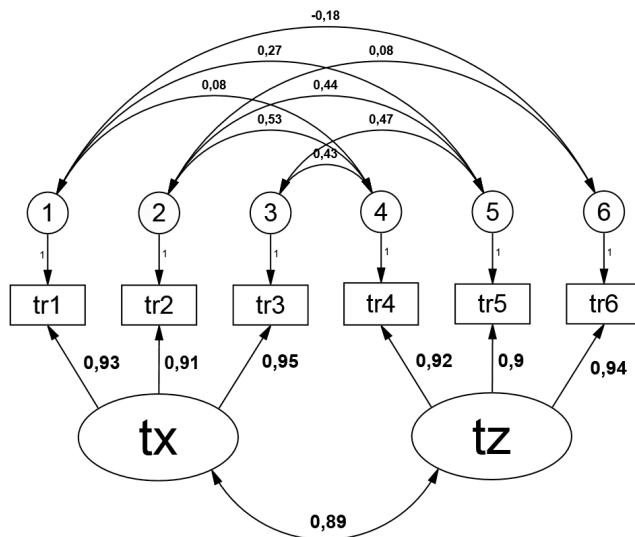
W modelach analizy ścieżkowej nie istnieje również powiązanie między resztami (epsilonami). Dopasowanie modelu analizy ścieżkowej do danych przez sprawdzanie wielkości testu  $\chi^2$  stało się dostępne, odkąd powszechnie pojawiły się programy do weryfikowania modeli za pomocą układów równań strukturalnych. Nieznajomość różnic między tymi dwiema procedurami analizy statystycznej pozwala się łudzić, jak dobre dopasowanie do danych osiągają te modele analizy ścieżkowej, w których dwa lub trzy stopnie swobody powodują najczęściej mały wzrost testu  $\chi^2$ , nie ma bowiem w tych modelach możliwości korelowania reszt (epsilonów). Tymczasem w modelach weryfikowanych za pomocą układów równań strukturalnych liczba stopni swobody zawsze jest większa, gdyż, poza zwolnionymi ścieżkami między zmiennymi latentnymi, uwolnionych jest wielokrotnie więcej związków między błędami pomiaru (zostanie to zaprezentowane dokładnie w dalszej części tekstu).

Miary dopasowania zostały zaproponowane dla modeli SEM m.in. po to, by sprawdzić, czy błędy pomiaru nie są ze sobą silnie skorelowane. W polskim opracowaniu analizy ścieżkowej (Gaul, Machowski, 1987) autorzy nawet nie wspominają o sprawdzaniu dopasowania modeli analizy ścieżkowej za pomocą testu  $\chi^2$ . Nadmienione to zostaje w tym miejscu, gdyż bardzo ważną rzeczą jest świadomość, że model nasycony w analizie ścieżkowej to ten, który ma wszystkie powiązania między zmiennymi obserwowalnymi, podczas gdy ten sam model nasycony w modelach strukturalnych, poza wszystkim powiązaniem między zmiennymi latentnymi, ma dodatkowo powiązania między błędami pomiarowymi. Dlatego modele weryfikowane za pomocą równań strukturalnych mają więcej stopni swobody, a także, najczęściej, przyjmują o wiele wyższe wartości testu  $\chi^2$  niż modele analizy ścieżkowej.

Model pomiarowy, pomimo że ma wszystkie powiązania między zmiennymi latentnymi, nie jest modelem nasyconym (*saturated*), ma bowiem uwolnione wszystkie związki między błędami pomiarowymi. W tym miejscu należy przybliżyć, czym jest model nasycony. Jego przykład został zaprezentowany na rysunku 3.

Przedstawiony wyżej przykładowy model nasycony ma do wyliczenia w sumie 27 parametrów. Są to: wariancje sześciu zmiennych obserwowalnych oraz sześć błędów pomiaru (epsilonów), sześć ładunków czynnikowych ( $\lambda - \lambda$ ), jedna korelacja między zmiennymi latentnymi, osiem korelacji między błędami pomiaru.

Model przedstawiony na rysunku 3 jest równoważny modelowi nasyconemu dlatego, że wylicza wszystkie możliwe parametry. W jego przypadku macierz modelu nasyconego w ogóle nie odbiega od macierzy modelu obliczanego pomiarowego, jako że zawiera on wszystkie możliwe parametry. Ponieważ nie uwolniono żadnych stopni swobody, obliczany model ma zero stopni swobody,  $df = 0$ , tyle samo, ile model nasycony (patrz tabela 1). Nie ma różnic między macierzą



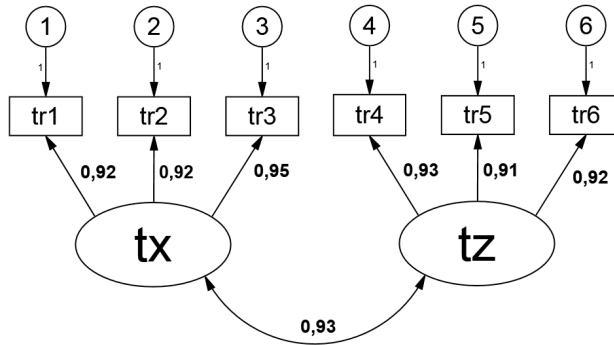
Rysunek 3. Model pomiarowy nasycony.

Tabela 1

Wartość testu  $\chi^2$  oraz liczba stopni swobody dla modelu zaprezentowanego na rysunku 3

| Model            | Liczba obliczanych parametrów | $\chi^2$ | $df$ | $p$   | $\chi^2/df$ |
|------------------|-------------------------------|----------|------|-------|-------------|
| Model obliczany  | 27                            | 0,000    | 0    |       |             |
| Model nasycony   | 27                            | 0,000    | 0    |       |             |
| Model niezależny | 12                            | 3048,738 | 15   | 0,000 | 203,249     |

Model przedstawiony na rysunku 4 jest klasycznym modelem pomiarowym, który ma już uwolnione stopnie swobody. Należy zauważyć, że uwolnionych zostało osiem związków między błędami pomiaru – na rysunku 4 nie ma strzałek z dwustronnym grotem między błędami pomiaru. Model ten ma 19 parametrów do obliczenia, tj.  $27 - 8 = 19$ .



Rysunek 4. Model pomiarowy.

Ponieważ model ten ma uwolnione stopnie swobody, nie jest to już model równoważny modelowi nasycionemu. Stopnie swobody (por. tabela 2) stanowią wolne od oszacowania parametry, czyli związki między epsilonami (błędami pomiaru, które nie podlegają wyliczaniu). Uwolnione stopnie swobody sprawiły, że wartość statystyki  $\chi^2$  wzrosła w stosunku do modelu nasycionego (dla którego wynosiła  $\chi^2 = 0$ ) do  $\chi^2 = 52,721$  (por. tabela 2, model obliczany). Model jest niedopasowany do danych.

Tabela 2

Wartość statystyki  $\chi^2$  oraz liczba stopni swobody dla modelu zaprezentowanego na rysunku 4

| Model            | Liczba obliczanych parametrów | $\chi^2$ | $df$ | $p$   | $\chi^2/df$ |
|------------------|-------------------------------|----------|------|-------|-------------|
| Model obliczany  | 19                            | 52,721   | 8    | 0,000 | 6,590       |
| Model nasyciony  | 27                            | 0,000    | 0    |       |             |
| Model niezależny | 12                            | 3048,738 | 15   | 0,000 | 203,249     |

Im większa liczba zmiennych obserwowalnych, tym więcej jest uwolnionych związków w modelu, wzrasta liczba stopni swobody, a tym samym wzrasta wartość testu  $\chi^2$ . Dla modeli należy się więc spodziewać wzrostu wartości  $\chi^2$ , odpowiednio do liczby zmiennych obserwowalnych. Przykładowo: dla modelu, w którym znalazłyby się 53 zmienne obserwowalne przyporządkowane do siedmiu zmiennych latentnych, wartość statystyki  $\chi^2$  mogłaby wynosić nawet 3260, gdyż model miałby 1304 stopnie swobody ( $df = 1304$ ). Wartość testu  $\chi^2$  będzie istotna statystycznie. Aby skorygować to przeszacowanie, wartość  $\chi^2$  dzieli się przez liczbę stopni swobody. Jeżeli wartość  $\chi^2/df$  nie przekracza wartości krytycznej, jaką jest 2,5, można uznać, że model dobrze pasuje do danych (Bartholomew, Steele, Moustaki, Galbraith, 2008; Hair i in.,

2006; Konarski, 2009). W naszym przykładzie  $3260/1304 = 2,5$ . W tabeli 2 wartość  $\text{Chi}^2/df$  wyniosła 6,590, dlatego zaprezentowany na rysunku 4 model jest niedopasowany do danych.

### WARUNKI POPRAWNOŚCI MODELU POMIAROWEGO

Jak już zostało wspomniane, model pomiarowy jest pierwszym etapem weryfikacji modelu za pomocą układu równań strukturalnych. Idea budowy każdego modelu pomiarowego – niezależnie od jego hierarchiczności czy jednopoziomowości itp. – polega na sprawdzeniu, czy zmienne latentne zostały dobrze zoperacjonalizowane. Aby tego dokonać, należy sprawdzić, czy zostały spełnione warunki poprawności modelu pomiarowego. Podane poniżej założenia formalne są uniwersalne dla każdego modelu pomiarowego.

Pierwszym warunkiem jest wielowymiarowa normalność zmiennych. Koniecznie trzeba w tym miejscu wspomnieć o wymogach danych dla modeli weryfikowanych za pomocą układów równań strukturalnych. Kwestia ta musi być przemyślana jeszcze przed przystąpieniem do badań, niniejszy opis znajduje się jednak w tym miejscu ze względu na silne powiązanie warunku wielowymiarowej normalności z danymi. Jest to warunek konieczny, aby można było zastosować większość estymatorów. Dopasowanie modelu dokonywane jest przez wykorzystanie jednej z metod dopasowania, w zależności od rodzaju danych: ML (*Maximum Likelihood*), GLS (*Generalized Least Squares*), ULS (*Unweighted Least Squares*), *Scale-Free Least Squares*, ADF (*Asymptotically Distribution Free*). Metoda największej wiarygodności (ML) wykorzystywana jest jedynie dla zmiennych ciągłych. Zbiór danych musi więc być wystarczająco duży, by można było określić, czy rozkłady nie odbiegają od normalnych, a dane powinny się wyrażać na skalach ciągłych. Jeżeli nie zostało pogwałcone założenie o wielowymiarowej normalności, a zmienne wyrażone są na skalach ciągłych, to można stosować estymator ML. Jeżeli natomiast założenie normalności zostało złamane, to należy zastosować jeden z estymatorów z rodziny *Asymptotically Distribution Free* (ADF).

Istnieje kilka estymatorów z rodziny ADF. Szczególnie godnym polecenia jest estymator WLSMV, który służy do szacowania modeli wówczas, gdy dane nie są ciągłe, a rozkłady odbiegają od normalnych. Estymator oprogramowany jest w pakiecie „lavaan” programu R+, obok wielu innych estymatorów odpornych<sup>6</sup>.

Silny warunek jakości modelu pomiarowego stanowi wysokość ładunków czynnikowych ( $\text{lambd} - \lambda$ ). Lambdy wyrażają związek zmiennych obserwowalnych ze zmienną latentną. Wartość ładunków czynnikowych ( $\text{lambd}$ ) nie powinna być mniejsza niż 0,7; w przeciwnym razie ich wartość podniesiona do

<sup>6</sup> Liczba estymatorów i możliwości ich wyboru to temat tak szeroki, że nadaje się na całkowicie osobny tekst. Tu jest jedynie wspomniany.

kwadratu daje wartość mniejszą niż 0,5 ( $0,7^2 = 0,49$ ), a to oznacza, że zmienność danej zmiennej obserwowalnej jest objaśniana w mniej niż 50% przez zmienną latentną.

Założenie to dotyczy pośrednio także błędów pomiaru. Im mniejsza wartość ładunków czynnikowych ( $\lambda$ ), tym większa wariancja błędów. Wraz ze wzrostem wariancji błędów wzrasta prawdopodobieństwo wystąpienia istotnych związków między błędami pomiaru i niedopasowania modelu. Dzieje się tak dlatego, że duża część zmienności całkowitej znajduje się w zmienności błędów. Doświadczenie pokazuje, że wzrasta wtedy prawdopodobieństwo odrzucenia modelu. Przyjęcie modelu, w którym ładunki czynnikowe zmiennych latentnych są niskie (co jest możliwe, jeżeli związki między błędami pomiaru okażą się nieistotne i niewysokie), stanowi i tak problem, ponieważ wątpliwa staje się wówczas rzetelność zmiennych latentnych. Niespełniony zostaje zatem kolejny warunek modelu pomiarowego.

Drugi warunek dotyczy konieczności obliczenia rzetelności zmiennych latentnych. Dokonuje się tego według wzoru 2:

$$(2) \quad CR = (\sum \lambda_i)^2 / [(\sum \lambda_i)^2 + (\sum \sigma_i)],$$

gdzie CR to skrót terminu *construct reliability* (rzetelność konstruktów), a  $\sigma_i^2$  jest wariancją błędu pomiaru.

Podobnie jak w przypadku rzetelności obliczanych dla skal psychometrycznych, wartości  $0,6 < CR < 0,7$  przyjmuje się za wartości akceptowalne, a wartości  $CR > 0,7$  ocenia się jako dobre (Hair i in., 2006). Można zauważyć, że im niższa wartość ładunków czynnikowych, a wyższa wariancja błędu, tym niższa rzetelność. Zmienna latentna jest rzetelna wówczas, gdy wartości  $\lambda$  przekraczają wartość 0,7. Warto również nadmienić, że w przeciwieństwie do miary zgodności wewnętrznej, jaką jest  $\alpha$ -Cronbacha, wprowadzanie do zmiennej latentnej pozycji o słabych ładunkach czynnikowych raczej nie podniesie rzetelności CR, tylko ją osłabi. Pomiar musi być dobrze przemyślany już na początku, w przeciwnym razie modele będą niedopasowane lub miary rzetelności zmiennych bardzo marne.

Podczas obliczania modelu pomiarowego należy również obliczyć wariancję zmiennych latentnych (*variance extracted*, VE) – jest to trzeci warunek modeli. Dokonuje się tego według wzoru 3:

$$(3) \quad VE = \sum \lambda_i^2 / n = \sum (\lambda_i - 0)^2 / n,$$

gdzie  $i$  jest numerem zmiennej obserwowalnej, będącej wskaźnikiem zmiennej latentnej,  $n$  jest liczbą zmiennych obserwowalnych.

Jak widać, jest to obciążony estymator wariancji wyjaśnionej, a zatem wartości  $\lambda_i$  muszą być duże, aby można było traktować je jako wyjaśnienie zmienności konkretnej zmiennej latentnej. VE powinien osiągnąć wartość bliską 0,5, aby można było uznać, że spełniony został warunek modeli. Wartość ta zbliży się do 0,5, gdy ładunki czynnikowe ( $\lambda$ ) będą bliskie wartości 0,7.

Czwarty warunek dotyczy sprawdzenia trafności różnicowej zmiennych latentnych. Zmienna latentna nie powinna być związana z inną zmienną latentną bardziej niż ze swoimi zmiennymi obserwowalnymi. Innymi słowy, zmienna latentna nie może objaśniać więcej zmienności drugiej zmiennej latentnej niż zmienności swoich zmiennych obserwowalnych. W przypadku modelu pomiarowego (ale również modelu obliczanego za pomocą układu równań strukturalnych) wymaga się również, aby zmienne obserwowalne (wskaźnikowe) były objaśniane przez swoją zmienną latentną, a zarazem by nie były objaśniane przez inne zmienne latentne. Chodzi więc o to, żeby nie występowała krzyżowość, czyli częściowa identyczność treści składających się na definicje różnych konstruktów.

Jeżeli wszystkie te warunki są spełnione, sprawdza się, czy model dobrze pasuje do danych. Jest to piąty i ostatni warunek modeli. Czyni się to przez oszacowanie miar dopasowania, takich jak np.  $\chi^2$ , RMSEA, CFI, a także  $\chi^2/df$ . Test  $\chi^2$  weryfikuje hipotezę zerową, mówiącą o tym, że model obliczany (*default model*) nie różni się od modelu nasyconego<sup>7</sup> (*saturated model*), wartość statystyki  $\chi^2$  powinna być zatem jak najmniejsza, a przy tym nieistotna statystycznie. W rzeczywistości dla modeli bardzo złożonych założenie to jest nierealne.

Jeżeli model pomiarowy nie jest dobrze dopasowany do danych, to kolejne uwolnienie stopni swobody, jakie następuje podczas obliczania modelu za pomocą układu równań strukturalnych, jest pozbawione sensu, gdyż może jedynie pogorszyć jego dopasowanie. Przy źle dopasowanym modelu pomiarowym trzeba przemyśleć budowę zmiennych latentnych i poprawić model pomiarowy. Nie przeprowadza się wówczas obliczania pełnego modelu SEM.

## MODEL SEM

Jeżeli model pomiarowy okazuje się dobrze dopasowany do danych empirycznych, to przechodzi się do weryfikowania modelu opisującego teorię za pomocą układu równań (model SEM, *Structural Equation Model*). W tym modelu, między zmiennymi latentnymi pozostawia się jedynie ścieżki przewidziane na poziomie teoretycznym. W modelu przedstawionym na rysunku 2 są to więc związki: R2.1, R3.2, R3.4, R4.1, R5.3, R5.6, R6.3.

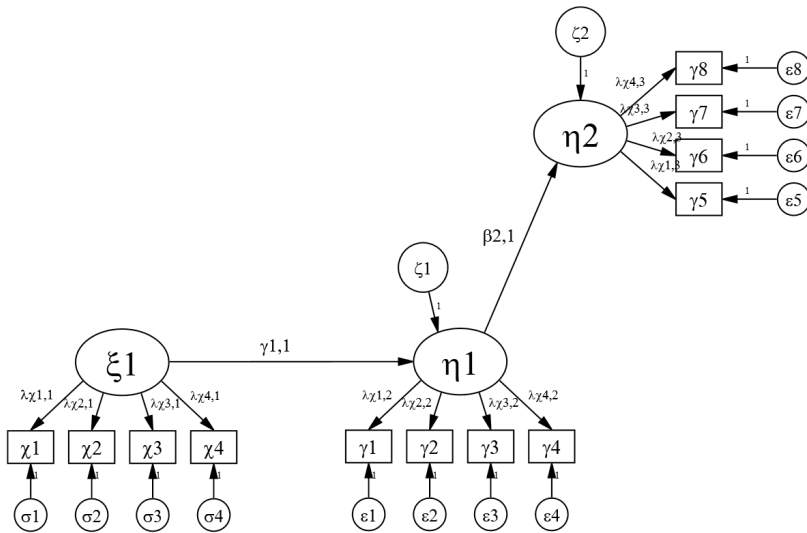
Jak widać na rysunku 1, w modelu opisującym strukturę nie pojawiają się wszystkie możliwe powiązania między zmiennymi, lecz siedem spośród możliwych 21 powiązań. Zostaje więc zwolnionych 14 związków między zmiennymi. Procedura uwalniania związków (stopni swobody) przyczyni się ponownie do wzrostu wartości  $\chi^2$  (bądź w bardzo rzadkich przypadkach nie ulegnie on

---

<sup>7</sup> Model nasycony (*saturated*) to model, który ma obliczane wszystkie możliwe parametry. Wartość testu  $\chi^2$  dla takiego modelu wynosi 0.

zmianie). Rysunek 5 przedstawia graf modelu strukturalnego, który reprezentuje teorię strukturalną i najczęściej jest przedstawiany w postaci diagramu.

To, czy zmienne latentne są endogeniczne (zależne) czy egzogeniczne (niezależne), zależy od weryfikowanej teorii strukturalnej. Zmienne endogeniczne oznaczane są symbolem  $\eta$  i towarzyszy im błąd (oznaczany symbolem  $\zeta$ ), zmienne egzogeniczne natomiast, czyli zmienne niezależne, oznaczane są symbolem  $\xi$  i nie definiuje się dla nich błędów (por. rysunek 5). Zależności między zmiennymi endogenicznymi a egzogenicznymi są oznaczone symbolem  $\gamma$ , a zależności między zmiennymi endogenicznymi – symbolem  $\beta$ . Zmienne obserwowalne dla zmiennych endogenicznych są oznaczane symbolem  $\gamma$ , a ich błędy – symbolem  $\epsilon$ , z kolei zmienne obserwowalne dla zmiennych egzogenicznych oznaczane są symbolem  $\chi$ , a ich błędy – symbolem  $\sigma$ . Model wizualny nazywa się *diagramem ścieżkowym* (*path diagram*).



Rysunek 5. Przykładowy diagram ścieżkowy dla modelu strukturalnego, gdzie:  $\lambda$  jest ładunkiem czynnikowym,  $\sigma$  i  $\epsilon$  są wariancjami błędów,  $\chi$  symbolizuje egzogeniczne zmienne obserwowalne,  $\gamma$  symbolizuje endogeniczne zmienne obserwowalne,  $\xi$  symbolizuje egzogeniczny konstrukt latentny,  $\eta$  symbolizuje endogeniczny konstrukt latentny,  $\gamma$  symbolizuje związek między konstruktem egzogenicznym a endogenicznym,  $\beta$  symbolizuje związek między konstruktem endogenicznymi.

Układ równań strukturalnych ma następującą postać:

$$(3.5.) \quad B\eta = \Gamma\xi + \zeta,$$

gdzie  $B$  to kwadratowa macierz współczynników stopnia  $[m \times m]$ ,  $\Gamma$  to prostokątna macierz współczynników stopnia  $[m \times n]$ , a  $\zeta$  to losowy wektor reszt (por. Aranowska, 2005). Równania strukturalne mogą być również zapisane



w postaci szczegółowych macierzy. Dla modelu strukturalnego przedstawionego na rysunku 5 macierz równań strukturalnych ma następującą postać:

$$\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} \gamma_{1,1} & 0 \\ 0 & \beta_{2,1} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \eta_1 \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}$$

Dla modelu pomiarowego macierze mają postać następującego układu równań:

1. Dla zmiennej egzogenicznej

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix} = \begin{bmatrix} \lambda_{x1,1} \\ \lambda_{x2,1} \\ \lambda_{x3,1} \\ \lambda_{x4,1} \end{bmatrix} \xi_1 + \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \end{bmatrix}$$

2. Dla zmiennych endogenicznych

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \\ Y_7 \\ Y_8 \end{bmatrix} = \begin{bmatrix} \lambda_{y1,2} & 0 \\ \lambda_{y2,2} & 0 \\ \lambda_{y3,2} & 0 \\ \lambda_{y4,2} & 0 \\ 0 & \lambda_{y1,3} \\ 0 & \lambda_{y2,3} \\ 0 & \lambda_{y3,3} \\ 0 & \lambda_{y4,3} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \end{bmatrix}$$

W testowaniu modelu strukturalnego kluczową rolę odgrywają dwa czynniki: ogólne dopasowanie modelu do danych (czyli do „macierzy wejściowej”) oraz wielkość, kierunek i istotność poszczególnych parametrów. Gdy istnieje zadowalający stopień dopasowania i gdy parametry wykazują istotność statystyczną, model uznaje się za dopasowany.

Aby skorygować przeszacowaną – z powodu braku odporności testu na dużą liczbę stopni swobody – wartość  $\text{Chi}^2$ , stosuje się inne miary, których zadaniem jest zweryfikowanie dopasowania modelu. Wśród nich dwiema najczęściej występującymi są statystyka CFI (wzór 1) oraz RMSEA (wzór 2). W obu wzorach test  $\text{Chi}^2$  jest wykorzystany do obliczenia wartości miary dopasowania.

RMSEA (*Root Mean Square Error of Approximation*) bazuje na rozkładzie zmiennej w populacji, dlatego lepiej niż inne statystyki oddaje to, na ile model jest reprezentatywny dla danej populacji. Jak wynika z wzoru 1, RMSEA jest miarą szacującą błąd, jest więc wielkością niedopasowania. Im niższe wartości przyjmuje statystyka RMSEA, tym lepsze dopasowanie modelu. Pożądane wartości to  $\text{RMSEA} < 0,01$ , w praktyce jednak najczęściej otrzymywane są

wartości  $0,02 < RMSEA < 0,08$ , które interpretuje się jako wskazujące na dobre dopasowanie modelu (Hair i in., 2006; Konarski, 2009).

$$(1) RMSEA = \sqrt{\max\left(\frac{Chi2/(N-1)}{df} - \frac{1}{N-1}\right)},$$

gdzie:  $Chi^2$  to wartość statystyki dla modelu wyliczanego,  $N$  – wielkość badanej próby,  $df$  – stopnie swobody wyliczanego modelu.

Postać formalna statystyki CFI przedstawiona została we wzorze 2. Przyjmuje ona wartości od 0 do 1. Wartości  $CFI > 0,9$  wskazują na dopasowanie modelu, a wartości  $CFI < 0,9$  uznaje się za wskazujące na niedopasowanie modelu.

$$(2) CFI = 1 - \frac{(Chi2-df)}{(Chi2n-dfn)},$$

gdzie:  $Chi^2_n$  to wartość statystyki dla modelu niezależnego,  $df_n$  to stopnie swobody modelu niezależnego.

RMSEA oraz CFI wskazują absolutnie poprawne dopasowanie obliczanego modelu do danych wówczas, kiedy  $Chi^2$  osiąga wartość równą 0. Jak widać z wzoru 1, kiedy  $Chi^2 = 0$ , wówczas statystyka RMSEA również osiąga wartość równą 0, a z wzoru 2 wynika, że statystyka CFI osiąga wówczas wartość równą 1.

Chociaż przyjęło się powyższe miary traktować jako odpowiednie do określenia dopasowania modelu, to należy pamiętać, że zarzuca się im nieujawnianie tego, jak bardzo model jest źle wyspecyfikowany (Saris, Satorra, van der Veld, 2009). Błąd złej specyfikacji polega na nietrafnym odtworzeniu prawdziwej wielkości szacowanego parametru.

### MODELE ALTERNATYWNE

Żaden model, nawet bardzo dobrze dopasowany do danych, nie jest jedynym modelem deskrybującym zjawisko. Najczęściej istnieje kilka możliwych modeli badanej teorii (Jonkisz, 1998). Rekonstruowanie tych alternatywnych modeli jest rzeczą wielce ważną.

Badacz powinien zatem stworzyć modele alternatywne. Jeżeli istnieje przypuszczenie, że analizowany proces może zostać objaśniony przez inne zmienne niż te, które są umieszczone w jednym modelu, wówczas modele alternatywne stanowią bardzo wartościowe rozwiązanie. Porównanie modeli alternatywnych i wybranie najlepszego jest opisywane w literaturze (Hair i in., 2006). Polega ono ogólnie na obliczeniu statystyki dopasowania modeli alternatywnych oraz wybraniu tego modelu, który ma najlepszy indeks dopasowania (spośród dwóch równie złożonych modeli). Drugą zasadą jest oszczędność. Przyjmuje się, że najlepszym jest ten model, który równie dobrze, jak inne modele, tłumaczy analizowane zjawisko, a równocześnie jest najbardziej oszczędny, tzn. najmniej złożony.

Nie należy jednak mylić budowy modeli alternatywnych wobec modeli rekonstruujących teorię z alternatywnymi modelami empirycznymi, zwanymi żartobliwie modelami *wykopanymi z danych*. Modele empiryczne mają zupełnie inny (niższy) status metodologiczny niż modele rekonstruujące teorię i nie są dla modeli rekonstruujących teorię żadną alternatywą. Jak podaje Gajda:

Modyfikowanie teorii na podstawie estymacji „może dać model, który pasuje do danych, a i z innych punktów widzenia wygląda dobrze, podczas gdy w gruncie rzeczy jest to marna aproksymacja prawdziwej struktury. Jeśli tylko szukać dość uporczywie, zwykle uda się zbudować coś, co wygląda na dobry model [...]. Poszukiwanie modelu w taki sposób nazywane jest czasem «wykopywaniem [modelu – przyp. JBG] z danych» (*data mining*), a czasem «poszukiwaniem specyfikacji» (*specification searches*), zależnie od nastroju [...]”. (1992, ss. 107–108)

Dla modeli rekonstruujących teorię jedyną alternatywą są inne modele, również rekonstruujące tę teorię. Modele empiryczne mogą być alternatywą jedynie dla innych modeli empirycznych. Nie mogą one weryfikować teorii, ponieważ powstały z tego, na co pozwalały dane. Aby modele empiryczne mogły mieć status modeli rekonstruujących teorię, należy najpierw na ich podstawie tę teorię opisać i na innej próbie przeprowadzić weryfikację (Hair i in., 2006). Wykorzystane w ten sposób, mogą do rozwoju teorii walenie się przyczynić. Jak podaje w innym miejscu Gajda (1992, s. 108):

Na szczęście istnieje sposób sprawdzenia, czy ktoś niewłaściwie „przekopał” dane, a są to testy [działania modelu – przyp. JBG] poza obrębem próby. Jeśli model jest marnie wyspecyfikowany, to nie powinien dobrze pasować do danych statystycznych niewchodzących do próby, jeżeli nawet dopasowanie wewnątrz próby wydaje się dobre (Fair, 1972, za: Gajda, 1992, s. 19).

Rekonstrukcja alternatywnych modeli empirycznych jest tym niemniej również dużą umiejętnością. Rekonstrukcję modelu empirycznego można rozpocząć od przeprowadzenia eksploracyjnej analizy czynnikowej, aby zdecydować, które pozycje wysycają jakie czynniki – może to ułatwić decyzję, które wskaźniki powinny definiować zmienną latentną. Kolejno uwalnia się pojedyncze ścieżki modelu, aby stwierdzić, czy ich uwolnienie nie pogarsza dopasowania modelu lub czy związki między pozostałymi zmiennymi w modelu nie ulegają zmianie. Może się tak dzieć, co wynika z prawa współskorelowania zmiennych (King, Minium, 2009). Niektóre programy mają funkcję umożliwiającą sprawdzanie, które związki w modelu są istotne statystycznie i powinny zostać w modelu<sup>8</sup>. Pokazują one nie tylko najsilniejsze związki między zmiennymi latentnymi,

<sup>8</sup> Np. funkcja Modification Indexes programu AMOS.

lecz także między błędami pomiaru. Niestety, nadużywanie tej funkcji prowadzi do tego, że bezmyślnie powstają modele, które nie mają już nic wspólnego ze swoim pierwotnym przeznaczeniem, jakim jest weryfikowanie założeń teoretycznych, i należy podkreślić, jeszcze raz mocno i zdecydowanie, że status metodologiczny tych modeli nie jest równy statusowi modeli weryfikujących założenia teoretyczne.

### PODSUMOWANIE

Opisane tu założenia formalne są obligatoryjne podczas obliczania modeli SEM. Należy pamiętać, że uwolnienie związków nieprzewidzianych na poziomie teoretycznym jest podstawową ideą modelowania. Pozostawienie tych związków zwiększa co prawda prawdopodobieństwo tego, że model będzie lepiej dopasowany do danych, ale uderza w podstawową ideę modelowania. Po pierwsze, dlatego, że modele, w których pozostawia się wszystkie związki, są równoważne modelom nasyconym, są więc w sposób naturalny dopasowane, co nie wynika z poprawności teorii, ale z faktu, iż nie uwolniono stopni swobody. Po drugie, może to wynikać z błędnej teorii – nie może być poprawnym założenie, że wszystko jest powiązane ze wszystkim. Taka teoria niczego nie wyjaśnia, co stoi w sprzeczności z głównym jej zadaniem, jakim jest deskrybowanie, wyjaśnianie i prognozowanie. Po trzecie, modele, które nie mają uwolnionych stopni swobody, są niezwykle skomplikowane; trudno jest zinterpretować związki, jeżeli wszystko jest powiązane ze wszystkim.

Niekiedy zdarza się, że badacze pozostawiają związki między błędami pomiaru. Takie rozwiązanie zawsze budzi duże kontrowersje. Jak już zostało napisane, uwolnienie tych związków jest bardzo silnym założeniem modeli SEM. Metodolodzy podkreślają, że można to robić jedynie wówczas, gdy zostanie przedstawione teoretyczne wyjaśnienie, dlaczego związek ten zostaje wyliczony (Hair i in., 2006); jeżeli takiego wyjaśnienia nie ma, związków tych zostawiać nie wolno. W przeciwnym razie przyczynia się to do przyjęcia modelu niepoprawnego, lub inaczej: zmniejsza to prawdopodobieństwo odrzucenia modelu błędnego. Innymi słowy, podnosi to prawdopodobieństwo popełnienia błędu drugiego rodzaju.

Analogicznie do uwolnienia parametrów w modelu wygląda sytuacja z doborem wielkości próby. Im bardziej złożony model, tym większa jest wymaga próba do jego weryfikacji. Jak podaje Aranowska: „zbyt mała próba nie umożliwia odrzucania modeli nietrafnych. Stosowanie małych prób do modeli równań strukturalnych jest niedopuszczalne z powodów formalnych; teoria statystyki stosowana do analizy macierzy kowariancji była do niedawna teorią rozwijaną wyłącznie dla dużych prób” (1996, s. 161).

Kwestią często podnoszoną w wielu dyskusjach jest normalność rozkładu zmiennych w modelach SEM. Jak podaje Aranowska: „Rozkłady normalne lub

wystarczająco normalne powodują homoscedastyczność wariancji regresji cechy Y względem X” (Aranowska, 1996). Dla dwóch zmiennych wystarczy więc, żeby rozkłady były zbliżone do rozkładów normalnych (wówczas *robust theory* dopuszcza możliwość użycia modeli SEM). Odpowiednia transformacja danych może się przyczynić do zsymetryzowania rozkładu zmiennej, ale może również zaburzyć relacje między pozostałymi zmiennymi w modelu (Meijerink, 1996). Stąd zalecenie, aby nie transformować zmiennych egzogenicznych modelu, a jedynie endogeniczne (Konarski, 2009, s. 95).

W końcu należy pamiętać, że związki między strukturami latentnymi w modelu pomiarowym – gdzie każda zmienna latentna jest powiązana ze wszystkimi innymi – mogą być związkami pozornymi. Wynika to ze współskorelowania zmiennych (Ferguson, Takane, 2002; King, Miniun, 2009). Aby ustrzec się przed interpretacją sieci wzajemnych powiązań między zmiennymi, uwalnia się część związków (zgodnie z założeniami teorii), co chroni przed interpretacją wielu (często pozornych) związków.

Jeżeli obliczany jest model, który zakłada wzajemne powiązania więcej niż dwóch zmiennych – podobnie jak w modelu zaprezentowanym na rysunku 1, gdzie zmienne X3, X5 i X6 są wzajemnie powiązane – to rekomenduje się zwalnianie poszczególnych związków, tak aby pokazać, w jaki sposób zmienna (w naszym przypadku X3) moderuje związek między zmiennymi X5 i X6.

Podkreślenia wymaga na koniec problem odrzucania modeli SEM po dokonaniu ich empirycznej weryfikacji. Jeżeli model okazał się niedopasowany do danych, to należy go odrzucić jako niepoprawny. Jeżeli jednak model okazał się dopasowany do danych, wówczas nie formułuje się zdania o „przyjęciu modelu”. Modele tych nie przyjmujemy, gdyż u ich podstaw weryfikowana jest hipoteza zerowa, która mówi, że model teoretyczny nie różni się od modelu empirycznego, a hipotezy zerowej nigdy się nie przyjmuje. Co więcej, w przypadku modeli SEM hipoteza zerowa jest zawsze fałszywa (Konarski, 2009). Każdy bowiem model teoretyczny stanowi tylko pewne przybliżenie rzeczywistości, a nie jej dokładny opis, nie można więc takiego modelu przyjmować. Oznaczałoby to stwierdzenie, że model nie różni się od rzeczywistości, a to byłoby nieprawdą. Można jedynie stwierdzać, że nie ma podstaw do odrzucenia modelu jako niepoprawnego.

## BIBLIOGRAFIA

- Aranowska, E. (1996). *Metodologiczne problemy zastosowań modeli statystycznych w psychologii. Teoria i praktyka*. Warszawa: Studio 1.
- Aranowska, E. (2005). *Pomiar ilościowy w psychologii [Quantitative measurements in psychology]*. Warszawa: Wydawnictwo Naukowe SCHOLAR.
- Bartholomew, D. J., Steele, F., Moustaki, I., Galbraith, J. I. (2008). *Analysis of multivariate social science data*. Boca Raton, FL: Chapman & Hall/CRC Press.

- Ferguson, G. A., Takane, Y. (2002). *Analiza statystyczna w psychologii i pedagogice*. Warszawa: Wydawnictwo Naukowe PWN.
- Figurska, E. (1993). Wybrana metoda sprawdzania poprawności modelu na przykładzie implementacji modelu pamięci semantycznej w Prologu. W: E. Aranowska (red.), *Psychologia matematyczna* (tom V, s. 155–169). Uniwersytet Jagielloński, Instytut Psychologii.
- Gajda, J. (1992). Modele strukturalne w naukach społecznych. W: E. Aranowska (red.), *Wybrane problemy metodologii badań* (s. 100–132). Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Garson, G. D. (2013). Introductory guide to HLM with HLM 7 software. W: *Hierarchical linear modeling: Guide and applications* (s. 55–96). Thousand Oaks, CA: Sage.
- Gaul, M., Machowski, A. (1987). Elementy analizy ścieżek. W: J. Brzeziński (red.), *Wielozmiennowe modele statystyczne w badaniach psychologicznych*. Warszawa–Poznań: Państwowe Wydawnictwo Naukowe.
- Gospodarek, T. (2009). Modelowanie w naukach o zarządzaniu oparte na metodzie programów badawczych i formalizmie reprezentatywnym. Wrocław: Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu. Pobrane z: <http://gospodarek.eu/publikacje/Modelowanie%20w%20naukach%20o%20zarzadzaniu.pdf>
- Hair, J. J., Black, W. C., Babin, B. J., Anderson, R. E., Tatham, R. L. (2006). *Multivariate data analysis*. New Jersey: Upper Saddle River.
- Heck, R. H., Thomas, S. L. (2009). *Introduction to multilevel modeling techniques*. New York: Routledge.
- Heck, R. H., Thomas, S. L., Tabata, L. N. (2010). *Multilevel longitudinal modeling with IBM SPSS*. New York, London: Routledge, Taylor & Francis Group.
- Jonkisz, A. (1998). *Ciągłość teoretycznych wytworów nauki. Ujęcie strukturalne*. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.
- King, B. M., Minium, E. W. (2009). *Statystyka dla psychologów i pedagogów*. Warszawa: Wydawnictwo Naukowe PWN.
- Konarski, R. (2009). *Modele równań strukturalnych*. Warszawa: Wydawnictwo Naukowe PWN.
- Meijerink, F. (1996). *A nonlinear structural relations model*. Leiden: DSWO Press, Leiden University.
- Milfont, T. L., Duckitt, J. (2004). The structure of environmental attitudes: A first- and second-order confirmatory factor analysis. *Journal of Environmental Psychology*, 24, 289–303.
- Nowak, S. (2007). *Metodologia badań społecznych*. Warszawa: Wydawnictwo Naukowe PWN.
- Rabe-Hesketh, S., Skrondal, A., Zheng, X. (2007). *Multilevel Structural Equation Modeling*, 1, 209–227.

- Raudenbush, S., Bryk, A., Cheong, Y. F., Congdon, R., Toit, M. (2011). *HLM7 Hierarchical Linear & Nonlinear Modeling*. Lincolnwood: Scientific Software International, Inc.
- Saris, W. E., Satorra, A., van der Veld, W. M. (2009). Testing Structural Equation Models or Detection of Misspecifications? *Structural Equation Modeling: A Multidisciplinary Journal*, 16(4), 561–582.
- Suppes, P. (1972). *Axiomatic Set Theory*. New York: Dover Publications.
- Tranmer, M., Elliot, M. (2007). *Multilevel Modelling Coursebook*. Manchester: Cathie Marsh Centre for Census and Survey Research.