

# Zagadnienie równoważności pomiarowej w badaniach różnic w postrzeganiu demokracji w UE.

Agnieszka Choczyńska

May 21, 2020

## **Abstract**

The text of your abstract. 200 or fewer words.

*Keywords:* measurement invariance, factor analysis, democracy, Europe

# Wprowadzenie

Słowo “demokracja” jest odmieniane przez wszystkie przypadki w deklaracjach polityków, na demonstracjach, w mediach i w codziennych rozmowach. Niejednokrotnie przedmiot tych wypowiedzi wychodzi daleko poza zwięzłą, przyjętą przez politologów definicję, tj. demokracja to system, w którym obywatele mogą przyjąć lub odrzucić elity mające nimi rządzić (Schumpeter 1942). Rozdźwięk między formalną definicją a społecznym postrzeganiem zjawiska nie jest niczym niezwykłym, jednak w tym przypadku nabiera szczególnego znaczenia. Choć Ziemia nie przestanie krążyć wokół Słońca, nawet gdyby cała ludzkość uważała inaczej, to kształt systemów demokratycznych zależy od tego, jak ich obywatele wyobrażają sobie demokrację.

Problem ma szczególne znaczenie w badaniach naukowych, w których dokonywana jest ocena demokratyczności państw lub nastawienia do demokracji grup czy całych społeczeństw, aby następnie porównywać je między sobą. Jest to zawsze odniesienie do jednej, eksperckiej koncepcji demokracji, więc nie pozwala na pełne zrozumienie wyborów ludzi, którzy kierują się funkcjonującymi w ich społecznościach koncepcjami, niekoniecznie “poprawnymi” i niekoniecznie ponadnarodowymi.

Dlatego jest tak ważne, żeby wciąż badać, czym demokracja w społecznym przekonaniu tak naprawdę jest. Ta praca stawia sobie na cel eksplorację przekonań mieszkańców Europy na temat istoty demokracji, wyrażonych w odpowiedziach na ankietę World Values Survey. Na tej podstawie zostanie podjęta próba skonstruowania społecznych koncepcji demokracji, pozwalających wyrazić liczbowo siłę tych przekonań.

Struktura pracy jest następująca: w pierwszym rozdziale przedstawiono pokrótce dotychczasowe sposoby definiowania i mierzenia demokracji, zarówno na podstawie wiedzy eksperckiej jak i dzięki analizie danych zebranych w badaniach społecznych. W rozdziale drugim omówiono model analizy czynnikowej, który zostanie wykorzystany w badaniu, oraz zagadnienie równoważności pomiarowej, która jest warunkiem przeprowadzania analizy międzygrupowej. W następnych rozdziałach opisano kolejne etapy badania wraz z wynikami i wnioskami z analizy.

# 1 Ilościowe badania nad demokracją w literaturze przedmiotu

## 1.1 Motywacje ilościowego badania demokracji

Potrzeba ilościowego wyrażenia stopnia zdemokratyzowania państwa narodziła się w drugiej połowie XX wieku w Stanach Zjednoczonych Ameryki. Według panującego wówczas paradygmatu, państwa demokratyczne charakteryzują się mniejszą skłonnością do prowadzenia wojen, zwłaszcza przeciwko innym demokracjom. Dla ludzi odpowiedzialnych za politykę zagraniczną USA, podwyższenie poziomu demokracji w kluczowych regionach było niezbędne do zapewnienia pokoju (Doorenspleet 2019).

Stąd zrodziła się potrzeba stworzenia wskaźnika demokratyczności, który pozwalałby porównywać państwa pod tym kątem, określać, jak daleko są od “idealnej” demokracji, oraz oceniać po czasie wyniki podjętych działań. Wyrażenie takiej własności liczbą otwiera również szerokie możliwości tworzenia modeli, w których stopień zdemokratyzowania państwa występuje jako zmienna objaśniana lub objaśniająca, co pozwala wyjaśniać przyczyny, skutki i własności demokracji w sposób dużo bardziej ścisły, poprawny metodologicznie oraz możliwy do zreprodukowania przez innych badaczy.

Między innymi dzięki tym możliwościom, bezpośredni związek stopnia demokratyzacji państwa z jego skłonnością do wojen został przez wielu badaczy podważony lub uznany za słabszy niż inne czynniki, np. wolnorynkowość czy stabilność struktur władzy (Doorenspleet 2019). Wciąż, demokracja jest przedmiotem intensywnych badań socjologów i politologów. A fakt, że badacze interesują się zrozumieniem zjawiska samego w sobie, a nie tylko kontekstem polityki zagranicznej danego państwa, wymusza bliższe przyjrzenie się kwestii obiektywności takich badań.

Zachowanie obiektywności nie jest w tym przypadku proste, z kilku powodów.

Po pierwsze, ustrój polityczny dotyczy tak wielkich jednostek jak państwa, wobec czego badania nad demokracją, które muszą uwzględnić wiele krajów, borykają się ze wszystkimi problemami dotyczącymi różnic kulturowych i językowych.

Po drugie, problematyczna jest sama kwestia definiowania demokracji. W klasycznym rozumieniu jej istotą była władza w rękach ludu, nadanie obywatelom prawa do decydowa-

nia o sprawach politycznych. W latach 40. XX w. zaczęło dominować alternatywne podejście, zaproponowane przez J. Schumpetera. Jego zdaniem, *“demokracja oznacza tylko tyle, że ludzie mają możliwość zaakceptować lub odrzucić tych, którzy mają nimi rządzić”* (Schumpeter 1942). Społeczeństwo jako całość nie musi w jakikolwiek czynny sposób angażować się w politykę; jedynie wyznaczać elity, które będą się tym zajmować. Partycypacja, kluczowa w klasycznej definicji, dla Schumpetera była nieistotna, lub nawet szkodliwa. Taką koncepcję nazywa się elektorálną.

Wymienia się jeszcze kilka innych, szeroko wykorzystywanych podejść, które przedstawia Tabela 1.

Tabela 1: Koncepcje demokracji. Źródło (Coppedge et al. 2011)

Typ demokracji	Postulaty
liberalna	wolności obywatelskie, prawa mniejszości, praworządność i przejrzystość, możliwość kontrolowania rządzących
większościowa	najważniejsza jest wola większości, silna, jednostkowa władza
bezpośrednia	jak największy udział obywateli w bezpośrednim podejmowaniu decyzji (w kontraście do wybierania przedstawicieli)
deliberatywna	decyzje podejmowane w drodze dyskusji i wypracowywania konsensusu (w kontraście do przegłosowania przeciwników)
egalitarna	równość w prawach, ale też dostępie do zasobów; każdy obywatel ma szansę w takim samym stopniu korzystać ze wspólnego państwa

Po trzecie wreszcie, badacze zauważają, że pojęcie demokracji jest silnie nacechowane emocjonalnie (Jasko & Kossowska 2008). Ludzie mają skłonność do uznawania demokracji nie za jedną z możliwych form systemu politycznego, ale raczej za stan idealny. Stąd tak wiele państw, które mają demokrację w nazwie, a są powszechnie uważane za autorytarne. Można by powiedzieć w uproszczeniu, że ludzie zawsze są za demokracją, tylko zazwyczaj ta demokracja ma niewiele wspólnego z którąkolwiek z definicji. Z jednej strony powoduje to przypisanie do czysto politycznego pojęcia demokracji licznych postulatów o charakterze gospodarczym czy etycznym. Z drugiej, przyjęcie jednej definicji, stworzonej przez człowieka ukształtowanego w określonej kulturze, zamyka na faktyczną różnorodność

zjawiska demokracji.

Według minimalistycznej definicji, istotą demokracji jest mechanizm wykorzystania argumentu siły bez używania przemocy. Akceptuje się wolę większości, nie dlatego, że większość ma rację, ale dlatego, że bunt wobec tej woli oznacza bunt wobec przeważającej siły - mimo, że współcześnie siła zależy bardziej od przewagi technicznej niż liczebności (Przeworski 1999). Głos jest walutą władzy; demokracja była rewolucją ustrojową na miarę wynalezienia pieniądza w gospodarce. Jak jednak zauważa dalej Przeworski, trwałość i jakość demokracji zależą od szeregu czynników, zarówno politycznych, jak i ekonomicznych. Dlatego warto spróbować określić istotę demokracji nie od strony mechanizmu przekazywania władzy, ale od strony oczekiwań, jakie ma ten mechanizm spełnić. Ponieważ właśnie to, a nie minimalistyczna definicja, może pomóc odpowiedzieć na pytanie, dlaczego demokracja bywa popierana albo odrzucana.

## 1.2 Wskaźniki demokracji

Zdecydowana większość wskaźników demokracji jest tworzona na podstawie oceny ekspertów. Obliczenie wartości polega na przyznaniu państwu punktów za każdy spełniony standard demokratyczny oraz pewną procedurę agregacji. To oznacza, że każdy przypadek porównuje się z pewną, wybraną przez twórców metody, koncepcją demokracji. Istnieją wskaźniki oceniające państwa według minimalistycznej koncepcji, w skali binarnej (jest demokracją lub nie jest), ale ze względu na swoją prostotę mają zastosowanie głównie przy badaniu ciągłości demokracji (Coppedge et al. 2011). Do celu porównywania ze sobą różnych państw demokratycznych powstały inne wskaźniki, oparte o węższe koncepcje.

Jeden z najczęściej używanych jest obliczany co roku przez amerykańską organizację Freedom House, powiązaną z CIA (Doorenspleet 2019). Freedom House Index (FH) wykorzystuje 7-stopniową skalę do oceny praw politycznych oraz obywatelskich (im niższa ocena, tym lepiej). Pomimo, że według intencji twórców nie miał być nawet wskaźnikiem demokracji, to jest jako taki powszechnie wykorzystywany przez badaczy (Coppedge et al. 2011), ponieważ wyraża własności obecnie dominującej koncepcji liberalnej. Inne wskaźniki mogą nieco różnić się konstrukcją oraz skalą, ale generalnie opierają się o tę samą metodykę.

Przedewszystkim zakładają, że przeciwieństwem demokracji jest system autorytarny, i

pozwalają umiejscowić każde badane państwo na osi autorytaryzm - demokracja. Przejście jest tonalne, nie ma ścisłego punktu, od którego kraj jest klasyfikowany jako któraś z tych dwóch opcji. Ponadto mniej więcej środek skali określa się jako odrębny typ ustroju, tak zwaną demokrację hybrydową, w której łączą się pewne rozwiązania z obu skrajnych typów (Doorenspleet 2019).

Wskaźniki demokracji weszły do powszechnego użytku, ale doczekały się też krytyki. Argumentowano, że zbyt mocno opierają się na procedurach i formalnym stanie prawnym, a za mało na stanie faktycznym, oraz że zbyt ogólne i niejednoznaczne kryteria pozwalają wystawiającym ocenę projektować swoje wyobrażenia o demokratyczności państw na liczbę punktów (Coppedge et al. 2011). Szczególnie wobec FH pojawiły się również zarzuty konserwatywnego i pro-rynkowego skrzywienia, oraz zaniżania oceny dla lewicowych i muzułmańskich rządów (Doorenspleet 2019). To by oznaczało, że wskaźnik nie ocenia wyłącznie stopnia spełnienia bardzo szerokiej, lecz precyzyjnej definicji, ale też nie uwzględnia całej istniejącej różnorodności systemów demokratycznych.

### 1.3 Społeczna koncepcja demokracji

Zupełnie innym podejściem do mierzenia demokracji jest oparcie wskaźnika o zdanie samych obywateli. Nie jest to alternatywny sposób na osiągnięcie tego samego celu, ale raczej próba badania demokracji takiej, jaka jest - niekoniecznie zgodnej z definicją, za to odzwierciedlającej zamysł ludzi, którzy ją wprowadzali i podtrzymywali.

Jak wskazuje Campbell (2019), u podstaw poparcia dla demokracji i demokratycznych rządów leży proces modernizacji, który dotknął każdej dziedziny ludzkiego życia, przez co zmiany stały się właściwie “sprawą każdego”. Jako wskazywane przez badaczy wyjaśnienia tego poparcia podaje: a) zaangażowanie w działalność społeczną, która “*zapewnia antidotum na uczucie apatii i alienacji, oraz projektuje atmosferę zaufania na organy rządowe*”<sup>1</sup>, b) czynniki ekonomiczne - sprzyjająca sytuacja powoduje wyższe poparcie, które spada w trudniejszych czasach, c) czynniki behawioralne, które powodują, że ludzie generalnie popierają wygrywających.

Badania przeprowadzone na polskim gruncie również wykazały utożsamianie demokracji

---

<sup>1</sup>Tłumaczenie własne

z realizowaniem określonego zestawu wartości (standard wolności) albo “zapewnieniem obywatelom bezpieczeństwa, dobrych warunków życia i rozwiązaniem problemów uboższych członków społeczeństwa” (standard opiekuńczości) (Jasko & Kossowska 2008).

Poparcie demokracji nie wynika więc głównie z jej własności ściśle politycznych, ale częściowo z tego, że jest opcją, która się przyjęła, a częściowo ze względu na poczucie wolności obywatelskiej i bezpieczeństwa ekonomicznego, z którymi się wiąże.

## 2 Modele teoretyczne - Międzygrupowa analiza czynnikowa

### 2.1 Model analizy czynnikowej

W naukach społecznych bardzo często badacze zajmują się własnościami, które nie są mierzalne. Ich istnienia można się domyślać, obserwując wpływ jaki wywierają na poziom innych, mierzalnych własności. Taka własnością jest również społeczna koncepcja demokracji, której kształt i wpływ można badać wyłącznie przez badanie poglądów poszczególnych członków społeczeństwa. Metoda została sformalizowana w postaci analizy czynnikowej przez psychologów, takich jak Spearman, Thomson, Thurstone i Burt (Everitt 1984).

Ogólna postać modelu czynnikowego została opisana układem równań 1. Czynniki (inaczej określane jako zmienne ukryte lub latentne)  $\lambda$  wpływają na zmienne mierzalne  $X$  z siłą wyrażoną ładunkami czynnikowymi  $\alpha$ , zgodnie z układem równań 1.

$$\begin{aligned}x_1 &= \alpha_0 1 + \alpha_1 \lambda_1 + \dots + \alpha_k \lambda_k + \epsilon_1 \\&\dots \\x_p &= \alpha_0 p + \alpha_1 p \lambda_1 + \dots + \alpha_k p \lambda_k + \epsilon_p\end{aligned}\tag{1}$$

Każdy element  $x$  oraz  $\epsilon$  jest  $n$ -elementowym wektorem. Zakładamy, że zmienne mierzalne są warunkowo niezależne. Oznacza to, że wzajemny związek dwóch dowolnych  $x$  wynika wyłącznie z uzależnienia każdego z nich od jednej lub więcej wspólnych zmiennych latentnych. Można to wyrazić wzorem:

$$\sigma_{ij} = \sum_{r=1}^k \alpha_i r \alpha_j r, i \neq j \quad (2)$$

gdzie  $\sigma_{ij}$  jest elementem macierzy wariancji i kowariancji  $x$ . Jeśli jako  $\psi_i$  oznaczmy wariancję błędu ( $\epsilon$ )  $i$ -tej zmiennej, wariancję zmiennych  $X$  można przedstawić w postaci:

$$\sigma_{ii} = \sum_{j=1}^k \alpha_i j^2 + \psi_i \quad (3)$$

gdzie pierwszy człon oznacza zmienność, którą  $x_i$  dzieli z innymi zmiennymi mierzalnymi poprzez wspólne zmienne latentne, a drugi to jej wariancja specyficzna.

Istotą estymacji jest więc znalezienie takich wartości parametrów, żeby między zmiennymi mierzalnymi występowały związki założone w strukturze modelu i nie występowały żadne inne. Model o błędnej specyfikacji może okazać się nieidentyfikowalny lub cechować się złymi własnościami, np. negatywnymi wariancjami zmiennych.

## 2.2 Identyfikowalność modelu analizy czynnikowej

Model jest identyfikowalny, jeśli dla danego rozkładu zmiennych mierzalnych istnieje jeden możliwy zestaw parametrów modelu. Jeżeli ten warunek nie jest spełniony i można znaleźć więcej zestawów parametrów, skutkujących tym samym rozkładem zmiennych mierzalnych, to parametrów nie można jednoznacznie zinterpretować [Katsikatsou, Kuha, ?]. Można zauważyć, że model ze zmiennymi latentnymi nigdy nie jest ściśle identyfikowalny.

Macierz wariancji i kowariancji, której elementy zostały wcześniej podane wzorami 2 i 3, można zapisać w postaci macierzowej:

$$\Sigma = AA' + \Psi \quad (4)$$

gdzie  $A$  jest macierzą ładunków czynnikowych o rozmiarze  $k \times n$ , a  $\Psi$  jest macierzą diagonalną wariancji specyficznych. Można zauważyć, że dla ortogonalnej macierzy  $T$  (tzn. takiej, że pomnożona przez swoją transpozycję daje macierz jednostkową) o rozmiarze  $n \times n$ , zachodzi równość:

$$AT(AT)' + \Psi = ATT'A' + \Psi = AA' + \Psi = \Sigma \quad (5)$$



Oznacza to, że ładunków czynnikowych nie da się jednoznacznie zidentyfikować; macierz  $A$  może być wyznaczona tylko z dokładnością do przemnożenia przez dowolną ortogonalną macierz (?). Ładunki czynnikowe nie mają więc bezpośredniej interpretacji, jak współczynniki w modelu regresji, i mogą zostać przeskalowane, tak żeby ustawić wartość oczekiwaną lub wariancję zmiennej latentnej na określonym poziomie.

O modelu analizy czynnikowej mówi się, że jest identyfikowalny, kiedy wszystkie możliwe zestawy parametrów wynikają z przeskalowania macierzy ładunków czynnikowych. Jeśli po wybraniu skali model nie jest w pełni identyfikowalny, musi zostać zmodyfikowany [Katsikatsou, Kuha, ?].

Warunkiem koniecznym identyfikowalności jest, aby liczba estymowanych parametrów nie przekraczała liczby wariancji i kowariancji zmiennych mierzalnych  $N_c$ , która dla  $p$  zmiennych mierzalnych wynosi:

$$N_c = \frac{p(p+1)}{2} \quad (6)$$

Dla ogólnego przypadku z układu równań 1, gdzie każda z  $p$  mierzalnych zmiennych może być objaśniana przez wszystkie  $k$  zmiennych latentnych, można określić liczbę stopni swobody modelu na poziomie:

$$df = \frac{(p-k)^2 - (p+k)}{2} \quad (7)$$

Model 1 jest identyfikowalny, jeśli  $df$  jest nieujemne. Każda restrykcja nałożona na model (w najczęstszym przypadku ustalenie części ładunków czynnikowych na 0, tak że poszczególne zmienne  $X$  są objaśniane tylko przez niektóre czynniki) zmniejsza liczbę parametrów do estymacji.

## 2.3 Estymacja modelu analizy czynnikowej

Analiza czynnikowa składa się zazwyczaj z dwóch etapów, niekoniecznie ściśle od siebie odseparowanych. Pierwszy, określany jako Eksploracyjna Analiza Czynnikowa, polega na wykryciu w danych wszystkich nielosowych wzorców, które wymagają wyjaśnienia. Te wzorce mogą być interpretowane jako przejawy istnienia ukrytych zmiennych. Drugi etap

wymaga postawienia hipotez na temat zależności między zmiennymi mierzalnymi i ukrytymi, i testowania ich za pomocą modelu; określany jest jako Konfirmacyjna Analiza Czynnikowa (Everitt 1984).

Mając realizacje zmiennych  $X$  można obliczyć macierz wariancji i kowariancji  $S$ , a następnie estymować parametry modelu, minimalizując różnicę między  $S$  i  $\Sigma$ . W najbardziej podstawowej Metodzie Najmniejszych Kwadratów minimalizuje się sumę kwadratów różnic między poszczególnymi elementami tych dwóch macierzy. Nie jest to jednak najlepsza miara, ponieważ elementy  $S$  są ze sobą na ogół skorelowane i mają nierówne wariancje (Everitt 1984).

Powszechnie wykorzystywana Metoda Największej Wiarygodności przyjmuje założenie, że obserwacje mają wielowymiarowy rozkład normalny. Funkcja przeciwna do log-wiarygodności, którą minimalizuje się w procesie estymacji, przyjmuje postać:

$$l(S, \Sigma) = \ln|\Sigma| - \ln|S| + \text{trace}(S\Sigma^{-1}) - p \quad (8)$$

Do zbadania istotności modelu można wykorzystać statystykę GFI (Goodness of Fit Index):

$$GFI = \frac{l(S, \Sigma)}{n-1} \chi^2\left(\frac{(p-k)^2 - (p+k)}{2}\right) \quad (9)$$

Dzięki tej statystyce można testować hipotezę zerową, że zakładana macierz wariancji i kowariancji jest identyczna z zaobserwowaną. Ta statystyka jest jednak zależna od  $n$  i dla dużych prób test jest zbyt czuły na drobne niedopasowanie do danych, żeby stosować go w praktyce [Cheung, Rensvold, 2002]. Istnieje wiele wskaźników dopasowania modelu do danych. Przyjmuje się, że wartości CFI (Comparative Fit Index) oraz TLI (Tucker-Lewis Index) powinny być w okolicach lub powyżej 0,95, natomiast RMSEA (Root Mean Square Error of Approximation) nie powinno przekraczać 0,06 (t. Hu & Bentler 1999).

W przypadku, gdy dane nie pochodzą z wielowymiarowego rozkładu normalnego, można wykorzystać estymator ADF (asymptotically distribution-free), ale tylko jeśli próba jest bardzo duża ( $n > 5000$ ). Dlatego częściej stosowanym wyjściem jest stosowanie MNW z pewnymi korektami. Złamanie założenia o normalności prowadzi do poprawnych estymacji, ale z mocno niedoszacowanymi błędami standardowymi i zawyżoną statystyką GFI. Można

to skorygować, stosując przeskalowaną statystykę Satorry-Bentlera oraz odporne błędy standardowe [Rosseel, 2012]. Również w sytuacji porównywania dwóch zagnieżdżonych modeli Testem Ilorazu Wiarygodności (Likelihood Ratio Test) stosuje się przeskalowanie statystyki testowej, tak by miała rozkład  $\chi^2$  [Satorra, Bentler, 2010].

## 2.4 Zagadnienie równoważności pomiarowej

Badania porównawcze stanowią podstawowy cel stosowania analizy czynnikowej. Wartości zmiennych latentnych zwykle nie mają interpretacji same z siebie, ale nabierają znaczenia dzięki porównaniu ich poziomów w innych grupach, np. wiekowych, płciowych, kulturowych, albo w innym punkcie czasowym (Pokropek 2013). Poważnym mankamentem modeli czynnikowych jest to, że jeśli badacz chce przeprowadzać analizy międzygrupowe - szczególnie, jeśli wiadomo, że grupy są pod wieloma względami zróżnicowane - musi mieć pewność, że relacja między zmiennymi mierzalnymi i ukrytymi, opisana modelem czynnikowym, jest we wszystkich grupach taka sama. Taką własność nazywa się równoważnością pomiarową (Chen et al. 2005).

W modelu czynnikowym poziom zmiennej ukrytej jest estymowany na podstawie zależności łączących mierzalne zmienne. Inaczej mówiąc, badacz uznaje, że ta zależność wynika z powiązania wszystkich<sup>2</sup> zmiennych mierzalnych w modelu ze zmienną ukrytą. Niespełnienie równoważności pomiarowej oznacza, że w niektórych grupach występuje inna, nieuwzględniona zależność, której wpływ zostanie zinterpretowany jako wpływ zmiennej ukrytej (Hirschfeld & von Brachel 2014).

Taką sytuację można by opisać, przyjmując, że rzeczywistym opisem zjawiska w grupie **A** jest układ równań 10, a w grupie **B** układ równań 1. Jeśli badacz nie zauważy problemu i zastosuje we wszystkich przypadkach model opisany układem równań 1, wpływ zmiennej  $\delta$ , występującej w grupie A, zostanie w całości przypisany zmiennym  $\lambda$ , które mogą oznaczać zupełnie inną własność rzeczywistego świata niż  $\delta$ .

---

<sup>2</sup>W modelu opisanym równaniem 1

$$\begin{aligned}
x_1 &= \alpha_0 1 + \alpha_1 1\lambda_1 + \dots + \alpha_k 1\lambda_k + \beta_1 \delta + \epsilon_1 \\
&\dots \\
x_p &= \alpha_0 p + \alpha_1 p\lambda_1 + \dots + \alpha_k p\lambda_k + \beta_p \delta + \epsilon_p
\end{aligned}
\tag{10}$$

Nieuwzględnione związki między zmiennymi mierzalnymi mogą mieć bardzo różne przyczyny, niekoniecznie należące do badanej dziedziny. W badaniach międzykrajowych może to być na przykład brak dosłownego tłumaczenia, różnice kulturowe, kontekst historyczny. Przykładowo, słowo oznaczające “sąsiedztwo” w języku angielskim oznacza najbliższą przestrzeń, podczas gdy w języku niemieckim odnosi się do bliskości w sensie społecznym - co okazało się problemem przy tłumaczeniu skali dystansu społecznego Bogardusa (Pokropek 2013).

Z tego powodu każdy model czynnikowy powinien opierać się na solidnej podstawie teoretycznej, a po wyestymowaniu zostać przetestowany pod kątem równoważności pomiarowej (Lubiewska & Glogowska 2018). W przypadku badań obejmujących wiele państw, badacze korzystają na ogół z danych zbieranych przez duże organizacje, jak World Values Survey (WVS), które dbają o standardy tłumaczenia i doboru próby. Mimo to, test równoważności pomiarowej dla bazującego na danych WVS modelu nastawienia do demokracji pokazał, że jest ona zachowana tylko w pewnym stopniu i nie dla wszystkich państw (Ariely & Davidov 2011). Podobne problemy występują w międzynarodowych badaniach zaufania do aparatu państwa i rządzących (Schneider 2016).

Różnice między grupami, ujawnione w badaniu równoważności pomiarowej, nie powinny być jednak traktowane wyłącznie jako problem, ograniczający możliwość badania (o ile niespełnienie równoważności nie wynika z błędu po stronie badaczy, np. w tłumaczeniu pytań). Mogą być źródłem informacji o tym, jak różne grupy postrzegają i konceptualizują rzeczywistość [Cheung, Rensvold, 2002].

## 3 Poziomy równoważności

### 3.1 Równoważność konfiguralna

Najniższym poziomem jest równoważność konfiguralna (równoważność konstruktów), która oznacza zgodność badanego konceptu w grupach. Można mówić o jej spełnieniu, gdy w

każdej grupie ten sam model czynnikowy jest identyfikowalny i dobrze dopasowany do danych. Równoważność konfiguralna zazwyczaj nie stanowi wyzwania w przypadku dobrze poznanych zjawisk, o ile grupy nie różnią się bardzo mocno (Lubiewska & Glogowska 2018). W przypadku badań międzynarodowych może być niespełniona na przykład z powodu błędów w tłumaczeniach, gdy konstrukty są oparte na specyficznym kulturowym kontekście lub mają inne znaczenie dla członków różnych grup [Cheung, Rensvold, 2002].

Niespełnienie tego poziomu równoważności oznacza, że nie w każdej grupie istnieje założona relacja między zmiennymi latentnymi i mierzalnymi, więc stosowanie modelu i tej postaci jest bezpodstawne.

### 3.2 Równoważność metryczna

Jeśli weryfikacja równoważności konfiguralnej przeszła pomyślnie, można zbadać równoważność metryczną. Oznacza ona, że w każdej grupie ładunki czynnikowe są na tym samym poziomie. Równoważność metryczną można zbadać, estymując osobno serię modeli dla każdej grupy, raz ze swobodą dopasowania ładunków czynnikowych do danych w każdej grupie, a drugi raz z restrykcjami równości odpowiadających ładunków w grupach. Jeśli model bez restrykcji okaże się istotnie lepszy, równoważność metryczna nie jest spełniona.

Jej spełnienie gwarantuje, że można przeprowadzić w grupach badanie, w którym zmienna latentna występuje jako zmienna objaśniająca. Relacje między zmiennymi latentnymi i mierzalnymi mają statystycznie równą siłę, wobec czego zmiana poziomu zmiennej mierzalnej o jednostkę powoduje taką samą zmianę poziomu zmiennej latentnej w każdej grupie.

Mając  $M$  grup, należy w każdej osobno wyestymować model postaci:

$$\begin{aligned} x_1 &= \alpha_0 1m + \alpha_1 1m\lambda_1 + \dots + \alpha_k 1m\lambda_k + \epsilon_1 m \\ &\dots \\ x_p &= \alpha_0 pm + \alpha_1 pm\lambda_1 + \dots + \alpha_k pm\lambda_k + \epsilon_p m \end{aligned} \tag{11}$$

gdzie  $m = 1, \dots, M$ . Następnie nałożyć na ładunki czynnikowe restrykcje równości we wszystkich grupach i jeszcze raz wyestymować modele:

$$\begin{aligned}
x_1 &= \alpha_0 1m + \alpha_1 1\lambda_1 + \dots + \alpha_k 1\lambda_k + \epsilon_1 m \\
&\dots \\
x_p &= \alpha_0 pm + \alpha_1 p\lambda_1 + \dots + \alpha_k p\lambda_k + \epsilon_p m
\end{aligned} \tag{12}$$

Do testowania można wykorzystać Test Ilorazu Wiarygodności. Ponieważ model 12 jest zagnieżdżony w modelu 11, jego funkcja wiarygodności (LR) może być tylko mniejsza lub równa od funkcji wiarygodności (L0), maksymalizowanej przy estymacji modelu 11. Są równe, jeśli narzucone restrykcje nie pogarszają modelu, to znaczy w rzeczywistości ładunki czynnikowe w grupach mają takie same wartości. Jeżeli LR jest istotnie niższe, ładunki czynnikowe w grupach są istotnie różne. Testowanie hipotez:

$$H_0 : L_0 = L_R \quad H_1 : L_0 > L_R$$

jest równoznaczne z testowaniem hipotez:

$$H_0 : R_{\text{w}} \text{nowanometryczna} \text{ jest speniona} \quad H_1 : R_{\text{w}} \text{nowanometryczna nie jest speniona}$$

### 3.3 Równoważność skalarna

Trzeci poziom równoważności zakłada spełnienie obu poprzednich oraz równość wyrazów wolnych w grupach. Analogicznie do równoważności metrycznej, można go zweryfikować, estymując modele z restrykcjami i bez restrykcji, a następnie testując istotność różnicy funkcji wiarygodności. Spełnienie tego poziomu daje możliwość bezpośredniego porównywania zmiennych latentnych między grupami.

Model z restrykcjami ma postać:

$$\begin{aligned}
x_1 &= \alpha_0 1 + \alpha_1 1\lambda_1 + \dots + \alpha_k 1\lambda_k + \epsilon_1 m \\
&\dots \\
x_p &= \alpha_0 p + \alpha_1 p\lambda_1 + \dots + \alpha_k p\lambda_k + \epsilon_p m
\end{aligned} \tag{13}$$

a model bez restrykcji jest postaci danej wzorem 11. Testowane są hipotezy:

$$H_0 : R_{\text{wnowanoskalarna}} \text{ jest speniona } H_1 : R_{\text{wnowanoskalarna}} \text{ nie jest speniona}$$

Spełnienie równoważności konfiguralnej daje w praktyce możliwość głębszego testowania (np. równości reszt) lub przeprowadzania szczegółowych testów, w których restrykcje są nałożone na wybrane parametry, podczas gdy pozostałe są pozostawione wolno. Wyniki takich testów mogą dostarczyć dodatkowych informacji o kształtowaniu się modeli badanego zjawiska w różnych grupach, ale nie dają dodatkowych możliwości statystycznych porównań międzygrupowych, dlatego w ogólnym przypadku rozważa się tylko te trzy poziomy.

## 4 Tytuł?

### 4.1 Zestaw danych i statystyki opisowe

World Values Survey jest międzynarodową organizacją, zajmującą się badaniem zmian wartości i przekonań, oraz ich wpływu na życie społeczne i polityczne (Inglehart et al. 2014). Jej ankieta, przeprowadzana co kilka lat w niemal stu krajach, pokrywa również zagadnienia związane z demokracją.

Ankietowanym zadano 10 pytań, z których każde było rozpięte na 10-stopniowej skali porządkowej. W pierwszych dziewięciu ankietowani mieli ocenić, na ile poszczególne cechy państwa są podstawowymi cechami demokracji (1 - nie jest podstawową cechą demokracji, 10 - zdecydowanie jest podstawową cechą demokracji):

1. Rząd nakłada podatki na bogatych i wspiera biednych;
2. Autorytety religijne mają wpływ na prawo;
3. Ludzie wybierają swoich przywódców politycznych w wolnych wyborach;
4. Bezrobotni otrzymują pomoc od państwa;
5. Wojsko przejmuje władzę, jeśli rząd jest niekompetentny;
6. Prawa obywatelskie chronią wolność ludzi;
7. Rząd wyrównuje dochody ludzi;
8. Ludzie są posłuszni tym, którzy nimi rządzą;
9. Kobiety mają takie same prawa jak mężczyźni.

W ostatnim pytaniu ankietowany miał wskazać, jak ważne jest, żeby kraj, w którym mieszka, był rządzony w sposób demokratyczny (1 - zupełnie nieważne, 10 - zdecydowanie ważne). Zmienne opisujące odpowiedzi na pytania będą w dalszej części oznaczane jako X1 - X10 w powyższej kolejności.

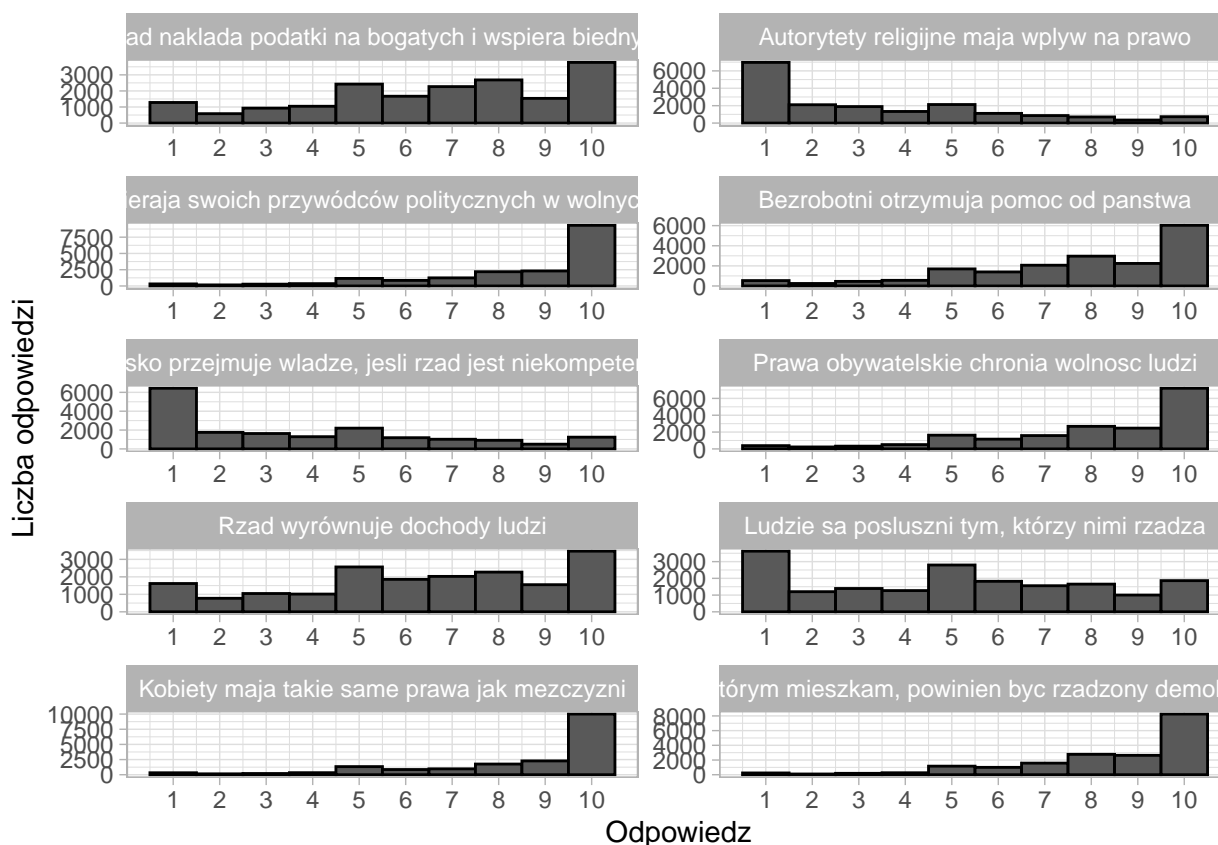
W badaniu wykorzystano dane zebrane w edycji szóstej (lata 2010-2014) ze wszystkich objętych badaniem krajów europejskich. Są to: **Białoruś, Estonia, Gruzja, Niemcy, Kazachstan, Holandia, Polska, Rumunia, Rosja, Słowenia, Hiszpania, Szwecja, Turcja, Ukraina.**

Rys. 1 przedstawia histogramy odpowiedzi na każde z pytań z osobna, łącznie dla wszystkich krajów. Na tej podstawie można podzielić cechy na trzy grupy:

- zdecydowanie istotne dla demokracji: wolne wybory, pomoc dla bezrobotnych, prawa obywatelskie i równość płci, gdzie modalną jest wartość 10, a wszystkie pozostałe występują o wiele rzadziej;
- zdecydowanie nieistotne dla demokracji: wpływ autorytetów religijnych na prawo i możliwość przejęcia władzy przez wojsko, o modalnej 1 i niskich częstościach występowania wyższych wartości;
- niespolaryzowane: wyższe opodatkowanie bogatych, wyrównywanie dochodów i posłuszeństwo wobec rządzących, których rozkłady są bardziej równomierne;

Rys. 2: Rozkład odpowiedzi na pytania o demokrację w ankiecie WVS. Źródło: opra-





cowanie własne

Już ta prosta analiza opisowa pokazuje, że elitystyczna definicja demokracji Schumpetera nie jest wystarczająca w europejskich społeczeństwach. Obywatele zgodnie uważają, że demokracja nie polega tylko na wolnych wyborach, ale też służy zapewnianiu praw i wolności obywatelskich, równości płci, oraz - co najbardziej zaskakujące - pomocy bezrobotnym. Zdaniem samych wyborców, do demokracji nie wystarczy sam fakt wybrania rządzących, ale rządzący ci muszą spełniać podstawowe potrzeby wolności, równości i ekonomicznego bezpieczeństwa rządzonych.

Istnieje też powszechne mniemanie, że system demokratyczny nie potrzebuje interwencji ze strony armii ani autorytetów religijnych. Natomiast opinia społeczna jest mniej jednogłówna, jeśli chodzi o kwestię redystrybucji i rozwarstwienia dochodowego. Choć większość uważa, że w demokracji państwo powinno się tym zajmować, przewaga najwyższej odpowiedzi nie jest aż tak wyraźna.

Ciekawym przypadkiem są tu odpowiedzi na pytanie o posłuszeństwo rządzącym. Koncepcja demokracji Schumpetera pośrednio zakłada raczej posłuszeństwo wobec rządzących, którzy zostali zaakceptowani, co jest też zresztą cechą jakichkolwiek systemów rządzenia

czy zarządzania.

## 4.2 Analiza korelacji

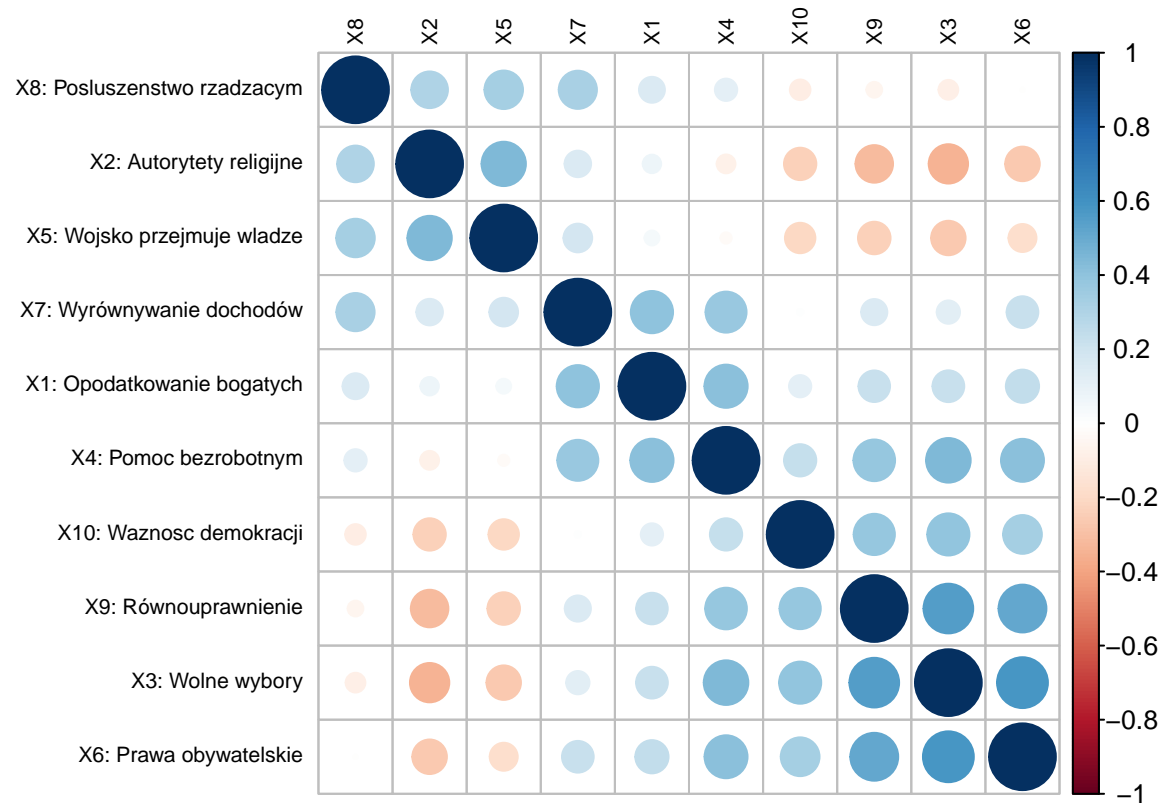
Kiedy ankietowany odpowiada na pytanie “w jakim stopniu [pewna cecha] jest podstawową cechą demokracji?”, musi odwołać się do jakiejś koncepcji demokracji i odnieść do niej cechę, której dotyczy pytanie. Jeżeli istnieje ogólnie przyjęta społeczna koncepcja demokracji, obejmująca część spośród zawartych w pytaniach cech, można się spodziewać, że odpowiedzi na te pytania będą silnie skorelowane. Ankietowani będą odpowiadać w podobny sposób, ponieważ posiadają tę samą koncepcję demokracji. Analogicznie, jeśli nie istnieje społeczna koncepcja demokracji (to znaczy: pogląd na demokrację jest bardzo indywidualną kwestią), można się spodziewać, że siła związków między odpowiedziami na różne pytania będzie nieistotna.

Siłę tych związków mierzono współczynnikiem korelacji Spearmana. Mimo, że dane mają charakter porządkowy, można je traktować jako ciągłe ze względu na dużą liczbę punktów skali [Revelle, 2019].

Koncepcję demokracji można opisać zmienną latentną, która wpływa na odpowiedzi (zmienne mieralne). Koncepcji może być kilka, ale nie więcej, niż pytań. Na odpowiedź na dane pytanie może wpływać więcej niż jedna koncepcja. Warto też zauważyć, że wyodrębnione z danych koncepcje są określone tylko w zakresie pytań, które zadano w ankiecie. Nie można na tej podstawie stwierdzić, czy koncepcja obejmuje też inne kwestie, np. bioetyczne. Teoretycznie mogą również istnieć koncepcje demokracji, które nie zawierają objętych pytaniami cech, wobec czego wymykają się badaniu.

Rys. 3 przedstawia macierz korelacji odpowiedzi na wszystkie pytania. Dla macierzy wykonano test sferyczności Bartletta, na podstawie którego stwierdzono, że jest istotnie różna od macierzy jednostkowej, zatem siła związków między zmiennymi jest istotna ( $\chi^2 = 46805,62$ ;  $df = 45$ ;  $p < 0,001$ ).

Rys. 3: Macierz korelacji odpowiedzi na pytania o demokrację. Źródło: opracowanie



własne

## References

- Ariely, G. & Davidov, E. (2011), 'Can we rate public support for democracy in a comparable way? cross-national equivalence of democratic attitudes in the world value survey', *Social Indicators Research* **104**(2), 243–257.
- Campbell, R. (2019), Theories of popular support for democracy, in 'Popular Support for Democracy in Unified Germany. New Perspectives in German Political Studies', Palgrave Macmillan, Cham.
- Chen, F. F., Sousa, K. H. & West, S. G. (2005), 'Teacher's corner: Testing measurement invariance of second-order factor models', *Structural Equation Modeling: A Multidisciplinary Journal* **12**:3, 471–492.
- Coppedge, M., Gerring, J., Altman, D., Bernhard, M., Fish, S., Hicken, A., Kroenig, M., Lindberg, S. I., McMann, K. M., Paxton, P., Semetko, H. A., Skaaning, S.-E., Staton,

- J. K. & Teorell, J. (2011), 'Conceptualizing and measuring democracy: A new approach', *Perspectives on Politics* **9**(2), 247 – 267.
- Doorenspleet, R. (2019), *Rethinking the Value of Democracy*, Palgrave Macmillan, Cham.
- Everitt, B. S. (1984), *An Introduction to Latent Variable Models*, Chapman and Hall.
- Hirschfeld, G. & von Brachel, R. (2014), 'Improving multiple-group confirmatory factor analysis in r – a tutorial in measurement invariance with continuous and ordinal indicators', *Practical Assessment, Research, and Evaluation* **19**(7).
- Inglehart, R., and A. Moreno, C. H., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E. & et al., B. P. (2014), 'World values survey: Round six - country-pooled datafile version'.
- URL:** [www.worldvaluessurvey.org/WVSDocumentationWV6.jsp](http://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp)
- Jasko, K. & Kossowska, M. (2008), 'Wplyw wartosci na rozumienie pojecia demokracji', *Psychologia spoleczna* **tom 3 3 (8)**, 243 – 257.
- Lubiewska, K. & Glogowska, K. (2018), 'Zastosowanie analizy rownowaznosci pomiarowej w badaniach psychologicznych', *Polskie Forum Psychologiczne* **23**(2), 330 – 356.
- Pokropek, A. (2013), Porownywalnosc w badaniach miedzynarodowych. przyklad wskaznika motywacji do nauki w badaniu pisa 2006, polska edukacja w swietle diagnoz prowadzonych z roznych perspektyw badawczych, in 'XIX Konferencja Diagnostyki Edukacyjnej'.
- Przeworski, A. (1999), *Minimalist Conception of Democracy: A Defense*, Cambridge University Press.
- Schneider, I. (2016), 'Can we trust measures of political trust? assessing measurement equivalence in diverse regime types', *Social Indicators Research* **133**, 963 – 984.
- Schumpeter, J. A. (1942), *Capitalism, Socialism, and Democracy*, 2 edn, Sublime Books.

- t. Hu, L. & Bentler, P. M. (1999), ‘Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives’, *Structural Equation Modeling* **6**(1), 1 – 55.