

Metodologiczne Inspiracje 2021

Badania Ilościowe w Naukach Społecznych — Wyzwania i Problemy

23–24 września, Jabłonna

23 września

Sesja 1. Prowadzący: Dariusz Przybysz (IFiS PAN)

BARBARA JANCEWICZ, OŚRODEK BADAŃ NAD MIGRACJAMI UW

Dlaczego rekrutacja badanych poprzez reklamy Facebooka nie jest dobrem losowym i co można z tym zrobić?

Facebook jest kanałem powszechnie stosowanym do rekrutacji badanych do wywiadów, focusów lub ankiet internetowych. Badacze zazwyczaj publikują posty z zaproszeniem do badania na wybranych grupach (np. Polacy w Londynie), a w artykułach opisują swój dobór próby jako wygodny (convenience) albo celowy. Jednak w ciągu ostatnich kilku lat popularność zyskuje używanie do rekrutowania narzędzi marketingowych Facebooka, a dokładniej nacelowanych reklam (mikrotargetowanie). Narzędzie to wygląda bardzo obiecująco, gdyż pozwala na selekcję respondentów ze względu na szereg cech wykraczających poza standardowe (wiek, płeć, miejsce pobytu): zainteresowania, poglądy polityczne, niedawne wydarzenia życiowe (ślub, narodziny dziecka, przeprowadzka). Dzięki temu możliwe jest precyzyjne dotarcie do grup, które innymi metodami są trudnodostępne, takich jak migranci. Niedawno przy użyciu reklam na Facebooku Potzschke i Weiss przeprowadzili globalne badanie niemieckiej diaspory, a uzyskane wyniki uznają za pozwalające na dokonywanie porównań między krajami. Niestety, o ile mikrotargetowanie reklam Facebooka teoretycznie pozwoliłoby na losowy dobór próby, o tyle algorytmy serwisu dążące do maksymalnej „efektywności” reklam sprawiają, że uzyskany dobór próby pozostaje jedynie celowy lub, w przypadku zastosowania odpowiednich nakierowań, kwotowy.

W wystąpieniu krótko opiszę, jak działają reklamy Facebooka, jak wygląda tworzenie kampanii reklamowej oraz jakie pułapki napotkamy rekrutując badanych tą metodą, czego (nie) możemy oczekiwać od uzyskanych wyników, a finalnie jak skonfigurować kampanię, aby uzyskać jak najlepszy dobór próby.

KAZIMIERZ M. SŁOMCZYŃSKI, INSTYTUT FILOZOFII I SOCJOLOGII PAN

ZBIGNIEW SAWIŃSKI, INSTYTUT FILOZOFII I SOCJOLOGII PAN

DOROTA LASKOWSKA, INSTYTUT FILOZOFII I SOCJOLOGII PAN

Kto, kiedy i dlaczego odpadł z wieloletniego panelu: Analiza badania POLPAN 1988-2018

Polskie Badanie Panelowe POLPAN 1988-2018 należy do niewielu badań na świecie, gdzie przez 30 lat starano się utrzymać kontakt z respondentami. Bazą wyjściową było badanie przeprowadzone w 1987/1988 roku na próbie mężczyzn i kobiet w wieku 21-65 lat (N = 5, 817). Następne edycje zostały przeprowadzone w odstępach pięcioletnich, zaczynając od 1993 roku, gdy dotarto do prawie połowy oryginalnej próby. Od 1998 roku próba panelowa była uzupełniana przez roczniki młodsze. W referacie przedstawiamy pełną historię uczestnictwa w naszym badaniu, koncentrując się na ubytkach i napływie nowych respondentów. O każdym, kto przynajmniej raz był badany mamy informację, czy pozostał w próbie do 2018 roku, a jeśli nie – z jakiego powodu w badaniu już nie uczestniczy. Zasadnicza część referatu dotyczy przyczyn ubytków. Podajemy te przyczyny w przekroju czasowym w powiązaniu z innymi zmiennymi, głównie społeczno-demograficznymi. W końcowej części nawiązujemy do ostatnich dyskusji metodologicznych o wazieniu prób panelowych na podstawie rodzajów i częstości ubytków w panelu w różnym czasie. Zwracamy uwagę na konsekwencje proponowanych wag dla substancywnych analiz karier zawodowych.

KONRAD KULIKOWSKI (SPOŁECZNA AKADEMIA NAUK W ŁODZI)

Wynagrodzenie sprawiedliwe, czyli jakie? – rozważania o pomiarze sprawiedliwości wynagrodzenia

Tematem wystąpienia będzie refleksja nad tym, jak dokonywać pomiaru sprawiedliwości wynagrodzenia. Zagadnienie to jest nie tylko ciekawe poznawczo, ale może mieć istotne znaczenie praktyczne. Brak sprawiedliwości wynagrodzenia to temat, który przejawia się w wielu społecznie ważnych kwestiach, takich jak np.: jawność wynagrodzeń, konflikty na linii pracodawca - pracownik czy różnice płacowe kobiet i mężczyzn. Mimo doniosłej społecznej roli sprawiedliwości wynagrodzenia, zagadnienie to wciąż jest niejasne. Co to bowiem znaczy, że wynagrodzenie jest sprawiedliwe i jak dokonać pomiaru sprawiedliwości wynagrodzenia? By odpowiedzieć na to pytanie, w ramach referatu planuję omówić różne formy rozumienia sprawiedliwości wynagrodzenia oraz sposoby jej pomiaru:

- a) sprawiedliwość wynagrodzenia jako rozbieżność między tym jakie wynagrodzenie jest a jakie być powinno oraz sposoby jej pomiaru, jako rozbieżność bezwzględna, logarytmiczna oraz procentowa;
- b) sprawiedliwość wynagrodzenia ujmowana jako sprawiedliwość dystrybutywna (sprawiedliwość stosunku własnych nakładów do zysków w porównaniu do stosunku nakładów do zysków innych, będących punktem odniesienia), proceduralna (sprawiedliwość reguł wynagradzania), informacyjna (sprawiedliwość otrzymywanych informacji o wynagrodzeniu) i interakcyjna (sprawiedliwość w relacji interpersonalnej z ustalającymi i wypłacającymi wynagrodzenie) oraz kwestionariuszowe pośrednie i bezpośrednie sposoby jej pomiaru;
- c) sprawiedliwość jako zadowolenie z wynagrodzenia oraz kwestionariuszowe jedno- i wielowymiarowe sposoby jej pomiaru (sprawiedliwość wysokości pensji, sprawiedliwość podwyżek, sprawiedliwość benefitów, sprawiedliwość administracji wynagrodzeń);

- d) obiektywne, niezależne od przekonań i wartościowań oceniającego sposoby pomiaru sprawiedliwości wynagrodzenia. Sprawiedliwość jako rozbieżność pomiędzy wynagrodzeniem otrzymywanym a obiektywnie ustalonym punktem odniesienia, np.: wynagrodzeniem minimalnym, wynagrodzeniem średnim, wynagrodzeniem zapewniającym godziwe życie, wynagrodzeniem wynikającym z regulaminu wynagradzania;
- e) sprawiedliwość wynagrodzenia jako takie samo wynagrodzenie za taką samą pracę oraz jej pomiar z wykorzystaniem regresji wielorakiej z logarytmizacją wynagrodzenia.

Podsumowując, celem planowanego wystąpienia jest ukazanie różnych sposobów rozumienia sprawiedliwości wynagrodzenia i próba pobudzenia dyskusji nad tym, jaka forma pomiaru sprawiedliwości wynagrodzenia może być najbardziej adekwatna.

ADAM RYBAK (WYDZIAŁ SOCJOLOGII UAM W POZNANIU)

Czy (wymuszone) odchodzenie od sondaży face-to-face to szansa, czy zagrożenie? Analiza na podstawie danych z międzykrajowych projektów sondażowych^a.

Konsekwencje pandemii rozpoczętej w 2020 roku obejmują zarówno odgórne ograniczenia w bezpośrednich kontaktach międzyludzkich, jak i oddolny strach przed takimi kontaktami. Skutkiem obydwu jest przyspieszenie przemian praktyk sondażowych, które w wolniejszym tempie dokonywały się już wcześniej w zachodniej Europie, północnej Ameryce i Australii z Nową Zelandią. W regionach tych stopniowa rezygnacja z przeprowadzania sondaży techniką wywiadu kwestionariuszowego prowadzonego przez ankietera w bezpośrednim kontakcie z respondentem wynikała dotychczas przede wszystkim z realiów ekonomicznych – rozwiązanie takie okazywało się w coraz większej liczbie przypadków nieakceptowalnie drogie. Nowe warunki spowodowały jednak, że nawet takie przedsięwzięcia jak Eurobarometr, czy Europejski Sondaż Społeczny, które dotychczas trzymały się zdecydowanie (w swojej standardowej praktyce) badań face-to-face rozpoczęły wdrażanie badań w innych technikach – lub w schemacie technik mieszanych.

W takiej sytuacji interesujące wydaje się zestawienie informacji na temat tego, gdzie i z jakim skutkiem sondaże w innych technikach były przeprowadzane dotąd w ramach dużych programów międzykrajowych. Dlatego zaprezentuję dystrybucję poszczególnych technik sondażowych dla krajów ze wszystkich kontynentów i na przestrzeni ostatnich około 25 lat. Dane pochodzą przede wszystkim z takich programów jak International Social Survey Programme, Comparative Study of Electoral Systems czy World Values Survey. Informacje te zostaną uzupełnione o współczynniki realizacji (*response rates*), oraz oszacowanie wielkości błędów sondażowych dla poszczególnych technik (oraz regionów świata), w tym za pomocą wielopoziomowej metaregresji.

Uzyskane wyniki pozwolą na ocenę, czy obecna przymusowa zmiana praktyk terenowych może być w większym stopniu widziana jako zagrożenie, czy jako szansa. Z jednej strony fakt niższych *response rates* (zazwyczaj) osiąganych przez sondaże przeprowadzane techniką pocztową czy internetową względem standardowych wywiadów kwestionariuszowych jest powszechnie znany. Z drugiej jednak strony nie ma bezpośredniej zależności pomiędzy odsetkiem zrealizowanych wywiadów, a błędem braków odpowiedzi (*nonresponse bias*). Dlatego analizy oparte na dużych zbiorach danych na temat fieldworku i jakości próby w dużych projektach międzykrajowych mogą prowadzić do zaskakujących rezultatów.

^aFinansowane z grantu NCN 2019/33/N/HS6/00322

SESJA 2 Prowadzący: Zbigniew Sawiński (IFiS PAN)

KAMIL FILIPEK, INSTYTUT SOCJOLOGII, UMCS W LUBLINIE

Skrzywienie algorytmiczne (SI): teorie, typy i metody zwalczania

W sferze prywatnej oraz publicznej coraz częściej spotykamy się rozwiązaniami wykorzystującymi algorytmy sztucznej inteligencji. Autonomiczne samochody, wirtualni asystenci, inteligentne systemy diagnozujące choroby, algorytmy rekomendacyjne to wybrane przykłady praktycznej implementacji metod sztucznej inteligencji w postaci algorytmów uczenia maszynowego oraz głębokiego. Niestety, rozwiązania te często działają wadliwie, co potwierdzają wyniki badań rozmaitych algorytmów prowadzone przez badaczy na całym świecie (Neff 20016; Geary, Danks 2019), w tym również w Polsce (Sztandar-Sztanderska, Kotnarowski, Zieleńska, 2021).

W ostatnich kilku latach algorytmy sztucznej inteligencji stały się polem badawczym współczesnych nauk społecznych (Joyce et al. 2021). Algorytmy często analizowane są jako: (i) mechanizm nadzoru (w tym element szeroko rozumianych struktur władzy), (ii) filtr informacyjny, (iii) autonomiczny system decyzyjny, (iv) źródło nierówności społecznych. Część badaczy koncentruje się na skutkach społecznych funkcjonowania algorytmów w przestrzeni prywatnej oraz publicznej. W literaturze nauk społecznych niewiele jest jednak prac, które analizowałyby problem skrzywienia algorytmicznego (ang. *algorithmic bias*) od strony teoretycznej i/lub metodologicznej (Garcia 2016; Kordzadeh, Ghasemaghaei 2021). Po pierwsze, brakuje prac teoretycznych (podejść lub teorii), które pozwalałyby formułować obiektywne założenia, oczekiwania oraz zasady działania algorytmów dedykowanych do optymalizacji wybranych procesów mających wpływ na kształt społeczeństw współczesnych. Próżnia teoretyczna sprawia, że wykrywanie anomalii algorytmicznych, a także kompleksowa ocena ich skuteczności stają się niezwykle trudne. Po drugie, niewiele jest prac metodologicznych oferujących rozwiązania służące zapobieganiu, redukowaniu oraz zwalczaniu skrzywienia algorytmicznego (zwłaszcza w obszarze algorytmów sztucznej inteligencji). Brak powszechnie uznanych metod zwalczania skrzywienia algorytmicznego sprawia, że metody sztucznej inteligencji są często kwestionowane i krytykowane. Wydaje się, że pośród dyscyplin społecznych socjologia dysponuje teoriami oraz metodami, które z powodzeniem wykorzystać można do walki ze skrzywieniem algorytmicznym.

Celem wystąpienia jest (i) prezentacja typów skrzywień algorytmicznych, które decydują o wadliwym działaniu algorytmów implementowanych w przestrzeni publicznej oraz prywatnej, (ii) przedstawienie wybranych metod redukowania skrzywienia algorytmicznego na przykładzie algorytmów wykorzystywanych na gruncie przetwarzania języka naturalnego (NLP). Oprócz metod koncentrujących się na jakości danych (np. metody równoważenia klas, metody imputacji danych), zaprezentowana zostanie metoda redukcji skrzywienia bazująca na danych pozyskanych z monitoringu działania algorytmu (NLP).

Joyce, K., Smith-Doerr, L., Alegria, S., Bell, S., Cruz, T., Hoffman, S. G., Noble, S. U., & Shestakofsky, B. (2021). Toward a Sociology of Artificial Intelligence: A Call for Research on Inequalities and Structural Change. *Socius*, 7, 2378023121999581.

Garcia, M. (2016). Racist in the machine: The disturbing implications of algorithmic bias. *World Policy Journal*, 33(4), 111-117.

Geary, T., & Danks, D. (2019, January). Balancing the benefits of autonomous vehicles. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 181-186).

Kordzadeh, N., & Ghasemaghaei, M. (2021). Algorithmic bias: review, synthesis, and future research directions. *European Journal of Information Systems*, 1-22.

Neff, G. (2016). Talking to bots: Symbiotic agency and the case of Tay. *International Journal of Communication*.

Sztandar-Sztanderska, K., Kotnarowski, M., & Zieleńska, M. (2021). Czy algorytmy wprowadzają w błąd? Metaanaliza algorytmu profilowania bezrobotnych stosowanego w Polsce. *Studia Socjologiczne*, 89-115.

ARTUR POKROPEK, INSTYTUT FILOZOFII I SOCJOLOGII PAN

ERNEST POKROPEK, DIVISION OF ROBOTICS, PERCEPTION, AND LEARNING, KTH ROYAL INSTITUTE OF TECHNOLOGY, STOCKHOLM, SWEDEN

Wykorzystanie głębokich sieci neuronowych do wykrywania błędnych specyfikacji modeli statystycznych. Przypadek równoważności pomiarowej

W ostatnich latach zaproponowano szereg nowych podejść statystycznych do modelowania różnic między grupami. Zdecydowana większość z tych modeli wyrasta z tradycji pomiaru cech ukrytych za pomocą confirmacyjnej analizy czynnikowej (ang. *Confirmatory Factor Analysis*, CFA). Poprawna estymacja różnic międzygrupowych wymaga, aby model pomiarowy został poprawnie wyspecyfikowany ze szczególnym uwzględnieniem równoważności pomiarowej wskaźników. Brak równoważności może wynikać z błędów tłumaczeń lub znaczących różnic kulturowych, a jego występowanie może w istotny sposób zaburzać porównania pomiędzy grupami. Dotychczasowe narzędzia służące do oceny spełnienia założenia równoważności pomiaru charakteryzują się umiarkowaną skutecznością (charakteryzują się dużą ilością błędnych ocen).

W niniejszym wystąpieniu przedstawiamy nowatorskie podejście do wykrywania błędnych specyfikacji założeń równoważności pomiaru przy użyciu głębokich sieci neuronowych (ang. *Deep Neural Networks*, DNNs). W proponowanym modelu skupiamy się na wykrywaniu, które parametry modelu pomiarowego nie spełniają założeń równoważności (tj. nie można założyć, że mają takie same wartości we wszystkich grupach). Proponowany model porównujemy z najpopularniejszymi tradycyjnymi metodami: a) wskaźnikami modyfikacji modelu (MI) i wskaźnikami oczekiwanej zmiany parametrów (EPC), b) analizą zróżnicowanego funkcjonowania pozycji testowej (*differential item functioning*; DIF) za pomocą regresji logistycznej oraz c) procedurą sekwencyjną zaproponowaną przez autorów metody tzw. optymalizacji wyrównującej (Muthen i Asparuhov 2014; ang. *alignment optimization*). Badania symulacyjne pokazują, że proponowana metoda przewyższa metody tradycyjne w prawie wszystkich rozpatrywanych scenariuszach lub jest co najmniej tak dokładna, jak pozostałe. Przedstawiamy również przykład empiryczny, wykorzystujący dane z Europejskiego Sondażu Społecznego (ESS), zawierający pozycje, co do których wiadomo, że zostały błędnie przetłumaczone. Pozycje te zostały poprawnie zidentyfikowane przez proponowaną przez nas metodę oraz wykrywanie DIF wykorzystujące regresję logistyczną.

Badanie to stanowi podstawę dla przyszłego rozwoju algorytmów uczenia maszynowego do wykrywania błędnych specyfikacji modeli statystycznych. Podczas prezentacji omówimy, jak prezentowane podejście może zostać wykorzystane dla wykrywania innych błędnych założeń w modelowaniu statystycznym.

Wpływ decyzji badacza na dokładność głębokich sieci neuronowych klasyfikujących opinie wyrażane na Twitterze

Celem referatu jest odpowiedź na pytanie, w jaki sposób decyzje metodologiczne podejmowane przez badacza jeszcze zanim rozpocznie się proces nadzorowanego uczenia maszynowego wpływają na dokładność głębokich sieci neuronowych w zadaniach klasyfikacyjnych. Problem jest ważny z dwóch powodów. Po pierwsze, celem trenowania modeli jest zazwyczaj uzyskanie jak najlepszych wyników działania. Stąd też istotne jest, aby decyzje sprzyjały tym dążeniom. Po drugie, najczęściej zespoły badawcze dysponują ograniczonymi środkami finansowymi i kadrowymi. Z tego względu ważne jest podejmowanie takich decyzji, które nie tylko doprowadzą od możliwie najlepszych wyników, ale też takich, które pozwolą je osiągnąć bez marnotrawienia funduszy.

W badaniu wykorzystano dane z autorskiego projektu badawczego, którego celem było stworzenie bazy 100 tysięcy ręcznie zakodowanych tweetów pod względem tego, czego dotyczą (przedmiot opinii) i stosunku autora wpisu do tej kwestii. Tweety zostały wylosowane z wypowiedzi polskich politycznych liderów opinii opublikowanych w 2019 roku (tj. wpisów kont regularnie publikujących na tematy społeczno-polityczne i komentujących aktualne wydarzenia, posiadających ponad 5 000 obserwujących, i których tweety były retweetowane przeciętnie ponad 100 razy; liczba kont: 983; liczba opublikowanych wpisów: 1 228 417).

Analizie poddane zostały trzy decyzje, przed którymi stoją badacze chcący wytrenować model klasyfikujący tekst: 1) Czy w ramach przygotowania danych należy dokonać lematyzacji słów (sprowadzić je do form podstawowych, np. „zrobiliśmy” → „zrobić”? 2) Czy uczenie powinno przebiegać w sposób pasywny (algorytm uczy się na losowej próbie przypadków) czy aktywny (algorytm uczy się na najtrudniejszych przypadkach)? 3) Czy lepiej weryfikować kodowanie (kodować każdy przypadek dwukrotnie), czy kodować więcej przypadków (zakodować dwa razy więcej danych kosztem utraty możliwości sprawdzenia ich poprawności)?

Wyniki tych decyzji zostały sprawdzone w trzech zadaniach klasyfikacyjnych o różnym stopniu trudności. Zadaniem sieci neuronowych była: 1) identyfikacja tego, czy tweet jest na temat aktualnych wydarzeń lub dotyczy spraw społeczno-politycznych (klasyfikacja binarna), 2) klasyfikacja pod względem tego, na jaki dokładnie temat jest tweet (klasyfikacja na podstawie 65 kodów) i 3) rozpoznanie stanowiska autora wpisu na dany temat (trzy kategorie: negatywne, pozytywne, neutralne/ambiwalentne). Kontrolowana była przy tym liczba przypadków, na których uczył się algorytm. Zaczynając od 1000 obserwacji liczba danych była stopniowo zwiększana o kolejny tysiąc aż do osiągnięcia poziomu 45000. W konsekwencji otrzymano do zbadania 1056 konfiguracji ($2 \times 2 \times 2 \times 3 \times 44$). Dane treningowe były za każdym razem losowane, dlatego aby wyeliminować efekt przypadkowo wylosowanego zbioru łatwych lub trudnych przypadków, w ramach każdej konfiguracji algorytm uczoneo 50 razy. W rezultacie baza do testów składała się z 52800 wyników (50×1056). W badaniu wykorzystano algorytm firmy Facebook – fastText.

Wnioski sformułowano na podstawie analizy regresji liniowej, w której zmienną zależną jest wskaźnik F ($F\text{-score}/F\text{-measure}$) a zmiennymi niezależnymi wielkość próby, rodzaj uczenia (aktywne/pasywne), lematyzacja (obecna/brak) i weryfikacja tweetów (obecna/brak). Dla każdego z trzech zadań klasyfikacyjnych sporządzono osobny model regresji. Na podstawie wyników zostaną na konferencji przedstawione rekomendacje a także omówione pułapki stojące przed badaczami społecznymi wykorzystującymi modele nadzorowane.

NATALIA BANASIK-JEMIELNIAK, AKADEMIA PEDAGOGIKI SPECJALNEJ IM. MARII GRZEGORZEWSKIEJ

MAREK MUSZYŃSKI, INSTYTUT FILOZOFII I SOCJOLOGII PAN

TOMASZ ŻÓŁTAK, INSTYTUT FILOZOFII I SOCJOLOGII PAN

ARTUR POKROPEK, INSTYTUT FILOZOFII I SOCJOLOGII PAN

O wyzwaniach w przekładalności testów językowych na przykładzie procesu adaptacji polskiej wersji testu Pathfinder do pomiaru inteligencji werbalnej

Przy przenoszeniu narzędzia badawczego do innego środowiska językowego i kulturowego, ważne jest przystosowanie jego oryginalnej wersji do specyfiki kontekstu, w którym ma ono funkcjonować. Ta sama cecha lub zdolność może mieć w różnych kulturach różne przejawy, a te same umiejętności lub preferencje mogą być przejawami różnych cech w różnych kulturach. Ważne jest, by znaleźć specyficzne kulturowo/językowo wskaźniki konstruktów, którymi mamy mierzyć (Matczak, 2011). W adaptacji narzędzi pomiaru często wyróżnia się dwa jej rodzaje: parafrazę i transkrypcję (Drwal, 1995). Podczas gdy parafraza polega na stworzeniu nowego zadania po przeprowadzeniu dokładnej analizy założeń, które leżą u podstaw oryginalnego narzędzia, transkrypcja jest stosunkowo wiernym oddaniem pierwowzoru przy zachowaniu ekwiwalencji zawartości i formy. W badaniach kompetencji językowych nie można sobie zwykle pozwolić na transkrypcję, konieczna jest parafraza. Okazuje się jednak, że wiąże się ona z wieloma wyzwaniami. Priorytetem jest zachowanie trafności treściowej i teoretycznej wyjściowego konstruktów. Jednak konieczny staje się również wybór cech związanych z formą i treścią materiału bodźcowego w taki sposób, by był on jak najbardziej zbliżony funkcjonalnie do oryginału. Innymi słowy, trzeba podjąć decyzję, które z tych cech są kluczowe i jaką mają rangę. Wybór ten staje się szczególnie trudny, gdy mierzymy zachowania mające być wskaźnikami złożonych procesów poznawczych. O ile w badaniu zasobu słownictwa (małe skomplikowanie poznawcze procesów) jest dość jasne, jakie cechy słów (elementów słownikowych) są wyznacznikiem ogólnego bogactwa leksykalnego danej osoby, o tyle w badaniach przetwarzania językowego, w których badani proszeni są np. o wskazanie brakującej litery uzupełniającej jednocześnie kilka brakujących wyrazów, nie jest jasne, które cechy bodźców werbalnych odpowiadają w największym stopniu za trudność tego zadania. Szczególnie że mamy tu do czynienia z wieloma procesami, takimi jak wydobywanie elementów z pamięci, dopasowywanie elementów do siebie, czy hamowanie poznawcze. Dodatkową trudność w adaptacji testu stanowią różnice związane z etymologią słów w języku polskim i angielskim, co przekłada się np. na lepszą rozpoznawalność dłuższych wyrazów w języku polskim przez natywnych użytkowników polskiego.

Wspomniane wyżej wyzwania omówimy na przykładzie testu inteligencji ogólnej Pathfinder (Malanchini et al., 2021). Jest to przypominający grę zestaw testów inteligencji, mierzący jej komponenty werbalne i niewerbalne. O ile w przypadku komponentów niewerbalnych adaptacja polskiej wersji narzędzia nie wymagała zmiany bodźców, o tyle przygotowanie komponentów werbalnych wymagało stworzenia, a następnie przepilotowania zestawu całkiem nowych bodźców. Zestaw ten musiał zarówno spełniać założenia oryginalnego zadania, jak również być dopasowanym do charakterystyki języka polskiego. Posługując się jednym z trzech z opracowanych podtestów werbalnych, przedstawimy proces adaptowania bodźców pod kątem lingwistycznym, a następnie ich weryfikację przy użyciu zaawansowanych metod psychometrycznych.

Celem wystąpienia jest nie tylko opowiedzenie o jednej konkretnej adaptacji, ale również zaproszenie do wypracowania zestawu dobrych praktyk przy tłumaczeniu i adaptacji

różnorodnych narzędzi pomiarowych do języka polskiego. To ostatnie wydaje się potrzebne, biorąc pod uwagę że sztandarowe polskie pozycje na temat adaptacji testów i kwestionariuszy są już dość wiekowe (np. Ciechanowicz, 1990; Drwał, 1995; Pawłowska, 1983), w związku z czym nie uwzględniają zmian społecznych, technologicznych i metodologicznych, które dokonały się od czasu ich napisania.

Ciechanowicz, A. (Ed.). (1990). *Kulturowa adaptacja testów: praca zbiorowa*. Polskie Towarzystwo Psychologiczne.

Drwał, R. Ł. (1995). *Adaptacja kwestionariuszy osobowości: wybrane zagadnienia i techniki*. Wydawnictwo Naukowe PWN.

Malanchini, M., Rimfeld, K., Gidziela, A., Cheesman, R., Allegrini, A. G., Shakeshaft, N., Schofield, K., Packer, A., Ogden, R., McMillan, A., Ritchie, S.J., Dale, P.S., Eley, T.C., von Stumm, S. & Plomin, R. (2021). Pathfinder: A gamified measure to integrate general cognitive ability into the biological, medical and behavioural sciences. *bioRxiv*.

Matczak, A. (2011). *Dlaczego trzeba adaptować testy?* Wystąpienie przedstawione podczas konferencji: Testy w biznesie: standardy a praktyka.

Pawłowska, B. (1983) *Zasady lingwistycznej adaptacji metod kwestionariuszowych*. W: Paluchowski, W. Z zagadnień diagnostyki osobowości. Wrocław: Ossolineum.

SESJA 3 Prowadząca: Katarzyna M. Wyrzykowska (IFiS PAN)

GRZEGORZ LISSOWSKI (WYDZIAŁ SOCJOLOGII UNIwersYTETU WARSZAWSKIEGO)

Szczepienia przeciwko COVID-19 jako problem sprawiedliwości dystrybucyjnej

Dystrybucja szczepionek przeciwko COVID-19 należy do problemów tzw. sprawiedliwości lokalnej, których celem jest rozwiązywanie praktycznych problemów alokacji dóbr lub obciążeń w konkretnych sytuacjach. Dystrybucja szczepionek może być przedstawiana jako alokacja zerojedynkowa (otrzymanie przez pretendenta szczepionki można oznaczyć jako 1, a nie otrzymanie – jako 0). Taka alokacja zerojedynkowa musi spełniać pewne warunki, przede wszystkim musi być oparta na relacji pierwszeństwa. Metodą jej wyznaczania może być system punktowy.

Problem dystrybucji szczepionek przeciwko COVID-19 wyróżnia to, że ma on dwa wymiary: indywidualny i społeczny. Wymiar indywidualny polega na zapewnieniu osobie pewnego poziomu odporności na zarażenie wirusem SARS-CoV-2 lub przynajmniej łagodniejszy przebieg spowodowanej przez niego choroby. Wymiar społeczny polega na stworzeniu odpowiedniego poziomu odporności zbiorowej, który redukuje możliwość zakażenia się tym wirusem.

Przedstawiony zostanie uproszczony model szczepień, relacja pierwszeństwa założona w Narodowym Programie Szczepień oraz hipotetyczny system punktowy, który wyznacza taką relację pierwszeństwa.

JOLANTA PEREK-BIAŁAS, UNIwersYTET JAGIELLOŃSKI/SZKOŁA GŁÓWNA HANDLOWA W WARSZAWIE

PAWEŁ GRYGIEL, UNIwersYTET JAGIELLOŃSKI

PAULINA SKÓRSKA, UNIwersYTET JAGIELLOŃSKI

Doświadczenia z unikalnych badań panelowych wśród młodzieży krakowskiej w czasie pandemii

Obecnie w Polsce nie ma za wiele badań prowadzonych w formule badań panelowych, które pozwalają na „śledzenie” losów osób *przez całe życie*. W marcu 2021r. zapoczątkowano badania dotyczące starzenia (*ageing*) przez całe życie w ramach dofinansowania z Programu Inicjatywa Badawcza UJ (POB Society of Future). Opracowano pilotażowo metodologię badań panelowych dla rocznika osób, które w 2020 osiągnęły pełnoletność (rocznik urodzenia 2002) i są mieszkańcami Krakowa (uczęszczają do szkół średnich w Krakowie). Rocznik osób urodzonych w 2002 roku jest dość wyjątkowy ze względu na fakt, że jest to rocznik doświadczony przerwami w edukacji (m.in. strajk w 2019 roku, pandemia w 2020 i w 2021 – rok zdawania matury). Przerwy w edukacji formalnej w 2019 (strajk), w 2020 i w 2021 (z powodu pandemii) w tym ograniczone kontakty społeczne bezpośrednio mają i będą mieć wpływ na dalsze funkcjonowanie tej grupy osób, dodatkowo jest to rocznik, który w 2021 będzie podejmował ważne decyzje życiowe dotyczące dalszej edukacji lub pracy zawodowej. Celem projektu jest „uchwycenie” kohorty, która w 2020 osiągnęła *status osoby dorosłej* i rozpoczęcie unikalnych w skali kraju badań interdyscyplinarnych (ze względu na tematykę nie tylko jeden obszar jak to bywa często w badaniach, np. rynek pracy, edukacja, ale dodatkowo połączeniem kontekstu psychologicznego, zdrowotnego, mobilności itp.). Dodatkowo projekt opiera się nie tylko na badaniach sondażowych (internetowe ankiety/wywiady telefoniczne), ale też z wykorzystaniem

danych z telefonów komórkowych, w tym danych geolokalizacyjnych respondentów. Zakres tematyczny badań rocznika 2002 w trybie pilotażu (II edycja: kwiecień i lipiec 2021) został ustalony przez interdyscyplinarny Zespół we współpracy z zagranicznymi doradcami projektu. W czasie prezentacji zostaną przedstawione wyzwania metodologiczne związane z projektem, jak i rozwiązania wykorzystane w jego realizacji. Omówiony zostanie przyjęty model rekrutacji i budowy bazy uczestników oraz sposoby optymalizowania stopy zwrotu (*response rate*). Kolejnym wyzwaniem będzie utrzymanie panelu, co zostanie także zaprezentowane na zakończenie wystąpienia.

ZBIGNIEW SAWIŃSKI, INSTYTUT FILOZOFII I SOCJOLOGII PAN

Podejście pragmatyczne w harmonizacji badań

Przedmiotem wystąpienia jest propozycja o charakterze teoretycznym, sformułowana w celu przełamania niektórych ograniczeń występujących w harmonizacji badań z różnych krajów. Propozycja, nazwana podejściem pragmatycznym, nawiązuje do koncepcji *kontekstualizacji* (Hantrais, 1999), wskazującej na potrzebę szukania wyjaśnień zarówno w prawach uniwersalnych, jak też w mechanizmach specyficznych dla krajów. Obecnie stosowane metody harmonizacji skupiają uwagę przede wszystkim na uniwersalnych aspektach badanych zjawisk, proponując dla wszystkich krajów takie same narzędzia badawcze, w postaci klasyfikacji czy zmiennych docelowych (target variables). Jednak wspólne narzędzia z zasady stanowią wypadkową obrazu zjawisk w różnych krajach, co prowadzi do zagrożeń dwojakiego rodzaju. Po pierwsze, wspólne narzędzia mogą łączyć w tych samych kategoriach sytuacje o odmiennym znaczeniu dla interpretacji zjawiska. Ogranicza to rozwój wiedzy, jak również powoduje, że zebrane w ten sposób dane okazują niewielką przydatność dla badaczy krajowych, gdyż ich wiedza i ich potrzeby sięgają głębiej. Wywołuje to wątpliwości co do sensu wykładania przez kraje środków na udział w międzynarodowym projekcie. Po drugie, zharmonizowane narzędzia implikują istnienie pewnych podziałów lub reguł, jakie w danym kraju mogą w ogóle nie występować, bądź nie wykazywać istotnego znaczenia. Podważa to zaufanie do naukowej rzetelności badań porównawczych (research integrity).

Istota podejścia pragmatycznego polega na wydzieleniu w narzędziach harmonizacji dwóch komponentów, z których jeden odpowiada wspólnym aspektom badanych zjawisk, zaś drugi uwzględnia aspekty specyficzne dla krajów. Pozwala to zachować wspólne ramy dla międzynarodowych porównań, a zarazem daje możliwość uzyskania oczekiwanej rzetelności w wykorzystaniu wyników badań dla potrzeb krajowych.

Rozważania zilustrowane zostaną przykładami dotyczącymi harmonizacji wykształcenia. W badaniach porównawczych najczęściej stosuje się do tego celu Międzynarodową Standardową Klasyfikację Wykształcenia ISCED, którą niestety adaptowano do badań nie uwzględniając jej niejednakowego dopasowania do systemów szkolnych. W rezultacie łatwo wskazać przykłady nieuzasadnionych wniosków, jakie powstają w wyniku stosowania tego samego schematu ISCED dla różnych krajów. W ramach podejścia pragmatycznego proponuje się, aby narzędzie harmonizacji wykształcenia zawęzić do podziałów najwyższego rzędu, w tym wypadku do podziału wykształcenia na podstawowe, średnie i wyższe, a następnie każdemu z krajów pozostawić możliwość dalszego rozdzielenia tych kategorii, w zależności od faktycznych różnicowań szkolnych. Dane Europejskiego Sondażu Społecznego pozwalają pokazać, że podejście pragmatyczne spełnia dwa sformułowane wcześniej wymogi harmonizacji. Po pierwsze, daje możliwość uwzględnienia różnicowań edukacyjnych, jakie występują tylko w niektórych krajach, a przez to trudno je wkomponować w dość sztywny schemat klasyfikacji ISCED. Po drugie, podejście pozwala ominąć niektóre przyjęte w ISCED podziały czy zasady, co

nie wymusza interpretacji nie znajdujących odbicia w wiedzy na temat specyfiki systemów szkolnych.

Podejście pragmatyczne wymaga pewnych zmian w zasadach współpracy między krajami. Dominujący w międzynarodowych projektach model, w którym centrala decyduje o kształcie narzędzi badawczych, wymaga zastąpienia modelem koordynacji i wsparcia dla działań podejmowanych przez kraje w celu uwzględnienia swojej specyfiki we wspólnym projekcie.

Terminy kluczowe: badania porównawcze, harmonizacja badań, rzetelność badawcza, Międzynarodowa Standardowa Klasyfikacja Wykształcenia ISCED, Europejski Sondaż Społeczny

KRZYSZTOF BULKOWSKI, WYDZIAŁ SOCJOLOGII UNIWERSYTETU WARSZAWSKIEGO

Projektowanie procedur w porównawczych badaniach sondażowych. Przypadek piątej edycji European Values Study

Sondażowe badania porównawcze powinny charakteryzować się wysokim stopniem szczegółowości planowania całego procesu badawczego. W badaniach międzynarodowych zarządzanie całym projektem i organizacja pracy poszczególnych zespołów konsorcjum międzynarodowego przekłada się m.in. spójność kolejnych etapów badawczych, zrozumienie i poprawne wdrożenie szczegółowych procedur przez zespoły krajowe, a w efekcie dobrą jakość danych zbieranych w poszczególnych krajach. Celem stojącym za projektowaniem takiego procesu badawczego jest minimalizowanie czynników, które mogą zaburzyć porównywalność wyników między krajami. Osiąga się to poprzez tworzenie zestawów procedur służących standaryzacji całego procesu badawczego oraz dobrym zdefiniowaniu zakresu działań, które każdy z zespołów krajowych może samodzielnie zaprojektować.

Na przykładzie piątej edycji badania *European Values Study* zostaną omówione szczegółowe problemy związane z porównywalnością wyników w kontekście zaprojektowanych procedur. Omówione zostanie podejście proceduralne do tworzenia schematu doboru próby, wybierania techniki zbierania danych oraz wyznaczania terminów realizacji badania. Przeanalizowane zostaną również efekty tych procesów skutkujące problemami w jednoznacznym stwierdzeniu, że dane zebrane w poszczególnych krajach mogą być ze sobą porównywane.

Szczegółowo zostanie omówiona niespójność pomiędzy dwoma związanymi ze sobą etapami – projektowaniem schematu doboru próby oraz wyznaczaniem wartości wag w ostatecznym zbiorze danych. Poprawne zaprojektowanie pierwszego z tych etapów skutkowało wymogami nałożonymi na każdy z krajów, zapewniającymi dobrą jakość badania. Tworzone zgodnie z procedurami krajowe plany doboru próby zawierały szczegółowe informacje dotyczące definicji populacji, określenia operatu losowania oraz przyjętego schematu losowania. Wiązało się to z dostarczeniem m.in.: opisu zmiennych użytych do warstwowania, liczebności poszczególnych warstw w populacji i w projektowanej próbie, wyliczeń prawdopodobieństw inkluzji jednostek na każdym poziomie losowania, podaniem wielkości wiązek oraz wynikającej z całego schematu doboru próby przewidywanej wielkości *design effect*. Drugi z analizowanych etapów polegał na obliczeniu wag w każdym zbiorze krajowym. W wystąpieniu zaprezentowany zostanie system ważenia zaimplementowany przez konsorcjum międzynarodowe, w którym nie wykorzystano szeregu informacji dostarczonych przez zespoły krajowe na etapie przygotowania badania i zbierania danych. System ten charakteryzował się niepożądanymi własnościami, np. ujemnymi wartościami wag dla niektórych jednostek obserwacji. Efekty tego sposobu ważenia, przekazane zespołom krajowym półtora roku po realizacji badania wraz z upublicznieniem pierwszej wersji zbioru danych z 30 krajów, zostaną porównane z poprawionym systemem ważenia, który zastosowany został w ostatecznej wersji zbioru danych z 36 krajów upublicznionej dziewięć miesięcy później.

W wystąpieniu przedstawione zostaną konsekwencje nieoptymalnego zarządzania procesem przetwarzania danych w kontekście pracy zespołów krajowych nad upublicznieniem wyników badania. Dokonana zostanie również próba zdiagnozowania przyczyn powstałych problemów poparta sformułowaniem rekomendacji mających na celu uniknięcie podobnych problemów w przyszłości.

ANNA KIERSZTYN, WYDZIAŁ SOCJOLOGII UW

KATARZYNA KOPYCKA, WYDZIAŁ SOCJOLOGII UW

STEFAN BIEŃKOWSKI, INSTYTUT FILOZOFII I SOCJOLOGII PAN

Jak mierzyć prekaryjne zatrudnienie w perspektywie porównań międzynarodowych?

Celem wystąpienia jest nakreślenie metodologii operacjonalizacji i pomiaru prekaryjnego zatrudnienia w kontekście międzykrajowych badań porównawczych. Pomimo rosnącej prekaryzacji stosunków pracy obserwowanej w krajach rozwiniętych a wynikającej z procesów globalizacyjnych, w badaniach społecznych nadal nie została wypracowana jasna definicja tego zjawiska. W literaturze dominuje przekonanie o jego wielowymiarowości, jednak w praktyce badawczej nad przyczynami i/ lub konsekwencjami prekaryjnego zatrudnienia często zostaje ono sprowadzone albo do specyficznej formy kontraktu (najczęściej jest to umowa na czas określony lub zbliżone formy), albo do subiektywnej oceny respondenta co do jakości stosunku pracy, w którym pozostaje. Oparcie definicji prekaryjności na subiektywnej ocenie pracownika niesie ze sobą ryzyko nieporównywalności. W zależności bowiem od indywidualnego kontekstu biograficznego jak również różnej wrażliwości na ryzyko ekonomiczne respondenci mogą w sposób zróżnicowany oceniać tę samą sytuację pracy. Pomimo, że wymiar subiektywny niewątpliwie odgrywa istotną rolę w doświadczeniu prekaryjności, nie jest on wystarczającą podstawą na przykład do kształtowania polityki publicznej w obszarze zatrudnienia.

Z uwagi na wielowymiarowość zjawiska prekaryzacji wielu badaczy skłania się do operacjonalizowania go za pomocą szeregu obiektywnych wskaźników sytuacji zatrudnienia. Nie ma jednak jasności co do wagi poszczególnych elementów charakteryzujących stosunek pracy dla stopnia jego prekaryjności. Ponadto, zastosowanie wielowymiarowych definicji ograniczone jest często problemem dostępności danych. Szczególnie wyraźnie problem ten zarysowuje się w przypadku wielotematycznych panelowych surveyów społeczno-ekonomicznych, które umożliwiają analizy wzdłużne.

Zdecydowanie najczęstszą praktyką badawczą staje się zatem określenie pewnych form zatrudnienia jako prekaryjnych. Takie rozwiązanie niesie ze sobą szereg problemów. Po pierwsze niestandardowe formy zatrudnienia mogą służyć jako „pomosty”, czy „trampoliny” prowadzące do stabilnej pracy i tym samym spełniać funkcję integracyjną. Określanie ich zatem a priori jako prekaryjnych wydaje się nadużyciem szczególnie w międzynarodowych badaniach porównawczych. W zależności od krajowych rozwiązań prawnych jak i specyfiki narodowego rynku pracy będą one bowiem oddziaływać bardziej integracyjnie lub bardziej segregacyjnie. Po drugie, mnogość typów umów niestandardowych jak i różny poziom szczegółowości zbierania danych w wielkoskalowych badaniach narodowych będących najczęstszym źródłem danych dla ilościowych analiz karier zawodowych skutkuje często zbyt daleko idącym ujednoczeniem.

Wychodząc od diagnozy obecnej praktyki badawczej, oraz biorąc pod uwagę dostępność danych ilościowych dla międzynarodowych badań porównawczych historii pracy referat przedstawia propozycję operacjonalizacji prekaryjnego zatrudnienia jako pewnego typu ścieżki kariery zawodowej, a nie charakterystyki sytuacji pracy w pojedynczym punkcie czasowym.

Definiujemy prekaryjność jako wzór kariery, w którym dominują krótkie epizody pracy przeplatane okresami bezrobocia lub bierności zawodowej, powiązany z niskim dochodem. W celu uchwycenia sekwencji statusów zatrudnienia i okresów niskiego dochodu za pomocą jednej zmiennej metrycznej wykorzystujemy najnowsze rozwiązania w dziedzinie analizy sekwencji. Proponowana miara prekaryzacji ma formę złożonego indeksu biorącego pod uwagę: a) złożoność/ poziom wewnętrznego skomplikowania kariery, b) występowanie i natężenie stanów niepożądanych, oraz c) ich rozkład w analizowanej sekwencji. Zachowanie indeksu obrazujemy na przykładach hipotetycznych karier zawodowych.

24 września

SESJA 4 Prowadzący: Henryk Domański (IFiS PAN)

MAREK MUSZYŃSKI, INSTYTUT FILOZOFII I SOCJOLOGII PAN

PIOTR JABKOWSKI, WYDZIAŁ SOCJOLOGII UAM W POZNANIU

Wpływ kontekstu wywiadu na jakość pomiaru: analiza stylów odpowiedzi na pytania Portretowego Kwestionariusza Wartości Shaloma Schwartza w Europejskim Sondażu Społecznym, 2008-2019

Najbardziej rozpowszechnioną techniką gromadzenia danych w większości wiodących międzykrajowych sondaży porównawczych (takich choćby jak EB, ESS, EVS, ISSP, PIAAC, czy też WVS) pozostaje wywiad osobisty realizowany "twarzą w twarz" w fizycznej obecności ankietera i respondenta. Uzyskiwane dane pozostają tym samym podatne na obciążenie efektami ankieterskimi (*interviewer effect*) oraz respondenckimi (*response bias*) (por. Curran, 2016; He et al., 2020; Khorramdel & van Davier, 2014; Maniaci & Rogge, 2014), gdyż specyficzna relacja zachodząca pomiędzy: a) ankieterem, b) respondentem oraz c) kontekstem realizowanego wywiadu prowadzi do sytuacji, w której pomiar warunkowany może być określonym stylem odpowiadania (*response styles*), przejawiającym się choćby w stosowaniu przez respondenta środka skali, potakiwaniu oraz wybieraniu skrajnych wartości skali, niezależnie od kontekstu zadawanego pytania. Warto jednocześnie zauważyć, że o ile w literaturze metodologicznej poświęcono stosunkowo wiele uwagi efektom ankieterskim, respondenckim oraz stylem odpowiadania, to zadziwiająco niewiele było dotąd studiów nad wpływem kontekstu wywiadu na style odpowiadania.

W ramach proponowanego wystąpienia dokonamy poszerzenia dotychczasowych stanu wiedzy (por. Jacobs et al., 2019; Loosveldt & Beullens, 2017; Kosyakova et al., 2021; Valdenplas et al., 2018) o wpływ kontekstu wywiadu na jakość danych. Przedmiotem analizy będą style dopowiadania w *Portretowym Kwestionariuszu Wartości Shaloma Schwartza*, a bazą źródłową wyniki sześciu odsłon Europejskiego Sondażu Społecznego z lat 2008-2019. Prezentowane analizy oparte zostaną na skumulowanych zbiorach danych ESSu, jak również gromadzonych w tym projekcie para-danych w ramach tzw. kwestionariuszy dla ankieterów oraz formularzy kontaktu. Przyjmując hierarchiczną strukturę danych (w której respondenci zagnieżdżeni są w krajowych realizacjach próby), przeanalizujemy także czynniki makrostrukturalne (krajowe charakterystyki realizacji próby) warunkujące występowanie określonych stylów odpowiadania w krajach objętych pomiarem.

Jako potencjalne korelaty stylów odpowiadania (wskaźniki *response style* skonstruowane zostaną zarówno w oparciu o zmienne obserwowalne, jak i latentne; por. Plieninger & Heck, 2018), wybraliśmy na poziomie indywidualnym zmienne opisujące cechy: a) ankietera przeprowadzającego wywiad (wiek, płeć, obciążenie pracą w projekcie), b) respondenta (płeć, poziom wykształcenia, wiek, gotowość do udziału w badaniu (kooperatywność i dostępność; por. Stoop, 2004)), c) relacje pomiędzy cechami ankieterów i respondentów (interakcja płci oraz dystans wieku dzielący ankietera i respondenta), oraz d) style odpowiadania w pytaniach poprzedzających skalę Schwartza (por. Kim et al., 2018).

Wyniki analiz opartych na wielopoziomowych równaniach regresji liniowych (w których miary stylów odpowiadania traktowano jako zmienne zależne, a charakterystyki wywiadu jako zmienne niezależne) prowadzą do ciekawych obserwacji ukazując, że style odpowiadania

respondentów na pytania w *Portretowym Kwestionariuszu Wartości Shaloma Schwartza* wskazują na istotny negatywny związek z czasem trwania wywiadu, oraz na znaczne natężenie efektów respondenckich u badanych niechętnych wywiadowi. Do czynników istotnie warunkujących style odpowiadania zaliczyć należy również relacje pomiędzy cechami ankietera i respondenta (relacje płci oraz dystans wieku), a także poziom wykształcenia respondentów. Charakterystyki metodologiczne sondaży pozostają z kolei w mniejszym stopniu powiązane ze stylami odpowiadania w krajach objętych analizą, poza gratyfikacją respondentów znacząco poprawiającą jakość danych oraz zaangażowanie respondentów w proces wywiadu.

NATALIA JÓZEFACKA, UNIWERSYTET PEDAGOGICZNY W KRAKOWIE / INSTYTUT BADAŃ EDUKACYJNYCH

Modelowanie Latentnych Krzywych Rozwojowych - jako metoda analizy rozwoju kompetencji

Nieustannie zachodzące zmiany, niemalże we wszystkich obszarach życia, wymuszają na ludziach konieczność stałego podnoszenia poziomu wykształcenia, ciągłego poszerzania swoich kompetencji, w celu przystosowania się do zmiennej rzeczywistości. Taka sytuacja sprawia, że naukowcy są zainteresowani uchwyceniem charakteru tego zjawiska, obserwacją jak przebiega zmiana, czego dotyczy, jakie czynniki ją przyśpieszają lub spowalniają. Badanie zmiany przez wiele lat było domeną psychologów rozwojowych, to oni analizowali dane aby określić jak przebiega zmiana np. myślenia u dzieci, czy też kompetencji społecznych. Współcześnie również w innych obszarach psychologii badanie zjawiska zmiany jest interesujące np. w kontekście postaw społecznych, czy też podążając za ideą uczenia się przez całe życie (LLL, *lifelong learning*) rozwoju kompetencji.

Badania prowadzone w celu analizy zmienności najczęściej wykorzystują schemat badań podłużnych, do którego analizy klasycznie wykorzystuje się analizę wariancji z powtarzaniem pomiarem, jednocześnie coraz częściej w publikacjach. W publikacjach z zakresu psychologii rozwojowej metoda Modelowania Latentnych Krzywych Rozwojowych (Latent Growth Curve Modeling, LGCM), weszła jako złoty standard analizy badań w schemacie podłużnym. W innych obszarach, jest to wschodząca metoda, właśnie dlatego warto przybliżyć możliwości wykorzystania metody LGCM, zarówno w kontekście psychologii rozwojowej, jak również w kontekście psychologii pracy i organizacji. Główną zaletą LGCM jest możliwość analizy nie tylko trendu grupowego, ale również różnic indywidualnych w przebiegu zmiany. Metoda ta ma oczywiście również minusy, dlatego prezentacja zostanie podsumowana informacjami kiedy warto skorzystać z takiej metody i jakie wnioski pozwoli ona wyciągnąć badaczowi.

TOMASZ ŻÓŁTAK, INSTYTUT FILOZOFII I SOCJOLOGII PAN

Style odpowiedzi w ESS - zróżnicowanie ze względu na kraj, skalę i edycję badania

Badania sondażowe narażone są na szereg trudności metodologicznych. Jedną z nich jest przejawianie przez respondentów zróżnicowanych stylów odpowiedzi (ang. *response styles*), definiowanych jako skłonność do wybierania pewnych pozycji na skali odpowiedzi niezwiązana bezpośrednio z cechą, którą ma mierzyć dane pytanie (Wetzel i in., 2016). Do tej pory zjawisko stylów odpowiedzi w badaniach sondażowych badane było przede wszystkim w odniesieniu do kwestionariuszy wypełnianych samodzielnie przez respondentów (w ostatnich latach w szczególności w formie ankiet internetowych), jednak wydaje się ono istotne również w kontekście wywiadów prowadzonych przez ankietera. Choć styl odpowiedzi jest tu wypadkową

wzajemnego oddziaływania na siebie respondentą i ankietera (np. Verduyssen i in., 2017), jednak konsekwencje jego występowania, w postaci zniekształcenia wyników pomiaru, pozostają takie same.

W tym kontekście najbardziej odpowiednim podejściem do modelowania stylów odpowiedzi wydają się tzw. modele losowych progów (ang. *random thresholds*), w których style odpowiedzi odwzorowuje się jako zróżnicowanie pomiędzy respondentami wartości parametrów modelu IRT opisujących trudności poszczególnych pozycji skali odpowiedzi. Możliwych jest tu wiele konkurencyjnych specyfikacji: w niektórych z nich parametry opisujące styl odpowiedzi posiadają jasno określoną interpretację (np. skłonność do udzielania skrajnych albo środkowych odpowiedzi, Falck i Cai, 2016; Plieninger, 2017), inne zaś umożliwiają bardziej elastyczne dopasowanie modelu do wzorów występujących w danych, ale kosztem trudności interpretacji wzorów odpowiedzi (tzw. Divide-by-Total Models, Henninger i Meiser, 2020a, 2020b).

W związku z tym ważne wydaje się udzielenie odpowiedzi na pytanie, które z takich specyfikacji najlepiej pasują do danych zbieranych w dużych socjologicznych badaniach porównawczych, a w szczególności, czy zachodzi w tym zakresie zróżnicowanie pomiędzy różnymi skalami, krajami i edycjami badania. Ewentualne występowanie takiego zróżnicowania (zwłaszcza pomiędzy krajami) bardzo utrudniałoby kontrolowanie wpływu stylów odpowiedzi w analizach porównawczych. Z drugiej strony, otwierało by pole do analiz uwarunkowań występowania poszczególnych wzorców stylów odpowiedzi (np. Rammstedt i in., 2017).

W niniejszym wystąpieniu przedstawione zostaną wyniki analiz z wykorzystaniem danych Europejskiego Sondażu Społecznego (ESS), w których konkurencyjne specyfikacje modeli losowych progów estymowane były oddzielnie w odniesieniu do trzech różnych skal we wszystkich krajach i rundach badania. Opisane zostanie zróżnicowanie najlepiej pasujących do danych modeli stylów odpowiedzi ze względu na skalę, kraj i rundę badania oraz związków wzorców stylów odpowiedzi z podstawowymi charakterystykami społeczno-demograficznymi respondentów. Spróbuję również odpowiedzieć na pytanie, czy daje się w ramach tego zróżnicowania zauważyć jakieś wzorce.

MAŁGORZATA TALIPSKA, WYDZIAŁ POLONISTYKI UNIwersYTET WARSZAWSKI

Wyzwania badań ilościowych przeprowadzanych spośród użytkowników polskiego języka migowego (PJM) - narzędzie do zbierania odpowiedzi w PJM

Głusi, natywni użytkownicy polskiego języka migowego, w okresie swojego życia natrafiają wielokrotnie na różnego typu badania społeczne (i nie tylko). Problemem dla członków społeczności Głuchych jest bariera językowa, gdyż język polski jest dla nich językiem obcym. W badaniach jakościowych procedurę można przeprowadzić w polskim języku migowym (PJM) z wykorzystaniem kamer rejestrujących badanie – najlepiej gdy badacz sam jest głuchy i biegle posługuje się polskim językiem migowym. Są to jednak bardzo kosztowne badania oraz ograniczają znacznie liczbę respondentów, którzy mogliby wypowiedzieć się na dany temat. Wyzwaniem natomiast są badania ilościowe, gdzie narzędziem badawczym jest kwestionariusz. Dotychczas w przeprowadzanych badaniach techniką PAPI często zdarzało się, że badacz wyjaśniał (nieraz w języku polskim, czasem z udziałem tłumacza) badanym, o co chodzi w danym pytaniu. Sprawiało to, że nie zostały stworzone jednakowe warunki do badania dla wszystkich jego uczestników. Drugim problemem był brak swobody wypowiedzi uczestników badań w języku polskim, co powodowało skąpe, a często fałszywe (ze względu na ograniczone możliwości swobodnej wypowiedzi w języku polskim) odpowiedzi. W chwili obecnej przy wykorzystaniu coraz bardziej popularniejszej techniki CAWI tłumaczy się pytania zawierające zamkniętą kafeterię na PJM. Wadą tego rozwiązania jest możliwość przeważnie potwierdzenia

dotychczasowych wniosków. Otwarte pytania często pozostają bez odpowiedzi ze względu na konieczność umieszczenia tekstu w języku polskim. Podobnie jest w przypadku kafeterii półotwartej, gdyż łatwiej wybrać odpowiedź, która już się znajduje na liście, aniżeli wytworzyć tekst w języku obcym. Instytut Spraw Głuchych podjął się stworzenia narzędzia, które umożliwi pobieranie odpowiedzi bezpośrednio w PJM. Takie odpowiedzi badacz znający PJM będzie mógł przetłumaczyć, a następnie dokonać kodowania wypowiedzi. W przypadku braku znajomości PJM, badacz będzie mógł zlecić tłumaczenie profesjonalnym tłumaczom, którzy będą mieli przyporządkowanie odpowiedzi do poszczególnych pytań. W trakcie konferencji zostaną zaprezentowane główne założenia pluginu do aplikacji limesurvey, która umożliwi wyciąganie rzetelnych i trafnych wniosków z (dostępnych dla Głuchych) badań, a Głuchym swobodę wypowiedzi.

SESJA 5 Prowadząca: Kinga Zawadzka (IFiS PAN)

MACIEJ ALIMOWSKI, SZKOŁA DOKTORSKA NAUK SPOŁECZNYCH, UNIwersYTET W BIAŁYMSTOKU

SŁAWOMIR BARTNICKI UNIwersYTET W BIAŁYMSTOKU

MACIEJ GÓRECKI, UNIwersYTET WARSZAWSKI

MICHAŁ PIERZGALSKI, UNIwersYTET ŁÓDZKI

Ocena efektów zmian wprowadzonych w kodeksie wyborczym w zakresie organizacji wyborów organów wykonawczych gmin w 2018 roku - analiza kontrfaktyczna z zastosowaniem metody syntetycznej jednostki kontrolnej (SCM)

Ostatnia nowelizacja kodeksu wyborczego wprowadziła wiele zmian dotyczących wyłaniania organu wykonawczego gminy w wyborach bezpośrednich, tak w zakresie mechaniki samego aktu głosowania (np. definicja krzyżyka), jak i w zakresie systematyki ustalania wyników (wprowadzenie dwóch obwodowych komisji wyborczych czy scedowanie wielu kompetencji będących dotychczas w gestii wójtów na komisarzy wyborczych). Deklaratywne cele zmian obejmują zwiększenie kontroli nad procesem głosowania i liczenia głosów oraz transparentność procesu wyborczego.

Jednocześnie, dowody empiryczne w zakresie wyborów bezpośrednich organu wykonawczego gmin w Polsce w schemacie badawczym regresji nieciągłej typu ostrego (*sharp regression discontinuity*, dalej RDD), sugerują brak ciągłości w przyporządkowaniu jednostek obserwacji wokół punktu cięcia wg różnych cech osobniczych. Taka obserwacja jest niezwykle rzadka, dotychczas w literaturze opisano jedynie cztery przypadki tzw. nieciągłej alokacji cech wyborczych przy punkcie cięcia: wybory Izby Reprezentantów USA (Snyder 2005; Grimmer i in. 2011; Caughey i Sekhon 2011; Eggers i in. 2015); wybory gminnej egzekutywy w USA (Vogl 2014), w Turcji (Aksoy 2016) i w Polsce (Bartnicki i in. 2021). Wyniki te sugerują, że powodem takich obserwacji może być kontekst wyborczy nie związany z manipulacjami wynikiem i oszustwami wyborczymi; manipulacje i oszustwa wyborcze lub te ostatnie wynikające z kontekstu wyborczego.

Celem badania jest odpowiedź na pytanie: czy zmiany w organizacji wyborów gminnych organów wykonawczych w 2018 roku wpływają na anomalie wyborcze i wybrane zmienne charakteryzujące sytuację wyborczą.

W projektowanym badaniu opieramy się na modelu wnioskowania przyczynowego Neymana-Rubina; efekt przyczynowy szacuje się jako różnicę między dwoma potencjalnymi wynikami—wynikiem zaobserwowanym i tzw. kontrfaktycznym (wynik nieobserwowalny). W szczególności, do oceny efektu przyczynowego wykorzystujemy metodę syntetycznej jednostki kontrolnej (SCM). SCM pozwala oszacować siłę oddziaływania bodźca jako różnicę wyników między przeciętnymi wynikami zaobserwowanymi dla jednostek poddanych oddziaływaniu i wynikiem dla tzw. jednostki syntetycznej (Abadie i Gardeazabal 2003; Xu 2017). Najogólniej, jednostka syntetyczna tworzona jest jako średnia ważona jednostek z grupy kontrolnej.

Grupa poddana oddziaływaniu bodźca charakteryzuje się anomaliami RDD w dotychczasowym przebiegu wyborów, spełnia strukturalne warunki dla wysokiego prawdopodobieństwa wystąpienia manipulacji wyborczych (Lehoucq 2003) oraz zawiera zidentyfikowane czynniki towarzyszące anomaliiom RDD (specyficzne cechy osobnicze). Grupę kontrolną tworzą gminy w których nie zaobserwowano problemów z walidacją RDD w dotychczasowym przebiegu wyborów, dodatkowo pozbawione są warunków strukturalnych towarzyszących manipulacjom

oraz nie posiada zmiennych wykazujących anomalie. Mimo, że zmiany w organizacji wyborów w 2018 roku dotyczyły wszystkich gmin w kraju w tym samym zakresie, przyjmujemy, że wskazywana grupa kontrolna odporna jest na oddziaływanie bodźca (zmian w organizacji wyborów w 2018 roku), a jego wpływ możliwy jest jedynie w zakresie grupy eksperymentalnej.

PRZEMYSŁAW KUREK, UNIWERSYTET WARSZAWSKI

MAJA PIEŃKOS, UNIWERSYTET WARSZAWSKI

Mierzenie religijności – problemy, wyzwania i propozycje potencjalnych rozwiązań

Religijność Polaków jest zjawiskiem o potencjalnie ogromnym znaczeniu dla kształtowania się wielu procesów społecznych i gospodarczych, a jednocześnie problematycznym do zmierzenia. Trudność w jednoznacznym zdefiniowaniu oraz wielowymiarowość tego zjawiska prowokują pytania: Jak mierzyć religijność? Jak interpretować dostępne miary? Do jakich dodatkowych źródeł można sięgnąć w celu poszerzenia zakresu i podniesienia jakości badań? Niniejsza praca porusza problematykę pomiaru religijności w ujęciu przestrzennym. Szczegółowej analizie zostają poddane dotychczas używane miary, bazujące na zmiennych *dominantes* i *communicantes*, pochodzących ze zbiorów Instytutu Statystyki Kościoła Katolickiego oraz miary religijności utworzone na podstawie zmiennych, pochodzących ze spisów powszechnych GUS. Dodatkowo zestawienie obejmuje autorskie miary religijności w Polsce, które nie znalazły dotychczas szerszego zastosowania w literaturze: liczby i odsetki występujących w gminach zabytków o charakterze religijnym oraz ulic o nazwach posiadających religijne konotacje, a także dane przestrzenne dotyczące urodzeń, zgonów oraz aktywności świętych na ziemiach polskich. Wszystkie wymienione miary są poddane wstępnej analizie pod kątem ich jakości, przydatności w badaniach oraz możliwości interpretacji i potencjalnych interpretacyjnych błędów

KAZIMIERZ CZARNOCKI, UTRECHT UNIVERSITY

Dyskontowanie dystansu społecznego, altruizm i kooperacja w pojedynczym dylemacie więźnia

Celem tego artykułu jest (1) rozszerzenie modelu preferencji uwzględniających innych, aby uwzględnił nie tylko orientację społeczną graczy, ale również dystans społeczny między graczami, (2) użycie tego rozszerzonego modelu do analitycznego uzyskania hipotez odnośnie zachowania graczy w pojedynczym dylemacie więźnia (DW), (3) oraz empiryczny test tych hipotez.

DW z indywidualistycznymi graczami ma jedną równowagę Nasha, wzajemną defekcję. Dzieje się tak, ponieważ indywidualistyczny gracz zawsze zyskuje defektując. Pomimo tych teoretycznych implikacji, badania empiryczne pokazują, że ludzie kooperują w DW (Balliet, Parks i Joireman, 2009; Cooper i Kagel, 2015). Pytanie, dlaczego? Jednym z wyjaśnień jest to, że ludzie troszczą się nie tylko o swoje wypłaty, ale również o wypłaty innych z altruistycznych lub instrumentalnych pobudek.

Literatura badająca altruistyczne pobudki używa gier pojedynczych z preferencjami uwzględniającymi innych. To podejście (znane jako *social value orientation*) powstało około 50 lat temu w psychologii społecznej (McClintock, 1972; Griesinger i Livingston, 1977) i od tego czasu zostało przetestowane i rozwinięte w licznych publikacjach (Pletzer et al., 2018). Kooperacja w pojedynczej grze jest interpretowana jako przejaw altruizmu, ponieważ gracz nie

ma z niej żadnej zewnętrznej gratyfikacji (zakładając anonimowość graczy). Kooperuje on ze względu na wypłaty innych graczy.

Literatura badająca instrumentalne pobudki używa gier iterowanych, żeby stworzyć sytuację przypominającą trwałą relację między graczami. Taka gra może być opisana jako zanurzona (ang. *embedded*) (Axelrod, 1984; Raub, Buskens i Corten, 2014). Ponieważ do interakcji między graczami dochodzi wiele razy, mogą karać lub nagradzać zachowania innych graczy (nawet jeśli do ponownych interakcji dochodzi pośrednio (Buskens i Weesie, 2000)). Dlatego kooperacja w grze iterowanej jest interpretowana jako motywowana instrumentalnie, a nie przez altruistyczne pobudki. Gracz kooperuje, ponieważ ma w tym swój własny zysk.

Próby połączenia tych dwóch ścieżek były teoretyczne (m. in. Dijkstra i van Assen, 2017) albo polegały na uwzględnieniu orientacji społecznej w grach powtarzalnych (m. in. Moisan et al., 2018). Ten artykuł wnosi wkład do literatury przedmiotu, podejmując próbę empirycznego i teoretycznego połączenia tych dwóch ścieżek, poprzez zanurzenie gry pojedynczej w społecznym kontekście i jednocześnie zapewnienie anonimowości graczy. Zadajemy pytanie, czy dystans społeczny wpływa na decyzję o kooperacji w DW, i jeśli tak, to czy hiperboliczny model dyskutowania dystansu społecznego daje adekwatny opis tego wpływu.

Żeby odpowiedzieć na te pytania:

1. Rozszerzyliśmy model preferencji uwzględniających innych poprzez dodanie do niego hiperbolicznej funkcji dyskutowania dystansu społecznego. Parametry rozszerzonego modelu utożsamiliśmy z altruistycznymi i instrumentalnymi pobudkami do kooperacji.
2. Rozwiązaliśmy DW, w celu analitycznego uzyskania hipotez odnośnie wpływu poszczególnych parametrów modelu.
3. Użyliśmy danych z badania ankietowego przeprowadzonego online do testu predykcji.
 - a. Parametry modelu preferencji uwzględniających innych zostały wyestymowane przy użyciu algorytmów Levenberg–Marquardt (Moré, 1978) and SAEM (Comets, Lavenu i Lavielle, 2017).
 - b. Hierarchiczna uporządkowana regresja logistyczna z efektami stałymi została użyta do testu hipotez.

Wyniki wskazują, że dystans społeczny oraz zmienne utożsamione z altruistycznymi i instrumentalnymi motywami do kooperacji, wpływają na prawdopodobieństwo kooperacji. Jednak ze względu na znaczącą liczbę niekonkluzywnych efektów, potrzeba więcej danych, żeby odpowiedzieć, czy hiperboliczny model dyskutowania dystansu społecznego daje adekwatny opis tego wpływu.

Literatura:

Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books

Balliet, D., Parks, C., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A metaanalysis. *Group Processes & Intergroup Relations*, 12(4), 533-547.

Buskens, V., & Weesie, J. (2000). Cooperation via social networks. *Analyse & Kritik*, 22(1), 44-74.

Comets, E., Lavenu, A., & Lavielle, M. (2017). Parameter estimation in nonlinear mixed effect models using saemix, an R implementation of the SAEM algorithm. *Journal of Statistical Software*, 80(1), 1-41.

Cooper, D., & Kagel, J. (2015). Other regarding preferences: A selective survey of experimental results. In J.H. Kagel, A. Roth, (Eds.), *The Handbook of Experimental Economics* (Vol. 2).

Dijkstra, J., & van Assen, M. A. L. M. (2017). Explaining cooperation in the finitely repeated simultaneous and sequential Prisoner's dilemma game under incomplete and complete information. *The Journal of Mathematical Sociology*, 41(1), 1-25.

- Griesinger, D. W., & Livingston, J. W. (1973). Toward a model of interpersonal motivation in experimental games. *Behavioral Science*, 18, 173-188.
- McClintock, C. G. (1972). Social motivation: A set of propositions. *Behavioral Science*, 17, 438-454.
- Moisan, F., ten Brincke, R., Murphy, R. O., & Gonzalez, C. (2018). Not all Prisoner's Dilemma games are equal: Incentives, social preferences, and cooperation. *Decision*, 5(4), 306-322.
- Moré, J. J. (1978). The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical Analysis* (pp. 105-116). Springer, Berlin, Heidelberg.
- Pletzer, J. L., Balliet, D., Joireman, J., Kuhlman, D. M., Voelpel, S. C., & Lange, P. A. M. V. (2018). Social Value Orientation, Expectations, and Cooperation in Social Dilemmas: A Meta-analysis. *European Journal of Personality*, 32(1), 62-83.
- Raub, W., Buskens, V., & Corten, R. (2014). Social Dilemmas and Cooperation. In *Handbuch Modellbildung und Simulation in den Sozialwissenschaften* (p. 30). Springer.

JACEK HAMAN, KATEDRA MODELI FORMALNYCH I METOD ILOŚCIOWYCH W SOCJOLOGII,
WYDZIAŁ SOCJOLOGII UNIWERSYTETU WARSZAWSKIEGO

Co wyniki badań mówią o stanie polskiego systemu wyborczego?

Na ostatniej konferencji *Metodologiczne Inspiracje* przedstawiłem referat *Badanie społeczne – badanie obywatelskie: czy badania mogą przywrócić wiarę w wybory?* poświęcony projektowi społecznej obserwacji wyborów samorządowych w 2018 roku – referat został wygłoszony 25 września 2018, a więc niespełna na cztery tygodnie przed wyborami 21 października. Referat przedstawiał projekt będący podówczas w początkowej fazie realizacji, wskazywał na jego założenia i cele. Uczestnikom kolejnej konferencji należy się ciąg dalszy: co z tego wynikło, czy cele zostały osiągnięte, czy akcja z udziałem społecznych aktywistów pozwoliła zebrać użyteczne (także w sensie naukowym) dane o funkcjonowaniu systemu wyborczego w 2018 roku; wreszcie – czego dowiedzieliśmy się o przebiegu wyborów samorządowych? Badania ilościowe – a taki charakter miała akcja obserwacyjna – wymagają szczególnej staranności metodologicznej, a ta się raczej nie kojarzy z oddolnymi działaniami rozproszonych aktywistów. Okazuje się jednak, że przy odpowiedniej koordynacji i podporządkowaniu oddolnych aktywności scentralizowanemu planowi możliwe jest spełnienie wysokich standardów metodologicznych. Sukces obserwacji wyborów oznacza więc również otwarcie nowej możliwości dla prowadzenia badań ilościowych dotyczących zagadnień ważnych dla bieżącej, publicznej debaty.

Obserwacja wyborów w 2018 to jednak tylko element szerszego planu badań nad aktualnym stanem systemu wyborczego w Polsce, podejmowanych systematycznie przez członków zespołu ekspertów wyborczych Fundacji im. Stefana Batorego. Poza szeregiem drobnych (lub mniej drobnych) analiz podejmowanych ad hoc, na szersze omówienie zasługuje także badanie urzędników wyborczych podjęte w 2019 roku (powołanie Korpusu Urzędników Wyborczych było istotnym *novum* wprowadzonym przez nowelizację Kodeksu Wyborczego w 2018 roku) – mimo zupełnie innego podejścia metodologicznego i objęcia nim innego elementu systemu wyborczego, wyniki tego badania prowadzą do wniosków pod pewnymi względami podobnych, jak wyniki obserwacji wyborów 2018.

Zaangażowanie naukowców w sprawy publiczne może przynieść korzyść, jeśli będzie opierało się na zasadach obowiązujących w nauce: a więc, z jednej strony, gdy głoszone opinie i wnioski będą się opierały na solidnych podstawach metodologicznych, z drugiej zaś, gdy zachowana będzie elementarna bezstronność, a działania będą podejmowane *sine ira et studio*. To, że nie są to wyłącznie hasła, można pokazać na przykładzie konkretnych, pozytywnych efektów omawianych wcześniej działań – i temu poświęcona będzie ostatnia część referatu.

A. JAN KUTYŁOWSKI, UNIVERSITY OF OSLO, NORWAY

Analityczne i praktyczne rozważania dotyczące badań preferencji i "wartości" w sondażach, ze szczególnym zastosowaniem do pojęcia "postmaterializmu"^a

(prezentacja będzie wygłoszona w języku polskim)

A very common fallacy follows from (typically) tacit reconstitution of variable(s) of interest, measured at the ordinal scale of measurement, into variable(s) at interval scale of measurement. Arbitrary elevation of the measurement scale from ordinal to interval, described elsewhere by the author as *Quantum Leap Fallacy* (QLF) puts the practice involving the unsound methods in jeopardy with respect to their grounding in scientific knowledge, and thus also its sincerity or legality (when it leads to policy advice).

Main part of the presentation shall deal QLF in the context of experimental schemes for elicitation of preferences underlying the award-winning concept of "postmaterialism", developed over years by R. Inglehart et al. (1990, 1997, 2003) and subsequently extended to a broader context of inter-temporal "changes in values" (Inglehart and Welzel 2005, Welzel 2014).

These schemes for elicitation of preference will be re-formulated and formalized in a probabilistic manner, so as to facilitate flexible modelling of a single or multiple ordinal categorical variables. Consequently, a sound test of the hypothesis of "postmaterialism" will be formulated.

It will be shown that this has not been recognized by the original authors and their epigones. Some extensions of the proposed formalizations will be discussed, in a substantively different context of M. Kohn's approach to differential patterns of socialization of children.

Time permitting, two general concepts related to celebrated distinctions by Steven's concerning levels of measurement will be put forward: *Steven's Razor* and *Steven's Ordinal Fallacy*, the latter arising from two different concepts of the ordinality of variables, as elucidated by results derived in the main part of the intended presentation.

REFERENCES

Inglehart, R. (1990). *Culture Shift in Advanced Industrial Society*. Princeton, NJ: Princeton University Press.

Inglehart, R. (1997). *Modernization and Postmodernization. Cultural, Economic, and Political Change in 43 Societies*. Princeton, NJ: Princeton University Press.

Inglehart, R., ed. (2003). *Human Values and Social Changes. Findings from the Values Surveys*. Leiden: Brill.

Inglehart, R., Welzel, C. (2005). *Modernization, Cultural Change, and Democracy. The Human Development Sequence*. Cambridge: Cambridge University Press.

Welzel, C. (2014). *Freedom Rising. Human Empowerment and the Quest for Emancipation*. Cambridge: Cambridge University Press.

^aResearch reported here originated in part in connection with financial support through Marie Curie Senior Fellowship for Transfer of Knowledge, funded by the European Commissari