

GRAŻYNA WIECZORKOWSKA-WIERZBIŃSKA

ZROZUMIEĆ DANE
I STARAĆ SIĘ ROZWIĄZAĆ PROBLEM

Zgadzam się, że jako czynni psychologowie, badacze i nauczyciele akademicy powinniśmy nie tylko dbać o nasze publikacje, ale także zastanawiać się nad kierunkiem rozwoju naszej dyscypliny. Stąd moje uznanie dla wysiłków prof. Brzezińskiego, których efekty on sam podsumował w swoim artykule opublikowanym w niniejszym numerze *Roczników Psychologicznych*. Wiele uwagi poświęcił on testowanym teoriom. Moje spostrzeżenia także wykraczają poza wąsko rozumianą analizę statystyczną.

Zajęci spełnianiem uznanych w środowisku standardów, ignorujemy **zmiany**, które zachodzą w naszym **otoczeniu informacyjnym**. Uważam, że już niedługo czeka nas zmiana jakościowa w naszej dyscyplinie, która dokona się (już dokonuje?) nie tyle pod wpływem nowych teorii, co rozwoju technologii, tak jak się to stało na przykład w medycynie. Zgadzam się całkowicie z opinią Profesora, że mamy zalew spełniających standardy, ale niewiele wnoszących do naszego rozumienia rzeczywistości publikacji. Przyjęte kryteria ilościowe oceny naszego dorobku skłaniają do niepohamowanego publikowania – bez względu na to, czy coś istotnego mamy do powiedzenia. Liczba przeprowadzonych badań przekracza możliwości integracji bez wprowadzenia zmian w **standardach publikacji**, które nie zmieniły się istotnie od ery przedkomputerowej. Podział publikacji na część merytoryczną i część techniczną umieszczoną w internecie pozwoliłby na wprowadzenie obowiązku załączania surowych danych, dzięki którym zainteresowani mogliby dokonać dużo lepszej integracji niż wymyślona w poprzedniej epoce **metaanaliza wartości statystyk**. Obecnie wyłonienie

ważnych myśli z rosnącej lawinowo liczby tekstów, które musimy przeczytać, wymaga ignorowania wielu informacji, np. nazwisk autorów podawanych w środku tekstu, a to redukuje zasoby poznawcze Czytelnika (Gilbert, Tafarodi, Malone, 1993). Za dużo jest w naszych publikacjach nic nie wnoszących do toku rozumowania odniesień. Przy takim skoku ilościowym sprawozdawanie rozwoju badań nad danym zagadnieniem powinno być pozostawione historykom psychologii. Więcej uwagi powinniśmy poświęcać rozwiązaniu problemu niż historii wcześniejszych prób. Niespotykany wcześniej dostęp do bezgranicznego oceanu informacji bardzo często nie wpływa niestety pozytywnie na głębokość ich przetwarzania.

Od lat głoszę tezę (Wieczorkowska, Wierzbiński, 2011), że **psychologia dla nauk społecznych i humanistycznych pełni taką samą rolę, jak matematyka dla nauk ścisłych**. Wielokrotnie na konferencjach na przykład zarządzania projektami, edukacji internetowej czy dietetyki irytował mnie całkowity brak wiedzy psychologicznej badaczy. Powodem, dla którego nie możemy „wymusić” u przedstawicieli innych dyscyplin (np. prawa, nauk politycznych czy historii) znajomości psychologii, jest **brak systematyzacji pojęć**. Nawet w publikacjach w fachowych czasopismach unikamy używania „żargonu” naukowego, co było przez lata wzmacniane obowiązkiem publikacji prac habilitacyjnych w formie książki dostępnej dla szerszej klasy czytelników. Prawdą jest, że posługujemy się **pojęciami naturalnymi**, więc ich definiowanie w sposób klasyczny nie jest możliwe i na szczęście prace przeglądowe różnych definicji zniknęły z pola naszych zainteresowań. Porównania wniosków muszą być jednak prowadzone na poziomie **operacjonalizacji**, których niestety wystandaryzować się nie da. Manipulacja zagrożeniem musi być inna, gdy jest stosowana w grupie studentów czy pacjentów szpitala. Także standardowe techniki pomiaru zależą od próby. Przykładem może być pozycja z kwestionariusza przedziałowości „Jestem nonszalanccki w traktowaniu szczegółów”, która zmieniła diametralnie swoją korelację ze skalą w grupie pracowników Wydziału Farmacji. Nonszalancja w traktowaniu szczegółów dla farmaceuty ma zupełnie inne znaczenie niż dla studenta. Nie możemy zapominać, że **obiekt naszych badań** jest dużo bardziej **skomplikowany i reaktywny** niż na przykład w naukach biologicznych, na których wzorujemy naszą metodologię. Aby zilustrować to studentom, na wykładzie z medycyny psychosomatycznej przeprowadziłam „trening relaksacyjny”. Prawie 300 osób siedzących na sali było proszonych o podążanie za instrukcjami podawanymi w tle relaksującej muzyki. Nie tylko obserwowałam reakcję, ale także poprosiłam ich o ocenę swojego stanu za pomocą wyboru jednej z sześciu odpowiedzi. Oprócz opcji oznaczających całkowity lub częściowy sukces w wykonaniu tego

zadania – zgodnie z moimi oczekiwaniami – 8% wybrało odpowiedź: „nie miałem ochoty się relaksować, więc nie próbowałem”, a 9% wykazało efekt kontrastu, stwierdzając: „próba relaksacji wywołała moją irytację”. Wariancja odpowiedzi umysłu jest dużo większa niż wariancja biologicznej odpowiedzi organizmu (podanie leków usypiających nie wywołałoby efektu kontrastu). To oznacza, że modele nie uwzględniające różnic indywidualnych i kontekstu sytuacyjnego nie mogą wyjaśniać wiele wariacji wyjaśnianych zmiennych. Koncentrując się w naszych analizach statystycznych na miarach tendencji centralnej, nie dostrzegamy często efektów kontrastu. Stąd tak niskie **wielkości efektów**, o których pisze Profesor. Powinniśmy zdecydowanie zmniejszyć poziom ogólności naszych teorii. Psychologowie zazdroszczą fizykom i chcieliby opracować tak proste reguły funkcjonowania umysłu, jak trzy zasady dynamiki Newtona, ignorując w ten sposób fakt, że nasz umysł rozwija się w odpowiedzi na pojawiające się w otoczeniu wyzwania – jest więc raczej zbiorem modułów wyuczonych reakcji na różne klasy (!) bodźców. Poważnym problemem jest też rozdział między badaniami naukowymi stosującymi głównie podejście **nomotetyczne** a praktyką, która wymaga wiedzy **idiograficznej**. Duże przyrosty naszej wiedzy nie mają – jak dotąd – przełożenia na zwiększenie skuteczności oddziaływań praktycznych. Dużo więcej wiemy, ale nadal niewiele więcej potrafimy. I tu znów czekamy na rozwój technologii, która umożliwi wielokrotny pomiar tej samej osoby (ilościowe porównawcze studia przypadków (Siarkiewicz, 2007) bez wywoływania jej irytacji. W tej chwili zmienna CZAS jest wielką nieobecną w badaniach psychologicznych. I tu jedyna uwaga polemiczna z Profesorem, który podkreśla konieczność przestrzegania **zasady niesprzeczności paradygmatycznej**. Uważam, że najważniejsze jest otrzymanie interesujących poznawczo wyników, więc nie mam nic przeciwko eklektyzmowi – łączeniu różnych szkół. Gdy zaczynałam przed 30 laty moją pracę naukową, mówienie o procesach nieświadomych było odrzucane przez psychologię poznawczą, dziś nikt nie ma wątpliwości, że przetwarzanie informacji odbywa się w przeważającej części poza świadomością. Zdecydowanie więc opowiadam się za koncentracją na rozwiązywaniu problemów i wykorzystywaniu tego, co wartościowe w różnych szkołach. Zgadzam się natomiast całkowicie z Profesorem na temat **wagi replikacji**. W psychologii za dużo mamy twórczości (ciągłego wprowadzania nowych pojęć), za mało rzemiosła. Nie jest tajemnicą, że najłatwiej osiągnąć „sukces”, wprowadzając nowe pojęcie i publikując narzędzie pomiaru. Jeżeli mamy szczęście, wiele osób zacznie korelować nowe narzędzie z istniejącymi, aż fala zainteresowania opadnie, bo wynik tych usiłowań można łatwo przewidzieć. Nawet studenci pierwszych lat studiów nakłaniani są, aby wymyślali sobie problem, operacjonalizację zmien-

nych... Na wykładach z metodologii powtarzam do znudzenia: replikować, replikować i jeszcze raz replikować, bo prawie żadna replikacja w psychologii nie jest banalna (Aronson, Wieczorkowska, 2001). Studenci najpierw powinni nauczyć się uzyskiwać przewidywalne – na podstawie innych badań – wyniki, a dopiero potem samodzielnie wychodzić dalej.

Jeśli chodzi o **analizę statystyczną danych**, to ograniczę swoje uwagi do **badania korelacyjnych**. O ile badania eksperymentalne mają dobre i oczywiste standardy analizy, to w badaniach korelacyjnych – jak słusznie zauważa Profesor – publikujemy sporo wyników o wątpliwej ważności. Wynika to przede wszystkim z niskiej trafności wewnętrznej, będącej immanentną własnością badań korelacyjnych i przejawiającej się w splataniu (*confounding*) wpływów różnych zmiennych. Prawdą jest jednak, że wielu zagadnień nie da się zbadać eksperymentalnie i powinniśmy bardzo intensywnie pracować nad nowym modelem analizy badań korelacyjnych. Jest to tym ważniejsze zadanie, że istnieją ogromne zbiory ogólnodostępnych badań prób reprezentatywnych, zgromadzonych dużym nakładem kosztów, jednak w niewielkim stopniu wykorzystywanych poznawczo. Zgadzam się w stu procentach z opinią Profesora, że **analizy wielozmiennowe są koniecznością**. Problem jednak tkwi w ich wrażliwości na trafność modelu. Jak to powtarzam do znudzenia moim studentom, „współczynniki regresji zależą od towarzystwa”. Tylko nierozumiejący istoty tych analiz mogą stosować krokową analizę regresji (pozostawiając decyzje merytoryczne algorytmom statystycznym), pisać z pełnym przekonaniem, że analiza czynnikowa, analiza skupień dowiodła istnienia *k* czynników czy skupień. Zbyt często usunięcie jednej zmiennej zmienia strukturę macierzy korelacji. Na moich studentach wielkie wrażenie robi analiza przykładu (Wieczorkowska, Wierzbiński, 2011), w którym stwierdzają na tych samych danych różne zależności między X a Y w zależności od tego, które zmienne dodatkowe zostały kontrolowane. Oczywiście tam, gdzie jest jasna teoria, lista zmiennych jest wyznaczona jednoznacznie. W analizach danych sondażowych na próbach reprezentatywnych, gdy kontrolowanie zmiennych socjodemograficznych jest koniecznością, często zapomina się o wprowadzaniu do modelu efektów interakcyjnych. Przykładowo, sprawdzając socjodemograficzne predyktory poziomu wykształcenia (operacjonalizowanego przez lata nauki) stwierdzimy brak istotnego związku z płcią, bo istotnym predyktorem jest interakcja wieku i płci. Przewaga edukacyjna starszych mężczyzn zamienia się w przewagę młodszych kobiet, dlatego efekt główny płci nie jest istotny. Podobną zależność stwierdzono w 14 z analizowanych 33 krajów (Wierzbiński, 2009). Ale to wymagało wprowadzenia do modelu regresyjnego efektów interakcyjnych, czego najczęściej badacze nie robią.

Ważny fragment artykułu Profesora dotyczy wielkości próby. O ile testy istotności (oparte przecież na przedziałach ufności) biorą poprawkę na wielkość próby, to sama wartość współczynnika determinacji wielkości próby nie uwzględnia. Dla ilustracji – studentów namawiam do losowania z całej próby PGSS ($N = 15\ 776$) podprób o wielkości 1% (około 150 osób), 2% itd. Mogą się wtedy przekonać, że w małych 150-osobowych próbach można uzyskać dużo większe (ale także i mniejsze) współczynniki determinacji, niż by to wynikało z zależności w „populacji”, gdzie $r^2 = (0,25)$. Pokazuje to im, że nasze zaufanie do statystyki zwiększa się wraz ze wzrostem liczebności próby, więc próba nie może być zbyt duża, choć sądzę, że duże próby najlepiej wykorzystywać do **bootstrapowej** (Davison, Hinkley, 1997) metody analizy, a nie testowania prostych zależności. Nie wiem, dlaczego nie jest stosowane na przykład zalecenie, aby równanie regresji ustalać na połowie dostępnych danych, a testować je na reszcie (Pedhazur, 1997). Pozostaję z nadzieją, że publikacja tego numeru *Roczników* zintensyfikuje nasze usiłowania poszukiwania nowych modeli testowania wielkich zbiorów danych.

LITERATURA CYTOWANA

- Aronson, E., Wieczorkowska, G. (2001). *Kontrola naszych myśli i uczuć. Rozdział 4. Jak możemy zdobyć odpowiedź na nurtujące nas pytania?* Warszawa: Jacek Santorski & Co.
- Davison, A. C., Hinkley, D. V. (1997). *Bootstrap methods and their application*. New York: Cambridge University Press. Cambridge Series in Statistical and Probabilistic Mathematics.
- Gilbert, D. T., Tafarodi, R. W., Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology*, 65, 221-233.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research*. Belmont: Wadsworth Publishing.
- Siarkiewicz, M. (2007). Zastosowanie metody analizy szeregów czasowych do badania zależności pomiędzy składem posiłków a subiektywnie odczuwanym poziomem energii. W: K. Winkowska-Nowak, A. Nowak, A. Rychwalska (red.), *Modelowanie matematyczne i symulacje komputerowe w naukach społecznych* (s. 15-20). Warszawa: Wydawnictwo Academica SWPS.
- Wieczorkowska-Wierzbińska, G. (2011). *Psychologiczne ograniczenia*. Warszawa: Wydawnictwo Naukowe Wydziału Zarządzania UW.
- Wieczorkowska, G., Wierzbiński, J. (2011). *Statystyka: od teorii do praktyki*. Warszawa: Wydawnictwo Naukowe Scholar.
- Wierzbiński, J. (2009). *Badanie zaufania do organizacji: problemy metodologiczne*. Warszawa: Wydawnictwo Naukowe Wydziału Zarządzania UW.