

16 lipca 2022

prof. UAM dr hab. Michał Klichowski
Pełnomocnik Rektora ds. Otwartej Nauki
Kierownik Pracowni Badań nad Procesem Ucznienia Sie
Wydział Studiów Edukacyjnych
Uniwersytet im. Adama Mickiewicza
w Poznaniu

RECENZJA

rozprawy doktorskiej Marcina Radziwołka, pt. *Skuteczność strategii profilaktyki szkolnej w zakresie zachowań ryzykownych młodzieży – częstochowskich szkół średnich*

1. Wprowadzenie

Dysertacja przygotowana przez Marcina Radziwołka pod kierunkiem Małgorzaty Przybysz-Zaremby, pt. *Skuteczność strategii profilaktyki szkolnej w zakresie zachowań ryzykownych młodzieży – częstochowskich szkół średnich*, jest pracą, którą nazwałbym klasycznym doktoratem z pedagogiki przygotowywanym w Polsce. Manuskrypt jest bowiem bardzo długi (ma blisko 450 stron!), zawiera bardzo rozbudowane tło teoretyczne (165 stron), znacząco szersze, niż problematyka realizowanych w ramach doktoratu badań, treści nie są raczej wizualizowane (część teoretyczna zawiera tylko 5 rycin, a część empiryczna zaledwie – co bardzo zaskakujące – 2 ryciny), ale są obudowane „przerazającą” liczbą tabel (w tej pracy jest ich 137!). Opis zastosowanych metod jest tu też swoisty – przypomina raczej traktat metodologiczny (rozważa się tu nawet to, czym jest metodologia), niż dogłębny protokół własnych badań; mamy tu też kilkustronicową listę problemów badawczych, przy bardzo zdawkowym opisie procedury analiz statystycznych. Doktorant bazuje też głównie na polskiej literaturze i najważniejsze kwestie definiuje opierając się na lokalnych pracach naukowych. Jednakże nic z tych wymienionych przeze mnie elementów nie stanowi zarzutu do Autora *sensu stricto*, ale pokazuje to tylko, że Autor jest silnie osadzony w przyjętym w polskiej pedagogice stylu uprawiania nauki i nie należy do grona śmiałków, którzy w ostatnim czasie przełamują ten model, piszą swoje doktoraty w bardziej uniwersalnej formie, co upodabnia nasze

pedagogiczne dysertacje do tych pisanych za granicą, lub w Polsce, ale przez np. psychologów czy kognitywistów.

Podjmując się merytorycznej oceny manuskryptu chciałbym odnieść się do dwóch aspektów: ulokowania rozważań w aktualnej debacie naukowej i poprawności przeprowadzonego wnioskowania.

2. Ulokowanie rozważań w aktualnej debacie naukowej

Praca Marcina Radziwołka znakomicie wpisuje się w aktualne potrzeby polskiej pedagogiki i jest w pedagogice doskonale osadzona. Autor z lekkością rekonstruuje aktualny dyskurs edukacyjny, z łatwością też zestawia ze sobą klasyczne dzieła pedagogiczne z nowymi opracowaniami. Potrafi także nie tylko odtwarzać wnioski innych autorów, ale też sprawnie tłumaczyć ich koncepcje, wyróżniać najważniejsze ich elementy i uwypuklać krytyczne lub kontrowersyjne aspekty bazowych dla swoich rozważań teorii. Autor napisał też swoją pracę z dbałością edytorską – nie ma w niej edycyjnych potknięć, wszystko jest czytelne, przejrzyste i – jestem o tym przekonany – bardzo rzetelnie sprawdzone. Ponadto, problem badawczy pracy jest ważny i aktualny. Badania Autora są więc uzasadnione i potrzebne. Doktorant wyraźnie poczuł „ważność” swoich analiz, bowiem oprócz obszernych konkluzji, przedstawił także praktyczne rekomendacje, które można niejako „od razu” użyć w codzienności edukacyjnej. Reasumując, w kontekście ulokowania rozważań w aktualnej debacie naukowej, wystawiam ocenę celującą.

3. Poprawność przeprowadzonego wnioskowania

Jeśli chodzi o poprawność przeprowadzonego wnioskowania, już na wstępie – nadal pozostając w takiej szkolno-oceniającej atmosferze – muszę wskazać, iż mogę ocenić je tylko na „czwórkę z plusem”. Już bowiem czytając – wspomniany powyżej, nieco zbyt zdawkowy – paragraf 3.6 (*Sposoby analizy materiału empirycznego*), napotykamy na pewien problem. Autor nie podaje, w jakim programie statystycznym dokonywał obliczeń, ale wskazuje wzór na χ^2 (po co?). Przed wzorem podaje, że przyjął wartość krytyczną p na poziomie 0.001 (czyli $\alpha = 1 - 0.999$ i błąd nie większy niż 0.1%). Powstaje pytanie: dlaczego? W naukach społecznych zwyczajowo przyjmuje się 0.05 (5%), ewentualnie 0.01 (1%). Piszą o tym np. George A. Ferguson i Yoshio Takane, w cytowanej przez Autora książce pt. *Analiza statystyczna w psychologii i pedagogice*. Ponadto, analizując interpretacje Autora, okazuje się, że przyjął on jednak wartość 0.05

za krytyczną, bowiem np. na stronie 280, przy $p = 0.047$, wynik uznaje za istotny statystycznie. Niemniej, wartości teoretyczne odczytuje dla $\alpha = 0.001$. Tu pojawia się kolejny problem. Autor, w punkcie 3.6 wspomina o poprawce na ciągłość i o tym, kiedy ją będzie stosował. Jednak później mylnie identyfikuje sytuacje, kiedy taka poprawka jest potrzebna. W zasadzie Autor powinien z niej korzystać we wszystkich swoich analizach, bowiem – choć nie wiem, dlaczego – cały czas dąży do doprowadzania danych do postaci tabeli 2×2 , czyli najbardziej specyficznej sytuacji z możliwych (tabela 2×2 w statystyce jest czymś podobnym do kwadratu w geometrii) i wymagającej podania statystyk chi-kwadrat (χ^2) z poprawką Yatesa. Tym samym, mamy tu do czynienia w gruncie rzeczy z analizami *post hoc*, ale niepoprzedzonymi analizami efektów głównych. I tak Autor wszystkie zmienne niezależne doprowadza – *post factum* – do rozkładu dwukategorialnego, czy, innymi słowy, powoduje, że mają one tylko dwa poziomy. Choć np. zmienna „Wiek” ma w jego badaniu 5 opcji (15, 16, 17, 18, 19 lat), to w analizach ma dwie („do 17 lat” oraz „18 i więcej”). Tak samo jest ze zmienną „Struktura rodziny” – badani mogli wybrać z czterech opcji, a Autor połączył odpowiedzi tak, by tworzyły dwie kategorie. Podobnie rzecz się ma ze zmiennymi zależnymi, tylko że Autor dokonał tu jeszcze dziwniejszego zabiegu. Otóż, jeśli badany mógł wybrać np. w pytaniu nr 6, jaki rodzaj alkoholu pije (piwo, wino, wódka), a Autor chciał sprawdzić, czy płeć różnicuje zwyczaj picia któregoś z nich, przeprowadził analizę 2×2 dla każdego typu alkoholu, zamiast zbudować tabelę 2×6 (płeć \times piję piwo/nie piję piwa, piję wino/nie piję wina, piję wódkę/nie piję wódki), która pozwoliłaby określić, czy płeć różnicuje preferencje odnośnie typów trunków, a jeśli tak, to później, w analizie *post hoc*, sprawdzić, względem których. Dalej, np. poprzez wielomianową regresję logistyczną, należałoby sprawdzić, czy płeć różnicuje picie wina w takim samym stopniu jak wiek i miejsce zamieszkania oraz struktura rodziny, czy może dla interakcji płeć \times wiek \times miejsce zamieszkania \times struktura rodziny, któryś z tych czynników dominuje. Trzeba jeszcze podkreślić, i jest to krytyczna kwestia, że Autor odczytuje błędne wartości stopni swobody – w tabeli 2×2 $df = 1$. Skąd zatem np. $df = 4$ przy analizie picia wina przez kobiety? Autor nieprawidłowo przyjmuje zatem też wartości krytyczne rozkładu χ^2 . W tym przykładzie przyjął 18.46, a powinien przyjąć 10.83, a w gruncie rzeczy 3.84 (dla poziomu istotności $\alpha = 1 - 0.95$). Przed publikacją wyników należy zatem rozważyć ponownie sensowność łączenia kategorii (tak, by zawsze były dwa poziomy zmiennej), przyjąć $\alpha = 0.05$ lub podać uzasadnienie dla $\alpha = 0.001$ (czasem stosuje się taki zabieg w badaniach medycznych) i „wrzucić” dane do jakiegoś programu statystycznego (np. darmowego i stosowanego obecnie na całym

świecie *jamovi* – <https://www.jamovi.org>), tak, by nie zestawiać „ręcznie” wyników testów z wartościami z tabel statystycznych, ale by program „sam” to zrobił, nie narażając Autora na błędne odczytanie liczby stopni swobody. Należy też wyczulić Autora, że opisując wyniki bazuje on na wartościach bezwzględnych, a jego grupy porównawcze nie są zbalansowane. Np. kobiet i mężczyzn nie jest tyle samo. Tyle samo nie jest też osób w grupie „do 17 lat” i „18 i więcej”. Tym samym, w takiej próbie badawczej, 228 pijących kobiet to nie to samo co 228 pijących mężczyzn. Koniecznie trzeba więc podać jakieś indeksy lub po prostu wartości procentowe z danej grupy, tak, by dane można było ze sobą porównywać. Autor powinien też pamiętać, że p nie może być < 0 , a Autor czasem tak podaje. I jeszcze jedna uwaga: Autor nazywa bardzo często analizy z zastosowaniem testu χ^2 analizami korelacji. Pisze np. „...potwierdzono istotną statystycznie korelację pomiędzy picią a spożywaniem alkoholu...” i podaje wartość χ^2 . To duże nieporozumienie, bo test χ^2 to test niezależności, a nie korelacji. Tytułu tabel przy analizach zależności to też prawie zawsze: „Wartości współczynników korelacji...”, a krytyczne interpretacje dotyczą analizy zależności/niezależności (nie skorelowania). Na koniec chciałbym zapytać Autora, dlaczego w kontekście miejsca zamieszkania wyróżnił tylko dwie kategorie: miasto i wieś? Z polskich badań wiemy, że najbardziej „nietypowe” wyniki uzyskuje się w małych miastach. Czemu więc ta kategoria została pominięta? Podsumowując, chcę mocno podkreślić, iż opisany w tym paragrafie problem nie wpływa znacząco na wnioski z pracy – wszystkie wyniki istotne, po nowych analizach, nadal będą istotne (sprawdziłem to).

4. Konkluzja

Raz jeszcze pragnę podkreślić, że powyższe uwagi nie obniżają w żadnym stopniu mojej oceny, która jest jednoznacznie wysoka (we wprowadzonej w tej recenzji szkolno-oceniającej stylistyce byłoby to pięć z małym plusem!). Jest to bowiem rozprawa silnie teoriiotwórcza i rozwiązuje ważne i aktualne problemy o charakterze poznawczo-edukacyjnym. Konkludując stwierdzam, że rozprawa doktorska Marcina Radziwołka, pt. *Skuteczność strategii profilaktyki szkolnej w zakresie zachowań ryzykownych młodzieży – częstochowskich szkół średnich*, przygotowana pod kierunkiem Małgorzaty Przybysz-Zaremby, **spełnia warunki określone w art. 13 ust. 1 Ustawy z dnia 14 marca 2003 roku o stopniach i tytule naukowym oraz stopniach i tytule w zakresie sztuki.**

Michał Klichowski