

Recenzja rozprawy doktorskiej mgr Katarzyny Zych
pt. Polska adaptacja Kwestionariusza MAI (Metacognitive Awareness Inventory)
służącego do pomiaru świadomości metapoznawczej

Prezentowana praca będąc raportem z badań nad adaptacją i standaryzacją narzędzia badawczego nie należy do często podejmowanych problemów rozpraw doktorskich. W tym sensie jest pracą oryginalną, wymagającą poruszania się po obszarze, w którym trudno znaleźć przykłady konstruowania rozpraw doktorskich. Z tym ograniczeniem Autorka dobrze sobie poradziła, starając się spełnić - uważane za standardowe - wymagania formalne stawiane strukturze doktoratów. W jednym miejscu można jednak odnieść wrażenie, że to staranie poszło zbyt daleko. Mam na myśli sformułowane hipotezy badawcze. Do tej kwestii wrócę jednak później. Kompozycja pracy jako całości nie budzi, pomimo tego, zastrzeżeń: jest logiczna w sensie formalnym, kolejność rozdziałów jest uzasadniona w sensie merytorycznym, a ich zawartość odzwierciedla poprawnie przeprowadzony program badawczy. Praca składa się z 8 rozdziałów z licznymi podrozdziałami. Do rozdziału 6 Autorka analizuje podstawy teoretyczne badań, definiuje świadomość metapoznawczą oraz zmienne, które się z nią łączą. Jako teoretyczny kontekst pracy występuje też analiza zasad adaptacji kulturowej narzędzi badawczych. 6 rozdział zawiera metodologiczny plan badań, charakteryzujący poszczególne kroki szacowania właściwości psychometrycznych MAI oraz opisujący zastosowane reguły tłumaczenia pozycji i ich adaptacji kulturowej. Przed połową rozprawy zaczynają się rozdziały prezentujące wyniki. Całość zamyka dyskusja podsumowująca osiągnięte rezultaty oraz opisująca polską wersję kwestionariusza MAI. Po części merytorycznej następują: abstrakt, bibliografia, spis tabel oraz załączniki prezentujące użyte narzędzia pomiarowe, część szczegółowych wyników analitycznych a także zgoda autora oryginalnej wersji testu na dokonanie polskiej adaptacji.

Przechodząc do bardziej szczegółowych analiz zastosuję postępowanie liniowe, pisząc o kolejnych rozdziałach pracy. Jednocześnie będę wskazywał słabsze i mocniejsze strony rozprawy.

„Teoretyczne podstawy badań” opisała Autorka próbując usytuować tytułową zmienną w szerszym kontekście jej korelatów psychologicznych i pedagogicznych. To bardzo dobra strategia zważywszy, że teoretyczne zaplecze pojęcia świadomości metapoznawczej nie jest szczególnie silnie rozbudowane w literaturze naukowej. Jej aktualny stan Autorka zrekonstruowała poprawnie, choć jeden wątek jej rozważań uważam za dyskusyjny. Otóż na s. 26 jest narracja dyskutująca status umiejętności metapoznawczych, jako ogólny versus domenowy. Pierwsze zdanie tej narracji wskazuje na brak zgodności teoretyków co do tego statusu. Następnie Autorka przytacza badania nad związkiem między studiowaniem przez badanych różnych kierunków akademickich a ich świadomością metapoznawczą. Pokazuje przy tym, że związki te nie są istotne statystycznie. I z tej analizy wyprowadza założenie, że „poziom świadomości metapoznawczej ma charakter ogólny, nie zależy od dziedziny wiedzy”. Wydaje się, że nie jest to uprawniony wniosek. Przytoczone badania nie analizowały bezpośrednio aktywności metapoznawczej podczas poruszania się badanych po zróżnicowanych dziedzinach wiedzy, a tylko kierunki ich studiowania. Takie badania nie rozstrzygają o statusie świadomości metapoznawczej jako dyspozycji ogólnej versus związanej z dziedziną wiedzy. Domyślam się, że Autorka chciała przyjąć takie założenie by uzasadnić badanie świadomości metapoznawczej kwestionariuszem. Nie było to jednak konieczne, ponieważ wcześniej zrobił to już autor oryginalnego narzędzia. W obecnej sytuacji pierwsze i ostatnie zdania omawianego fragmentu tekstu, sugerują, że Autorka ustaliła to, nad czym inni zastanawiają się od lat, i to zrobiła na podstawie zastanych wyników badań. Przy okazji chciałbym zwrócić uwagę na nieprecyzyjne posługiwanie się pojęciem „wpływu” w odniesieniu do układu zmiennych: kierunek studiów – wynik pomiaru świadomości metapoznawczej. To pojęcie jest w metodologii zarezerwowane dla badań eksperymentalnych, badań opartych na zachowaniu kanonu jedynej różnicy. Należało raczej użyć pojęcia „różnic między średnimi”, co odpowiada wykonanym analizom w cytowanych badaniach. Wskazany przeze mnie wątek dyskusyjny jest jednym z wielu wątków, których analizę prowadziła Autorka starannie, rzeczowo, konkluzywnie i trafnie.

Rozdział o badaniu świadomości metapoznawczej odzwierciedla stan literatury przedmiotu, stanowi spójne i trafne charakterystyki poszczególnych narzędzi badania. Jest przy tym napisany dynamicznie i w sposób, który określiłbym jako oszczędny w treści, co uważam za zaletę nie tylko tej części, ale całej rozprawy.

Rozdział 4 stanowi już wprowadzenie w problematykę MAI, a 5 zawiera informacje ogólne dotyczące badań psychometrycznych, ich koniecznych parametrów. Autorka definiuje te parametry odwołując się do ustaleń innych autorów oraz do stosowanych wcześniej praktyk badawczych. To również rzeczowa analiza literatury, która pokazuje swobodę poruszania się Autorki w obrębie psychometrii związanej ze standaryzacją narzędzi pomiarowych. Nie mam uwag krytycznych w odniesieniu do tych tekstów. Zastanawiam się tylko, czy nie lepiej byłoby zamienić kolejność tych rozdziałów. Rozdział 5 bowiem zawiera wiedzę ogólną stosowaną do podjęcia rozstrzygnięć, o których mowa w rozdziale 4.

Zazwyczaj bardzo ważnym rozdziałem rozprawy doktorskiej jest ten poświęcony planowaniu badań od strony metodologicznej. Przyjmowane założenia, a szczególnie ich trafność w odniesieniu do przedmiotu badania rozstrzygają o wartości później prezentowanych wyników badań. Od razu zaznaczę, że Autorka nie popełniła w tym rozdziale takich błędów, które rzucałyby cień na uzyskane wyniki. Jednak nie jest to rozdział idealny. Na początek wróć do kwestii formułowanych hipotez. Można odnieść wrażenie, że powstały one nieco „na siłę”, by były na wszelki wypadek. Wszak w rozprawie doktorskiej powinny być hipotezy - może się mylę. Jednak w metodologii hipotezą jest wniosek logiczny z twierdzenia teoretycznego odniesiony do rzeczywistości empirycznej. Zdania prognozujące, że polska wersja MAI będzie trafna lub rzetelna nie są hipotezami. Zresztą hipotezy należy formułować w badaniach weryfikujących teorię. Tutaj nie mamy do czynienia z takim przypadkiem. Lepszym rozwiązaniem byłoby pozostanie przy pytaniach badawczych, choć sądzę, że wystarczyło poprzestać na celach badania, którymi są oszacowanie trafności, rzetelności, itd. Druga kwestia natury metodologicznej to pomieszanie hipotez badawczych ze statystycznymi. Na s. 53 i 54 są hipotezy zerowe i alternatywne. Hipotezy statystyczne to jednak nie hipotezy badawcze, operacjonalizujące badanie weryfikacyjne. Hipotezy zerowe odnoszą się bowiem do konkretnych testów statystycznych. Trzecia kwestia to z kolei przyjęta klasyfikacja zmiennych. Autorka dzieli je na niezależne i zależne. Przy czym w dalszych częściach pracy wprowadza je m.in. do modelu korelacji. A korelacja to związek między zmiennymi losowymi. Nie ma w nim mowy o zmiennych zależnych i niezależnych. Zapewne chodziło Autorce o zmienne ustalone i losowe. W modelu korelacyjnym występują losowe a w modelu istotności różnic ustalone (czynniki) i losowe. Jeśli zmienną zależną traktuje się jak skutek a niezależną jak przyczynę, to staje się jasne, że w omawianych badaniach zostały przywołane nietrafnie,

ponieważ korelacja nie jest związkiem przyczynowo-skutkowym, i nie jest nim też szacowanie istotności różnic między średnimi, czy innymi pomiarami w schemacie quasieksperymentalnym. Dodam jednak na usprawiedliwienie, że to częste błędy w rozprawach doktorskich i habilitacyjnych, i że nie wpływają one na wartość uzyskanych rezultatów a jedynie psują nieco spójność i czystość języka metodologicznego pracy. Podobny charakter ma używanie przez Autorkę terminu związek w sformułowanych pytaniach i hipotezach. Otóż termin ten nic nie znaczy w sensie metodologicznym. De facto jest on dopiero operacjonalizowany przez: relację przyczynowo-skutkową (schemat eksperymentalny), współwystępowanie (schemat korelacyjny), istotność różnic (schemat quasieksperymentalny). Na s. 117 dopiero podczas analizy wyników badań Autorka precyzuje rodzaj związku. Warto jednak takie kwestie rozstrzygać już na etapie planowania badania, by nie narażać czytelników na wątpliwości.

Przyglądając się w dalszym ciągu decyzjom metodologicznym Autorki mam wątpliwości czy słusznie do badań normalizacyjnych zrezygnowała z doboru losowego na rzecz doboru celowego-kwotowego. Moje wątpliwości podzieliła Autorka przywołując na s. 70 uwagi Brzezińskiego na ten temat. Rozumiem, że to był jednak konieczny, niezbyt lubiany przez Autorkę kompromis.

Poza tymi kilkoma uwagami wysoko oceniam rozdział metodologiczny. Autorka szczegółowo zaplanowała każdy krok badawczy, rozstrzygnęła we właściwym czasie poziomy pomiaru by potem ze spokojem zastosować odpowiednie testy statystyczne, wszechstronnie opisała badane próbki, pokazała zasady losowania i doboru celowego. Należy też docenić szczegółowe i trafne zaplanowanie badania trafności teoretycznej MAI, pokazanie wszystkich wykonanych kroków analitycznych i uzasadnienie szczegółowych wyborów, a także użycie obu rodzajów analizy czynnikowej: eksploracyjnej i confirmacyjnej, testujących możliwe układy czynników. Jeśli chodzi o trafność kryterialną to również adekwatnie korelowano wyniki między MAI i SES, choć można jeszcze było to samo zrobić z pomiarem poczucia skuteczności, stosując bandurowskie badanie mikroanalityczne lub któryś z dostępnych w Polsce kwestionariuszy. W tym miejscu przechodzi mi na myśl, że z poczuciem skuteczności jest ten sam problem co ze statusem świadomości metapoznawczej, tzn. również dyskutuje się czy ma ono charakter zgeneralizowany (układ dyspozycji osobowości) czy też jest sądem odniesionym do rodzaju zadania. W przypadku poczucia skuteczności prowadzi się dwa programy badań:

mikroanalityczne (sąd jednostkowy odniesiony do konkretnego zadania) i kwestionariuszowe mierzące ogólną dyspozycję u tych samych osób badanych. Rezultaty sugerują stopniowo, choć jeszcze nie w pełni jednoznacznie, że badanie mikroanalityczne daje zróżnicowane wyniki u tej samej osoby w zależności od rodzaju zadania wobec którego poczucie skuteczności jest szacowane. Być może jest to pomysł na weryfikowanie obu ujęć świadomości meta poznawczej.

Część prezentująca wyniki badań właściwych została opracowana bardzo starannie a niekiedy nawet użyto rzadko pojawiające się w takich badaniach procedury, które jednak wzmacniały wartość rezultatu i eliminowały ewentualne pytania krytyczne. Mam na myśli np. zastosowanie metody odpowiedzi dwujęzycznych, która nie pozostawia już wątpliwości co do równoważności lingwistycznej obu wersji kwestionariusza. Zresztą cała procedura przygotowania i zweryfikowania polskich pozycji budzi podziw i uznanie dla Autorki.

Przyglądając się właściwościom psychometrycznym trzeba docenić wykonanie analizy mocy dyskryminacyjnej pozycji, co coraz częściej bywa pomijane w takich badaniach z uwagi – jak się to wyjaśnia – na stosowanie analizy czynnikowej. Jednak uważam, że sposób postępowania Pani mgr Katarzyny Zych jest właściwy. Policzone współczynniki Pearsona dają obraz mocy dyskryminacyjnej, choć uważam, że równie dobre rezultaty dają inne współczynniki, jak np. korelacji punktowo-czteropolowej i dwuseryjnej. Raz, że wymagają one sprowadzenia odpowiedzi do postaci dwupunktowej a dwa, że po użyciu np. współczynników korelacji dwuseryjnej można te dane wykorzystać do szacowania rzetelności metodą Spearmana-Browna z poprawką Gaylorda, zyskując alternatywny w stosunku do Alfya Cornbacha wskaźnik rzetelności testu. Niemniej jednak użycie „Pearsona” było w pełni uzasadnionym i trafnym, autorskim posunięciem.

Z punktu widzenia wartości teoretycznej testu ważne są wyniki analiz czynnikowych. Autorka dostarczyła argumentów by uznać zgodność wyników (EFA) z modelem teoretycznym, oraz dzięki wykonaniu CFA wykazała dobre dopasowanie tego modelu do danych empirycznych. Także badania struktury wewnętrznej narzędzia MAI potwierdziły zgodność interkorelacji wszystkich skal i podskal (2 i 8) z podstawami teoretycznymi. Tę część badań uważam za najtrudniejszą i wzorcowo wykonaną. Jeśli chodzi o trafność kryterialną to samoocena była dobrze dobranym kryterium, a jak już wspomniałem zabrało szacowania relacji pomiędzy MAI i poczuciem skuteczności. Czytając o poczuciu skuteczności jako

kluczowym korelacje MAI w części teoretycznej rozprawy, można było się tej analizy spodziewać. Sądzę, że na tym etapie i wobec wykonanych innych analiz trafności teoretycznej, brak ten można pominąć. W innym badaniu jednak warto tym się zająć.

Patrząc całościowo na analizę wyników badań warto wskazać także pewne niedociągnięcia. Niektóre są konsekwencją połączenia szczegółowych wyników z częścią metodologiczną, a zwłaszcza decyzjami, które przy jej ocenie uznałem za błędne, np. wyodrębnianie zmiennych zależnych i niezależnych a następnie liczenie pomiędzy nimi współczynników korelacji. Inna kwestia to prezentacja wyników jednoczynnikowej Anova, podczas której nie testowano homogeniczności wariancji. To ważna informacja, ponieważ wynik jest kryterium stosowania testu parametrycznego. W innych miejscach Autorka pokazuje poziomy istotności, np. 0.000 (s.129), pod tabelką pisząc $*p<0.05$. Przy czym gwiazdka nie ma odniesienia a oba wyniki mogą mylić. W jednym miejscu jest przynajmniej jednadziesiąt tysięcy a w drugim pięćsetnych. Jak wiadomo oba wyniki w zróżnicowany sposób odnosi się do prawdopodobieństwa popełnienia błędu pierwszego (odrzućenie prawdziwej hipotezy zerowej) i drugiego rodzaju (utrzymanie fałszywej hipotezy zerowej). Warto zatem jednoznacznie pokazać czy założony przed badaniem poziom istotności (jak sądę tu $p<0.05$) został osiągnięty. Im mniejsza jest alfa tym większe prawdopodobieństwo popełnienia błędu II rodzaju. Tak więc zapis z alfą 0.000 jest czymś innym niż z alfą 0.05. A jeśli widać dwa zapisy w okolicy tabeli pojawiają się wątpliwości.

Na zakończenie nasuwa mi się uwaga dotycząca techniki pisania. Szczególnie w pierwszych rozdziałach nie zawsze trafnie do prowadzonej narracji stosowane są akapity. Niekiedy pojawia się też wers bez akapitu przy zostawionym w połowie wersie uprzednim. Poza tym jednak praca jest napisana poprawnym, precyzyjnym językiem, adekwatnym do prowadzonej analizy, ścisłym i oszczędnym. Przypisy w tekście niekiedy powinny być opatrzone odniesieniem do stron, zwłaszcza przy cytatach, jak np. na s.15 w pierwszym akapicie. Jest tam zdanie opisujące inteligencję zaczerpnięte od Strelaua. Brak strony utrudnia dotarcie do niego. Zdanie wygląda na definicję inteligencji, jednak jest tak ogólne, że w zasadzie może także opisywać (definiować) równie dobrze osobowość. I jeśli jest to definicja (numer strony pozwoliłby to łatwo sprawdzić), to warto by Autorka odniosła się do niej krytycznie.

Podsumowując należy podkreślić, że badania zostały zaplanowane i przeprowadzane starannie. Wykonane analizy wymagały sporej orientacji w pedagogice, psychologii, psychometrii i statystyce. Zadanie adaptacji narzędzia badawczego należy do trudnych przedsięwzięć naukowych w obszarze nauk społecznych. Mgr Katarzyna Zych zaprezentowała się jako dojrzała badaczka o dużej wrażliwości analitycznej. Sama rozprawa jako raport z badań zdała z tej pracy czytelną relację. Rezultat jakim jest udostępnienie polskiemu uczonemu narzędzia do pomiaru świadomości metapoznawczej, o wysokich właściwościach psychometrycznych i teoretycznych, jest bez wątpienia dużym osiągnięciem naukowym. Prowadząc badania i pisząc rozprawę doktorską mgr Katarzyna Zych przedstawiła oryginalne (autorski plan adaptacji narzędzia pomiarowego) rozwiązanie problemu naukowego, wykazała wiedzę teoretyczną z zakresu pedagogiki, a także psychometrii i statystyki oraz umiejętność samodzielnego zaplanowania i realizacji procesu badawczego w naukach społecznych. Dlatego bez cienia wątpliwości rekomenduję Radzie Naukowej Wydziału Pedagogicznego UW dopuszczenie mgr Katarzyny Zych, zgodnie z Ustawą z dnia 14 marca 2003 r. o Stopniach Naukowych i Tytule Naukowym oraz o Stopniach i Tytule w zakresie Sztuki, do kolejnych etapów przewodu doktorskiego.

