

KRYTERIA OCENY RYGORU METODOLOGICZNEGO BADAŃ W NAUKACH O ZARZĄDZANIU

Wprowadzenie

Kluczowym, choć nie jedynym i nie najważniejszym kryterium oceny badań naukowych jest rygor metodologiczny. Przesądza on bowiem jednocześnie o jakości warsztatu badacza, jakości badań oraz wiarygodności rezultatu. Ocena rygoru jawi się więc jako istotne wyzwanie dla każdego odbiorcy pracy naukowej: czytelnika, badacza, praktyka czy recenzenta. Na podstawie tej oceny są podejmowane decyzje o finansowaniu badań oraz ich rozliczeniu, a także decyzje o awansach naukowych. Jakkolwiek sposoby podejmowania tych decyzji są utrwalone i oparte o kompetencje recenzenta, to jednak debata o rygorze metodologicznym wydaje się potrzebna. Warunkuje ona podnoszenie jakości badań, przyczyniając się do konstruowania uniwersalnego języka nauki, którym nie jest matematyka, a metodologia właśnie. Z tym większym entuzjazmem trzeba witać pojawiające się ostatnio coraz częściej publikacje, konferencje i dyskusje w tym zakresie.

Naukom o zarządzaniu przyświeca dokładnie taki sam cel jak każdej innej dziedzinie nauki, a mianowicie dostarczanie wiarygodnej czy solidnej wiedzy [16]. Wiarygodność zaś jest uwarunkowana rygorem [11]. Rygorystyczne postępowanie badawcze to takie, które jest ścisłe, jednoznaczne i daje gwarancję osiągnięcia wyniku [2]. Najwyższym kryterium oceny wiarygodności jest zatem prawda [9]. Jednakże zgoda co do istoty prawdy, dopuszczalnych sposobów jej dociekania oraz oceny tych zabiegów poznawczych jest uwarunkowana paradygmatycznie. Badacze tworzą, uzgadniają i stosują pewne zbiory założeń co do sposobu uprawiania nauk

* Prof. dr hab. Wojciech Czakon – Zakład Teorii Zarządzania, Uniwersytet Ekonomiczny w Katowicach.

o zarządzaniu. Rezultatem są wspólnoty spójne, ale też zamknięte na odmienność, a nawet ją zwalczające. Obserwacja tego zjawiska przyniosła metaforę plemiennosci nauk o zarządzaniu [7]. Wynika stąd, niestety, że kryteria oceny rygoru będą zmienne oraz liczne. Wyodrębniam spośród nich paradygmat pozytywistyczny oraz społecznego konstruktywizmu jako główną oś podziału, ze świadomością wszelako istnienia mniej popularnych, choć ważnych paradygmatów krytycznego zarządzania oraz postmodernizmu [14].

Ocena rygoru badań jest kategorią szerszą, bowiem obejmuje całość zabiegów podejmowanych przez badacza, w tym wkładu w zbiór istniejącej już wiedzy. Niektórzy autorzy wyróżniają trzy obszary oceny rygoru badań: adekwatność konceptualną, rygor metodologiczny oraz zgromadzony materiał empiryczny [13]. Adekwatność konceptualna odnosi się do sposobu, w jaki badacz osadził swoje badania w dostępnych ujęciach teoretycznych. Oznacza to potrzebę doboru właściwych koncepcji, ujawnienia problemów badawczych, gruntu badań empirycznych oraz metod badawczych. Ocenę tego elementu pozostawiam poza zakresem opracowania.

Celem niniejszego artykułu jest przedstawienie kryteriów oceny rygoru stosowanych w naukach o zarządzaniu. Wbrew formułowanemu gdzieniegdzie przekonaniu [2; 13] istnieje daleko idący konsensus co do kryteriów oceny w ramach paradygmatów nauk o zarządzaniu. Ujawnia go systematyczny przegląd literatury, a także badania metodologiczne nad szczegółowymi dyscyplinami.

W pierwszej części opracowania przedstawiam główne założenia paradygmatu pozytywistycznego oraz wynikające zeń wprost kryteria oceny rygoru metodologicznego. Zalicza się do nich powszechnie wewnętrzną i zewnętrzną trafność, rzetelność oraz trafność miar. Drugą część poświęciłem kryteriom oceny badań w paradygmacie społecznego konstruktywizmu, tj.: wierność, przechodniość, solidność i potwierdzalność. Podkreślam tu zarówno linię podziału dominujących paradygmatów, jak i warianty przyjmowane przez mniej liczne wspólnoty postmodernistów czy krytycznych badaczy zarządzania. Trzecia część stanowi pewną syntezę opartą o zestaw kryteriów stosowanych zbiorczo w literaturze w ocenie rygoru metodologicznego.

1. Klasyczne kryteria oceny rygoru – pozytywizm

Pozytywiści zakładają obiektywne istnienie rzeczywistości, a w niej zjawisk, procesów oraz związków przyczynowych. Obiektywność oznacza w tym przypadku, że rzeczywistość jest stabilna i niezależna od badacza. Rolą badacza jest odkrycie istniejących w takim świecie prawidłowości, a dalej formułowanie praw dotyczących jego funkcjonowania i całych teorii wyjaśniających rzeczywistość. W procesie odkrywania podstawowe znaczenie ma ochrona badanego obiektu przed wpływem

badacza, zarówno bezpośrednim na przebieg badanych zjawisk, jak i pośrednim na ich uzyskany obraz.

Badania nauk o zarządzaniu sytuujące się w tym paradygmacie polegają na formułowaniu hipotez oraz na empirycznym testowaniu każdej z nich czy modeli porządkujących określone ich zbiory. Wobec tego można oczekiwać realizacji określonej sekwencji zabiegów badawczych zadeklarowanej przez badacza. Ta procedura obejmuje zwykle:

1. określenie istniejącego stanu wiedzy,
2. identyfikację luk poznawczych,
3. określenie problemu badawczego,
4. sformułowanie hipotez,
5. dobór przedmiotu badań,
6. dobór próby,
7. gromadzenie danych empirycznych,
8. analizę danych empirycznych,
9. wnioskowanie o odrzuceniu/braku podstaw do odrzucenia hipotez.

Procedura badawcza może zostać sformułowana w sposób bardziej szczegółowy lub bardziej ogólny. Jednakże ocena rygoru metodologicznego musi obejmować ścisłość, rozumianą jako wykonanie przez badacza dokładnie takiej, a nie innej ścieżki postępowania. Następnie ocena rygoru dotyczy każdego z określonych etapów.

Kluczowym problemem dla rygoru metodologicznego jest ochrona obiektu badań przed subiektywizmem badacza [8]. Zagrożenie to pojawia się już w fazie opisywania istniejącego stanu wiedzy, a następnie selekcji badanej rzeczywistości, gromadzenia danych, kodowania danych oraz doboru kategorii. Dlatego stosowane są cztery kryteria oceny rygoru metodologicznego:

- wewnętrzna trafność (*internal validity*), nazywana także trafnością logiczną, odnosi się do związków przyczynowo-skutkowych między zmiennymi a wynikami;
- zewnętrzna trafność (*external validity*), czy też zdolność do uogólniania odnosi się do oczekiwania, że ustalenia badawcze będą znajdowały potwierdzenie w innych sytuacjach empirycznych;
- trafność miar (*construct validity*), odnosi się do jakości operacjonalizacji pojęć czy konceptów. W tym sensie pozwala ustalić, czy badanie rzeczywiście mierzy to, co powinno, opisuje więc fazę gromadzenia danych i projekt badania;
- rzetelność (*reliability* lub *generalizability*) odnosi się do braku błędów losowych, pozwalając innym badaczom dojście do tych samych rezultatów, jeśli zastosują te same kroki postępowania.

Trafność wewnętrzna odnosi się do najtrudniejszego, ale też najważniejszego zadania badań naukowych, tj. ustalania związków przyczynowych. Istniejący w momencie podejmowania badań stan wiedzy pozwala zwykle formułować przypuszczenia dotyczące związków przyczynowych. Stąd alternatywna nazwa trafności logicznej. Badania

eksperymentalne będą oceniane w inny sposób, jeśli chodzi o trafność wewnętrzną, aniżeli badania na populacjach niekontrolowanych. Eksperymenty stwarzają bowiem badaczowi możliwość manipulowania zmiennymi zależnymi po to, aby stwierdzić istnienie związków przyczynowych. Zidentyfikowano wiele typowych zjawisk, które obciążają wnioskowanie o przyczynowości w warunkach eksperymentalnych.

Najlepsze wyniki pod względem rygoru oraz wynikającej zeń wiarygodności daje badanie całej populacji, czyli wszystkich obiektów interesujących badacza. Dostępność danych, koszt, a także inne względy powodują jednak, że najczęściej takie badanie całej populacji jest niemożliwe. Dobór próby badawczej powinien wobec tego być reprezentatywny, gdyż chroni to przed subiektywizmem badacza. Warto podkreślić, że wymóg ten należy stosować zarówno do badań literaturowych, jak i empirycznych. W przeciwnym przypadku powstaje ryzyko obciążenia badań empirycznych subiektywizmem, zanim zostaną one zaprojektowane. Procedury statystyczne losowania próby pozwalają ocenić badania pod tym względem. Obecny stan debaty nad reprezentatywnością wprawdzie docenia to kryterium przy projektowaniu próby, ale zgłaszany jest też szereg wątpliwości: czy rzeczywiście uzyskana próba, zwykle różna od próby projektowanej, nadal spełnia to kryterium; a także czy sama reprezentatywność jest w ogóle możliwa. Dobór próby może też być teoretyczny, czyli celowy ze względu na założenia referencyjnej perspektywy poznawczej. Wynikające zeń rezultaty nie będą jednak reprezentatywne, wobec czego badanie takie nie może spełnić kryterium zewnętrznej trafności. Podobnie nisko trzeba ocenić takie badania, w których próba jest nieokreślona. Dlatego oceniając dorobek badań nad rozwojem organizacyjnym, zastosowano ocenę zero-jedynkową – badania wykorzystujące próbę reprezentatywną otrzymują ocenę 1, a inne – ocenę 0 [17].

Oprócz sposobu doboru próby ze względu na kryterium reprezentatywności znaczenie ma także często stosowane kryterium wielkości próby. Wpływa ona na statystyczną istotność, dlatego na etapie planowania badań badacze powinni ustalić minimalną wielkość próby dla uzyskania określonego prawdopodobieństwa stwierdzenia małych, średnich i dużych badanych efektów [4]. Podobnie jak w przypadku reprezentatywności wielkość próby często utożsamiana jest z rygorem, co powoduje dążenie do zwiększania liczebności prób. Silny oddźwięk ma jednak pogląd mówiący o tym, że wielkość próby pozwala jedynie maskować niedostatki dzięki uzyskiwanej w ten sposób statystycznej istotności.

Zewnętrzna trafność ma kluczowe znaczenie w naukach społecznych, a w tym naukach o zarządzaniu. Jeżeli bowiem istnieją prawidłowości, które można odkryć, to wówczas inni badacze czy praktycy, znajdujący się w innej sytuacji, powinni móc się spodziewać takich samych efektów jak ustalono w rygorystycznych badaniach. Kryterium trafności zewnętrznej dotyczy więc związków przyczynowych oraz ich niezależności od czasu, miejsca, osób czy kontekstu. To kryterium budzi jednak istotną dyskusję epistemologiczną. Z jednej strony bowiem proponuje się, aby każde badanie

oceniać pod względem trafności zewnętrznej, ale z drugiej zaś wyraźnie rysuje się stanowisko falsyfikacjonistów [3]. Wynika z niego, że również badania pozbawione trafności zewnętrznej przyczyniają się do postępu nauki, nawet w większym stopniu niż badania spełniające to kryterium. Skoro nie można potwierdzić teorii, a jedynie ją sfalsyfikować, to obserwowane przez badacza tylko w określonym kontekście, czyli nietrafne zewnętrznie wyniki przyczyniają się do falsyfikacji wówczas, gdy są sprzeczne z dotychczasowym stanem wiedzy.

Trafność miar ma kluczowe znaczenie w badaniach empirycznych, dotyczy bowiem zgodności pomiaru z rzeczywistą wariacją badanego zjawiska. Stąd niektórzy autorzy wspominają o pionowej korespondencji między nieobserwowalnym z natury teoretycznym konstruktom a jego proponowaną miarą [12]. Najprościej biorąc, chodzi o to, aby miara mierzyła ten i tylko ten konstrukt, który jest przedmiotem badania. Ścisła definicja tego kryterium wymaga, aby miara pozwalała ocenić skalę oraz kierunek wszystkich cech konstruktów i tylko cech tego konstruktów. Bardziej realistyczne kryterium odnosi się do stopnia, w którym miara pozwala ocenić skalę i kierunek cech konstruktów na próbie reprezentatywnej oraz stopień, w którym miara nie jest skażona innymi konstruktami lub obciążona błędem.

Obciążenia miar wpływają w oczywisty sposób na wnioski badań. Wyróżnia się obciążenia systematyczne, związane z metodą pomiaru, oraz obciążenia losowe. Pośród systematycznych obciążeń pomiaru wymienia się najczęściej tendencyjność informatorów, stronniczość danych archiwalnych, uprzedzenia, oczekiwania czy pragnienia [1]. Aby ocenić trafność pomiaru, stosuje się procedurę testowania stopnia zbieżności wariacji dwóch lub więcej miar tego samego zjawiska. Jeśli stopień zbieżności jest wysoki, to ocenia się miary jako trafne. I odwrotnie, trafność dyskryminacyjna powinna powodować, że miary dwóch różnych konstruktów nie będą ze sobą istotnie skorelowane. Znacznie bardziej zaawansowaną techniką oceny trafności pomiaru jest CFA (ang. *confirmatory factor analysis*).

Warto podkreślić założenia leżące u podstaw oceny trafności miar, prowadzą bowiem one do dwóch rodzajów trafności i powiązanych z nimi dwóch sposobów oceny. Trafność cech zakłada bowiem badanie w swoistej próżni teoretycznej, w której wyodrębniono konstrukt. Celem oceny trafności, a także rzetelności miar jest określenie systematycznej wariacji miar oraz siły korelacji tej wariacji z innymi miarami konstruktów i miarami innych konstruktów. Nawet trafna miara musi jednak być użyteczna. Stąd trzeba także oceniać trafność nomologiczną, która skupia się na teoretycznych związkach między konstruktami oraz empirycznych związkach miar tych konstruktów. Tego rodzaju procedura w niewielkim stopniu różni się jednak od walidacji teorii.

Kluczowe znaczenie trafności miar ujawnia się także przez związek między trafnością a wielkością próby. Niska trafność miar obniża jakość pomiaru efektów badanego zjawiska, prawdopodobieństwo ustalenia statystycznie istotnych wyników, a dalej

szanse na publikację [4]. I odwrotnie, przy próbach o znacznej liczebności nawet niska trafność miar jest maskowana statystyczną istotnością ustaleń, wynikającą właśnie z wielkości próby. Oznacza to w szczególności, że wybiórcze stosowanie kryterium statystycznej istotności bez związku z oceną trafności miar prowadzi do wniosków o niskiej wartości.

Warto zauważyć, że różne rodzaje trafności mają kluczowe znaczenie dla oceny rygoru metodologicznego i badań w ogóle. Wewnętrzną trafność można sprowadzić do stopnia korespondencji między konceptualizacją a rzeczywistością [10]. Zewnętrzną trafność z kolei można sprowadzić do replikacji i generalizowalności. Porządek pojęciowy oceny trafności pozwala skierować uwagę na szczegółowe kryteria oceny badań prowadzonych w paradygmacie pozytywistycznym. Badacz powinien przedstawić sposób doboru próby badawczej, narzędzie pomiaru, zmienne oraz techniki analizy danych. Następnie powinien określić stopień ich zobiektywizowania. Wiarygodność uzyskanych wyników zależy bowiem od tego, na ile: procedura badawcza była ściśle realizowana; zapewniono maksymalną korespondencję między konstrukatami a miarami; uzyskano istotność statystyczną.

2. Kryteria oceny badań jakościowych – społeczny konstruktywizm i nie tylko

Pozytywizm w nauce jest postawą epistemologiczną tyleż rozpowszechnioną co, jak twierdzą niektórzy, archaiczną. Warto w tym miejscu przypomnieć, że fundamentalne dla nauk o zarządzaniu badania E. Mayo w zakładach Hawthorne ujawniły przede wszystkim efekt wpływu badacza na wydajność pracownic. To atencja prowadzących eksperyment okazała się zmienną wpływającą na wzrost wydajności, bez względu na wartości innych badanych zmiennych. Z punktu widzenia oceny rygoru badań eksperymentalnych i podstawowych założeń leżących u podstaw paradygmatu pozytywistycznego należałoby uznać, że te badania przeprowadzono nieprawidłowo. Według pozytywistów należy bowiem chronić badaną rzeczywistość przed wpływem subiektywizmu badacza. Tymczasem w naukach o zarządzaniu chodzi nie tylko o ujawnienie prawidłowości, lecz także o zdolność jej kształtowania. Aplikacyjny charakter naszych nauk pozostaje więc w fundamentalnej sprzeczności z założeniami pozytywizmu. Co więcej, ortodoksyjne podejście do stosowania wyników nauk o zarządzaniu nie kończy się na oczekiwaniu uzyskiwania wartości przeciętnych czy ustaleniu prawdopodobieństwa wpływu drobiazgowo wyizolowanych zmiennych wyjaśniających na zmienną zależną, jak np.: efektywność, rentowność, innowacyjność czy przewaga konkurencyjna. Przeciwnie, menedżerowie oczekują wręcz recept sukcesu i choć większość badaczy wyraża przekonanie, iż takich recept

nie ma i być nie może, to jednak ich implikacją metodologiczną jest poszukiwanie skrajnie dodatnich, a nie średnich efektów.

Widoczny sprzeciw wobec pozytywizmu przyjmuje jednak różne postaci. Jedna grupa badaczy odrzuca wymóg falsyfikacji teorii w badaniach empirycznych na rzecz indukcji, zachowując jednak przekonanie o potrzebie gromadzenia danych w sposób obiektywny i chroniony przed woluntaryzmem badacza [8]. Takie podejście można nazywać neoempiryzmem. Wyrasta zeń osobny zestaw kryteriów oceny rygoru metodologicznego alternatywny wobec pozytywizmu. Proponowane są bowiem następujące kryteria: wierność, transferowalność, solidność i potwierdzalność. W istocie więc zastąpiono cztery klasyczne kryteria oceny pewnym alternatywnym zestawem (por. tabela 1).

Tabela 1. Zestawienie kryteriów oceny rygoru w podejściach pozytywistycznym i neoempirycznym

Kryterium pozytywistów	Kryterium neoempiryków	
Wewnętrzna trafność	Wierność	Dokładne odzwierciedlenie przekonań uczestników badanych zjawisk
Zewnętrzna trafność	Transferowalność	Możliwość wykorzystania wyników przez innych badaczy w ich własnych studiach
Rzetelność	Solidność	Ścisłość procesu prowadzenia badań, w tym protokoły gromadzenia danych oraz ich analiza
Obiektywność	Potwierdzalność	Możliwość uzyskania podobnych wyników przez innych badaczy

Źródło: [5, s. 16].

Jakkolwiek propozycja wyodrębnienia osobnych kryteriów do wyraźnie odrębnej metody, osadzonej w wyraźnie innym podejściu epistemologicznym, wydaje się uzasadniona, to jednak stosowanie tego alternatywnego zestawu kryteriów pozostaje nadal ograniczone. Niektórzy badacze wskazują wręcz, że nawet realizując studia przypadków, *par excellence* wykraczające poza paradygmat pozytywistyczny, stosuje się kryteria klasyczne [6]. Trzeba tu zgłosić wątpliwość, czy autorami nie kierowała raczej pragmatyka publikowalności badań, a więc wymóg przekonania recenzentów i redaktorów, a nie ortodoksja metodologiczna.

Bardziej radykalne stanowisko wobec pozytywizmu w nauce przyjmują przedstawiciele krytycznej teorii zarządzania, odrzucają oni bowiem istnienie jednego, stabilnego i obiektywnego świata na rzecz przekonania o społecznej konstrukcji rzeczywistości. Najprościej biorąc, nie chodzi więc o to, co jest, ale o to, co postrzegają różni uczestnicy tej rzeczywistości. Wówczas kryterium prawdy spełnia autentyczność, czyli przedstawienie uzgodnionego przez uczestników, nie zaś wniesionego przez badacza spojrzenia na rzeczywistość. Ponownie zatem pojawia się wymóg wiernego

odzwierciedlenia tego, jak konkretni uczestnicy postrzegają ich konkretne zachowania, zdarzenia i decyzje zachodzące w im właściwym kontekście. Takie założenia poznawcze kierują w stronę etnografii czy strukturalizmu, a więc metod badawczych rozwiniętych na gruncie antropologii. Ułatwia to ocenę rygoru tylko pozornie, bowiem metody te czynią z subiektywizmu zaletę, co stoi w jaskrawej sprzeczności z pozytywistycznym czy potocznym wymogiem intersubiektywnej sprawdzalności. Wprawdzie subiektywizm jest w tym przypadku atrybutem uczestników, a badacz ich obserwatorem lub interpretatorem, to jednak zmienność poglądów właściwa subiektywizmowi zdaje się być sprzeczna z pozytywistycznym oczekiwaniem istnienia trwałych prawidłowości. Pozostaje więc skupić się na wierności zapisu dyskursu, którym posługują się uczestnicy badanej rzeczywistości.

Porzucenie paradygmatu pozytywistycznego otwiera drogę do bardziej zaawansowanej metodologii, która nad wyrafinowanie ilościowych technik obróbki danych przedkłada zrozumienie zjawisk, doznań i przekonań ich uczestników. W najbardziej zaawansowanej postaci oznacza to identyfikację i analizę dyskursów, otwierając pole oceny na strukturalizm lingwistyczny i antropologiczny, a także na dekonstrukcję. Im bardziej współczesne metody badań, im bardziej odległe od przekonań pozytywistów, tym bardziej eksponują dążenie do odzwierciedlenia wielości znaczeń – heteroglosji. W rezultacie pojawia się pułapka oceny rygoru metodologicznego. Niezbędna staje się ocena spójności dokonań badacza względem stanowiska poznawczego, które wybrał.

3. Ocena rygoru wielu badań – wielowymiarowe systemy

Nie sposób oprzeć się wrażeniu, iż pozytywistyczny paradygmat oferuje znacznie większy komfort nie tylko badaczowi, lecz także odbiorcy badań. Ocena rygoru metodologicznego jest znacznie łatwiejsza, szybsza i prostsza, gdyby przyjąć stanowisko metodologii pozytywnej, odzwierciedlonej w potocznej nazwie badań ilościowych. Ta łatwość prowadzi niektórych badaczy do oceny rygoru całych nurtów badawczych według kryteriów właściwych pozytywistom, bez względu na stopień zaawansowania tego nurtu badawczego czy stanowisko epistemologiczne w nich przyjmowane. Na przykład, by zdać sprawę z rygoru metodologicznego badań nad rozwojem organizacji, zastosowano następujące kryteria [15]:

- strategia badawcza, deskryptywna lub testująca;
- dobór próby, reprezentatywny lub nie;
- zastosowanie grup kontrolnych;
- cechy pomiaru;
- zaawansowanie technik statystycznej obróbki danych.

Późniejsze powtórzenie tych badań [17] wskazało na potrzebę zastosowania bardziej precyzyjnego zestawu kryteriów, dodając do pięciu wymienionych wcześniej wymiarów jeszcze: wielkość próby, rodzaj zmiennej zależnej, analizę czynnikową oraz miary rzetelności i trafności. W rezultacie otrzymano dziewięć kryteriów, które pozwalają stosunkowo szybko zakwalifikować całe zbiory badań, bowiem wymienionym zmiennym łatwo przyporządkować wartości zero-jedynkowe. Dalej z kolei stwierdzono, że każde badanie można ocenić w skali punktów, gdzie 9 uzyskały badania najbardziej rygorystyczne, a 0 – najmniej rygorystyczne.

Łatwość takiej procedury kryje kilka pułapek, które ostrożny badacz powinien mieć na uwadze. Po pierwsze, waga tych kryteriów nie jest jednakowa, dlatego dodawanie punktów, by w sposób zagregowany ocenić pojedynczą pracę, jest czynnością obciążoną sporym błędem. Po drugie, kryteria te są pomocne w naukach eksperymentalnych, ale w mniejszym stopniu dla badania bardzo złożonych układów zmiennych. Po trzecie, wszystkie te kryteria odzwierciedlają pogląd pozytywistyczny na poznanie naukowe, całkowicie marginalizując zrozumienie, odczucie człowieka i relacje między ludźmi. Instytucjonalizują wręcz monizm metodologiczny przy pomocy rzekomo obiektywnej oceny rygoru metodologicznego. Tymczasem samo założenie obiektywizmu staje się zmienną, gdyby porównać stanowiska epistemologiczne. Dlatego zbiorcze próby oceny rygoru metodologicznego powinny uwzględniać stanowiska, z których badacz rozpoznaje rzeczywistość.

Zakończenie

Klasykne kryteria ewaluacji rygoru metodologicznego milcząco promują paradygmat pozytywistyczny. Odrzucenie monizmu metodologicznego na rzecz dopuszczalności innych podejść wymaga stosowania zróżnicowanego zbioru kryteriów do oceny tak prowadzonych badań. Niezależnie jednak od sposobu, w jaki je prowadzono, niezbędne jest dochowanie zasad przejrzystości metodologicznej po to, aby krytyczna ocena rygoru była możliwa.

Rygor metodologiczny dotyczy bowiem oceny sposobu, w jaki przeprowadzono badanie. Nie odnosi się do trafności wyboru metod badawczych. Pomija także dobór ram teoretycznych, w których badanie zostało przeprowadzone. Skupia się wyłącznie na tym, czy badacz ściśle, systematycznie według założonej procedury przeprowadził swoje badanie. Tam, gdzie badania wykorzystują pomiar, ocena rygoru uzupełniona będzie oceną jakości pomiaru oraz rozpoznania związków między zmiennymi. Tam, gdzie badania sięgają do zrozumienia i społecznej konstrukcji rzeczywistości, istotne znaczenie będzie miała ocena stopnia wierności, z którą badacz odzwierciedlił to, co zastał.

Debata o ocenie rygoru metodologicznego jest potrzebna dlatego, że warunkuje on wiarygodność wyników badań. Na najwyższym poziomie ogólności oznacza to ocenę badań ze względu na koncepcję prawdy, za którą opowiedział się badacz. Z kolei na najbardziej technicznym poziomie ogólności oznacza to ocenę rzetelności gromadzenia i obróbki danych. Pole do wypełnienia nie tylko kryteriami, lecz także praktyką oceny rygoru metodologii jest dalekie od wyczerpania. Przeciwnie, zdaje się ciągle rozszerzać.

Bibliografia

- [1] Bagozzi R.P., Yi Y., Phillips L.W., *Assessing Construct Validity in Organizational Research*, "Administrative Science Quarterly" 1991, Vol. 36, No. 4, pp. 421–458.
- [2] Blaug M., *The Trade-Off Between Rigor and Relevance: Sraffian Economics as a Case in Point*, "History of Political Economy" 2009, Vol. 41, No. 2, pp. 219–247.
- [3] Calder B.J., Phillips L.W., Tybout A.M., *The concept of external validity*, "Journal of Consumer Research" 1982, Vol. 8, No. 2, pp. 240–244.
- [4] Combs J.G., *Big Samples and Small Effects: Let's Not Trade Relevance and Rigor for Power*, "Academy of Management Journal" 2010, Vol. 53, No. 1, pp. 9–13.
- [5] Czakon W., *Mity o badaniach jakościowych w naukach o zarządzaniu*, „Przegląd Organizacji” 2009, nr 9, s. 13–17.
- [6] Gibbert M., Ruigrok W., Wicki B., *What Passes as a Rigorous Case Study?*, "Strategic Management Journal" 2008, Vol. 29, No. 13, pp. 1465–1474.
- [7] Gulati R., *Tent Poles, Tribalism, and Boundary Spanning: The Rigor-Relevance Debate In Management Research*, "Academy of Management Journal" 2007, Vol. 50, No. 4, pp. 775–782.
- [8] Johnson P., Buehring A., Cassell C., Symon G., *Evaluating Qualitative Management Research: Towards a Contingent Criteriology*, "International Journal of Management Reviews" 2006, Vol. 8, No. 3, pp. 131–156.
- [9] Kieser A., Leiner L., *Why the rigour-relevance gap in management research is unbridgeable*, "Journal of Management Studies" 2009, Vol. 46, No. 3, pp. 516–533.
- [10] McGrath J.E., Brinberg D., *External validity and the research process: A comment on the Calder/Lynch dialogue*, "The Journal of Consumer Research" 1983, Vol. 10, No. 1, pp. 115–124.
- [11] Niemczyk J., *Chciałbyś coś odkryć?*, „Przegląd Organizacji” 2009, nr 10, s. 3–5.
- [12] Peter J.P., *Construct validity: a review of basic issues and marketing practices*, "Journal of Marketing Research" 1981, Vol. 18, No. 2, pp. 133–145.
- [13] Shrivastava P., *Rigor and Practical Usefulness of Research in Strategic Management*, "Strategic Management Journal" 1987, Vol. 8, No. 1, pp. 77–92.

- [14] Sułkowski Ł., *Paradygmaty nauk społecznych w zarządzaniu zasobami ludzkimi*, „Zarządzanie Zasobami Ludzkimi” 2009, nr 6, s. 121–134.
- [15] Terpstra D.E., *Evaluating selected organization development interventions: The state of the art*, “Group & Organization Management” 1982, Vol. 7, No. 4, pp. 402–417.
- [16] Thompson J.D., *On Building an Administrative Science*, “Administrative Science Quarterly” 1956, Vol. 1, No. 1, pp. 102–111.
- [17] Woodman R.W., Wayne S.J., *An Investigation of Positive-Findings Bias in Evaluation of Organization Development Interventions*, “Academy of Management Journal” 1985, Vol. 28, No. 4, pp. 889–913.

KRYTERIA OCENY RYGORU METODOLOGICZNEGO BADAŃ W NAUKACH O ZARZĄDZANIU

Streszczenie

Problem oceny badań naukowych jest kluczowy w procedurach oceny prac badawczych, finansowaniu projektów badań, awansach naukowych, ale dla wiarygodności nauki w ogóle. Artykuł podejmuje w nim wątek kryteriów oceny rygoru metodologicznego. Przedstawiono kryteria oceny w paradygmacie pozytywistycznym, tj. trafność i rzetelność, wraz z dyskusją o ich ograniczeniach. Następnie wskazano, że choć paradygmat konstruktywizmu społecznego wykształcił własne kryteria oceny rygoru metodologicznego, to jednak badacze ulegają presji posługiwania się kryteriami klasycznymi. Przedstawiono punktowe systemy oceny rygoru badań, ze wskazaniem uproszczeń i słabości takiej agregacji oceny. Podsumowując, stwierdzono, że rygor metodologiczny ocenia sposób przeprowadzenia badań. Tam, gdzie badania wykorzystują pomiar, ocena rygoru będzie uzupełniona oceną jakości pomiaru oraz rozpoznania związków między zmiennymi. Tam, gdzie badania sięgają do zrozumienia i społecznej konstrukcji rzeczywistości, istotne znaczenie będzie miała ocena stopnia wierności, z którą badacz odzwierciedlił to, co zastał.

SŁOWA KLUCZOWE: METODOLOGIA, RYGOR, TRAFNOŚĆ, RZETELNOŚĆ

CRITERIA FOR RIGOR EVALUATION IN MANAGEMENT RESEARCH

Abstract

Research evaluation remains a key issue for the publication process, financing research projects, individual researcher's career development, and for the credibility of knowledge at large. This paper focuses on methodological rigor evaluation criteria. Firstly, criteria stemming from the positivist paradigm, are discussed. Validity and reliability are presented, and recent concerns regarding their use are outlined. Secondly, it is demonstrated that while the social constructivist paradigm offers its own rigor evaluation criteria: credibility, transferability, dependability and confirmability, researchers tend to stick with the traditional ones. Next, an aggregate approach to rigor evaluation is discussed, including the pitfalls of rigor scoring. Hence, methodological rigor addresses the process of how research has been conducted. If a study uses measurements, then rigor assessment will be coupled with the quality of measurements and relationships between variables. If a study reaches for understanding a socially constructed reality, a crucial importance is attributed to the degree of how accurately has it been reported.

KEYWORDS: METHODOLOGY, RIGOR, VALIDITY, RELIABILITY