

EMANUEL KULCZYCKI, EWA A. ROZKOSZ, ANETA DRABEK*

Ocena ekspercka jako trzeci wymiar ewaluacji krajowych czasopism naukowych

1. Wprowadzenie

Publikacje w czasopismach naukowych są jednym z najważniejszych kanałów komunikacji naukowej. W ten sposób czasopisma stają się nie tylko narzędziami upowszechniania wiedzy, ale również nadają rangę publikowanym w nich treściom. Dlatego autorzy tych publikacji oceniani są coraz częściej nie przez pryzmat tego, *co* opublikowali, lecz *gdzie*, tj. w jakim czasopiśmie (Kulczycki, Drabek, Rozkosz, 2015).

W ten sposób zarówno artykuły, jak i autorzy „świecą światłem odbitym” czasopism naukowych. Oznacza to, że o naukowości i ważności danego artykułu czy naukowca świadczyć ma jakość czasopisma. Ta natomiast rozumiana jest w przeróżny sposób: najczęstszym jest pośrednie utożsamianie jakości z „cytowalnością” danego periodyku. Na tym opiera się właśnie fenomen popularności wskaźnika wpływu – Impact Factor (IF) – wyliczanego przez firmę Thomson Reuters. Pozwala on bowiem, pod pewnymi zastrzeżeniami, tworzyć ranking czasopism z różnych dziedzin nauki (Sangster, 2015; Vanclay, 2011; Wheeler, 2011). O jakości czasopism mówić ma również to, przez jakie bazy są indeksowane (np. Web of Science™ Core Collection [WoS] czy SCOPUS), co świadczy o spełnianiu wielu kryteriów formalnych i jakościowych. Ostatnim wyznacznikiem jakości, który zdaniem wielu naukowców powinien być wymieniany na pierwszym miejscu, jest prestiż i uznanie czasopisma przez wspólnotę badaczy. Jednakże w odróżnieniu od cytowania i indeksowania czasopism prestiż jest trudny do zmierzenia, a przygotowanie rankingu w oparciu o ocenę ekspercką jest nie tylko czaso- i pracochłonne, ale również niezwykle problematyczne (Bornmann, 2013; Jabłicka, 1997). Dlatego skonstruowanie wytycznych rankingowania czasopism wymaga uwzględnienia różnic w tzw. kulturze cytowań i standardów wydawniczych na poziomie dziedzin czy poszczególnych dyscyplin.

W Polsce od dwóch dekad przygotowuje się „Wykazy czasopism punktowanych” (WCP), czyli rankingi czasopism. Artykuły naukowe publikowane w czasopismach z tych

* Dr Emanuel Kulczycki, Uniwersytet im. Adama Mickiewicza w Poznaniu, e-mail: emek@amu.edu.pl; mgr Ewa A. Rozkosz, Dolnośląska Szkoła Wyższa we Wrocławiu, e-mail: ewa.rozkosz@dsw.edu.pl; dr Aneta Drabek, Biblioteka Uniwersytetu Śląskiego w Katowicach, e-mail: aneta.drabek@us.edu.pl

wykazów są uznawane w procesie parametrycznej oceny jednostek naukowych. W nadchodzącej ocenie, która zostanie przeprowadzona w 2017 r., będą uwzględniane publikacje z lat 2013-2016. Dotychczas – dla tego okresu oceny – zostały opublikowane trzy WCP¹, według których będą oceniane publikacje w czasopismach: wykaz z 17 grudnia 2013 r., 31 grudnia 2014 r. oraz 18 grudnia 2015 r.

W 2011 roku została wdrożona nowa procedura tworzenia WCP, która w istotny sposób zmieniła dotychczasowe wytyczne. Kulczycki (2014) zrekonstruował zmiany w zasadach budowy WCP oraz pokazał, na czym polega przełomowość zmian wprowadzonych w 2011 r. W roku 2015 r. regulacje uległy dalszym istotnym modyfikacjom. Jedną z najważniejszych zmian wprowadzonych przez Komunikat Ministra Nauki i Szkolnictwa Wyższego z dnia 2 czerwca 2015 r. w sprawie kryteriów i trybu oceny czasopism naukowych (zwany dalej Komunikatem MNiSW z 2015 r.) jest dodatkowy czynnik oceny, tj. ocena ekspercka krajowych czasopism naukowych.

Celem niniejszego artykułu jest omówienie wyników oceny czasopism naukowych przeprowadzonej w 2015 r., wskazanie problemów związanych z wyliczaniem cytowań krajowych czasopism naukowych z części B na WCP oraz przedstawienie analizy przeprowadzonej oceny eksperckiej. Prezentując zmiany w regulacjach, będziemy odnosić się do wszystkich trzech części WCP. Natomiast szczegółowa analiza oceny formalnej, bibliometrycznej i eksperckiej krajowych czasopism będzie koncentrować się na części B na WCP, która zawiera tylko krajowe czasopisma niemające wyznaczonego wskaźnika IF lub nieindeksowane² na liście European Reference Index for the Humanities (ERIH).

¹ Podstawą tworzenia wszystkich wykazów (od 2013 r.) było Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 13 lipca 2012 r. w sprawie kryteriów i trybu przyznawania kategorii naukowej jednostkom naukowym. Wykaz z 2013 r. powstał według zasad oceny opublikowanych w Komunikacie MNiSW z dnia 29 maja 2013 r. w sprawie kryteriów i trybu oceny czasopism naukowych i został opublikowany w Komunikacie Ministra Nauki i Szkolnictwa Wyższego z dnia 17 grudnia 2013 r. w sprawie wykazu czasopism naukowych wraz z liczbą punktów przyznawanych za publikacje w tych czasopismach. Wykaz z 2014 r. powstał na bazie tego samego komunikatu, co wykaz z 2013 r., tj. z 29 maja 2013 r., lecz została jedynie zaktualizowana część A „Wykazu...” (w oparciu o aktualną wówczas bazę Journal Citations Reports). Opublikowano go w Komunikacie Ministra Nauki i Szkolnictwa Wyższego z dnia 31 grudnia 2014 r. w sprawie wykazu czasopism naukowych wraz z liczbą punktów przyznawanych za publikacje w tych czasopismach. Dodatkowo Minister sprostował błędy w wykazie, publikując Obwieszczenie Ministra Nauki i Szkolnictwa Wyższego z dnia 25 marca 2015 r. o sprostowaniu błędów. Wykaz z 2015 r. powstał według zasad oceny opublikowanych w Komunikacie MNiSW z dnia 2 czerwca 2015 r. w sprawie kryteriów i trybu oceny czasopism naukowych i został opublikowany w Komunikacie Ministra Nauki i Szkolnictwa Wyższego z dnia 18 grudnia 2015 r. w sprawie wykazu czasopism naukowych wraz z liczbą punktów przyznawanych za publikacje w tych czasopismach.

² Czasopisma krajowe indeksowane na liście ERIH mogły poddać się w 2015 r. ocenie według zasad dla części B na WCP i potencjalnie uzyskać większą liczbę punktów, aniżeli miały zagwarantowane w części C na WCP.

Przedmiotem naszych analiz będą ministerialne wytyczne zawarte w Komunikacie MNiSW z 2015 r., informacje o wynikach przeprowadzonej oceny formalnej, bibliometrycznej oraz eksperckiej dla wszystkich czasopism zgłoszonych na część B na WCP.

Sformułowaliśmy cztery pytania badawcze, które stały się podstawą naszych analiz.

- 1) Jakie były zasady tworzenia części B na WCP w 2015 r. oraz w jaki sposób przeprowadzano ocenę formalną, bibliometryczną i ekspercką?
- 2) Jaka była charakterystyka wyników oceny: formalnej, bibliometrycznej i eksperckiej?
- 3) Czy wprowadzenie oceny eksperckiej do zasad ewaluacji zmieniło pozycję rankingową czasopism w części B na WCP?
- 4) Czy wyniki oceny eksperckiej wiążą się z: wynikami oceny formalnej i bibliometrycznej, wartością wskaźnika cytowalności, wartościami wskaźników umiędzynarodowienia?

Wszystkie obliczenia wykonaliśmy przy użyciu pakietu IBM SPSS Statistics 22.

Przeprowadzone analizy pokazują, że wyniki oceny eksperckiej pozostają w silnym związku z oceną formalną i bibliometryczną w grupie nauk technicznych, ścisłych, medycznych, przyrodniczych (TZ) oraz społecznych (S), natomiast związek oceny eksperckiej z oceną formalną i bibliometryczną jest przeciętny dla czasopism z nauk humanistycznych (H). Dlatego można rozważyć rezygnację z oceny eksperckiej w grupach nauk TZ i S. Jednocześnie wyniki ukazują, że cytowalność czasopism koreluje z oceną ekspercką we wszystkich grupach nauk, co może być istotnym argumentem w dyskusji na temat stosowania wskaźników bibliometrycznych w ocenie czasopism.

Struktura niniejszego tekstu jest następująca. W sekcji drugiej przedstawiamy, jak wyglądały zasady budowy części A, B i C na WCP z 2015 r. W sekcji trzeciej prezentujemy wyniki analizy oceny formalnej, oceny bibliometrycznej oraz eksperckiej z części B na WCP. W sekcji czwartej analizujemy związki między oceną ekspercką a innymi kryteriami (oceną formalną i bibliometryczną, wskaźnikiem cytowalności oraz kryteriami wskazującymi na umiędzynarodowienie). Ostatnia sekcja jest podsumowaniem zaprezentowanych analiz.

2. Zasady budowy „Wykazu czasopism punktowanych” w 2015 roku

Zasady konstrukcji WCP określa Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 13 lipca 2012 roku³. Następnie są one doprecyzowywane komunikatem ministra w roku przeprowadzania oceny. WCP z 2015 r. został zbudowany na podstawie Komunikatu MNiSW z 2015 r.

WCP z 2015 r. składa się z trzech części:

³ Zob. omówienie ogólnych ram wykazów wynikających z rozporządzenia oraz analizę zasad budowy trzech części WCP w tekście *Zasady oceny czasopism humanistycznych i ich rola w parametryzacji jednostek naukowych* (Kulczycki, 2014).

- 1) część A – zawiera listę czasopism naukowych znajdujących się w bazie Journal Citation Reports (JCR) i posiadających współczynnik wpływu Impact Factor (IF) oraz przypisaną im liczbę punktów za publikacje.
- 2) część B – zawiera listę krajowych czasopism naukowych nieposiadających współczynnika wpływu Impact Factor (IF), których wydawcy działają zgodnie z ustawą 26 stycznia 1984 r. Prawo prasowe, oraz przypisaną im liczbę punktów za publikacje.
- 3) część C – zawiera listę czasopism naukowych znajdujących się w bazie European Reference Index for the Humanities (ERIH) i nieujętych w bazie Journal Citation Reports (JCR) oraz przypisaną im liczbę punktów za publikacje.

2.1. Konstrukcja części A

Część A jest jedyną częścią, której zasady budowy nie uległy zmianie. Tworzona jest na bazie produktu JCR firmy Thomson Reuters. Ranking konstruowany jest na podstawie znormalizowanego pięcioletniego IF (a w przypadku jego braku dwuletniego IF) i na tej podstawie przypisuje się czasopismu od 15 do 50 punktów. Normalizacja prowadzona jest osobno dla każdej grupy tematycznej (ang. *subject category*).

2.2. Konstrukcja części B

Zasady tworzenia części B uległy istotnym modyfikacjom. Utrzymany został podział na trzy grupy nauk – nauki techniczne, ścisłe, medyczne, przyrodnicze (TZ); społeczne (S); humanistyczne (H) – który różnicował progi graniczne w ocenie formalnej i bibliometrycznej.

Kluczową zmianą jest rozszerzenie zakresu punktowego, który przed zmianami wynosił od 1 do 10 pkt, teraz natomiast czasopisma mogą mieć przypisane maksymalnie 15 pkt. Dodatkowe 5 punktów czasopisma mogą uzyskać poprzez ocenę ekspercką. Część B została ograniczona tylko do czasopism krajowych. Krajowość periodyku definiowało się następująco: wydawca lub redakcja powinni mieć siedzibę w Polsce.

Przekształcenia zasad tworzenia części B można zestawić w następujące grupy:

- a) Zmianom uległy kryteria wstępne i kryteria oceny formalnej oraz bibliometrycznej na liście B – zmieniły się nie tylko wymogi, ale również skala punktacji za poszczególne parametry.
- b) Została dodana ocena ekspercka, która stanowi trzeci wymiar oceny obok oceny formalnej, np. liczby artykułów, i oceny bibliometrycznej, np. wysokości wskaźnika Predicted Impact Factor (PIF).
- c) Czasopisma z bazy ERIH mogły złożyć⁴ ankietę aplikacyjną na część B – jeśli wynik

⁴ Ankietę złożyły 94 czasopisma z bazy ERIH. Z tej liczby 63 czasopisma zostały przeniesione z części C na część B. Wśród tych 63 są dwa czasopisma, które „straciły punkty”. „Kwartalnik Historii Nauki i Techniki”: 6 punktów (poprzednio na części C: 10 punktów) oraz „Medycyna

punktowy na części B był wyższy niż na części C, wówczas czasopismo znajdowało się na liście B.

- d) Ocena bibliometryczna czasopism została rozszerzona. Oprócz wskaźnika PIF, którego definicja została zmieniona (zob. szczegółową analizę w sekcji 3.2.1 niniejszego tekstu) został dodany wskaźnik Scimago Impact Factor (SIF) obliczany jako iloraz współczynnika Total Cites (3 years) publikowanego w bazie Scimago Journal & Country Rank (dane do tego rankingu pochodzą z bazy SCOPUS) i liczby wszystkich artykułów naukowych opublikowanych w czasopiśmie w okresie od trzech lat do roku poprzedzających rok wypełnienia ankiety.
- e) Usunięto stałą przeniesienia (tzw. punkty za zasiedzenie, czyli indeksowanie na poprzednich WCP), lecz zostały dodane punkty za staż czasopisma.
- f) Zapowiadany od 2013 r. Polski Współczynnik Wpływu (PWW) nadal nie został wyliczony. Jednakże czasopisma aplikujące do części B mogły wprowadzić dane bibliograficzne ze swoich czasopism do bazy POL-index. Za uzupełnienie pełnych rekordów bibliograficznych czasopismo mogło uzyskać 1 lub 2 dodatkowe punkty. W założeniach wprowadzone dane posłużą w przyszłości do wyliczenia PWW.

2.3. Konstrukcja części C

Część C na WCP nadal jest tworzona w oparciu o tzw. stary ERIH, a nie na podstawie bazy będącej jego kontynuacją, czyli ERIH Plus. Część C odgrywa istotną rolę nie tylko wśród przedstawicieli nauk humanistycznych i społecznych, ale również nauk przyrodniczych, gdyż publikacje w czasopismach zawartych na tej części WCP są uwzględniane w kategoryzacji jednostek naukowych we wszystkich grupach nauk (Kulczycki, Rozkosz, Drabek, 2015).

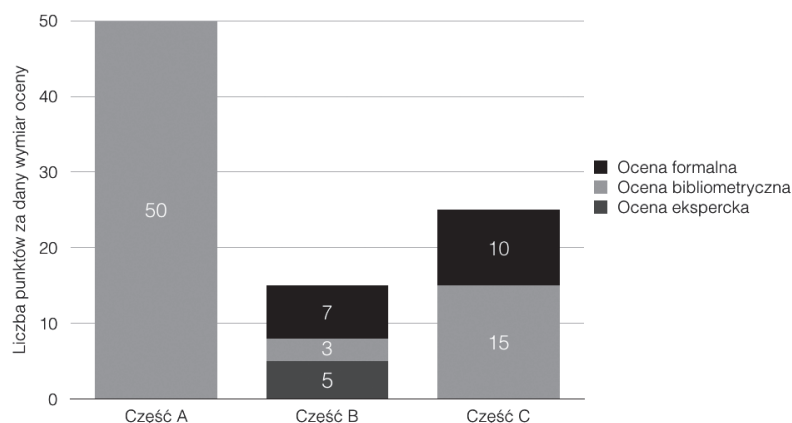
W 2015 r. zostały wprowadzone dwie istotne zmiany:

- 1) został wskazany zakres punktowy wyznaczający, ile punktów dane czasopismo może uzyskać – od 10 do 25 pkt (na poprzednim WCP wszystkie czasopisma z części C miały 10 pkt);
- 2) dodatkowe punkty mogły uzyskać czasopisma, które oprócz indeksowania na ERIH-u są rejestrowane w bazie SCOPUS służącej za źródło danych do bazy Scimago. W tym celu został zaproponowany współczynnik Scimago Impact Factor 2 (SIF2) obliczany jako iloraz współczynnika Total Cites (3 years) i współczynnika Citable Docs. (3 years) publikowanych w bazie Scimago Journal & Country Rank.

Hybrydowa metoda ustalania liczby punktów na części C, polegająca na tym, że kryte-

Nowożytna. Studia nad kulturą medyczną”: 4 punkty (poprzednio na części C: 10 punktów), siedem kolejnych czasopism zachowało *status quo* (10 punktów), następnie: 11 czasopism uzyskało 11 punktów, 18 czasopism uzyskało 12 punktów, 11 czasopism uzyskało 13 punktów, 11 czasopism uzyskało 14 punktów i 3 czasopisma otrzymały 15 punktów.

rium doboru czasopism zawarte jest w jednej bazie (ERIH), a źródło danych koniecznych do wyliczenia wskaźnika w drugiej bazie (Scimago), budzi wiele wątpliwości natury bibliometrycznej i naukometrycznej.



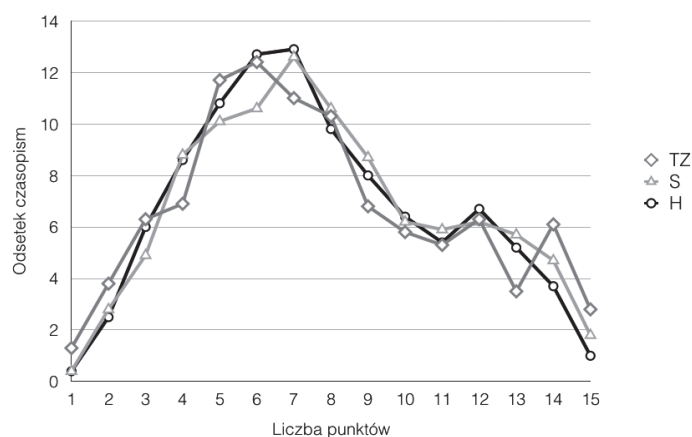
Ryc. 1. Porównanie maksymalnej liczby punktów do zdobycia przez czasopismo na podstawie oceny w trzech wymiarach (ocenie kryteriów formalnych, ocenie bibliometrycznej i ocenie eksperckiej) w poszczególnych częściach „Wykazu czasopism punktowanych”. W części B punkty za wymiary zostały uśrednione dla trzech grup nauk. W części C ocena formalna została utożsamiona z indeksowaniem w bazie European Reference Index for the Humanities

Rycina 1 pokazuje, że na WCP z 2015 r. wymiar, jakim jest ocena bibliometryczna czasopisma, ma znaczenie w konstrukcji każdej z części WCP, chociaż waga tego wymiaru jest różna. Warto podkreślić, że WCP z 2015 r. jest pierwszym wykazem tworzonym według nowego modelu oceny, w którym ocenę bibliometryczną stosuje się we wszystkich częściach.

3. Rekonstrukcja zasad tworzenia oraz charakterystyka części B „Wykazu czasopism punktowanych” z 2015 roku

Redakcje czasopism złożyły 2412 ankiet aplikacyjnych na część B na WCP. Ostatecznie zostało uwzględnionych 2212 czasopism. Rycina 2 pokazuje rozkład punktów przyznanych czasopismom w podziale na poszczególne grupy nauk TZ, S i H.

Analizie poddaliśmy dane wszystkich czasopism z części B: (a) pochodzące z ankiet ewaluacyjnych złożonych w Polskiej Bibliografii Naukowej, (b) dane o spełnianiu lub niespełnianiu poszczególnych kryteriów ewaluacyjnych, (c) wyniki oceny formalnej, bibliometrycznej i eksperckiej. Podstawą analizy były dane źródłowe uzyskane z Ośrodka Przetwarzania Informacji oraz Index Copernicus International. Nie stwierdziliśmy braku danych, co umożliwiło przeprowadzenie analizy na pełnej zbiorowości.



Ryc. 2. Odsetek czasopism z Wykazu czasopism punktowanych z określoną liczbą punktów w poszczególnych obszarach nauk: technicznych, ścisłych, medycznych, przyrodniczych (TZ); społecznych (S); nauk humanistycznych (H)

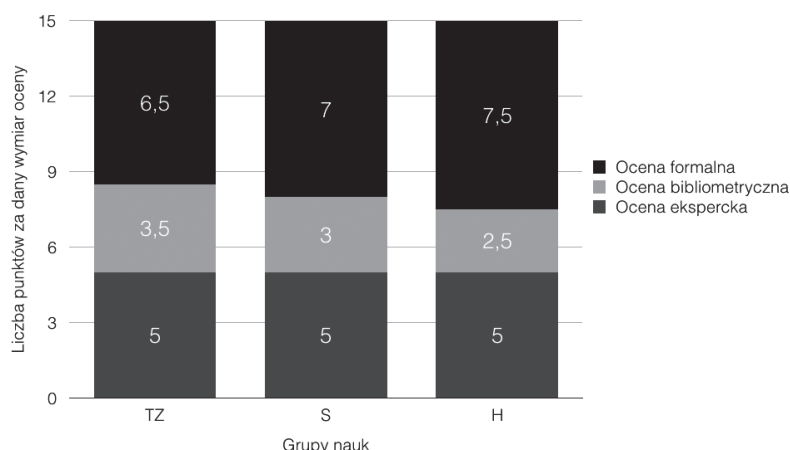
Ocena czasopism aplikujących do części B przebiegała dwuetapowo. W pierwszym etapie oceniano spełnianie następujących „kryteriów wstępnych”:

- 1) udostępnianie listy recenzentów publikacji nie rzadziej niż raz w roku na stronie internetowej czasopisma lub w numerze wydrukowanym;
- 2) stosowanie udostępnionej na stronie internetowej czasopisma lub w wersji drukowanej procedury recenzowania;
- 3) posiadanie przez czasopismo naukowe czynnej i aktualnej strony internetowej;
- 4) co najmniej 50% recenzentów oceniających publikacje naukowe stanowią recenzenci zewnętrzni;
- 5) zamieszczenie w każdym artykule naukowym opublikowanym w okresie dwóch lat poprzedzających rok wypełnienia ankiety co najmniej tytułu i streszczenia w języku angielskim;
- 6) stabilność wydawnicza w okresie dwóch lat poprzedzających rok wypełnienia ankiety, potwierdzona ciągłością wydawania czasopisma (brak opóźnień większych niż 6 miesięcy).

Do drugiego etapu przechodziły czasopisma, które spełniły co najmniej 5 z 6 kryteriów wstępnych⁵. W drugim etapie przyznawano punkty za poszczególne parametry oceny. Wyznaczono takie same parametry dla trzech grup nauk. Wszystkie czasopisma zostały

⁵ Ostatecznie decyzją ministerialnego zespołu ds. oceny czasopism, na ostatnim etapie przygotowania WCP, zmniejszono liczbę wymaganych kryteriów wstępnych do trzech. W ten sposób na część B na WCP zostało zakwalifikowanych dodatkowo 176 czasopism. Gdyby zostały w pełni utrzymane założenia z Komunikatu z 2015 r., w części B znalazłoby się 2036 czasopism.

poddane trzem rodzajom oceny⁶: formalnej, bibliometrycznej oraz eksperckiej. Rycina 3 pokazuje wpływ poszczególnych rodzajów oceny na ostateczną punktację czasopisma, czyli liczbę możliwych do uzyskania punktów za maksymalny wynik w danym rodzaju oceny.



Ryc. 3. Porównanie maksymalnej liczby punktów do zdobycia przez czasopismo na podstawie oceny formalnej, bibliometrycznej i eksperckiej w poszczególnych grupach nauk: technicznych, ścisłych, medycznych, przyrodniczych (TZ); społecznych (S); nauk humanistycznych (H). Ocena bibliometryczna obejmuje wsakźnik Predicted Impact Factor lub Scimago Impact Factor oraz wprowadzenie danych do bazy POL-index

W poszczególnych rodzajach oceny wskazano następujące parametry:

I. Ocena formalna

- 1) Odsetek autorów z zagraniczną afiliacją.
- 2) Indeksacja w bazach danych.
- 3) Liczba artykułów publikowanych na rok.
- 4) Odsetek recenzentów z zagraniczną afiliacją.
- 5) Częstotliwość wydawania.
- 6) Język publikacji.
- 7) Odsetek członków rady naukowej z zagraniczną afiliacją.
- 8) Wersja on-line.
- 9) Okres funkcjonowania czasopisma.

⁶ Niniejszy trójelementowy podział oceny nie jest wskazany w Komunikacie z 2015 r. Mowa jest jedynie o dwóch etapach oceny. My natomiast na owym drugim etapie wyróżniamy trzy wymiary przeprowadzanej oceny. Czynimy tak, aby uporządkować rekonstrukcję zasad WCP oraz wskazać, że ocena bibliometryczna jest możliwa do wyróżnienia na każdej z części WCP, a ocena ekspercka jedynie na części B.

II. Ocena bibliometryczna

- 10) Cytowalność czasopisma.
- 11) Wprowadzenie danych do bazy POL-index.

Za spełnienie parametrów oceny formalnej i bibliometrycznej czasopismo mogło uzyskać maksymalnie 10 punktów.

III. Ocena ekspercka

Dodatkowe 5 punktów mogło być przyznane czasopismu na podstawie oceny eksperckiej. Komunikat MNiSW z 2015 r. wskazuje, że te dodatkowe punkty przydziela zespół ds. oceny czasopism, zasięgając opinii innych zespołów eksperckich lub komitetów (np. Polskiej Akademii Nauk) dotyczącej poszczególnych czasopism z dziedzin nauki właściwych dla opiniujących zespołów.

3.1. Ocena formalna

Ocena formalna polegała na analizie deklaracji przedstawionych przez redakcje w ankiecie i na tej podstawie przypisaniu czasopismu określonej liczby punktów według zasad ustalonych odrębnie dla każdej grupy nauk w Komunikacie MNiSW z 2015 r.

Tabela 1 pokazuje, że różnicowanie punktacji za spełnianie kryteriów formalnych dla różnych grup nauk odnosiło się do minimalnych wartości parametrów (tzw. progów granicznych) lub puli punktów do zdobycia za dane kryterium. Nie wprowadzono zróżnicowania tylko przy dwóch kryteriach: odsetek recenzentów z zagraniczną afiliacją oraz wersja online.

Zestaw kryteriów formalnych zastosowanych w ewaluacji przeprowadzonej w roku 2015 różnił się od zestawu z 2013 r. Odnotowaliśmy trzy rodzaje różnic: dodanie i usunięcie kryteriów formalnych, zmiana wielkości progów granicznych i zmiana punktacji za spełnianie kryteriów. Zasady ewaluacji przyjęte w 2015 r. nie uwzględniały jednego kryterium z 2013 r., tj. „posiadanie przez czasopismo redaktorów językowych, tematycznych lub redaktora statystycznego”, natomiast zostało dodane nowe kryterium, tj. „okres funkcjonowania czasopisma”. Jeżeli chodzi o zmianę wielkości progów granicznych to dotyczyła ona kryteriów: „odsetek autorów z zagraniczną afiliacją” (zmniejszenie z 25 do 10 w TZ, zwiększenie z 5 do 10 w S), „liczba artykułów publikowanych na rok” (zmniejszenie z 24 do 12 dla roczników we wszystkich grupach nauk), „odsetek recenzentów z zagraniczną afiliacją” (zmniejszenie z 50 do 10 w TZ), „język publikacji” (zmniejszenie ze 100 do 20 w TZ), „odsetek członków rady naukowej z zagraniczną afiliacją” (wprowadzenie dodatkowego wymogu o liczbie członków w TZ). Porównanie punktów przyznawanych za kryteria formalne jest trudne do interpretacji z uwagi na odmienną skalę oraz odmienne kryteria.

Tabela 1. Kryteria formalne wraz z progami granicznymi oraz punktacją dla poszczególnych grup nauk

Kryterium formalne	Próg graniczny			Punktacja		
	TZ	S	H	TZ	S	H
Odsetek autorów z zagraniczną afiliacją [%]	≥ 10	≥ 10	≥ 5	0,5	0,5	1
Indeksacja w bazach danych [liczba baz]	SCOPUS lub WoS ≥ 2	≥ 2 = 1	≥ 2 = 1	1 0,5	1 0,5	1 0,5
Liczba artykułów publikowanych na rok	≥ 12 (rocznik) ≥ 24 (inne)	≥ 12 (rocznik) ≥ 24 (inne)	≥ 12 (rocznik) ≥ 24 (inne)	0,5	0,5	1
Odsetek recenzentów z zagraniczną afiliacją [%]	≥ 10	≥ 10	≥ 10	0,5	0,5	0,5
Częstotliwość wydawania [liczba numerów/rok]	≥ 4	≥ 4	≥ 4	0,5	1	0,5
Język publikacji [%]	≥ 20	≥ 20 ≥ 5 i < 20	≥ 5	1	1 0,5	0,5
Odsetek członków rady naukowej z zagraniczną afiliacją [%]	≥ 50 + min. 5 osób z zagraniczną afiliacją	≥ 10	≥ 10	0,5	0,5	0,5
Wersja on-line [%]	100	100	100	1	1	1
Okres funkcjonowania czasopisma [liczba lat]	≥ 20 ≥ 10 i < 20	≥ 10 ≥ 5 i < 10	≥ 20 ≥ 10 i < 20 ≥ 5 i < 10	1 0,5	1 0,5	1,5 1 0,5

Analiza zmian pozwala na sformułowanie kilku wniosków: 1) dodane w 2015 r. kryterium miało na celu docenienie czasopism o dłuższym stażu i było jednocześnie niekorzystne dla prężnie rozwijających się młodych czasopism naukowych, 2) największa zmiana w wielkości progów punktowych objęła grupę nauk TZ i odnosiła się do kryteriów związanych z umiędzynarodowieniem; w trzech kryteriach znacząco obniżono wymogi, w przypadku jednego nieznacznie je zaostrzono⁷, 3) progi punktowe w grupie nauk H i S niemal nie uległy zmianie, za wyjątkiem podniesienia w grupie S minimalnego „odsetka autorów z zagraniczną afiliacją”, 4) dokonane zmiany zbliżyły wymogi w grupie TZ do grup S i H, tym samym nowe zasady bardziej doceniały czasopisma o charakterze lokalnym, aniżeli przewidywały to zasady obowiązujące w 2013 r.

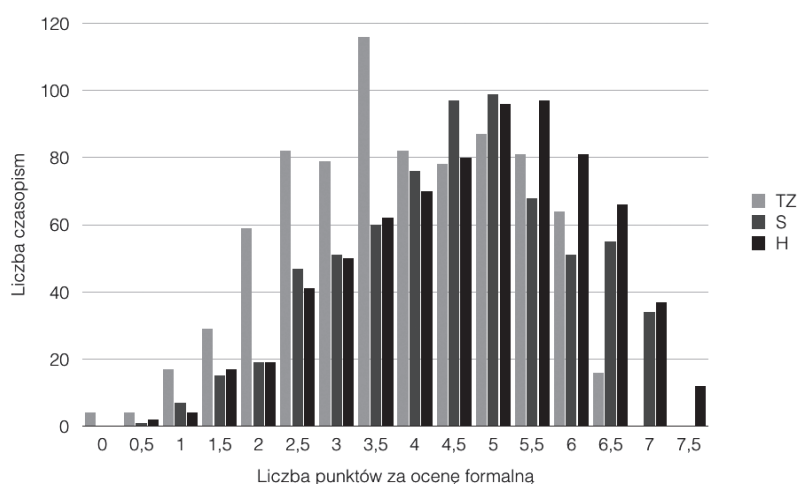
3.1.1. Charakterystyka wyników oceny formalnej

W pierwszej części analizy odpowiadaliśmy na pytanie o charakterystykę wyników oceny formalnej. Charakterystykę wyników oceny formalnej zoperacjonalizowaliśmy jako wynik oceny formalnej stanowiący sumę punktów uzyskanych przez czasopisma za poszczególne kryteria formalne.

⁷ Wprowadzono wymóg minimalnej liczby członków rady naukowej z zagraniczną afiliacją ($N=5$).

Przeanalizowaliśmy rozkład wyników dla czasopism z każdej z grup nauk oddzielnie. Obliczyliśmy w danej grupie wartość średniego wyniku oceny formalnej (M) oraz odchylenie standardowe, czyli rozrzut wyników wokół tej średniej (SD). Uwzględniliśmy 798 czasopism z grupy TZ ($M = 3,85$; $SD = 1,42$), 680 czasopism z grupy S ($M = 4,5$; $SD = 1,42$) oraz 734 czasopism z grupy H ($M = 4,7$; $SD = 1,47$).

Rozkład wyników oceny formalnej jest lewoskośny w grupach nauk: S (skośność = $-0,234$; błąd standardowy skośności = $0,094$) oraz H (skośność = $-0,346$; błąd standardowy skośności = $0,09$), co oznacza, że w tych grupach większość czasopism uzyskiwała wyniki powyżej średniej. Rozkład wyników jest platykurtyczny w grupie nauk TZ (kurtoza = $-0,724$; błąd kurtozy = $0,173$), co oznacza duże rozproszenie wyników od średniej w tym zbiorze czasopism. Wyniki oceny formalnej były wyższe w przypadku czasopism z grup H i S aniżeli czasopism z grupy TZ. Świadczy o tym zarówno wyższa wartość średniej (M), jak i odmienny kształt histogramu obrazującego rozkład wyników (zob. ryc. 4).



Ryc. 4. Liczba czasopism, które otrzymały punkty za ocenę formalną w poszczególnych grupach nauk: technicznych, ścisłych, medycznych, przyrodniczych (TZ); społecznych (S); nauk humanistycznych (H)

Analiza wyników oceny formalnej wykazała, że najniższy wynik (0 pkt) uzyskały cztery czasopisma z grupy TZ. Czasopisma z grupy nauk humanistycznych i społecznych częściej uzyskiwały wartości wyższe niż wartości średnich obliczonych dla wyników w tych dwóch grupach, o czym świadczy lewoskośny rozkład. Czasopisma z grupy nauk technicznych, ścisłych, medycznych, przyrodniczych uzyskiwały bardziej rozrzucone od średniej wyniki, o czym świadczy platykurtyczny rozkład.

3.1.2. Charakterystyka wyników za kryteria określające umiędzynarodowienie

W drugiej części analizy odpowiadaliśmy na pytanie o wartość kryteriów formalnych określających umiędzynarodowienie oraz na to, czy wartość ta pozwala na spełnienie tych kryteriów. Spośród kryteriów formalnych uznaliśmy za kryteria określające umiędzynarodowienie czasopism: a) odsetek autorów z zagraniczną afiliacją, b) odsetek recenzentów z zagraniczną afiliacją, c) język publikacji, d) odsetek członków rady naukowej z zagraniczną afiliacją. Posłużyliśmy się dla czterech zakwalifikowanych kryteriów wartościami zadeklarowanymi w ankiecie przez redakcje oraz punktami przyznanymi za ich spełnianie.

Część B ma za zadanie indeksować przede wszystkim czasopisma, które nie są rejestrowane w międzynarodowych bazach bibliometryczno-bibliograficznych. Jednakże konstrukcja kryteriów i parametrów oceny odzwierciedla jednocześnie przyjmowane wartości w polskiej polityce naukowej. Zatem kryteria ewaluacji mają stymulować i przyspieszać osiągnięcie przez polskie czasopisma odpowiedniego poziomu umiędzynarodowienia, które rozumiane jest jako włączanie czasopism do obiegu ponadnarodowego oraz angażowanie zagranicznych naukowców w komunikację naukową realizowaną przez polskich badaczy. Należy jednak mieć na uwadze, że powszechnie, tj. zarówno w literaturze przedmiotu, jak i w krajowych systemach ewaluacji nauki, podkreśla się diametralne zróżnicowanie poziomu umiędzynarodowienia i – wynikającej z tego cytowalności – między tzw. naukami twardymi a miękkimi, czyli naukami przyrodniczymi (TZ) a naukami społecznymi (S) i humanistyką (H). Takie też założenie o istotnym zróżnicowaniu tych typów nauk będziemy przyjmować w interpretacji wyników.

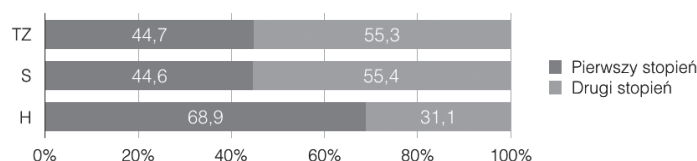
Wprowadziliśmy do analizy stopnie spełnienia kryteriów, co umożliwiło nam przedstawienie wyników bez konieczności normalizacji punktów przyznawanych za dane kryterium. Porównaliśmy odsetek czasopism w poszczególnych grupach spełniających wymienione kryteria na jednym ze stopni. Przypomnijmy, że taka konieczność wynikała ze zróżnicowanej skali punktowej za spełnienie danego parametru w różnych obszarach nauk. Pierwszy stopień oznacza największą liczbę punktów możliwych do uzyskania, drugi albo trzeci (w zależności od wskaźnika) określa 0 pkt – a dokładniej: niespełnienie kryterium.

Porównanie odsetka czasopism z TZ, S i H, które spełniały dane kryterium (związane ze wskaźnikiem umiędzynarodowienia), pozwoliło nam wnioskować o różnicach między wynikami w poszczególnych grupach nauk. Przeprowadziliśmy dodatkowy zabieg polegający na zrównaniu progów granicznych, aby porównać, jak kształtowałyby się rozkład wyników, gdyby przyjąć takie same zasady oceny dla trzech grup nauk (TZ, S i H). Ten zabieg pozwolił na porównanie wyników między grupami opartymi na odmiennych założeniach (o równości oczekiwań względem nauk twardych i miękkich).

Umożliwiło to nam ocenę, czy uzyskane różnice oznaczają, że czasopisma z danego obszaru nauki są bardziej umiędzynarodowione od czasopism z innego obszaru, czy też różnica jest raczej następstwem odmiennych progów punktowych.

Odsetek autorów z zagraniczną afiliacją

Rycina 5 pokazuje odsetek czasopism, które spełniły kryterium „odsetek autorów z zagraniczną afiliacją”. Widać znaczące zróżnicowanie między wynikiem uzyskanym przez czasopisma z grupy TZ i S a wynikami osiągniętymi przez czasopisma z grupy H. Tę dysproporcję może tłumaczyć różnica między wielkością progów granicznych. W przypadku TZ i S próg wynosił 10%, natomiast w przypadku H: 5% autorów z zagraniczną afiliacją. Sprawdziliśmy, jaki odsetek czasopism z grupy H spełniałby kryterium, a tym samym uzyskał punkty, gdyby zastosowano ten sam próg, co w przypadku TZ i S. Wynosiłby on 53, byłaby więc to wartość nadal wyższa od wyników w pozostałych grupach nauk.



Ryc. 5. Odsetek czasopism, które uzyskały punkty za spełnienie kryterium „odsetek autorów z zagraniczną afiliacją” w poszczególnych grupach nauk: technicznych, ścisłych, medycznych, przyrodniczych (TZ); społecznych (S); nauk humanistycznych (H)

Posłużyliśmy się wartością z ankiet ewaluacyjnych, tj. odsetkiem autorów z zagraniczną afiliacją, aby sprawdzić, jakie są różnice między wartościami kryterium w poszczególnych grupach nauk. Wykorzystaliśmy test *t*Studenta dla prób niezależnych. Analiza wykazała, że wartość kryterium różni się istotnie statystycznie w zależności od tego, czy czasopismo należy do grupy nauk TZ czy też do S, $t(1466) = 3,179$; $p < 0,005$, d Cohena = 0,166⁸). Średnia wartość kryterium „odsetek autorów z zagraniczną afiliacją” jest wyższa w grupie nauk TZ ($M = 17,93$; $SD = 24,671$) niż w grupie nauk S ($M = 14,29$; $SD = 19,401$). Wartość kryterium różni się również istotnie statystycznie w zależności od tego, czy czasopismo należy do grupy nauk H czy do S, $t(1408) = 3,83$; $p < 0,001$, d Cohena = 0,204). Średnia wartość kryterium „odsetek autorów z zagraniczną afiliacją” jest wyższa w grupie nauk H ($M = 18,51$; $SD = 22,008$) niż średnia wartość tego kryterium w grupie nauk S. Jednocześnie wielkość efektu jest w obu opisanych porówna-

⁸ Wartość współczynnika d Cohena = 0,2 oznacza słaby efekt, d Cohena = 0,5 przeciętny efekt i d Cohena = 0,8 silny efekt (Cohen, 1992).

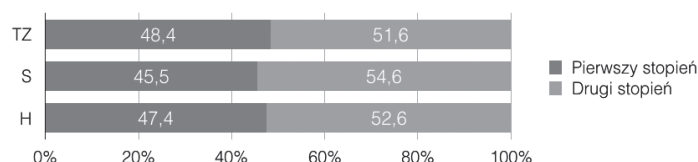
niach bardzo mała oraz mała, co w tym przypadku oznacza nieznaczną różnicę między wartościami kryterium. Nie stwierdziliśmy istotnych statystycznie różnic przy porównaniu wartości kryterium w grupie nauk TZ i H.

Zaskakujący jest niski wynik w grupie nauk TZ, w szczególności, po porównaniu go z wynikiem czasopism z grupy nauk H. W pierwszej grupie punkty otrzymało 44,7% czasopism, a w drugiej 68,9%. Ponadto, jeżeli spojrzeć na wartości kryterium, nie ma istotnych statystycznie różnic między odsetkiem autorów z zagraniczną afiliacją w tych dwóch grupach nauk. Jest to wynik odmienny od oczekiwanego, gdzie zarówno wartość kryterium, jak i wynik (spełnianie kryterium) są wyższe w TZ aniżeli w H. Czasopisma w grupie nauk TZ wydają się najbardziej predestynowane do umiędzynarodowienia z uwagi na charakter tzw. twardej nauki, która z założenia ma charakter bardziej międzynarodowy niż humanistyka czy nauki społeczne. Z drugiej strony w tej grupie nauk wśród autorów (z zagranicznych jednostek) znaczenie mają przede wszystkim publikacje w czasopismach indeksowanych w JCR, co może być przeszkodą w zachęcaniu ich do publikowania w polskich czasopismach lokalnych. Jednocześnie zastanawiająca jest wysoka wartość kryterium oraz wysoki wynik w grupie H. Może on odzwierciedlać odsetek przekładów tekstów zagranicznych autorów publikowanych w polskich czasopismach humanistycznych, prezentowanych w ankiecie jako oryginalne artykuły naukowe. Jest to jednak tylko przypuszczenie nieodnoszące się do całej analizowanej grupy, które jednak pozwalamy sobie postawić, bazując na przeprowadzonych przez nas badaniach nad publikacjami polskich humanistów w czasopismach z listy ERIH (zob. Kulczycki, Rozkosz, Drabek 2015).

Odsetek recenzentów z zagraniczną afiliacją

Rycina 6 pokazuje odsetek czasopism, które spełniły kryterium „odsetek recenzentów z zagraniczną afiliacją”. Wyniki uzyskiwane w poszczególnych grupach nauk są zbliżone, co oznacza, że podobny odsetek czasopism w grupach TZ, S i H uzyskał punkty za spełnienie omawianego kryterium. Próg graniczny był taki sam dla wszystkich grup i wynosił 10, co oznacza, że zgodnie z deklaracją (uznaną w procesie ewaluacji) ponad 45% wszystkich czasopism miało w składzie recenzentów co najmniej 10% osób z zagranicznej instytucji naukowej.

Wykorzystaliśmy dane z ankiet, tj. odsetek recenzentów z zagraniczną afiliacją, aby przyjrzeć się deklarowanemu udziałowi recenzentów z zagranicy w czasopismach z poszczególnych grup nauk. Zastosowaliśmy test *t*Studenta dla prób niezależnych, aby porównać wartość kryterium w grupach TZ, S i H. Analiza wykazała, że wartość kryterium różni się istotnie statystycznie w zależności od tego, czy czasopismo należy do grupy nauk TZ, czy też do S, $t(1472) = 3,613$; $p < 0,001$, d Cohena = 0,19). Średnia wartość kryterium „odsetek recenzentów z zagraniczną afiliacją” jest wyższa w grupie nauk TZ ($M = 20,56$; $SD = 27,171$) niż w grupie nauk S ($M = 15,94$; $SD = 21,988$).

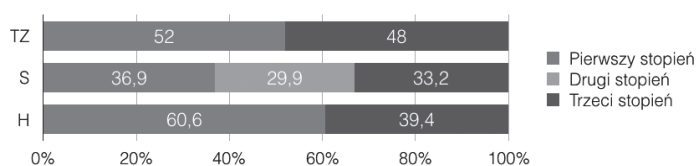


Rycina 6. Odsetek czasopism, które uzyskały punkty za spełnienie kryterium „odsetek recenzentów z zagraniczną afiliacją” w poszczególnych grupach nauk: technicznych, ścisłych, medycznych, przyrodniczych (TZ); społecznych (S); humanistycznych (H)

Siła efektu jest jednak mała, o czym świadczy wartość współczynnika d Cohena. Nie stwierdziliśmy istotnych statystycznie różnic przy porównaniu wartości kryterium w grupach nauk TZ i H ($M = 18,36$; $SD = 25$) oraz S i H. Wyniki po raz kolejny pokazały duże podobieństwo zarówno wartości kryterium, jak i wyników między grupami TZ i H, co jest rezultatem odmiennym od oczekiwanego (zob. wyniki analizy kryterium „odsetek autorów z zagraniczną afiliacją”).

Język publikacji

Rysunek 7 pokazuje odsetek czasopism, które spełniły kryterium „język publikacji”. Kryterium to oznaczało procentowy udział artykułów naukowych opublikowanych w pełnej wersji w językach kongresowych (angielskim, francuskim, hiszpańskim, niemieckim, rosyjskim lub włoskim) w stosunku do ogólnej liczby opublikowanych artykułów naukowych. Przewidziano odrębne skale dla poszczególnych grup nauk (TZ: 1; S: 1, 0,5; H: 0,5), co utrudnia interpretację wyników. Przyjęliśmy zasadę, że czasopismo spełnia kryterium, gdy uzyskało jakiekolwiek punkty za przekroczenie progu granicznego. Analiza pokazała, że osiągnięte w poszczególnych grupach wyniki znacząco się różnią.



Ryc. 7. Odsetek czasopism, które uzyskały punkty za spełnienie kryterium „język publikacji” w poszczególnych grupach nauk: technicznych, ścisłych, medycznych, przyrodniczych (TZ); społecznych (S); humanistycznych (H)

Najwyższy odsetek czasopism spełniających kryterium jest w grupie S, niższy w grupie H i najniższy w TZ. Podobnie jak w przypadku pierwszego kryterium lepsze wyniki w grupie S i H tłumaczy wyższy próg graniczny w TZ: 20%, aniżeli w S i H: 5%. Sprawdziliśmy, jaki odsetek czasopism z grup S i H spełniałby kryterium, a tym samym uzys-

kał punkty, gdyby zastosowano ten sam próg, co w przypadku TZ. Wynosiłby on dla czasopism z grupy S i H odpowiednio: 30 oraz 34,4. Byłaby więc to wartość znacznie niższa od uzyskanej przez czasopisma z grupy TZ.

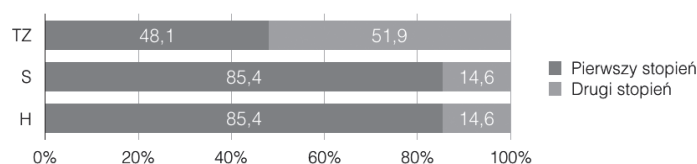
Sprawdziliśmy wartość kryterium z ankiet. Odnosiła się ona do odsetka artykułów naukowych opublikowanych w językach kongresowych. Wykorzystaliśmy test *t*Studenta dla prób niezależnych, aby porównać wartość kryterium w grupach TZ, S i H. Analiza wykazała, że wartość kryterium różni się istotnie statystycznie w zależności od tego, czy czasopismo należy do grupy nauk TZ czy też do grupy S, $t(1455) = 9,325$; $p < 0,001$, d Cohena = 0,49, a także różni się istotnie statystycznie w zależności od tego czasopisma należy do grupy TZ czy do grupy H, $t(1493) = 9,081$; $p < 0,001$, d Cohena = 0,46. Średnia wartość kryterium „język publikacji” jest wyższa w grupie nauk TZ ($M = 44,99$; $SD = 44,778$) niż w grupie nauk S ($M = 25,9$; $SD = 33,774$) oraz w wyższa w grupie nauk TZ niż w grupie nauk H ($M = 26,39$; $SD = 35,145$). Wartość współczynnika d Cohena w obu porównaniach świadczy o przeciętnym efekcie (zróznicowaniu wartości kryterium). Nie stwierdziliśmy istotnych statystycznie różnic przy porównaniu wartości kryterium w grupach nauk S i H. Odnotowaliśmy zatem spodziewaną różnicę między wartościami kryterium w grupie nauk twardych (TZ) a grupach nauk miękkich (S i H). Znacząco większy odsetek artykułów naukowych w czasopismach z TZ publikowany był w językach kongresowych aniżeli w czasopismach z H i S.

Odsetek członków rady naukowej z zagraniczną afiliacją

Rycina 8 pokazuje odsetek czasopism, które spełniły kryterium „odsetek członków rady naukowej z zagraniczną afiliacją”. Widoczna jest znacząca dysproporcja między wynikami w grupie TZ a wynikami w grupach H i S. W TZ wynik jest niemal o połowę niższy niż w pozostałych grupach. Ma to związek z różnicami w wartościach progów granicznych. W grupie nauk H i S wynoszą one 10%. W przypadku TZ spełnienie kryterium wiąże się z przekroczeniem progu granicznego 50%, jak również dodatkowego warunku, aby w radzie zasiadało co najmniej pięć osób z zagraniczną afiliacją. Sprawdzenie, jaki odsetek czasopism z grupy S i H spełniałby kryterium, a tym samym uzyskał punkty, gdyby zastosowano ten sam próg, co w przypadku TZ, nie jest w pełni możliwe z uwagi na brak pełnych danych (brak informacji o liczbie członków rady naukowej z zagraniczną afiliacją). Jednakże wstępna symulacja oparta na samym odsetku członków rady naukowej z zagraniczną afiliacją pozwala określić, że przy podwyższeniu progu kryterium spełniałoby co najwyżej 54,1% czasopism z grupy S i co najwyżej 44,9% czasopism z grupy H, byłby więc to wynik zbliżony do uzyskanego przez czasopisma z grupy TZ.

Podobnie jak w przypadku wcześniejszych kryteriów związanych z umiędzynarodowieniem sprawdziliśmy wartość kryterium z ankiet dla poszczególnych grup nauk. Wartość ta odnosiła się do odsetka członków rady naukowej z zagraniczną afiliacją. Za pomo-

cą testu *t*Studenta dla prób niezależnych przeanalizowaliśmy, czy wartości kryterium w różnych grupach różnią się między sobą. Nie stwierdziliśmy istotnych statystycznie różnic. Średnia wartość kryterium „odsetek członków rady naukowej z zagraniczną afiliacją” jest podobna w grupie nauk TZ ($M = 39,56$; $SD = 27,535$), S ($M = 40,16$; $SD = 24,78$) oraz H ($M = 38,1$; $SD = 24,417$).



Ryc. 8. Odsetek czasopism, które uzyskały punkty za spełnienie kryterium „odsetek członków rady naukowej z zagraniczną afiliacją” w poszczególnych grupach nauk: technicznych, ścisłych, medycznych, przyrodniczych (TZ); społecznych (S); humanistycznych (H)

Zaprezentowane wyniki wskazują na spełnianie kryteriów formalnych określających umiędzynarodowienie oraz ich wartość (odsetek autorów z zagraniczną afiliacją, odsetek recenzentów z zagraniczną afiliacją, język publikacji, odsetek członków rady naukowej z zagraniczną afiliacją). W grupie nauk TZ podobny lub niższy niż w grupach nauk S i H odsetek czasopism uzyskiwał punkty za spełnienie kryteriów umiędzynarodowienia. Oznacza to, że czasopisma z TZ nie zyskały na obecności (jak można byłoby się spodziewać) kryteriów umiędzynarodowienia w zasadach oceny czasopism.

3.2. Ocena bibliometryczna

W związku z oceną bibliometryczną czasopism krajowych, dla których nie wylicza się IF, wprowadzono wskaźniki PIF oraz SIF. Jednocześnie, zdając sobie sprawę, że PIF i SIF – ze względu na swoją konstrukcję – w zasadzie nie nadają się do oceny polskich czasopism z nauk humanistycznych i społecznych, zaproponowano w 2013 roku jeszcze jeden wskaźnik, tj. Polski Współczynnik Wpływu (PWW). Ma on bazować na danych ze źródeł polskich. Jego wzór i definicję sformułowano w Komunikacie Ministra Nauki i Szkolnictwa Wyższego z dnia 29 maja 2013 r. w sprawie kryteriów i trybu oceny czasopism naukowych (zwanego dalej Komunikatem MNiSW z 2013 r.), nie rezygnując jednocześnie z PIF. Komunikat zapowiadał obowiązywanie PWW od 2014 r. Jednakże ze względu na brak danych, czyli brak polskiego indeksu cytowań, postulat tego nie udało się do tej pory zrealizować.

W 2015 roku w ocenie bibliometrycznej na części B na WCP pojawiły się dwa kryteria: (1) cytowalność oraz (2) wprowadzenie danych do POL-indexu.

3.2.1. Pierwsze kryterium: cytowalność

Cytowalność ustalana była w oparciu o dwa wskaźniki:

A) Predicted Impact Factor (PIF) – wyliczany na podstawie danych wyszukanych w ba-

zie WoS oraz podanej w ankiecie liczby artykułów naukowych. Jego definicja zmieniła się względem tej ogłoszonej w Komunikacie Ministra Nauki i Szkolnictwa Wyższego z dnia 14 września 2012 r. w sprawie kryteriów i trybu oceny czasopism naukowych i Komunikacie MNiSW z 2013 r. Wskaźnik ten kalkulowany był dla czasopism krajowych, które aplikowały do części B na WCP.

B) Scimago Impact Factor (SIF) – wyliczany na podstawie danych z Scimago Journal & Country Rank (tworzonej w oparciu o dane z bazy SCOPUS) oraz danych przekazywanych przez redakcje w ankiecie czasopisma. Dotyczy on periodyków polskich indeksowanych w bazie SCOPUS, które złożyły na ankietę ewaluacyjną.

Definicja wskaźnika PIF brzmi następująco: PIF „obliczany [jest] jako iloraz liczby cytowań publikacji danego czasopisma naukowego w bazie WoS za okres od trzech lat dorobku poprzedzającego rok wypełnienia ankiety i liczby wszystkich artykułów naukowych opublikowanych w czasopiśmie w tym okresie” (Komunikat MNiSW z 2015 r.). Niestety definicji tej nie towarzyszy wzór ani dokładniejsze wyjaśnienie, dlatego w dyskusjach w środowisku naukowym można było spotkać jego różne interpretacje. Ostatecznie przyjęto następujące doprecyzowanie: PIF to iloraz liczby cytowań publikacji danego czasopisma (niezależnie od daty wydania tej publikacji) w bazie WoS w latach 2012-2014 i liczby artykułów naukowych opublikowanych w latach 2012-2014 (zgodnie z danymi zgłoszonymi w ankiecie).

Definicja wskaźnika SIF, prostsza i niebudząca wątpliwości, brzmi następująco: „Scimago Impact Factor (SIF) obliczany [jest] jako iloraz współczynnika Total Cites (3 years) publikowanego w bazie Scimago Journal & Country Rank dostępnej na dzień 31 sierpnia roku przeprowadzania oceny czasopism naukowych i liczby wszystkich artykułów naukowych opublikowanych w czasopiśmie w okresie od trzech lat do roku poprzedzających rok wypełnienia ankiety”. Ze względu na dane źródłowe, wskaźnik ten mógł być wyliczany tylko dla czasopism indeksowanych w bazie SCOPUS.

Wszystkie czasopisma dzielone były na trzy części zgodnie z zadeklarowaną w ankiecie grupą nauk. Następnie dla każdej z nich układane były dwie listy rankingowe obejmujące czasopisma z wyliczonym PIF lub SIF i uszeregowane według malejącej wartości tego wskaźnika (osobno dla każdego z nich). Punkty miało szansę dostać 20% czasopism z listy z najwyższą wartością wskaźnika (odpowiednio: 1,5 punktu dla grupy nauk TZ; 1 punkt dla czasopism z grupy nauk S i 0,5 dla grupy nauk H). Kolejne 30% czasopism mogło otrzymać odpowiednio 1 punkt, 0,5 lub 0,25. W sytuacji, gdy czasopismo na obu listach miało odmienną pozycję rankingową, liczyła się ta wyższa lokata i punktacja.

Wyliczenie obu współczynników, przygotowanie list rankingowych, a następnie ostatecznej punktacji zostało przeprowadzone przez Ośrodek Przetwarzania Informacji. Dane o cytowaniach czasopism w WoS opracowane zostały, na zlecenie Ministerstwa Nauki i Szkolnictwa Wyższego, przez bibliotekarzy z Uniwersytetu Śląskiego w Kato-

wicach. Podstawowym kryterium wyszukiwawczym była nazwa czasopisma (tytuł oryginalny) i jej warianty (w tym także podany w ankiecie tytuł w języku angielskim). Wzięto pod uwagę wszystkie odszukane cytowania w latach 2012-2014, niezależnie od daty wydania cytowanego artykułu. Do wyszukania cytowań w bazie WoS wykorzystano formularz wyszukiwawczy „Cited Reference Search”. Wśród zestawu sprawdzanych czasopism były również te, które ostatecznie nie znalazły się w części B na WCP opublikowanego 23 grudnia 2015 r. Warto zauważyć, że wyszukiwanie cytowań czasopism po tytule nie jest łatwe. Co prawda opisy bibliograficzne w bazie od kilku lat nie są już skracane, ale sposoby zapisu tytułu czasopisma wydają się być niezliczone. Problem ten został dostrzeżony przez Chadaj i Turecką (2007), które poszukując cytowań czasopisma „Geologia”, znalazły 57 wariantów w bazie WoS, czy też przez Bemke-Świtlik i Drabek (w druku), które także zwróciły uwagę na wielowariantowość tytułów.

Biorąc pod uwagę specyfikę bazy WoS oraz to, że tytuły czasopism z bibliografii załącznikowej są identyfikowane tylko w przypadku, gdy dany periodyk jest indeksowany w tej bazie (choć i tu zdarzają się pomyłki), łatwo dojść do wniosku, że tytuł czasopisma nie jest idealnym kryterium wyszukiwawczym. Jednak w sytuacji, gdy trzeba w dość krótkim czasie sprawdzić dużą liczbę czasopism, jest to jedyne kryterium, jakie można wykorzystać. W związku z tym warto zwrócić uwagę, że istnieje sporo czynników, które mogą wpłynąć na niezbyt dokładny wynik. Wśród nich można wymienić: wiele wariantów tytułu, zmiany tytułu (czasem kilkukrotne w ciągu kilku lat), błędy w tytułach, zgłoszenie serii monograficznej jako czasopisma, błędy w opisach bibliograficznych, czy istnienie różnych czasopism o takiej samej nazwie, wreszcie autorskie lub redakcyjne pomyłki, do jakich czasem dochodzi przy sporządzaniu bibliografii załącznikowej. Im bardziej skomplikowany tytuł, tym trudniej odszukać wszystkie cytowania tego czasopisma. Im więcej członów zawiera tytuł, tym częściej autorzy (lub redakcje) będą skłonni do stosowania różnych skrótów, co później przełoży się na trudności w wyszukaniu i identyfikacji takiego czasopisma. Na niektórych internetowych stronach czasopism można znaleźć wzór cytowania lub prośbę o używanie pełnego tytułu. To dobra praktyka, niestety nie zawsze można się do niej stosować, jeśli w instrukcji dla autorów znajdują się zupełnie inne wytyczne.

3.2.2. Drugie kryterium: wprowadzenie danych do POL-indexu

Do POL-indexu należało wprowadzić pełne bibliografie załącznikowe z publikowanych artykułów naukowych. W przypadku braku bibliografii redakcje miały wprowadzić opisy bibliograficzne cytowanych w artykułach prac zebrane z przypisów dolnych lub końcowych. Akceptowane były wyłącznie dane kompletne. Zbiór danych obligatoryjnych obejmował:

- tytuł artykułu,
- typ artykułu,

- tom/numer,
- rok,
- język artykułu,
- imiona i nazwiska autorów,
- wykaz cytowanej literatury (bibliografia załącznikowa albo jej opracowanie na podstawie przypisów dolnych lub końcowych).

Dodatkowo można było podać informacje fakultatywne, takie jak: tytuł artykułu w innych językach, strony od – do, afiliacje autorów, role autora w tworzeniu artykułu, języki abstraktów, dziedziny i dyscypliny dotyczące artykułu, słowa kluczowe oraz DOI. Informacje wprowadzone do POL-indexu zostały sprawdzone pod kątem kompletności danych przez operatora merytorycznego bazy. Jako niekompletne uznane były zbiory danych pozbawione jednego z trzech elementów: autora, tytułu lub bibliografii. Sprawdzono również, czy liczba artykułów naukowych podana w ankiecie ewaluacyjnej czasopisma jest taka sama, jak liczba artykułów podanych w bazie POL-index. Punkty otrzymały te czasopisma, dla których poziom błędów dla wprowadzonych danych nie przekroczył 20%.

Twórcy niektórych bibliograficznych baz danych (m.in. BazTech, BazEkon czy Central European Journal of Social Sciences and Humanities [CEJSH]) wsparli redakcje indeksowanych czasopism w przygotowaniu pliku w formacie xml z danymi bibliograficznymi pochodzącymi z tych baz. Redakcje czasopism importowały plik do bazy POL-index, uzupełniały i poprawiały ewentualne braki i błędy, a następnie zatwierdzały dane. Redakcje czasopism niedysponujące plikiem xml z danymi bibliograficznymi wprowadzały dane do POL-indexu za pomocą formularza udostępnionego na stronie tej bazy. Za dostarczenie poprawnych danych czasopisma mogły otrzymać 2 punkty (za pełne dane z lat 2009-2014) lub 1 punkt (za pełne dane z lat 2013-2014).

3.2.3. Charakterystyka wyników oceny bibliometrycznej. Kryterium cytowalności

W pierwszej części analizy sprawdziliśmy, czy wprowadzenie wskaźnika SIF pozwoliło spełnić kryterium cytowalności czasopismom, które nie spełniały tego kryterium na podstawie wartości PIF. Analizy pokazały, że wartość SIF nie miała istotnego wpływu na spełnienie tego kryterium. Uwzględnienie SIF w ocenie podwyższyło punktację zaledwie czterech czasopism z grupy nauk TZ (z 1 pkt do 1,5 pkt) i nie podwyższyło oceny żadnego czasopisma z grupy nauk S i H.

PIF był uwzględniany w ocenie spełnienia kryterium cytowalności: $N = 794$ czasopism z grupy nauk TZ, $N = 680$ z grupy nauk S oraz $N = 732$ czasopism z grupy nauk H.

Tabela 2 prezentuje średnią wartość PIF, odchylenie standardowe oraz minimalną wartość PIF warunkującą uzyskanie pierwszego, drugiego i trzeciego stopnia cytowa-

ności. W analizie wyników posłużyliśmy się miarami stopni, aby uniknąć normalizowania punktów przyznawanych za spełnienie kryterium dla poszczególnych grup nauk.

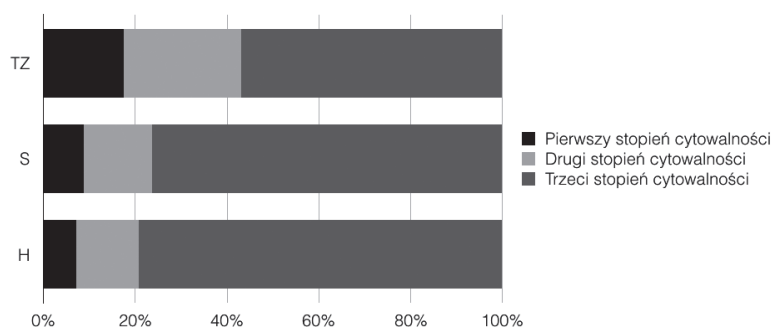
Tabela 2. Średnie (*M*), odchylenia standardowe (*SD*) oraz wartość minimalna (*Min*) dla PIF obliczone dla czasopism z trzech grup nauk: technicznych, ścisłych, medycznych, przyrodniczych (TZ); społecznych (S); humanistycznych (H)

Grupa nauk	Pierwszy stopień cytawalności			Drugi stopień cytawalności			Trzeci stopień cytawalności		
	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>M</i>	<i>SD</i>	<i>Min</i>
TZ (<i>N</i> = 794)	5,393	7,615	1,657	0,838	0,355	0,405	0,12	0,116	0
S (<i>N</i> = 680)	0,686	0,691	0,195	0,086	0,04	0,039	0,006	0,01	0
H (<i>N</i> = 732)	0,778	0,617	0,268	0,139	0,045	0,085	0,01	0,1	0

Adnotacja. W obliczeniach nie uwzględniono dwóch czasopism z grupy nauk H (z uwagi na brak danych) oraz czterech czasopism z grupy nauk TZ (z uwagi na nieuwzględnienie wartości PIF przy ustalaniu punktów za spełnienie kryterium cytawalności). Pierwszy stopień cytawalności (TZ: 1,5 pkt, S: 1 pkt, H: 0,5 pkt), drugi stopień cytawalności (TZ: 1 pkt, S: 0,5 pkt, H: 0,25 pkt), trzeci stopień cytawalności (0 pkt).

Pierwszy stopień cytawalności oznacza uzyskanie: 1,5 pkt (TZ), 1 pkt (S) lub 0,5 pkt (H). Drugi stopień cytawalności oznacza uzyskanie: 1 (TZ), 0,5 (S) lub 0,25 (H). Trzeci stopień cytawalności oznacza uzyskanie 0 pkt we wszystkich grupach nauk.

Rycina 9 prezentuje rozkład wyników za kryterium cytawalności. Wielkości są przedstawione w procentach, co pozwala porównać wyniki uzyskane w poszczególnych grupach nauk.



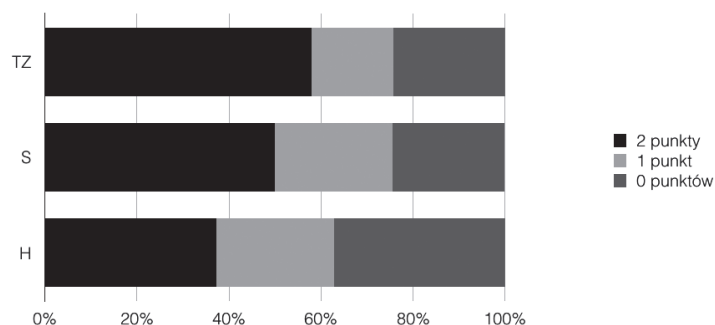
Ryc. 9. Odsetek czasopism z poszczególnych grup nauk – technicznych, ścisłych, medycznych, przyrodniczych (TZ); społecznych (S); humanistycznych (H) – które uzyskały punkty za pierwszy stopień cytawalności (TZ: 1,5 pkt, S: 1 pkt, H: 0,5 pkt), za drugi stopień cytawalności (TZ: 1 pkt, S: 0,5 pkt, H: 0,25 pkt) i za trzeci stopień cytawalności (0 pkt)

Analiza wyników odnoszących się do kryterium cytawalności wykazała, że znacząco większy odsetek czasopism z grupy nauk TZ spełniało to kryterium i tym samym uzyskało punkty (43,1%) aniżeli odsetek czasopism z grup S i H (odpowiednio: 23,7% oraz

20,7%). Potwierdza to różnice międzydyscyplinarne, wynikające z odmiennej kultury cytowań (Slyder i in., 2011; Wróblewski, 2002), małą reprezentację czasopism z nauk społecznych i humanistycznych w WoS, a także różne oddziaływanie omawianego parametru na wyniki ewaluacji czasopism. Wniosek ten potwierdza również analiza odsetka czasopism, których wartość $PIF = 0$ i wartość $SIF = 0$. W grupie nauk TZ taki warunek spełniało 10,1% czasopism, natomiast w grupach nauk S i H, odpowiednio 51,6% oraz 52,4%.

3.2.4. Charakterystyka wyników oceny bibliometrycznej. Kryterium „wprowadzenie danych do POL-index”

Dane zostały poprawnie wprowadzone do systemu POL-index przez 1578 z 2212 czasopism, z czego 1074 (48,6%) otrzymało po 2 pkt, a 504 (22,8%) po 1 pkt. Pozostałe czasopisma w liczbie 634 (28,7%) nie otrzymały punktów ze względu na niezgłoszenie danych do POL-indexu lub umieszczenie w bazie niepełnych danych.



Ryc. 10. Odsetek czasopism, które uzyskały punkty za wprowadzenie danych do POL-indexu z poszczególnych grup nauk: technicznych, ścisłych, medycznych, przyrodniczych (TZ); społecznych (S); humanistycznych (H)

Rycina 10 pokazuje odsetek czasopism z poszczególnych grup nauk, które spełniły kryterium „wprowadzenie danych do POL-indexu”. Punkty otrzymało 75,7% czasopism z grupy nauk TZ, 75,5% czasopism z grupy nauk S oraz 62,8% czasopism z grupy nauk H. Wynika z tego, że znacznie większy odsetek czasopism z grup TZ i S dostarczył dane aniżeli z grup H. Na ten wynik mogły mieć wpływ trudności z przetworzeniem opisów bibliograficznych z przypisów dolnych na bibliografię załącznikową.

3.3. Ocena ekspercka

Do oceny eksperckiej zostały dopuszczone wszystkie czasopisma, dla których złożono ankietę aplikacyjną do część B i spełniły co najmniej 5 z 6 kryteriów wstępnych (por. przypis 5). Dodatkowe punkty mogły być przyznane, zgodnie z komunikatem określającym

jącym tryb i sposób oceny, na podstawie oceny czasopisma naukowego przez środowisko naukowe, spełnianie standardów etycznych, wydawniczych, ocenę wkładu w naukę polską oraz naukę światową, jak i innych kryteriów oceny eksperckiej.

W maju 2015 r. sekretarz stanu w MNiSW skierował do prezesa Polskiej Akademii Nauk (PAN) pismo z prośbą o przeprowadzenie przez komitety naukowe i problemowe PAN oceny eksperckiej krajowych czasopism naukowych. Minister wskazał, że komitety PAN są demokratycznie wybraną najszerszą reprezentacją środowisk naukowych. Jednocześnie minister zaproponował, aby – jeśli zajdzie taka potrzeba – komitety zapraszały do współpracy przy ocenie przedstawicieli wybranych jednostek posiadających kategorię naukową A+ lub A. Komitety zostały poproszone o wyróżnienie czasopism o szczególnie wysokim poziomie naukowym: „dany komitet może zaproponować przyznanie 5 punktów nie więcej niż 10% czasopism z ocenianego obszaru i 3 punktów dla kolejnych 10% najwyżej ocenianych czasopism. W sumie premiowanych jest 20% czasopism z ocenianej listy” (Ratajczak, 2015). Jednocześnie zostało podkreślone, że „biorąc pod uwagę olbrzymie zróżnicowanie charakteru, problematyki czy profilu czasopism, (...) zespół [ds. oceny czasopism – przyp. E.K., E.A.R., A.D.] odstąpił od formułowania szczegółowych kryteriów, którymi mają kierować się komitety. Podstawowym i wspólnym kryterium ma być jakość i ranga naukowa czasopisma” (Ratajczak, 2015).

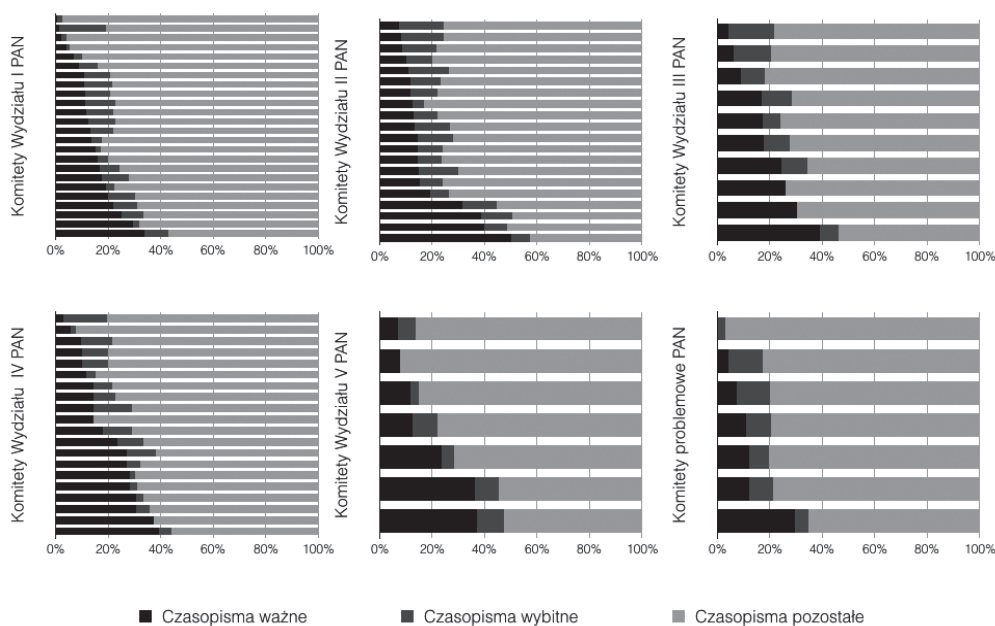
Podczas składania ankiety na część B na WCP redakcje czasopism musiały wskazać jedną lub dwie podstawowe oraz maksymalnie trzy dodatkowe dyscypliny naukowe, które charakteryzują dane czasopismo. Można było wybrać spośród wykazu określonego w Rozporządzeniu Ministra Nauki i Szkolnictwa Wyższego z dnia 8 sierpnia 2011 r. w sprawie obszarów wiedzy, dziedzin nauki i sztuki oraz dyscyplin naukowych i artystycznych. Na tej podstawie Index Copernicus International – jeden z operatorów technicznych procesu oceny czasopism – przyporządkował poszczególne czasopisma do komitetów PAN w oparciu o algorytm zestawiający zadeklarowane przez czasopisma dyscypliny z tymi, które reprezentują komitety PAN. Wstępne zestawienie zostało przesłane do poszczególnych komitetów, aby ich przewodniczący mogli wnieść uwagi i zastrzeżenia. Każdy z komitetów PAN mógł zadeklarować dowolną liczbę dyscyplin naukowych, w których się specjalizuje. W ten sposób dane czasopismo było oceniane przez kilka lub kilkanaście komitetów. Jednocześnie eksperci komitetu mogli dokonać oceny periodyków spoza listy czasopism przypisanych do danego komitetu. Rozwiązanie to miało pozwolić na poprawienie efektów prac algorytmu, gdyby istotne czasopismo nie zostało przyporządkowane do danego komitetu.

W ten sposób dwa komitety PAN otrzymały do oceny ponad 500 czasopism, a 41 komitetów ponad 100 czasopism. Zaledwie 5 komitetów miało przypisanych mniej niż 10 czasopism. Oceny eksperckiej dokonały 94 komitety PAN, w tym wszystkie komitety Wydziału I, III i IV oraz niemalże wszystkie komitety Wydziału II (dwa komitety nie

dokonały oceny) i Wydziału V (pięć komitetów nie dokonało oceny). Również siedem komitetów problemowych PAN przesłało swoją propozycję oceny.

Komitety PAN powoływały ekspertów, którzy oceniali czasopisma. Liczba ekspertów wahała się od jednego do kilkunastu na komitet. Mogli oni każde z przypisanych do komitetu czasopism ocenić jako: (1) wybitne, (2) ważne, (3) pozostałe lub podjąć decyzję o rezygnacji z oceny w przypadku występowania konfliktu interesów. Na podstawie przygotowanych ocen dany komitet PAN tworzył ranking czasopism. Wyniki rankingów po ich akceptacji przez przewodniczących komitetów PAN zostały przekazane do ministerialnego zespołu ds. oceny czasopism naukowych, który dopiero na tym etapie zaproponował algorytm przypisywania punktów na podstawie wskazań ekspertów.

Rycina 11 pokazuje, że większość komitetów nie dotrzymała wskazanych kryteriów oceny, czyli przekroczyła odsetek możliwych czasopism do wyróżnienia. Załedwie 14 (na 94) komitetów dokonało oceny zgodnie ze wskazaną procedurą i oceniło do 10% czasopism przypisanych do danego komitetu jako ważne oraz do 10% za wybitne. Jeden z komitetów uznał, że 50% przypisanych czasopism to czasopisma ważne (41 na 82 czasopisma), a 7,3% to wybitne (6 na 82 czasopisma).



Ryc. 11. Odsetek czasopism uznanych za ważne, wybitne i pozostałe przez poszczególne komitety Polskiej Akademii Nauk (podział na komitety w ramach wydziałów oraz komitety programowe). Odsetek czasopism ważnych i wybitnych nie powinien być przekroczyć łącznie 20%. Na wykresie wyniki uszeregowane są według odsetka czasopism ważnych. Uwzględniono wyniki tych komitetów, którym przypisano co najmniej po 10 czasopism do oceny

Ze względu na otrzymane wyniki od komitetów PAN ministerialny zespół ds. oceny czasopism zmodyfikował zasady przypisywania czasopismom dodatkowych punktów i skalę trójstopniową (0, 3, 5 punktów) zamieniono na skalę sześciostopniową (0, 1, 2, 3, 4, 5 punktów). O przypisaniu czasopismu odpowiedniej liczby punktów decydowała liczba komitetów, które dane czasopismo wyróżniły oraz liczba ocen („wybitne”, „ważne”). Ostatecznie niemalże wszystkie czasopisma, które zostały wskazane jako ważne lub wybitne przez choćby jeden komitet, otrzymały dodatkowe punkty za ocenę ekspercką.

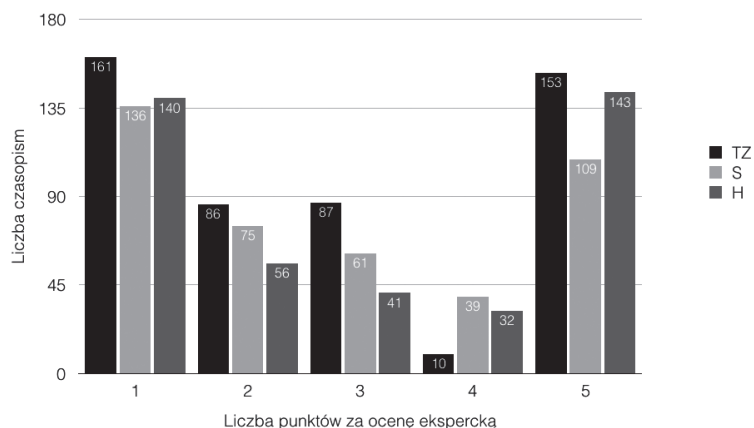
Ekspertci powołani przez komitety PAN oddali łącznie 9972 głosy na wyróżnione czasopisma, w tym 7843 na czasopisma uznane za ważne, a 2129 na czasopisma wybitne. Aż 1367 czasopism zostało wyróżnionych przez co najmniej jeden komitet (wskazywano również czasopisma, które ostatecznie zostały umieszczone w części A albo C na WCP). Tabela 3 ilustruje rozkład punktów uzyskanych za ocenę ekspercką. Ostatecznie aż 60,1% ($N = 1329$) czasopism umieszczonych w części B uzyskało dodatkowe punkty za oceną ekspercką, a nie jak wstępnie zakładano 20% ($N = 442$).

Tabela 3. Rozkład punktów uzyskanych za ocenę ekspercką przez czasopisma znajdujące się w części B na Wykazie czasopism punktowanych ($N = 2212$)

Liczba punktów	N	%	% skumulowany
5	405	18,3	18,3
4	81	3,7	22
3	189	8,5	30,5
2	217	9,8	40,3
1	437	19,8	60,1
0	883	39,9	100

Rysunek 12 pokazuje rozkład punktów przyznanych za ocenę ekspercką w poszczególnych grupach nauk. Blisko jedna piąta czasopism otrzymała najwyższą ocenę, czyli 5 punktów. Dodatkowe punkty uzyskało najwięcej czasopism z grupy nauk TZ – 497 czasopism (69,5% tej grupy), następnie z grupy nauk S – 420 czasopism (64,8%), najmniej natomiast z grupy nauk H – 412 czasopism (59,3%).

Rozkład we wszystkich grupach nauk: TZ, S i H jest platykurtyczny (odpowiednio: kurtoza = -1,137; błąd standardowy kurtozy = 0,182; kurtoza = -1,045; błąd standardowy kurtozy = 0,192; kurtoza = -1,117; błąd standardowy kurtozy = 0,185) i prawoskośny (odpowiednio: skośność = 0,573; błąd standardowy skośności = 0,091; skośność = 0,65; błąd standardowy skośności = 0,096; skośność = 0,714; błąd standardowy skośności = 0,093). Oznacza to, że we wszystkich grupach nauk wyniki były bardziej rozproszone wokół średniej, a jednocześnie rozkład miał charakter asymetryczny (więcej czasopism uzyskało wyniki poniżej średniej).



Ryc. 12. Liczba czasopism, które otrzymały dodatkowe punkty za ocenę ekspercką w poszczególnych grupach nauki: technicznych, ścisłych, medycznych, przyrodniczych (TZ); społecznych (S); humanistycznych (H)

Tabela 4. Wady i zalety oceny eksperckiej krajowych czasopism z części B Wykazu czasopism punktowanych

Zalety	Wady
• Włączenie oceny eksperckiej zrealizowało liczne postulaty środowiska naukowego, które uważało ocenę formalną i bibliometryczną za niewystarczającą. • Zaangażowanie przedstawicieli środowiska naukowego do oceny eksperckiej. • Ocena ekspercka umożliwiła docenienie czasopism, które nie spełniają części parametrów oceny związanych np. z umiędzynarodowieniem czasopism – istotne szczególnie dla nauk humanistycznych i społecznych.	• Brak jasnych kryteriów wyznaczających sposób oceny czasopism przez ekspertów. • Zwrócenie się o ocenę ekspercką jedynie do komitetów Polskiej Akademii Nauk – pominięcie stowarzyszeń i towarzystw naukowych. • Algorytm przypisywania czasopism do komitetów sprawił, że niektóre komitety miały do oceny kilkaset czasopism. • Zbyt wiele wyróżnionych czasopism przez większość komitetów (istotnie więcej niż zakładane 20%). • Ponad połowa czasopism została ostatecznie uznana za „ważne” lub „wybitne” i otrzymała dodatkowe punkty z oceny eksperckiej. • Brak uzasadnień wyników oceny. • Brak możliwości odwołania od oceny. • Podniesienie zakresu punktowego (poprzez dodanie oceny eksperckiej) części B na „Wykazie czasopism punktowanych” z 10 do 15 punktów. Doszło do zrównania najlepszych lokalnych czasopism z czasopismami posiadającymi wyliczony współczynnik Impact Factor z części A, a przez to doprowadziło do „inflacji punktów”.

Najbardziej docenionym czasopismem w ocenie eksperckiej było czasopismo „Journal of Plant Protection Research”, które otrzymało wyróżnienie od 13 komitetów PAN, w tym dziewięć oceniło to czasopismo jako wybitne, a cztery jako ważne.

Tabela 4 pokazuje, że ocena ekspercka odpowiadała licznym postulatom środowiska naukowego i umożliwiała docenienie lokalnych czasopism. Oprócz zalet można wskazać wady oceny eksperckiej: zarówno na poziomie samych zasad, jak i ich realizacji, np. brak jasnych kryteriów czy wyróżnianie zbyt wielu czasopism przez komitety PAN niż zakładano. Dlatego jednoznaczna ocena tego procesu nie jest łatwa. Ocena ekspercka została przeprowadzona po raz pierwszy, dlatego można powiedzieć, iż był to sprawdzian zarówno dla samych zasad oceny, jak i środowiska naukowego, które oceniało czasopisma.

Wyniki przedstawione w sekcji 3 odpowiadają na dwa pytania badawcze, czyli na pytanie nr 1 „Jakie były zasady tworzenia części B na WCP w 2015 r. oraz w jaki sposób przeprowadzano ocenę formalną, bibliometryczną i ekspercką?” oraz na pytanie nr 2 „Jaka była charakterystyka wyników oceny: formalnej, bibliometrycznej i eksperckiej?”.

4. Ocena ekspercka a pozostałe wymiary oceny

Postanowiliśmy sprawdzić, czy wprowadzenie oceny eksperckiej do zasad ewaluacji zmieniło pozycję rankingową czasopism w części B na WCP. W ten sposób oceniamy zasadność wprowadzenia oceny eksperckiej do ewaluacji. Zbadaliśmy, czy eksperci doceniali bardziej czasopisma z wysokimi wartościami wskaźnika cytowalności oraz wskaźników umiędzynarodowienia, czy może czasopisma z niskimi wartościami wskaźnika cytowalności i wskaźników umiędzynarodowienia. W tym celu skonstruowaliśmy trzy pytania badawcze odnoszące się do czasopism z każdej grupy nauk: 1) czy istnieje związek między wynikami oceny eksperckiej a wynikami oceny formalnej i bibliometrycznej, 2) czy istnieje związek między wynikami oceny eksperckiej a wartością wskaźnika cytowalności, 3) czy istnieje związek między wynikami oceny eksperckiej a wartościami wskaźników umiędzynarodowienia.

Podstawą założenia o związku między oceną formalną i bibliometryczną a oceną ekspercką jest sposób postrzegania we współczesnej komunikacji naukowej wskaźników cytowań jako pośredniego wskaźnika jakości tekstu naukowego (Elsaie, Kammer, 2009; Eyre-Walker, Stoletzki, 2013). Takie uproszczenie stosuje się przy analizie bardzo wielu publikacji naukowych (np. publikacje wszystkich pracowników jednego uniwersytetu). Jednocześnie tego samego założenia nie powinno się przyjmować w odniesieniu do oceny pojedynczej pracy: tę można ocenić jedynie poprzez lekturę, czyli ocenę ekspercką. Przyjmujemy zatem, że czasopismo może mieć wyliczone cytowania (ocena bibliometryczna) tylko wtedy, kiedy będzie indeksowane w bazie bibliometryczno-bibliograficznej. To natomiast jest możliwe po spełnieniu odpowiednich kryteriów (ocena formalna), które są wyznacznikiem profesjonalizmu czasopisma. Spełnienie tych warunków jest jednocześnie pośrednim wyznacznikiem jakości czasopisma naukowego. Dlatego wyko-

rzystanie połączonych wyników oceny bibliometrycznej i formalnej może umożliwić przewidywanie wyników oceny eksperckiej.

4.1. Ocena ekspercka a ocena formalna i bibliometryczna

Wykorzystaliśmy zmienną stanowiącą sumę punktów przyznanych w wyniku oceny formalnej i bibliometrycznej (maksymalnie 10 pkt) – określamy tę zmienną jako wynik oceny formalnej i bibliometrycznej – aby sprawdzić związek wyniku tej oceny z wynikiem oceny eksperckiej. W przypadku wyniku będącego liczbą nienaturalną (np. 7,5 pkt) zaokrągliliśmy wartość w górę (np. z 7,5 pkt do 8 pkt).

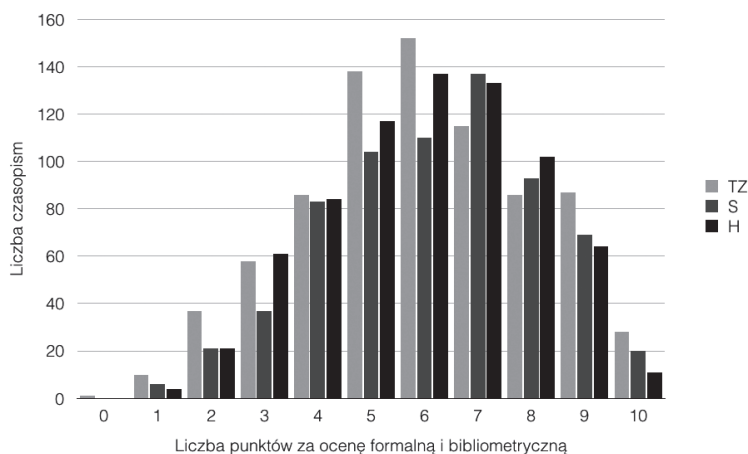
Tabela 5 pokazuje wyniki oceny formalnej i bibliometrycznej dla wszystkich grup nauk łącznie. Większość czasopism uzyskała bardzo wysokie wyniki za ocenę formalną i bibliometryczną: 60,7% ($N = 1344$) otrzymało ponad 5 pkt. Najwyższą liczbę punktów (10 pkt) otrzymało tylko 2,7% czasopism ($N = 59$). Najniższą (0 pkt) otrzymało poniżej 0,1% ($N = 1$) czasopism. Przypadek czasopisma, które nie spełniało żadnego kryterium formalnego, jak również nie uzyskało żadnych punktów w wyniku oceny bibliometrycznej, pokazuje, że eksperci mogą nie brać pod uwagę żadnej z cech czasopisma, uwzględnionej w zasadach oceny formalnej i bibliometrycznej, a jednak uznać czasopismo za ważne (czasopismo otrzymało 1 pkt w ocenie eksperckiej).

Tabela 5. Rozkład punktów uzyskanych za ocenę formalną i ocenę bibliometryczną przez czasopisma znajdujące się w części B w Wykazie czasopism punktowanych ($N = 2212$)

Liczba punktów	N	%	% skumulowany
10	59	2,7	2,7
9	220	9,9	12,6
8	281	12,7	25,3
7	385	17,4	42,7
6	399	18	60,7
5	359	16,2	76,9
4	253	11,4	88,3
3	156	7,1	95,4
2	79	3,6	99
1	20	0,9	99,9
0	1	0,1	100

Sprawdziliśmy, jak kształtował się rozkład punktacji w poszczególnych grupach. Rycina 13 pokazuje wyniki oceny formalnej i bibliometrycznej w grupach nauk TZ, S i H. Wartość średniej dla wyniku oceny formalnej i bibliometrycznej (M) i odchylenie standardowe od tej średniej (SD), wyniosło w grupach nauk: TZ ($M = 5,96$; $SD = 2,117$), S ($M = 6,17$; $SD = 1,989$) oraz H ($M = 6$; $SD = 1,93$). Rozkład w grupie nauk TZ i H okazał się być platykurtyczny (odpowiednio: kurtioza = $-0,569$; błąd standardowy kurtiozy

= 0,173; kurtoza = -0,629; błąd standardowy kurtozy = 0,18), natomiast w grupie nauk S platykurtyczny i jednocześnie lewoskośny (skośność = -0,226; błąd standardowy skośności = 0,094; kurtoza = -0,521; błąd standardowy kurtozy = 0,187).



Ryc. 13. Liczba czasopism, które otrzymały punkty za ocenę formalną i bibliometryczną w poszczególnych grupach nauki: technicznych, ścisłych, medycznych, przyrodniczych (TZ); społecznych (S); humanistycznych (H)

Postawiliśmy hipotezę zakładającą istnienie silnego związku między wynikami oceny eksperckiej oraz wynikami oceny formalnej i bibliometrycznej. Odrzucenie tej hipotezy (brak związku lub słaby związek) potwierdziłoby zasadność wprowadzenia oceny eksperckiej, natomiast potwierdzenie (silny związek) – brak zasadności jej wprowadzenia. Obliczyliśmy współczynnik korelacji r -Pearsona między wynikami oceny eksperckiej oraz wynikami oceny formalnej i bibliometrycznej. Korelacja w TZ wynosi $r = 0,547$; $p < 0,001$, d Cohena = 1,307, w S wynosi $r = 0,498$; $p < 0,001$, d Cohena = 1,148; w H wynosi $r = 0,369$; $p < 0,001$, d Cohena = 0,794. Wyniki okazały się istotne statystycznie. We wszystkich trzech grupach nauk, wraz ze wzrostem wyników za ocenę formalną i bibliometryczną, wzrastają wyniki za ocenę ekspercką. Na podstawie wartości wskaźnika d Cohena ustaliliśmy, że związek ten jest silny w grupie nauk TZ i S oraz przeciętny w grupie nauk H. Oznacza to, że w przypadku czasopism z grup nauk TZ i S można rozważyć odrzucenie hipotezy i podważyć zasadność wprowadzenia do zasad ewaluacji czasopism oceny eksperckiej, natomiast w przypadku czasopism z grupy nauk H wyniki nie pozwalają na odrzucenie hipotezy, co oznacza, że ocena ekspercka zmieniła pozycje rankingowe tych czasopism w części B.

4.2. Ocena ekspercka a wskaźnik cytowalności

Z oceny bibliometrycznej wyodrębniliśmy kryterium cytowalności jako jedyne w tym wymiarze oceny oparte na cytowaniach (pozycji czasopisma na liście rangowej utworzo-

nej na podstawie wartości PIF lub SIF, zob. 3.2.1). Sprawdziliśmy, czy recepcja czasopism wyrażona miarą cytowań pokrywać będzie się z oceną ekspercką. Postawiliśmy hipotezę zakładającą istnienie związku między liczbą punktów przyznanych za spełnienie kryterium cytowalności a wynikami oceny eksperckiej.

Obliczyliśmy współczynnik korelacji r Pearsona między wynikami oceny eksperckiej oraz liczbą punktów przyznaną za spełnienie kryterium cytowalności. Korelacja w TZ wynosi $r = 0,454$; $p < 0,001$, d Cohena = 1,019, w S wynosi $r = 0,442$; $p < 0,001$, d Cohena = 0,985, w H wynosi $r = 0,405$; $p < 0,001$, d Cohena = 0,886. Wyniki okazały się istotne statystycznie. Na podstawie wartości współczynnika d Cohena ustaliliśmy, że związek między analizowanymi zmiennymi we wszystkich trzech grupach nauk jest silny. Wraz ze wzrostem wyników za spełnienie kryterium cytowalności wzrastają wyniki za ocenę ekspercką.

Potwierdzenie istnienia związku między spełnianiem kryterium cytowalności i wynikami oceny eksperckiej stanowi istotny argument w dyskusji między zwolennikami i przeciwnikami stosowania wskaźników opartych na cytowaniach. Obala mit o braku przełożenia wielkości tych wskaźników na wyniki oceny jakościowej (eksperckiej). Trzeba przy tym pamiętać, że przedstawiony rezultat ma również swoje źródło w metodzie wyliczania wartości wskaźników cytowalności, tworzenia rankingów według wartości wskaźników i na podstawie tego uznawania, w jakim stopniu kryterium zostało spełnione (por. 3.2.1).

4.3. Ocena ekspercka a wskaźniki umiędzynarodowienia

Sprawdziliśmy, czy wartość wskaźników umiędzynarodowienia ma związek z wynikami oceny eksperckiej, a więc czy eksperci docenili czasopisma umiędzynarodowione lub aspirujące do zmiany profilu na międzynarodowy.

Obliczyliśmy współczynnik korelacji r Pearsona między wynikami oceny eksperckiej oraz liczbą punktów przyznaną za spełnienie poszczególnych kryteriów związanych ze wskaźnikami umiędzynarodowienia czasopism: a) odsetek autorów z zagraniczną afiliacją, b) odsetek recenzentów z zagraniczną afiliacją, c) język publikacji, d) odsetek członków rady naukowej z zagraniczną afiliacją.

Ocena ekspercka a odsetek autorów z zagraniczną afiliacją

Korelacja w TZ wynosi $r = 0,294$; $p < 0,001$, d Cohena = 0,615. Wyniki okazały się istotne statystycznie. W przypadku nauk TZ istnieje słaby związek (efekt przeciętny) między wynikami oceny eksperckiej a odsetkiem autorów z zagraniczną afiliacją. W przypadku grup nauk S i H siła związku była zbyt niska, a efekt zbyt mały (odpowiednio: $r = 0,122$; $p < 0,001$, d Cohena = 0,246; $r = 0,134$; $p < 0,001$, d Cohena = 0,27), aby uznać, że zmienne wyraźnie korelują ze sobą. Wyniki oceny eksperckiej w naukach miękkich nie wzrastały wraz z wielkością odsetka autorów z zagraniczną afiliacją. A za-

tem jedynie w grupie nauk TZ zauważalna była tendencja oferowania większej liczby punktów przez ekspertów czasopismom zabiegającym o autorów z zagranicy.

Ocena ekspercka a odsetek recenzentów z zagraniczną afiliacją

Korelacja w TZ wynosi $r = 0,275$; $p < 0,001$, d Cohena = 0,572. Jest to wynik zbliżony do poprzedniego (sprawdzającego związek wyników oceny eksperckiej z odsetkiem autorów z zagraniczną afiliacją). W grupie nauk TZ związek między wynikami oceny eksperckiej a odsetkiem recenzentów z zagraniczną afiliacją jest przeciętny. W przypadku grup nauk S i H siła związku i jego efekt były również zbyt niskie (w obu: $r = 0,175$; $p < 0,001$, d Cohena = 0,355), aby potwierdzić istnienie znaczącego związku między zmiennymi. Wyniki oceny eksperckiej w czasopismach z nauk miękkich nie wzrastały wraz z wielkością odsetka recenzentów z zagraniczną afiliacją.

Ocena ekspercka a język publikacji

Sprawdzenie współzmienności oceny eksperckiej i języka publikacji przyniosło odmienny (od powyżej zaprezentowanych) wynik. Korelacja w TZ wynosi $r = 0,41$; $p < 0,001$, d Cohena = 0,899, a w S wynosi $r = 0,234$; $p < 0,001$, d Cohena = 0,481. Wyniki okazały się istotne statystycznie. W przypadku nauk TZ i S istnieje odpowiednio średni i słaby związek między wynikami oceny eksperckiej a odsetkiem autorów z zagraniczną afiliacją. Przy czym siła związku jest w TZ duża, a w S przeciętna. W przypadku grupy nauk H siła związku okazała się również zbyt niska ($r = 0,138$; $p < 0,001$, d Cohena = 0,278), aby uznać, że zmienne znacząco korelują ze sobą. Oznacza to, że wyniki oceny eksperckiej w czasopismach humanistycznych nie wzrastały wraz z odsetkiem artykułów opublikowanych w językach kongresowych.

Ocena ekspercka a odsetek członków rady naukowej z zagraniczną afiliacją

Sprawdzenie współzmienności ostatniego wskaźnika umiędzynarodowienia przyniosło efekt zbliżony do wyników uzyskanych dla pierwszych dwóch wskaźników umiędzynarodowienia. Korelacja w TZ wynosi $r = 0,29$; $p < 0,001$, d Cohena = 0,606. Wynik jest istotny statystycznie i oznacza słaby związek między zmiennymi, a także przeciętną siłę tego związku. W przypadku grupy nauk S i H po raz kolejny siła związku oraz jego efekt okazała się zbyt niskie (odpowiednio: $r = 0,113$; $p < 0,001$, d Cohena = 0,227; $r = 0,11$; $p < 0,001$, d Cohena = 0,221), aby uznać, że zmienne korelują ze sobą. Wyniki oceny eksperckiej nie wzrastały wraz z odsetkiem członków rady naukowej z zagraniczną afiliacją.

Odnotowaliśmy istotne różnice między grupami nauk a korelacją wyników oceny eksperckiej ze wskaźnikami umiędzynarodowienia. W przypadku grupy nauk TZ zmienne te korelowały ze sobą w sposób wyraźny. W przypadku grupy nauk S wynik oceny eksperckiej korelował tylko z „językiem publikacji”. W przypadku grupy nauk H nie odnotowaliśmy, aby wynik oceny eksperckiej korelował z którymkolwiek ze wskaźników

umiędzynarodowienia. Te wyniki pokazują, że eksperci doceniają czasopisma dbające o wszystkie wskaźniki umiędzynarodowienia jedynie w grupie nauk twardych. Przedstawiciele nauk społecznych doceniali czasopisma publikujące w językach kongresowych. W humanistyce dbałość o wskaźniki umiędzynarodowienia nie miała związku z oceną ekspercką.

Wyniki przedstawione w sekcji 4 odpowiadają na dwa pytania badawcze, czyli na pytanie nr 3 „Czy wprowadzenie oceny eksperckiej do zasad ewaluacji zmieniło pozycję rankingową czasopism w części B na WCP?” oraz na pytanie nr 4 „Czy wyniki oceny eksperckiej wiążą się z: wynikami oceny formalnej i bibliometrycznej, wartością wskaźnika cytawalności, wartościami wskaźników umiędzynarodowienia?”.

Wprowadzenie oceny eksperckiej miało wpływ na pozycję rankingową czasopism w części B na WCP w grupie nauk H i nie miało istotnego wpływu na pozycję rankingową w grupie nauk TZ i S.

5. Podsumowanie

Wykaz czasopism punktowanych odgrywa ważną rolę nie tylko w procesie kategoryzacji jednostek naukowych, ale również w wyznaczaniu strategii publikacyjnej dla większości polskich naukowców i uczelni. Dlatego też każda aktualizacja WCP jest mocno dyskutowana w całym środowisku akademickim. Niniejszym tekstem chcieliśmy dostarczyć wnioski do analiz umożliwiających kompleksową debatę oraz realizację polityki naukowej opartej na dowodach.

Analizując całościowo zmiany dokonane we wszystkich trzech częściach WCP, można uznać, że zwiększenie liczby punktów przyznawanych czasopismom umieszczonym w części B spowodowało zrównanie najlepszych krajowych czasopism nieposiadających wyliczonego IF z wieloma prestiżowymi czasopismami międzynarodowymi. Trzeba bowiem pamiętać, że ze względu na algorytm normalizacyjny w części A niektórych dyscyplin naukowych najwyższej punktowane czasopisma mają przypisane 20 punktów.

Poniżej, w formie listy najważniejszych wniosków, prezentujemy podsumowanie analiz obejmujących zasady tworzenia części B i oceniania czasopism aplikujących do tej części WCP w 2015 r.:

- 1) W 2013 r. w ewaluacji czasopism można było wyróżnić dwa wymiary oceny: formalny i bibliometryczny. W 2015 r. wprowadzono nowy wymiar: ekspercki.
- 2) W ocenie formalnej wprowadzono zmiany, które obniżały wymogi odnoszące się do wskaźników umiędzynarodowienia w grupie nauk TZ. Skutkiem tej zmiany może być słabsza motywacja czasopism z nauk twardych we włączanie się w naukę światową.
- 3) Wprowadzenie zmian do zasad oceny spowodowało, że WCP z 2015 r. zachowuje większy balans między czasopismami z trzech grup nauk (TZ, S i H) niż wcześniej.

sze WCP. Większy balans oznacza, że podobny odsetek czasopism z poszczególnych grup nauk posiada określoną liczbę punktów. Ma to istotne znaczenie dla jakości narzędzia ewaluacyjnego, jakim jest WCP.

- 4) W ocenie formalnej czasopisma z grup nauk S i H częściej uzyskiwały wyniki wyższe od średniej niż niższe od średniej. Rozkład wyników w grupie nauk TZ był bardziej rozproszony wokół średniej.
- 5) Wyniki oceny formalnej ujawniły, że w przypadku kryteriów formalnych związanych z umiędzynarodowieniem, tj. odsetkiem osób z zagraniczną afiliacją wśród autorów, recenzentów i członków rady naukowej, odsetkiem artykułów naukowych napisanych w językach kongresowych, czasopisma z grupy nauk TZ uzyskiwały punkty z podobną częstotliwością lub rzadziej niż czasopisma z grup nauk S i H.
- 6) Wprowadzenie do oceny bibliometrycznej wskaźnika SIF okazało się nietrafione, gdyż wpłynął on na zmianę wyniku jedynie 4 z 2212 czasopism.
- 7) Wynik oceny bibliometrycznej ujawnił, że dane wprowadzone do POL-indexu, który ma być podstawą obliczenia PWW, wprowadziło 71,3% czasopism z części B. Daje to nadzieję na przeprowadzenie analizy bibliometrycznej w kolejnych latach także na podstawie polskiego indeksu cytowań.
- 8) Wyniki oceny eksperckiej pokazały, że w wyniku niedostosowania się przez komitety PAN do zasad wyznaczonych przez zespół ds. oceny czasopism dodatkowe punkty za ocenę ekspercką otrzymało 60,1% czasopism, a nie jak zakładano 20% czasopism.
- 9) Ocena ekspercka mogła zmienić ocenę czasopism, które otrzymały 0 pkt w wyniku oceny formalnej i bibliometrycznej, co pokazuje przykład jednego czasopisma z grupy nauk TZ. W ten sposób czasopismo znalazło się na WCP, mimo że nie spełniło żadnego kryterium formalnego i bibliometrycznego.
- 10) Analiza wyników ocen czasopism pokazuje, że można rozważyć rezygnację z oceny eksperckiej w grupach nauk TZ i S z uwagi na liniowy związek liczby punktów przyznanych w ramach oceny formalnej i bibliometrycznej z liczbą punktów przyznanych w ramach oceny eksperckiej. W przypadku grupy nauk H związek nie jest dostatecznie silny, aby uzasadnić odrzucenie oceny eksperckiej.
- 11) Cytowalność koreluje z oceną ekspercką we wszystkich grupach nauk, co może być istotnym argumentem w dyskusji na temat stosowania wskaźników bibliometrycznych w ocenie czasopism.
- 12) Wyniki oceny eksperckiej pokazały, że czasopisma w grupie nauk TZ nie uzyskały oczekiwanych, tj. wyższych od czasopism z grupy S i H, wyników za spełnienie uwzględnionych w ocenie formalnej wskaźników umiędzynarodowienia.

Zaprezentowane i zinterpretowane wyniki pokazują, że zmiany zasad oceny czasopism powinny być ugruntowane na wcześniej przeprowadzonych analizach naukowe-

trycznych i bibliometrycznych, które pozwoliłyby na użyteczną modyfikację dotychczasowych reguł. Jednocześnie warto podkreślić, że nie wszystkie z wymaganych kryteriów czy parametrów są potrzebne do przeprowadzenia oceny. Szczególnie, że dotychczasowe sposoby zgłaszania informacji przez redakcje uniemożliwiają rzetelną ocenę niektórych aspektów pracy wydawniczej. Jak pokazują nasze analizy, aktualne zasady oceny czasopism krajowych na części B są przeregulowane. Jednocześnie redakcje nie mają obowiązku przysyłania zeszytów czasopism z okresu, który podlega ocenie. Dlatego oprócz modyfikacji (uproszczenia) zasad powinien zostać wprowadzony obowiązek przysyłania kopii zeszytów (w formie papierowej lub elektronicznej). To oczywiście wymaga daleko idących modyfikacji samego sposobu przeprowadzania oceny. Uważamy jednak, że są to kroki niezbędne, aby środowisko naukowe mogło w pełni zaufać prowadzonej ocenie czasopism.

Podziękowania

Artykuł powstał w ramach prac grupy badawczej *Scientometrics. Polish Research Group* (scientometrics.amu.edu.pl) realizującej projekt pt. „Współczesna polska humanistyka wobec wyzwań naukometrii” finansowany ze środków Narodowego Programu Rozwoju Humanistyki, numer decyzji 0057/NPHR3/H11/82/2014. Dziękujemy Alicji Bodych i Piotrowi Brzezińskiemu z Index Copernicus International oraz Sebastianowi Fijałkowskiemu z Ośrodka Przetwarzania Informacji za dostarczenie danych źródłowych o spełnianiu przez czasopisma kryteriów ewaluacji. Za lekturę pierwszej wersji tekstu dziękujemy Małgorzacie Gamian-Wilk, Markowi Hołowieckiemu, Przemysławowi Korytkowskiemu oraz Tomaszowi Zarębskiemu.

Bibliografia

- Bemke-Świtlik M., Drabek A. (w druku). *A citation analysis of the Research Reports of the Central Mining Institute. Mining and Environment using Web of Science, Scopus, BazTech and Google Scholar: a case study*. Journal of Scientometric Research, 4(3).
- Bornmann L. (2013). *Evaluations by Peer Review in Science*. Springer Science Reviews, 1(1), 1-4. doi: 10.1007/s40362-012-0002-3
- Cohen J. (1992). *A Power Primer*. Psychological Bulletin 112(1), 155-159. doi: 10.1037/0033-2909.112.1.155
- Chadaj A., Turecka D. (2007). *Analiza cytowań artykułów kwartalnika AGH Geologia na podstawie bazy SCI Expanded*. Geologia, 33(4), 109-140.
- Elsaie M.L., Kammer J. (2009). *Impactitis: The impact factor myth syndrome*. Indian Journal of Dermatology, 54(1), 83-85. doi: 10.4103/0019-5154.48998
- Eyre-Walker A., Stoletzki N. (2013). *The Assessment of Science: The Relative Merits of Post-Publication Review, the Impact Factor, and the Number of Citations*. PLoS Biology, 11(10), e1001675-8. doi: 10.1371/journal.pbio.1001675
- Jablecka J. (1997). *Peer Review in Poland: Practical Solutions and Possible Improvements*. [W:]

- M.S. Frankel, J. Cave, *Evaluating Science and Scientists: An East-West Dialogue on Research Evaluation in Post-Communist Europe* (s. 96-111). Budapest: Central European University Press.
- Komunikat Ministra Nauki i Szkolnictwa Wyższego z dnia 14 września 2012 r. w sprawie kryteriów i trybu oceny czasopism naukowych.*
- Komunikat Ministra Nauki i Szkolnictwa Wyższego z dnia 29 maja 2013 r. w sprawie kryteriów i trybu oceny czasopism naukowych.*
- Komunikat Ministra Nauki i Szkolnictwa Wyższego z dnia 17 grudnia 2013 r. w sprawie wykazu czasopism naukowych wraz z liczbą punktów przyznawanych za publikacje w tych czasopismach.*
- Komunikat Ministra Nauki i Szkolnictwa Wyższego z dnia 31 grudnia 2014 r. w sprawie wykazu czasopism naukowych wraz z liczbą punktów przyznawanych za publikacje w tych czasopismach.*
- Komunikat Ministra Nauki i Szkolnictwa Wyższego z dnia 2 czerwca 2015 r. w sprawie kryteriów i trybu oceny czasopism naukowych.*
- Komunikat Ministra Nauki i Szkolnictwa Wyższego z dnia 18 grudnia 2015 r. w sprawie wykazu czasopism naukowych wraz z liczbą punktów przyznawanych za publikacje w tych czasopismach.*
- Kulczycki E. (2014). *Zasady oceny czasopism humanistycznych i ich rola w parametryzacji jednostek naukowych*. Nauka, 3, 117-140.
- Kulczycki E., Drabek A., Rozkosz E.A. (2015). *Publikacje a zgłoszenia ewaluacyjne, czyli zniekształcony obraz nauki w Polsce*. Nauka, 3, 35-58.
- Kulczycki E., Rozkosz E., Drabek A. (2015). *Publikacje polskich badaczy w czasopismach z list ERIH w kontekście ewaluacji jednostek naukowych*. Kultura i Edukacja, 107(1), 149-172. doi: 10.15804/kie.2015.01.08
- Obwieszczenie Ministra Nauki i Szkolnictwa Wyższego z dnia 25 marca 2015 r. o sprostowaniu błędów.*
- Ratajczak, M. (2015, maj). [Oficjalny list do Jerzego Duszyńskiego – prezesa Polskiej Akademii Nauk]. Kopia w posiadaniu Zespołu ds. oceny czasopism.
- Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 8 sierpnia 2011 r. w sprawie obszarów wiedzy, dziedzin nauki i sztuki oraz dyscyplin naukowych i artystycznych* (2011). Dz. U. 2011, nr 179, poz. 1065.
- Rozporządzenie Ministra Nauki i Szkolnictwa Wyższego z dnia 13 lipca 2012 r. w sprawie kryteriów i trybu przyznawania kategorii naukowej jednostkom naukowym* (2012). Dz. U. 2012, poz. 877.
- Sangster A. (2015). *You Cannot Judge a Book by Its Cover: The Problems with Journal Rankings*. Accounting Education, 24(3), 175-186. doi: 10.1080/09639284.2015.1055929
- Slyder J.B., Stein B.R., Sams B.S., Walker D.M., Beale B.J., Feldhaus J.J., Copenheaver C.A. (2011). *Citation pattern and lifespan: a comparison of discipline, institution, and individual*. Scientometrics, 89(3), 955-966. doi: 10.1007/s11192-011-0467-x
- Vanclay J.K. (2011). *An evaluation of the Australian Research Council's journal ranking*. Journal of Informetrics, 5(2), 265-274. doi: 10.1016/j.joi.2010.12.001
- Wheeler B. (2011). *Journal 'Ranking' Issues and the State of the Journal in the Humanities*. Journal of Scholarly Publishing, 42(3), 323-381. doi: 10.3138/jsp.42.3.323
- Wróblewski A.K. (2002). *Bibliometryczna trylogia*. Zagadnienia Naukoznawstwa, 38(1-2), 7-29.

Peer review as a third dimension of the evaluation of national scientific journals

The article discusses the peer-review evaluation of the national scientific journals in the Polish research evaluation system. Our study examines all journals ($N = 2212$) that are indexed on the 2015 Polish Journal Ranking (the so-called „ministerial list”). It is the first edition of the Polish Journal Ranking in which the peer-review evaluation has been implemented. We have analyzed three dimensions of the evaluation (the formal, the bibliometric, and the peer-review one) in each of the three groups of sciences in which the journals are classified. We have found out that the journal evaluation that is based on the peer-review has fundamentally changed the final assessment of the journals in the group of the humanities. The results of the peer-review evaluation in this group are not related to the results of the formal and bibliometric evaluation. In the group of natural sciences and in the group of social sciences, the results of the peer-review evaluation correlate with the results of the formal and bibliometric evaluation. Moreover, the journals citation correlate with the results of the peer-review evaluation in the three groups of sciences: the natural sciences, the social sciences, and the humanities. Our findings show that adding the third dimension to the evaluation of Polish scientific journals, i.e. peer-review, has resulted in significant transformations of the assessment results only for the humanities journals.

Key words: journal rankings, Ministry of Science and Higher Education in Poland, peer-review, science policy, scientometrics