

**Aleksandra Jasińska**

Instytut Badań Edukacyjnych

Zespół Badawczy EWD

## **Pomiar gotowości szkolnej uczniów za pomocą skali quasi-obszewacyjnej**

### **Wprowadzenie**

W badaniach edukacyjnych do pomiaru pewnych cech lub dyspozycji uczniów wykorzystuje się często skale quasi-obszewacyjne. Rozwiązanie to ma niewątpliwie swoje zalety. Opracowanie takiego narzędzia jest zasadniczo prostsze niż przygotowanie testu kompetencyjnego, a także samo wykorzystanie go w badaniu pociąga za sobą mniej problemów. Narzędzie takie pozwala bowiem w niedługim czasie oraz bez zapewniania specjalnych warunków zebrać informacje o wielu badanych osobach. Do najważniejszych jego wad natomiast możemy zaliczyć mniej bezpośredni sposób pomiaru danej cechy, który jest odzwierciedleniem opinii osób udzielających odpowiedzi na pytania.

Zespół EWD w pierwszym etapie prowadzonych przez siebie badań podłużnych w szkołach podstawowych<sup>1</sup> zdecydował się zastosować skalę quasi-obszewacyjną do pomiaru gotowości szkolnej badanych uczniów<sup>2</sup>. Planował on wykorzystać zebrane w ten sposób dane jako jedną z miar potencjału uczniów na progu szkoły podstawowej, która mogłaby posłużyć do modelowania wskaźnika EWD dla pierwszego poziomu edukacyjnego. Artykuł ten przedstawi wyniki i wnioski z analiz, które miały na celu odpowiedź na pytanie, czy wykorzystana w badaniu skala jest wartościowym narzędziem pomiaru potencjału uczniów „na wejściu” dla modelowania wskaźnika EWD.

### **Kwestionariusz opisu ucznia – charakterystyka narzędzia**

Na potrzeby wspomnianego badania R. Dolata opracował kwestionariusz zawierający 40 stwierdzeń dotyczących umiejętności ważnych z punktu widzenia gotowości szkolnej dzieci. Był on przeznaczony do wypełniania przez nauczyciela nauczania zintegrowanego prowadzącego badaną klasę, którego zadaniem było określenie na czteropunktowej skali szacunkowej (od *zdecydowanie tak* do *zdecydowanie nie*), na ile dobrze każde ze stwierdzeń opisuje ucznia w pierwszych tygodniach nauki szkolnej. Dopuszczano także odpowiedź *trudno powiedzieć*, która ostatecznie została potraktowana jako brak danych. Pozostałym odpowiedziom przypisano wartości od 4 (dla *zdecydowanie tak*) do 1 (dla *zdecydowanie nie*).

<sup>1</sup> Dokładny opis badań czytelnik znajdzie w: Z. Lisiecka „Badania panelowe w szkołach podstawowych i gimnazjach a rozwój wskaźnika EWD”. Badania te są realizowane w ramach projektu Centralnej Komisji Egzaminacyjnej *Badania dotyczące rozwoju metodologii szacowania wskaźników edukacyjnej wartości dodanej (EWD)*, współfinansowanego przez Unię Europejską w ramach Europejskiego Funduszu Społecznego – Program Operacyjny *Kapitał Ludzki*, Priorytet III *Wysoka jakość systemu oświaty*, Działanie 3.2 *Rozwój systemu egzaminów zewnętrznych*.

<sup>2</sup> Opiswane tu narzędzie jest dostępne na stronie [www.ewd.edu.pl](http://www.ewd.edu.pl) w zakładce *Badania podłużne*.

## Analiza właściwości psychometrycznych narzędzia

Dzięki zastosowaniu *Kwestionariusza opisu ucznia* każde badane dziecko można opisać za pomocą 40 zmiennych (wskaźników). Aby móc jednak efektywnie wykorzystywać w analizach te informacje, warto uprościć dane, wyznaczając odpowiednią liczbę czynników (wymiarów, zmiennych ukrytych), które dobrze reprezentowałyby wszystkie ważne aspekty zróżnicowania badanych uczniów. Zazwyczaj, stosując odpowiednie metody statystyczne, można uprościć wielozmiennowy zbiór danych do kilku czynników, dzięki czemu jedną osobę można opisać ze względu na jej wyniki uzyskane w kilku wymiarach, a nie za pomocą kilkudziesięciu zmiennych. Ułatwia to wykorzystywanie danych w dalszych analizach, interpretację wyników, a także podwyższa rzetelność pomiaru. Przeprowadzone analizy miały na celu właśnie taką redukcję danych i ocenę, czy wyznaczone czynniki dobrze opisują zróżnicowanie badanych jednostek, jak również analizę trafności zastosowanego narzędzia wykorzystującą dostępne zmienne zebrane niezależnie od odpowiedzi wychowawcy klasy.

Informacje o badanych dzieciach, uzyskane za pomocą szacunków nauczyciela, mają specyficzny charakter. Należy na nie patrzeć jako na opinie wychowawców klas na temat poziomu umiejętności ich uczniów (dane są zbierane na poziomie nauczyciela, natomiast jednostką analizy jest uczeń). Analizując właściwości psychometryczne zastosowanego narzędzia, należy wziąć więc pod uwagę m.in. to, że nauczyciele, wypełniając kwestionariusze swoich uczniów, mogli oceniać ich relatywnie (np. jedni bardziej surowo niż inni). Można jednak założyć, że jeśli dany nauczyciel dokonywał ocen w swoisty sposób, to do każdego ucznia przykładał tę samą miarę.

Wariancja wyników między klasami może być więc spowodowana różnicami w sposobie oceniania przez nauczycieli. Oczywiście za część wariancji międzyoddziałowej odpowiada także rzeczywiste zróżnicowanie umiejętności uczniów z różnych klas. Ponieważ jednak nie można jednoznacznie określić, jakie jest źródło wariancji międzyoddziałowej, na potrzeby testowania modelu zdecydowano się ją wytracić. Dokonano tego w następujący sposób: wartości każdej z 40 zmiennych (pozycji kwestionariusza) zestandaryzowano w obrębie każdej klasy oraz liniowo przekształcono, tak aby przedstawić je na skali o średniej 100 i odchyleniu standardowym 15, zgodnie z poniższym wzorem.

$$z = \frac{x - \mu}{\sigma} \cdot 15 + 100$$

gdzie:

$x$  – wartość zmiennej przed standaryzacją

$\mu$  – średnia zmiennej w danej klasie

$\sigma$  – odchylenie standardowe zmiennej w danej klasie

Zdarzało się, że nauczyciele, wypełniając kwestionariusze o swoich uczniach, wszystkich oceniali tak samo ze względu na pewne stwierdzenia, co powodowało, że odchylenie standardowe dla takiego pytania było równe 0. W takiej sytuacji danej zmiennej przypisano wartość średnią, czyli 100.

**Struktura i rzetelność.** Tak przygotowane dane poddano analizie statystycznej<sup>3</sup>. Przeprowadzono eksploracyjną analizę czynnikową metodą największej wiarygodności z rotacją ukośnokątną (*Oblimin* z normalizacją Kaisera), która dopuszcza pewien stopień skorelowania pomiędzy czynnikami. Rotację tę zastosowano ze względu na to, że niepoprawne wydaje się przyjęcie założenia o nieskorelowaniu ze sobą wymiarów opisujących dojrzałość szkolną ucznia. Gdyby się jednak okazało, że czynniki nie są ze sobą skorelowane, ukośna metoda rotacji wygenerowałaby rozwiązanie świadczące o braku wspomnianego związku. Z analizy wyłączono wszystkie obserwacje z brakami danych. Liczbę czynników wyznaczono na podstawie kryterium Kaisera-Guttmana (kryterium wartości własnej większej niż jeden). Posługując się macierzą ładunków czynnikowych oraz miarami dopasowania modelu do danych (RMSEA i NNFI<sup>4</sup>), usunięto wskaźniki nieistotne, czyli te, które w żadnym wymiarze nie miały ładunku czynnikowego o wartości bezwzględnej większej niż 0,3, wskaźniki niejednoznaczne, czyli te, które miały istotne (najczęściej jednak niewysokie) ładunki na kilku czynnikach oraz wskaźniki trudne w interpretacji w kontekście danego czynnika, lub te, które miały względnie niewysokie ładunki czynnikowe, jeśli poprawiało to dopasowanie modelu do danych. W ten sposób wyeliminowano 12 pytań. Wskaźnik KMO dla pozostałych 28 stwierdzeń wyniósł 0,959, co świadczy o bardzo dużej sile powiązań pomiędzy zmiennymi.

Analiza czynnikowa, w której uwzględniono 28 pytań, wykazała, że na podstawie wzorów odpowiedzi nauczycieli na pytania kwestionariusza można wyodrębnić cztery spójne wewnętrznie wymiary opisujące ważne aspekty gotowości szkolnej dziecka. Czynniki te wyjaśniają 64,9% wariancji całkowitej. Uzyskany model dobrze odzwierciedla macierz kowariancji ocen, za czym przemawiają wartości miar dopasowania: RMSEA = 0,068<sup>5</sup>, NNFI = 0,923<sup>6</sup>.

Uzyskane wyniki świadczą o tym, że na gotowość szkolną dziecka (w zakresie mierzonym omawianym tu kwestionariuszem) składają się następujące wymiary: umiejętności językowe, umiejętności matematyczne, dojrzałość społeczna oraz przystosowanie i zaangażowanie szkolne. Każdy z tych wymiarów jest wskaźnikowany przez kilka pytań. O tym, w jak dużym stopniu pytanie jest nasycone danym czynnikiem, informują ładunki czynnikowe.

Pierwszy wymiar – umiejętności językowe – skupił 9 pytań, które opisują poziom gotowości dziecka do nauki czytania i pisanania. Pozycje te zostały przytoczone w tabeli 1. Skala ma wysoką rzetelność ( $\alpha$ -Cronbacha = 0,936).

<sup>3</sup> Wszystkie opisane w tym artykule analizy wykonano przy uwzględnieniu wag związanych ze sposobem losowania próby do badań.

<sup>4</sup> Root Mean Square Error of Approximation (RMSEA), Tucker Lewis Index or Non-normed Fit Index (NNFI).

<sup>5</sup> Uznaje się, że dla modelu dobrze dopasowanego wartość RMSEA wynosi 0,05 lub mniej, natomiast wartość powyżej 0,1 świadczy o słabym dopasowaniu.

<sup>6</sup> Wartość wskaźnika NNFI równą 0,9 uznaje się za wystarczającą, by mówić o dobrze dopasowanym modelu.

**Tabela 1. Pytania reprezentujące umiejętności językowe**

Treść pozycji skali W pierwszych tygodniach nauki szkolnej dziecko...	Ładunki czynnikowe
potrafiło przeczytać całe zdania	0,937
umiało przeczytać proste wyrazy	0,890
znało wszystkie litery	0,875
podejmowało udane próby przeczytania nawet dłuższych tekstów	0,818
sprawnie pisało litery, potrafiło napisać prosty wyraz	0,630
umiało napisać całe zdanie, a nawet kilkuzdaniowy tekst	0,576
znało nazwy dni tygodnia i miesięcy	0,523
rozumiało wysłuchany tekst, potrafiło o nim opowiedzieć	0,481
rozumiało pojęcia związane z przestrzenią: nad, pod, za, na górze, na dole, z lewej, z prawej	0,476

**Tabela 2. Pytania reprezentujące umiejętności matematyczne**

Treść pozycji skali W pierwszych tygodniach nauki szkolnej dziecko...	Ładunki czynnikowe
dodawało i odejmowało w zakresie 100	0,940
wykonywało działania dodawania i odejmowania na liczbach przekraczających 100	0,832
potrafiło przeliczać do 100	0,553

Umiejętności matematyczne reprezentują trzy pytania przedstawione w powyższej tabeli. Tak mała liczba pozycji wynika z jednej strony z faktu, że *Kwestionariusz opisu ucznia* zawierał mniej pytań dotyczących kompetencji matematycznych niż np. językowych, z drugiej natomiast strony dwa pytania, które – zgodnie z przewidywaniem autora – mogłyby wiązać się z tym wymiarem, okazały się niewystarczająco z nim skorelowane. Stwierdzenie *znało się na zegarku* zostało wyeliminowane z analizy ze względu na to, że miało bardzo słaby, choć istotny ładunek czynnikowy na dwóch wymiarach (umiejętnościach językowych i matematycznych). Stwierdzenie *rozumiało pojęcia związane z przestrzenią: nad, pod, za, na górze, na dole, z lewej, z prawej* okazało się natomiast istotnie nasycone czynnikiem umiejętności językowych, a nie zauważono jego związku z wymiarem umiejętności matematycznych. Skala ta, z uwagi na niewielką liczbę pytań wchodzących w jej skład, ma niższą rzetelność niż pozostałe ( $\alpha$ -Cronbacha = 0,846). Aby podnieść jej rzetelność, należałoby w przyszłości dodać więcej pozycji wskaźnikujących omawiany wymiar.

**Tabela 3. Pytania reprezentujące dojrzałość społeczną**

Treść pozycji skali W pierwszych tygodniach nauki szkolnej dziecko...	Ładunki czynnikowe
było śmiałe i odważne w kontaktach z dorosłymi, nie bało się obcych osób	0,932
otwarcie i swobodnie rozmawiało z nauczycielem	0,899
było śmiałe, pewne swojej wartości	0,785
chętnie przewodziło, organizowało zabawy, wciągało inne dzieci do wspólnych działań	0,769
chętnie recytowało wiersze i śpiewało piosenki, bez lęku występowało na forum klasy	0,733
było towarzyskie, chętnie bawiło się z innymi dziećmi	0,678

W tabeli 3. przedstawiono pytania reprezentujące dojrzałość społeczną. Czynniki ten dotyczy zachowań i umiejętności dziecka w zakresie komunikowania się i współdziałania z rówieśnikami oraz osobami dorosłymi. Opisuje cechy ucznia, który jest śmiały, otwarty i łatwo nawiązuje kontakt z innymi. Skala ta charakteryzuje się wysoką rzetelnością ( $\alpha$ -Cronbacha = 0,925).

Czwarty wymiar – przystosowanie i zaangażowanie szkolne – skupił stwierdzenia opisujące zachowania osoby, która potrafiła podporządkować się zasadom, normom postępowania i wartościom obowiązującym w szkole. Opisuje on postawę zaangażowania w szkolne działania, pozytywnego do nich nastawienia oraz gotowości do wytrwałego przewycięzania trudności pojawiających się na drodze do wykonania zadań stawianych przez nauczyciela. Także ta skala ma wysoką rzetelność ( $\alpha$ -Cronbacha = 0,930).

**Tabela 4. Pytania reprezentujące przystosowanie i zaangażowanie szkolne**

Treść pozycji skali W pierwszych tygodniach nauki szkolnej dziecko...	Ładunki czynnikowe
bez ociągania wykonywało polecenia nauczyciela	0,786
gdy musiało czekać na swoją kolej, cierpliwie czekało i nie okazywało złości	0,746
potrafiło pracować w skupieniu nawet, gdy coś mu przeszkadzało	0,725
z dużym zaangażowaniem wykonywało zadania zlecone przez nauczyciela	0,717
nawet niezbyt ciekawe zadania wykonywało do końca	0,704
zgodnie bawiło się z dziećmi, nie wdawało się w bójki, nie przeszkadzało innym w pracy	0,701
gdy napotykało trudności, cierpliwie starało się je przewyciężyć	0,691
szybko i chętnie przystępowało do działań proponowanych przez nauczyciela	0,688
śluchało w skupieniu nawet dłuższej opowieści	0,624
było życzliwe dla rówieśników, chętnie pomagało innym, pożyczało swoje rzeczy	0,542

Wyodrębnione czynniki są ze sobą silnie skorelowane (tabela 5.).

Tabela 5. Macierz korelacji czynników

czynnik	um. język.	um. matem.	dojrz. społ.
um. matem.	0,592		
dojrz. społ.	0,454	0,323	
przyst. i zaang. szk.	0,515	0,300	0,261

Najsilniejszy związek zaobserwowano między umiejętnościami językowymi i matematycznymi, a także między umiejętnościami językowymi oraz przystosowaniem i zaangażowaniem szkolnym. Najślabszą, lecz nadal istotną, zależność można zauważyć między dojrzałością społeczną a przystosowaniem i zaangażowaniem szkolnym. Ponieważ czynniki te opisują różne wymiary gotowości szkolnej, ich skorelowanie nie powinno dziwić. Za część tego związku może również odpowiadać fakt, że oceny ucznia we wszystkich wymiarach dokonywał ten sam nauczyciel. Ocenienie danego dziecka pozytywnie ze względu na pewne cechy mogło spowodować większą skłonność do przypisania mu pozytywnych ocen także w innych pytaniach (tzw. *efekt halo*).

**Trafność.** Analiza trafności omawianej skali polegała na zbadaniu siły zależności ocen czynnikowych – czyli wyników uczniów dla każdego wymiaru – z takimi zmiennymi, jak: wiek uczniów (wykluczono dzieci, które rozpoczęły naukę wcześniej lub później niż w roku, w którym kończą 7 lat), płeć, wynik testu Ravena, wykształcenie rodziców (mierzone liczbą lat nauki<sup>7</sup>), wskaźnik ISEI<sup>8</sup> rodziców oraz liczba lat spędzonych przez dziecko w przedszkolu. Wykazała ona istnienie umiarkowanie silnych pozytywnych zależności między wynikami na czterech wymiarach wyznaczonych dla danych standaryzowanych wewnątrz klas a zmiennymi opisującymi uczniów lub ich rodziców (tabela 6.).

Tabela 6. Współczynniki korelacji\* liniowej Pearsona – model dla danych standaryzowanych

	Oceny czynnikowe			
	um. język.	um. matem.	dojrz. społ.	p. i z. szk.
Wynik testu Ravena	0,365	0,299	0,210	0,239
L. lat w przedszkolu	0,144	0,100	0,184	0,020
Wiek	0,125	0,078	0,090	0,111
Płeć**	0,130 (K)	0,030 (M)	0,129 (K)	0,298 (K)
Wykształcenie matki	0,257	0,212	0,188	0,152
Wykształcenie ojca	0,202	0,173	0,133	0,112
ISEI matki	0,218	0,165	0,122	0,115
ISEI ojca	0,136	0,110	0,073	0,063

\* W tabeli podano korelacje istotne na poziomie istotności statystycznej  $p < 0,001$

\*\* Dla zmiennej *płeć* wyznaczono korelację punktowo-dwuseryjną. W nawiasie podano płeć, na korzyść której przemawia zależność.

<sup>7</sup> Poszczególnym kategoriom wykształcenia przypisano standardową liczbę lat nauki, podczas których można było uzyskać dany poziom wykształcenia. Np. wykształcenie podstawowe = 8 lat, wykształcenie średnie (technikum) = 12 lat.

<sup>8</sup> International Socio-Economic Index of Occupational Status.



Najsilniejszy związek zaobserwowano między poziomem inteligencji ucznia, mierzonym testem Ravena, a jego umiejętnościami językowymi i matematycznymi. Zgodnie z intuicją, pozaintelektualne wymiary prezentowanej skali, czyli dojrzałość społeczna oraz przystosowanie i zaangażowanie szkolne, okazały się pozostawać w słabszym związku z inteligencją. Liczba lat spędzonych przez dziecko w przedszkolu jest słabo, choć pozytywnie związana z umiejętnościami językowymi, matematycznymi i dojrzałością społeczną. Nie zaobserwowano natomiast zależności tej zmiennej z przystosowaniem i zaangażowaniem szkolnym dziecka. Wyniki uczniów na czterech wymiarach skali okazały się być nieznacznie, lecz pozytywnie skorelowane z ich wiekiem wyrażonym w dniach. Związek z płcią także pozostał w zgodzie z intuicją. Dziewczynki uzyskiwały średnio wyższe oceny z umiejętności językowych i dojrzałości społecznej. Znacząco przewyższyły także chłopców w zakresie przystosowania i zaangażowania szkolnego. Na korzyść chłopców przemawia związek z umiejętnościami matematycznymi, jednak korelacja w tym przypadku jest bliska zeru.

Biorąc pod uwagę zmienne związane ze statusem rodziny, możemy mówić o umiarkowanie silnym pozytywnym związku wykształcenia rodziców z intelektualnymi wymiarami gotowości szkolnej oraz słabszej, choć także pozytywnej zależności z dojrzałością społeczną dziecka oraz jego przystosowaniem i zaangażowaniem szkolnym. Podobnie można opisać zależność wyników na poszczególnych wymiarach ze wskaźnikami statusu społeczno-ekonomicznego ISEI rodziców, choć związki te są nieco słabsze. Widoczny jest także silniejszy związek wymiarów gotowości szkolnej dziecka z wykształceniem i statusem matki niż ojca.

Podsumowując, analiza zależności między ocenami czynnikowymi i zmiennymi mierzonymi niezależnie od odpowiedzi nauczyciela, które powinny być związane z umiejętnościami dzieci, wykazała istnienie pozytywnych korelacji. Oznacza to, że nauczyciele w obrębie swoich klas potrafili w taki sposób różnicować oceny różnych aspektów gotowości szkolnej swoich uczniów, by pozostały one w związku z rzeczywistym potencjałem dzieci. Wyniki te przemawiają za trafnością analizowanej skali jako narzędzia do różnicowania umiejętności uczniów wewnątrz klas.

### **Kwestionariusz opisu ucznia jako wartościowa miara poziomu umiejętności na progu szkoły podstawowej**

Jeżeli przedstawiane tu narzędzie ma służyć do pomiaru umiejętności uczniów różnych nauczycieli, wyniki uzyskane przy jego pomocy powinny dobrze odzwierciedlać nie tylko względny dystans między uczniami w danej klasie, lecz także rzeczywiste różnice w poziomie umiejętności uczniów z różnych klas. Wykorzystując wyniki otrzymane na podstawie analizy czynnikowej przeprowadzonej na danych standaryzowanych wewnątrz klas, tracimy informację o zróżnicowaniu uczniów między oddziałami. Korzystając z wyników surowych, dobrze opisywalibyśmy zróżnicowanie uczniów niezależnie od podziału na klasy, ale tylko pod warunkiem, że wszyscy nauczyciele, oceniając ucznia, w taki sam sposób używają omawianych skal szacunkowych i kotwiczą je w tym samym punkcie odniesienia. Warto więc postawić następujące pytanie: Czy narzędzie to można uznać za obiektywną miarę poziomu umiejętności uczniów? Lub inaczej pytając, czy skala ta ma charakter tylko lokalny (wewnątrz klas), czy także uniwersalny?

Aby na nie odpowiedzieć, wykonano analizę czynnikową dla tych samych 28 pytań, stosując jednakową metodę i rotację na danych surowych (niestandardyzowanych) oraz danych zagregowanych do poziomu klasy. Dane zagregowane uzyskano, licząc wartości średnie danych zmiennych dla każdej klasy. Jednostką analizy w tym przypadku była więc klasa. Analiza danych zagregowanych pozwoliła przyrzeć się zależnościom na poziomie międzygrupowym bez udziału zmienności na poziomie indywidualnym. Przyjęto następujące założenie: skoro na poziomie indywidualnym zostały zaobserwowane zależności między ocenami czynnikowymi i innymi zmiennymi opisującymi ucznia (np. inteligencją), to jeśli wszyscy nauczyciele potrafili zakotwiczyć skalę w uniwersalnym punkcie odniesienia, zaobserwuje się podobne zależności także na poziomie grupowym. Dodatkowy warunek mówi o tym, że niezbędne jest zróżnicowanie klas ze względu na interesujące nas zmienne. Tak więc, ukazując to na przykładzie, jeśli skala ma charakter uniwersalny, powinno być tak, że klasy, które uzyskały średnio wyższy wynik w teście Ravena, powinny uzyskać średnio wyższe oceny na tych wymiarach, które są związane z aspektem poznawczym (umiejętności językowe i matematyczne), skoro taką zależność zaobserwowano na poziomie indywidualnym.

Wyniki analiz czynnikowych przeprowadzonych na danych surowych i zagregowanych dały zadowalające rezultaty. W obu sytuacjach analiza wykazała istnienie czterech – takich samych jak w przypadku analizy przeprowadzonej na danych standaryzowanych wewnątrz klasy – czynników, które wyjaśniały odpowiednio 67,43% (dla danych surowych) i 70,13% (dla danych zagregowanych) wariacji całkowitej. Poszczególne wskaźniki nieznacznie różniły się wielkością ładunków czynnikowych. Cztery pytania dla danych zagregowanych oraz jedno dla danych surowych miały istotne, choć niewielkie ładunki czynnikowe także na innym czynniku niż w przypadku analizy na danych standaryzowanych. Miary RMSEA i NNFI wskazały na słabsze dopasowanie tych modeli do danych. Wyniosły one odpowiednio dla danych surowych: 0,085 i 0,892, a dla danych zagregowanych: 0,120 i 0,819. Na ich podstawie można wysnuć wniosek, że model wyznaczony dla danych zagregowanych nie opisuje dobrze zróżnicowania klas.

Analiza wariacji wykazała natomiast, że klasy różnią się między sobą pod względem wyników na czterech wymiarach gotowości szkolnej<sup>9</sup>. Przynależność uczniów do danego oddziału wyjaśnia od 20% do 29% zróżnicowania danych wyników w populacji<sup>10</sup>. Dla porównania, w przypadku inteligencji przynależność do klasy wyjaśnia 27% zróżnicowania, a w przypadku wykształcenia matki 28%. Pozostaje zapytać, czy różnice w średnich wynikach klas dla czterech wymiarów omawianej skali są pochodną rzeczywistych różnic w umiejętnościach uczniów, czy wynikają z faktu, że nauczyciele, oceniając uczniów, mieli różny punkt odniesienia. Odpowiedź na to pytanie przybliży analiza korelacji ocen czynnikowych dla danych zagregowanych ze zmiennymi opisującymi potencjał danej klasy.

<sup>9</sup> Analiza uwzględniająca oceny czynnikowe dla danych surowych.

<sup>10</sup> Dla umiejętności matematycznych jest to 50%, jednak z powodu małej liczby pytań reprezentujących ten czynnik nie powinno się przywiązywać zbyt dużej wagi do tej różnicy.



Zgodnie z przyjętym wcześniej założeniem, skoro zaobserwowano opisane zależności wewnątrz klas, to jeśli omawiana skala jest dobrym narzędziem do różnicowania uczniów w populacji, powinny być one także zauważalne między klasami. Dane przedstawione w tabeli 7. pokazują, że tak jednak nie jest. Prawa część tabeli zawiera współczynniki korelacji liniowej Pearsona między ocenami czynnikowymi wyznaczonymi na podstawie analizy czynnikowej na danych zagregowanych do poziomu klasy a zmiennymi, takimi jak: średni wynik testu Ravena w klasie, średnia liczba lat nauki w przedszkolu, średnia liczba lat nauki matki i ojca dzieci z jednej klasy oraz średnia statusu społeczno-ekonomicznego matki i ojca.

Okazuje się, że związki zaobserwowane między ocenami czynnikowymi wyznaczonymi na podstawie analizy czynnikowej na poziomie indywidualnym i inteligencją nie utrzymują się na poziomie grupowym. Czyli mimo iż wewnątrz klas daje się zauważyć zależność mówiącą o tym, że im uczeń ma wyższy potencjał intelektualny, tym lepiej został oceniony przez nauczyciela, nie dostrzega się zależności świadczącej o tym, że klasy o średnio wyższym potencjale intelektualnym otrzymały średnio wyższe oceny. Zasadniczo zanika także zależność ze średnią liczbą lat spędzonych w przedszkolu. Natomiast związek z wykształceniem rodziców i statusem społeczno-ekonomicznym, poza wymiarem dojrzałości społecznej, utrzymuje się.

**Tabela 7. Porównanie współczynników korelacji liniowej Pearsona – modele dla danych standaryzowanych i zagregowanych**

	Oceny czynnikowe							
	Dane standaryzowane (poziom indywidualny)				Dane zagregowane (poziom grupowy)			
	um. język.	um. matem.	dojrz. społ.	p. i z. szk.	um. język.	um. matem.	dojrz. społ.	p. i z. szk.
Wynik testu Ravena	0,365	0,299	0,210	0,239	ns****	ns	-0,13 ***	ns
L. lat w przedszkolu	0,144	0,100	0,184	0,020	ns	0,151***	ns	ns
Wykształcenie matki**	0,257	0,212	0,188	0,152	0,177	0,174	ns	0,161
Wykształcenie ojca**	0,202	0,173	0,133	0,112	0,216	0,193	ns	0,147 ***
ISEI matki	0,257	0,212	0,188	0,152	0,140 ***	0,161	ns	0,126 ***
ISEI ojca	0,202	0,173	0,133	0,112	0,145 ***	0,170	ns	ns

\* Umieszczone w tabeli korelacje są istotne na poziomie istotności statystycznej  $p < 0,001$ , chyba że zaznaczono inaczej. \*\* Mierzone liczbą lat nauki. \*\*\* Korelacja istotna na poziomie  $p < 0,01$ .

\*\*\*\* Związek nieistotny statystycznie.

Uzyskane wyniki przemawiają więc za tym, że wątpliwe jest traktowanie wyników uczniów na opisywanej skali jako obiektywnej miary poziomu ich umiejętności. Być może nauczycielskie oceny są zrelatywizowane do ich uprzednich

doświadczeń. Wynikiem tego może być np. sytuacja, w której nauczyciele pracujący zazwyczaj z uczniami o wyższym poziomie intelektualnym zaniżają oceny swoich uczniów, bowiem porównują ich z im znanym przeciętnym poziomem umiejętności. Ci sami uczniowie mogliby uzyskać wyższe wyniki, gdyby ocenił ich nauczyciel pracujący na co dzień ze słabszymi uczniami.

W związku z tym, że status rodziny pochodzenia ucznia i związane z nim cechy dziecka są łatwiej dostrzegalne niż poziom ogólnych zdolności poznawczych, nauczyciele, mając świadomość tego, że poziom gotowości szkolnej jest związany z cechami środowiskowymi, mogą trafniej dostosowywać swoje oceny do wyznaczników statusu rodziny dziecka. Nauczyciel pracujący z dziećmi z niekorzystnych środowisk ma przypuszczalnie większą skłonność do przypisywania niższych ocen swoim uczniom, gdyż dostrzega to, że jego klasa – na tle innych w kraju – jest słabsza. Z tego też względu korelacje z wykształceniem i statusem rodziców mogą się utrzymywać na poziomie zagregowanym.

Odpowiadając na podstawowe pytanie postawione w tym artykule, warto jeszcze przyrzeć się temu, czy wraz ze wzrostem zróżnicowania uczniów w klasie pod względem intelektualnym rośnie zróżnicowanie nauczycielskich ocen (taka zależność przemawiałaby na korzyść skali). W tym celu porównano wskaźnik zróżnicowania uczniów w klasie pod względem poziomu ogólnych zdolności poznawczych (odchylenie standardowe wyniku testu Ravena w klasie) ze wskaźnikiem zróżnicowania ocen uczniów ze względu na ich umiejętności językowe i matematyczne (odchylenie standardowe ocen czynnikowych umiejętności językowych i matematycznych w klasie). Z analizy tej wynika, że nie obserwuje się istotnego statystycznie związku między wspomnianymi wskaźnikami dla umiejętności językowych, a dla umiejętności matematycznych związek ten jest słaby ( $r = 0,150$ ;  $p < 0,01$ ). Tak więc poziom zróżnicowania klasy pod względem potencjału intelektualnego nie przełożył się na poziom zróżnicowania ocen nauczycielskich.

## Podsumowanie

Artykuł podjął problematykę oceny wykorzystanej w badaniu podłużnym prowadzonym przez zespół EWD skali quasi-observacyjnej jako narzędzia pomiaru poziomu umiejętności uczniów na progu szkoły podstawowej. Przeprowadzone analizy statystyczne pozwoliły na wyodrębnienie czterech spójnych wewnętrznie wymiarów gotowości szkolnej, które mogą być mierzone za pomocą prezentowanego narzędzia. Przedstawiane wyniki świadczą o tym, że skala może być wykorzystywana do różnicowania uczniów wewnątrz klas. Świadczy o tym spójny i dobrze dopasowany model czynnikowy, a także pozytywne i istotne statystycznie związki wymiarów gotowości szkolnej ze zmiennymi opisującymi potencjał uczniów. Prezentowanej skali nie można jednak uznać za dobre narzędzie do różnicowania uczniów niezależnie od podziału na klasy. Wyniki przedstawionych analiz wskazują na to, że nauczyciele oceniają swoich uczniów relatywnie i nie potrafią zakotwiczyć skali w uniwersalnym dla wszystkich punkcie odniesienia. Dlatego też ten sposób pomiaru nie prowadzi do uzyskania wartościowych danych z punktu widzenia modelowania wskaźników EWD dla szkół podstawowych.