

**Radosław Kaczan**

**dr Piotr Rycielski**

Instytut Badań Edukacyjnych

Pracownia Szkolnych Uwarunkowań Efektywności Kształcenia

Zespół Wczesnej Edukacji

## **Diagnoza umiejętności dzieci 5-, 6- i 7-letnich za pomocą Testu Umiejętności na Starcie Szkolnym TUnSS**

### **1. Wprowadzenie**

#### **1.1. Koncepcja narzędzia diagnostycznego dla dzieci na starcie szkolnym.**

Przeciągający się proces obniżania wieku szkolnego do wieku 6 lat sprawił, że jesteśmy świadkami wielu zmian w obszarze edukacji przedszkolnej i wczesnoszkolnej. Obecnie, ze względu na planowane zmiany i dotyczące ich protesty rodziców, skład grup ostatnich roczników wychowania przedszkolnego i pierwszych klas szkół podstawowych jest bardzo zróżnicowany pod względem wieku biologicznego uczęszczających tam dzieci. Dodatkowo skład wiekowy grup edukacji wczesnoszkolnej różni się w gminach ze względu na decyzje polityczne samorządów dotyczące np. zaprzestania finansowania oddziałów zerowych w przedszkolach (przypadek Warszawy). W sytuacji tak dużego zróżnicowania wiekowego dzieci w grupach szkolnych i przedszkolnych paląca wydaje się potrzeba rzetelnej i trafnej analizy zróżnicowania poziomu ich umiejętności (por. Karwowski, 2011). Do przeprowadzenia takiej analizy nie ma jednak standaryzowanych na dużych próbach narzędzi. Test Umiejętności na Starcie Szkolnym (TUnSS) powstał właśnie, by zaspokoić tę potrzebę. Dzięki temu narzędziu możliwe będzie oszacowanie poziomu zróżnicowania kompetencji między dziećmi na różnych etapach edukacji wczesnoszkolnej, a także - dzięki badaniom z udziałem TUnSS - oszacowanie wpływu różnego rodzaju ścieżek edukacyjnych na rozwój edukacyjny dzieci 5-, 6- i 7-letnich.

TUnSS został przygotowany jako test wykonania, mierzący umiejętności poprzez próbki zachowania dziecka. Została w nim wykorzystana metodologia testowania adaptatywnego, dzięki której możliwe jest bardzo precyzyjne określenie poziomu umiejętności dziecka (Weiss, 2011). Podczas badania kolejne rozwiązywane przez dziecko zadania dobierane są na podstawie wcześniej udzielonych odpowiedzi. Co za tym idzie algorytm testowanie adaptatywnego (CAT) dopasowuje trudność kolejnych zadań do poziomu umiejętności dziecka. Dzięki zastosowaniu CAT można zapewnić sytuację, w której badane dziecko nie będzie rozwiązywać zadań zbyt łatwych lub zbyt trudnych w stosunku do swojego poziomu umiejętności (Hurek, Szejnberg; 2010). Daje to szansę na redukcję negatywnych efektów związanych ze znużeniem (zadania zbyt proste) lub lękiem i utratą motywacji (zadania zbyt trudne). Ze względu na adaptatywny charakter testu konieczne było przygotowania i przeprowadzenie go w formie komputerowej. Jednocześnie, biorąc pod uwagę fakt, że część

badanych dzieci nie byłaby w stanie poradzić sobie z zadaniami, w których konieczne jest samodzielne czytanie instrukcji i obsługę klawiatury komputera i/lub myszy, zdecydowano się zastosować urządzenia z ekranami dotykowymi. Wszystkie polecenia konieczne dla zrozumienia stawianych przed dzieckiem zadań zostały nagrane przez profesjonalnego lektora.

Należy zwrócić również uwagę na to, że TUnSS nie jest testem gotowości szkolnej, ponieważ przedmiotem pomiaru są jedynie kompetencje poznawcze, nie zaś charakterystyki osobowościowo-motywacyjne czy emocjonalne. Koncentracja na kompetencjach poznawczych daje możliwość diagnozy umiejętności dziecka niezbędnych do nauki czytania i pisanie oraz matematyki. Dzięki temu, że przygotowany test uwzględnia zarówno te umiejętności, które nabywane są przez dziecko w trakcie rocznego pobytu w zerówce, jak i te, które będą rozwijane dopiero w klasie pierwszej szkoły podstawowej, możliwa jest diagnoza zróżnicowania dzieci rozpoczynających naukę oraz ocena przyrostu tych umiejętności po pierwszym roku nauki w szkole podstawowej. Przygotowany test może być wykorzystywany zarówno w celach naukowych, jak i podczas sporządzania indywidualnej diagnozy dziecka.

## 1.2. Autorzy prac w projekcie TUnSS

Realizacja Testu Umiejętności na Starcie Szkolnym trwała ponad rok. Prace nad TUnSS toczyły się w dwóch zespołach.

Pracownia Szkolnych Uwarunkowań Efektywności Kształcenia (SUEK, Instytut Badań Edukacyjnych) - odpowiada za koncepcję testu adaptatywnego do badań dzieci z zastosowaniem urządzeń z ekranem dotykowym wraz z badaniami standaryzacyjnymi: pilotażowym i normalizacyjnym. Instytut Badawczy Millward Brown SMG/KRC wraz ze współpracownikami - to wykonawcy projektu, którzy odpowiadają za realizację aplikacji testującej, algorytmu testowania, banku zadań, badań standaryzacyjnych i wytworzenie podręcznika do testu oraz wszelkich materiałów badawczych i pomocniczych.

## 1.3. Konstrukcja testu

Test ma złożoną konstrukcję. Składa się z aplikacji testującej - obsługującej algorytm testowania adaptatywnego (CAT) i banku wystandaryzowanych zadań. Test realizowany jest na tabletach z ekranami dotykowymi.

W badaniu TUnSS implementacja algorytmu CAT prowadzona jest równoległe dla zadań z trzech obszarów: matematyka, pisanie, czytanie. Każdy z tych obszarów ma własny bank zadań i punkty startowe ( $\theta$ -start) dla grup wiekowych i płci. Zadania prezentowane są naprzemiennie dla każdego z obszarów i wybierane na podstawie wartości ich funkcji informacyjnej. Jeżeli dla któregoś z obszarów badanie zostanie zakończone (przez osiągnięcie zakładanego poziomu błędu standardowego oszacowania poziomu danej umiejętności), zadania z tego obszaru nie są już prezentowane, ale nie wpływa to na realizację algorytmu CAT w pozostałych obszarach. Poziom umiejętności ucznia szacowany jest przy pomocy estymatora największej wiarygodności<sup>1</sup>.

<sup>1</sup> Szczegółowe informacje na temat konstrukcji algorytmu CAT oraz sposobu jego działania znajdują się w publikacji przygotowanej w ramach prezentowanego badania: *Test Umiejętności na Starcie Szkolnym. Podręcznik* (2012). Warszawa: Instytut Badań Edukacyjnych.

W badaniu pilotażowym bank składał się z 420 zadań, spośród których 146 mierzyło umiejętności w zakresie matematyki, 136 w zakresie pisania i umiejętności istotnych z punktu widzenia pisania i 138 umiejętności w zakresie czytania. Każde dziecko w badaniu pilotażowym rozwiązywało 36 zadań (po 12 z każdej ze skal, wymieszanych w teście), co zajmowało przeciętnie około 25 minut. Do badania normalizacyjnego na podstawie przeprowadzonych analiz zakwalifikowano 336 zadań, w tym 127 mierzących umiejętności matematyczne, 106 mierzących umiejętności niezbędne do nauki pisania oraz 103 związane z nauką czytania. W badaniu normalizacyjnym liczba zadań rozwiązywanych przez dziecko zależna była od uzyskanej dokładności pomiaru (wielkości błędu standardowego). Założono, że minimalnie będzie to 5, a maksymalnie 10 zadań dla każdej z trzech skal. Średni czas trwania badania wynosił na etapie normalizacyjnym około 22 minut.

## 2. Badania standaryzacyjne

### 2.1. Pilotaż

Badania standaryzacyjne TUnSS odbyły się w dwóch etapach. W pierwszym etapie pilotażu zadań wzięło udział 5239 dzieci w wieku od 5 do 7 lat. Badanie odbywało się w okresie pomiędzy lutym a majem 2012 roku. Dzieci rekrutowane były z oddziałów przedszkolnych, szkolnych oddziałów przedszkolnych (zerówek) i pierwszych klas szkół podstawowych. Etap pilotażowy miał na celu określenie parametrów zadań z zakresu matematyki, pisania i czytania. Zadania te miały później zasilić bank zadań testu. W badaniu pilotażowym zadania zostały podzielone na 73 zeszyty, a ich dystrybucja wśród uczniów zapewniała, że każde zadanie zostało rozwiązane przez co najmniej 400 dzieci. Na podstawie uzyskanych wyników można było określić miary dopasowania krzywej charakterystycznej zadania do modelu teoretycznego i moc dyskryminującą zadania.

Skalowanie zadań dychotomicznych odbywało się przy zastosowaniu dwuparametrycznego modelu logistycznego IRT (2PL) (Hambleton i Swaminathan, 1985). Dla każdego zadania szacowana była moc dyskryminacyjna (współczynnik a) oraz trudność (współczynnik b) zadań. W przypadku zadań ocenianych na skali 0-1-2 zastosowano generalized partial credit model (GPCM) (Penfield i Bergeron, 2005). Skalowanie zrealizowane zostało przy wykorzystaniu programu XCalibre w wersji 4.6.1 (Test Umiejętności na Starcie Szkolnym. Podręcznik, 2012).

**Tabela 1. Liczebności grup w badaniu pilotażowym ze względu na wiek dziecka i status edukacyjny**

Grupa:	7-latki	6-latki	5-latki	Suma:
przedszkole	0	281	2286	2567
zerówka w przedszkolu	2	636	384	1022
zerówka w szkole	33	520	270	823
klasa I szkoły podstawowej	666	160	1	827
Suma:	701	1597	2941	5239

## 2.2. Normalizacja

W badaniu normalizacyjnym wzięło udział 3769 dzieci w wieku od 5 do 7 lat. Dzieci rekrutowane były z oddziałów przedszkolnych, szkolnych oddziałów przedszkolnych (zerówek) i pierwszych klas szkół podstawowych. Badanie zostało przeprowadzone w okresie pomiędzy majem a czerwcem 2012 roku. W badaniu normalizacyjnym po raz pierwszy zastosowano pełną wersję TUnSS - z ostatecznym bankiem standaryzowanych zadań i aplikacją testującą wykorzystującą algorytm testowania adaptatywnego. Badania normalizacyjne TUnSS pozwoliły na sprawdzenie działania testu w wersji ostatecznej oraz ustalenie norm wykonaniowych dla dzieci w różnym wieku i o różnym statusie edukacyjnym.

**Tabela 2. Liczebności grup w badaniu normalizacyjnym ze względu na wiek dziecka i status edukacyjny**

Grupa:	7-latki	6-latki	5-latki	Suma:
przedszkole	3	110	1138	1251
zerówka w przedszkolu	0	742	224	966
zerówka w szkole	6	366	294	666
klasa I szkoły podstawowej	703	183	0	886
Suma:	712	1401	1656	<b>3769</b>

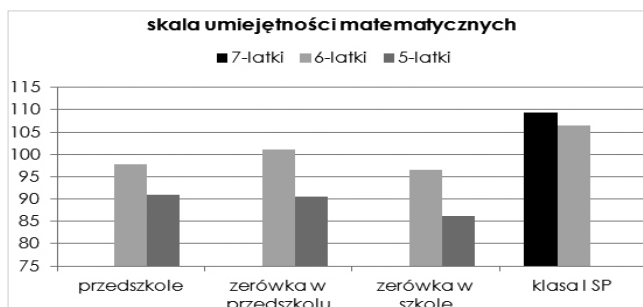
## 3. Wyniki TUnSS na podstawie badania normalizacyjnego

Poniżej przedstawiono wyniki badania normalizacyjnego. Dla większej klarowności wyników ze zbioru usunięto 9 zbadanych 7-latków, którzy z różnych przyczyn nie uczęszczali do klasy pierwszej szkoły podstawowej. Wszystkie wyniki TUnSS prezentowane są na trzech skalach: umiejętności matematycznych, pisania i czytania. Wyniki zostały wystandaryzowane - średnia dla wszystkich zbadanych dzieci wynosi 100, a odchylenie standardowe 15. Przy standaryzacji wyników zostały zastosowane wagi uwzględniające rozkład wieku, płci, województwo oraz typ placówki, do której uczęszczało dziecko.

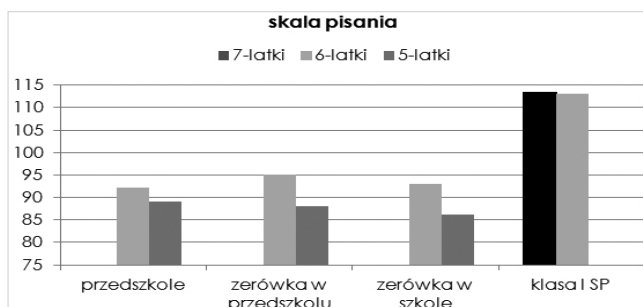
### 3.1. Zróznicowanie wyników ze względu na wiek i status edukacyjny

Pierwsze wykonane przez nas analizy dotyczyły zróznicowania wyników zbadanych dzieci ze względu na ich wiek oraz rodzaj placówki, do której uczęszczało dziecko. Rysunki 1., 2. i 3. przedstawiają różnice w średnich wynikach na skali umiejętności matematycznych, pisania i czytania wśród dzieci 5-, 6- i 7-letnich w zależności od ich statusu edukacyjnego. Najwyższe wyniki zawsze osiągają dzieci 7-letnie, co wydaje się być wynikiem oczekiwanym (por. Test Umiejętności na Starcie Szkolnym. Podręcznik, 2012). Są to uczniowie pierwszych klas szkół podstawowych, najstarsza grupa biorąca udział w badaniu. Badanie realizowane było na koniec roku szkolnego - test pokazał poziom umiejętności, które dzieci trenowały przez pierwszy rok nauki, a wcześniej także w zerówce. Interesujące są również wysokie wyniki 6-latków w szkołach podstawowych (por. Kopik, 2007). Gdyby decyzja o wcześniejszym pójściu do pierwszej klasy była dokonywana w sposób losowy, można by stwierdzić, że nauka szkolna znacznie lepiej rozwija dzieci niż przedszkole czy zerówka.

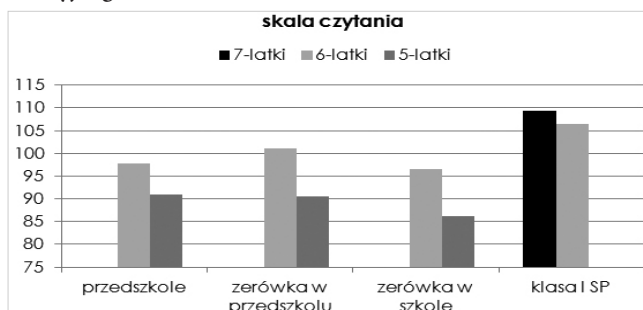
Tak jednak nie było, a decyzja o wcześniejszym posłaniu dzieci do szkoły podstawowej była często podejmowana na podstawie mniej lub bardziej całościowej diagnozy gotowości szkolnej dziecka. Można śmiało założyć, że większość sześciolatków w szkole podstawowej to dzieci, których rodzice uznali, że są one gotowe na rozpoczęcie nauki szkolnej. Jest to skrzywienie próby, przez które niemożliwe jest jednoznaczne określenie przyczyn sukcesu sześciolatków w szkole w porównaniu z innymi sześciolatkami w przedszkolu czy zerówce. Najniższe osiągnięcia obserwujemy u 5-latków, co również wydaje się być wynikiem oczekiwanym - jest to najmłodsza grupa dzieci.



Rysunek 1. Średnie wyniki na skali umiejętności matematycznych wśród dzieci 5-, 6- i 7-letnich w zależności od ich statusu edukacyjnego



Rysunek 2. Średnie wyniki na skali pisania wśród dzieci 5-, 6- i 7-letnich w zależności od ich statusu edukacyjnego

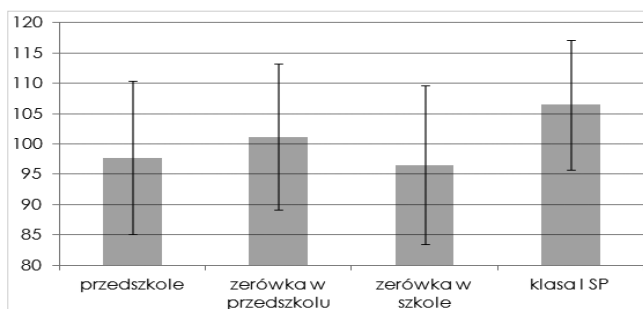


Rysunek 3. Średnie wyniki na skali czytania wśród dzieci 5-, 6- i 7-letnich w zależności od ich statusu edukacyjnego

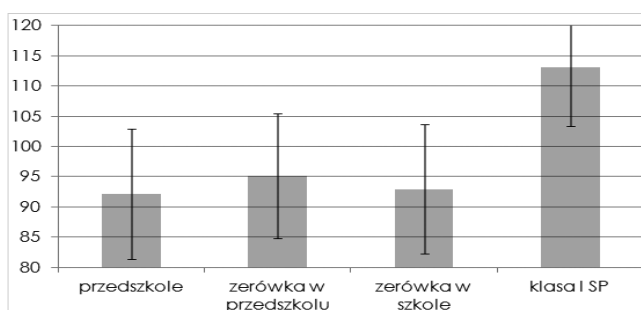
### 3.2. Sytuacja 6-latków

Ze względu na toczące się dyskusje na temat obniżenia wieku szkolnego oraz pytania pojawiające się w kontekście rozpoczęcia nauki w klasie pierwszej przez dzieci 6-letnie, przeanalizowaliśmy wyniki tej grupy bardziej szczegółowo. Poniżej przedstawiono na trzech wykresach (rysunki 4., 5. i 6.) wyniki 6-latków w różnych placówkach edukacyjnych. Wąsy na każdym z wykresów przedstawiają zakres jednego odchylenia standardowego powyżej i poniżej średniej. Wykresy te obrazują bardzo wysoki poziom zróżnicowania kompetencji małych dzieci właściwie we wszystkich placówkach edukacyjnych. Wartości plus/minus jednego odchylenia od średniej obejmują swoim zakresem 68% badanych dzieci - nie jest to zatem cała grupa. Nawet po usunięciu z wizualizacji 16% najmniej zdolnych i 16% najbardziej zdolnych dzieci widzimy, że umiejętności dzieci są niesłychanie zróżnicowane wewnątrz grup. Skutkuje to sytuacją dydaktyczną, w której część dzieci z przedszkola funkcjonuje całkiem nieźle na poziomie pierwszej klasy szkoły podstawowej - a część na poziomie młodszych grup przedszkolnych. Aktualna sytuacja edukacyjna w Polsce - ze względu na obniżenie wieku szkolnego - wydaje się być bardzo złożona i niezbędny jest jej monitoring.

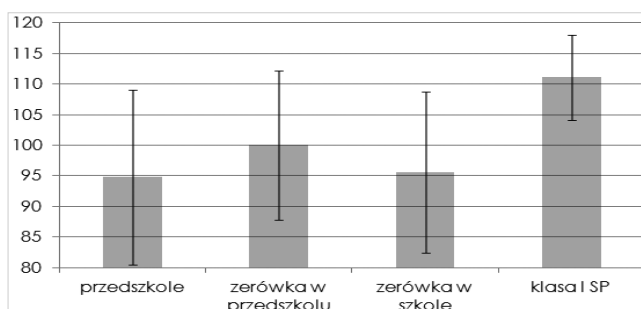
Gdy przyjrzymy się losom dzieci sześciolletnich o różnych statusach edukacyjnych, zauważymy duże różnice w ich kompetencjach w zależności od wybranej dla nich ścieżki edukacyjnej. Najlepiej rozwinięte są dzieci w pierwszej klasie szkoły podstawowej. Szczególnie silnie efekt ten widoczny jest w przypadku skali pisania, gdyż te umiejętności są istotnie częściej ćwiczone w pierwszych klasach niż w zerówkach. Warto także zauważyć wyraźne różnice między dziećmi sześciolletnimi uczęszczającymi do zerówki przyszkolnej a dziećmi sześciolletnimi uczęszczającymi do zerówek przedszkolnych. Dzieci w zerówkach przyszkolnych wypadły w testach nieco gorzej niż te, które pozostały w przedszkolach. Być może można to tłumaczyć korzystnym efektem pozostawienia dzieci w tym samym środowisku edukacyjnym. Być może do zerówek przyszkolnych trafiły dzieci, które zdiagnozowano jako jeszcze niegotowe do nauki szkolnej - stąd gorszy ich wynik.



Rysunek 4. Średnie wyniki na skali matematyki dla dzieci sześciolletnich w zależności od ich statusu edukacyjnego. Wąsy reprezentują jedno odchylenie standardowe powyżej i poniżej średniej.



Rysunek 5. Średnie wyniki na skali pisania dla dzieci sześciolletnich w zależności od ich statusu edukacyjnego. Wąsy reprezentują jedno odchylenie standardowe powyżej i poniżej średniej.

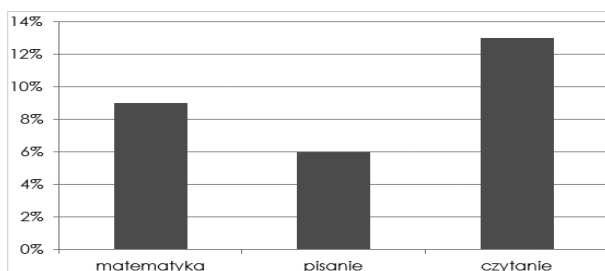


Rysunek 6. Średnie wyniki na skali czytania dla dzieci sześciolletnich w zależności od ich statusu edukacyjnego. Wąsy reprezentują jedno odchylenie standardowe powyżej i poniżej średniej.

Powyższe różnice występują również po wprowadzeniu do analiz wielkości miejscowości zamieszkania oraz statusu socjoekonomicznego rodziny pochodzenia dziecka (określanego przez wykształcenie rodziców dziecka) jako zmiennych kontrolowanych. Wielkość miejscowości zamieszkania po wprowadzeniu do modelu statusu socjoekonomicznego nie wpływa istotnie na umiejętności dziecka. Natomiast status socjoekonomiczny, zgodnie z oczekiwaniami (por. Dolata, 2008; Konarzewski, 2007), wpływa pozytywnie na obserwowane poziomy umiejętności. Jednak dla wszystkich trzech skal jest on słabiej związany z wynikami, jakie uzyskują uczniowie niż typ placówki, do której uczęszczają. Warto zauważyć, że obserwowane wzory różnic - najwyższe wyniki w pierwszych klasach, najniższe w zerówkach przyszkolnych i przedszkolach nie zmieniają swojego kształtu po wprowadzeniu zmiennych kontrolowanych.

Wykształcenie rodziców istotnie statystycznie pozwala przewidywać wyniki dziecka w przeprowadzonych testach osiągnięć. Poniżej przedstawione (rysunek 7.) zostały wartości współczynników  $R^2$  (modele regresji liniowej), obliczone dla wszystkich 6-latków znajdujących w różnych placówkach, wyjaśniające kolejno umiejętności matematyczne, pisania i czytania. Zmienną wyjaśniającą był standaryzowany wskaźnik wykształcenia rodziców dziecka.





Rysunek 7. Procent wariacji wyników dzieci 6-letnich na trzech skalach TUnSS wyjaśnianej przez poziom wykształcenia rodziców

Jak widać na powyższym rysunku, wykształcenie rodziców w sposób najlepszy pozwala przewidywać u dziecka sześciolatniego umiejętność czytania (14%), następnie umiejętności matematyczne (9%), a w najmniejszym stopniu umiejętność pisania (6%).

#### 4. Dyskusja wyników

Najważniejszym rezultatem badania standaryzacyjnego TUnSS (oprócz oczywiście określenia parametrów samego testu) jest możliwość określenia skali różnic w poziomach umiejętności dzieci na pierwszych etapach edukacyjnych. Wnioski z przeprowadzonych badań mają bardzo istotne znaczenie dla konstrukcji programów dydaktycznych na pierwszych etapach kształcenia. W pierwszej klasie szkoły podstawowej uczą się wspólnie dzieci, z których część jest już na poziomie drugiej klasy, a część ma umiejętności niewykraczające ponad poziom przedszkola. Opisany przez nas poziom zróżnicowania w zakresie trzech istotnych na progu edukacyjnym umiejętności stanowi duże wyzwanie dla nauczycieli pracujących z takimi grupami. Mamy jednak nadzieję, że przygotowanie precyzyjnego i łatwego w użyciu narzędzia oraz możliwość korzystania z wyników prowadzonych przy jego pomocy badań przyczynią się do wsparcia zarówno nauczycieli, jak i decydentów w podejmowaniu jak najlepszych decyzji dotyczących startu szkolnego zarówno dzieci 6-letnich, jak 7-letnich.

#### Bibliografia:

1. Dolata, R. (2009). *Szkoła, segregacja, nierówności*. Warszawa: Uniwersytet Warszawski.
2. Hambleton, R. K. i Swaminathan, H. (1985). *Item Response Theory: Principles and applications*. Boston: Kluwer.
3. Hurek, J. & Szejnberg, A. (2010). *Doskonalenie komputerowego pomiaru testowego*. Opole: Wydawnictwo Uniwersytetu Opolskiego.
4. Karwowski, M. (2011). *Wczesna opieka i edukacja*, [w:] Federowicz M. i Sitek M. (red.), *Spółczesność w drodze do wiedzy. Raport o stanie edukacji 2010*. Warszawa: Instytut Badań Edukacyjnych.
5. Konarzewski, K. (2007). *PIRLS 2006. Jak czytają dzieci w Polsce?* Warszawa: Centralna Komisja Egzaminacyjna.
6. Penfield, R. D. i Bergeron, J. M. (2005). Applying a Weighted Maximum Likelihood latent trait estimator to the Generalized Partial Credit Model. *Applied Psychological Measurement*, 29, 218-233.
7. Weiss, D. J. (2011). Better data from better measurement using computerized adaptive testing. *Journal of Methods and Measurement in the Social Sciences*, 2, 1-23.