

Janusz Mulawa, Dariusz Sroka

Instytut Badań Kompetencji w Wałbrzychu

Trafność pomiarowa empirycznej normy osiągnięć

Diagnozowanie osiągnięć uczniów wymaga stawiania im zadań. Jakość ich rozwiązania pozwala na ocenę tego, co wiedzą, umieją i potrafią. Ta ocena może mieć jednak dwa punkty odniesienia. Z jednej strony może być odniesiona do tego, ile uczniowie wiedzą z tego, co wiedzieć powinni, a zatem odniesiona do zakresu przedmiotowych treści kształcenia, z drugiej natomiast może wynikać z porównania tego, co umieją na tle populacji, czyli z zestawienia wskaźników łatwości zadań.

W dzisiejszych uwarunkowaniach społecznych istnieje tylko jeden sposób oceniania – wewnątrzszkolny, oparty na subiektywnych sądach nauczycieli. Instytut Badań Kompetencji, prowadząc badania, dąży do dostarczenia nauczycielom, czyli wewnątrzszkolnemu systemowi oceniania, **empirycznych norm osiągnięć** polskiej szkoły.

Osiągnięcia uczniów są wypadkową działania wielu czynników. Najważniejsze z nich ilustruje poniższy schemat.

Schemat 1.



Wszystkie te czynniki jak mechanizmy ogromnej maszyny oddziałują codziennie na rzecz generowania osiągnięć uczniów. Działanie to przebiega w warunkach pełnej gotowości, determinacji i motywacji wszystkich jego uczestników do ciągłego podnoszenia jakości.

Uczniowie **chcą** poznawać świat. Rodzice **chcą** stwarzać swoim dzieciom coraz lepsze warunki do kształcenia i rozwoju. Władza samorządową różnych szczebli **chce**, by ich lokalne społeczności rozwijały się, a różnice kulturowe, gospodarcze i cywilizacyjne niwelowały na rzecz wyrównywania szans. Szkoły **chcą** być dobrze wyposażone, a nauczyciele **chcą** uczyć lepiej, nowocześniej, z pełnym zaangażowaniem.

Gdy wszyscy **chcą**, należy postawić pytanie o efektywność i wydajność takiego działania.

Czy osiągnięcia uczniów, jako efekt pracy systemu, mierzone łatwością stawianych im zadań, mają społeczne optima i czy wskaźniki te, na skutek czynionych nieustannych wysiłków, ulegają zmianom.

Czteroletnie obserwacje i prowadzone przez Instytut Badań Kompetencji badania osiągnięć uczniów pozwalają na postawienie kilku hipotez i próbę ich dowodzenia już nie tylko w oparciu o czysto teoretyczne wnioskowanie, lecz także o liczne empiryczne przykłady.

Hipoteza I: Istnieją empiryczne normy osiągnięć uczniów na każdym poziomie kształcenia.

Hipoteza II: Empiryczne normy osiągnięć na danym poziomie kształcenia są niezmiennie w czasie.

Hipoteza III: Wartość dodana (VAT) łatwości zadań na kolejnych poziomach kształcenia jest niewielka.

Hipoteza I

W badaniach naukowych, w których bierze udział kilka tysięcy uczniów, obserwowanie wskaźnika łatwości zadania prowadzi do stwierdzenia, że po wstępnych wahaniach stabilizuje się on na pewnym poziomie, który można uznać za wystarczająco trafną normę do rozstrzygnięcia o poziomie osiągnięć innych grup uczniów.

Dołączenie do tego zbioru wyników kolejnej grupy uczniów, różniących się nawet bardzo od ukształtowanej normy, nie wpłynie na jej zmianę.

Hipoteza II

Analizując pięć czynników decydujących w efekcie o osiągnięciach uczniów, można zauważyć, że wszystkie one w sensie socjologicznym są słabo zmienne w czasie.

Rozkład „obdarowanych genetycznie” w społeczeństwie jest z roku na rok podobny. Dobrostan, wydolność, poziom wykształcenia polskiej rodziny to społeczna

zmienna raczej pokoleniowa. Środowiska lokalne, choć zmieniane materialnie, w ostatnich latach nie powodują szybkich zmian mentalnych.


Materialny rozwój szkół podejmowany z ogromnym wysiłkiem organizacyjnym i finansowym, może dawać efekty tylko wówczas, gdy inwestycje te pójdą w parze z doskonaleniem zawodowym nauczycieli. To jednak proces ciągły i wieloletni.

Czy w warunkach niezmiennych socjologicznie generatorów osiągnięć uczniów mogą ulegać zmianie normy osiągnięć? Badania potwierdzają, że odkrywane empirycznie poziomy osiągnięć w ostatnich latach nie uległy w zasadzie zmianie. Poruszone zagadnienia dowodzą, że, w różnych zakresach badanych umiejętności, wieloletnie obserwacje łatwości zadań wyznaczają empiryczną normę osiągnięć. Może ona przekładać się na wielopoziomową ocenę osiągnięć ucznia w wewnątrzszkolnym systemie kształcenia.

Charakterystycznym jest również i to, że w ciągu czterech lat nie uległ zmianie kształt krzywych łatwości tych zadań na staninowej skali rozkładu poziomów kompetencji.

W latach 2000–2003 w badaniach osiągnięć uczniów klas III szkół podstawowych powtarzano, w identycznym brzmieniu, cztery zadania. Zmieniło się jedynie miejsce zadania w arkuszu testowym. Oto zestawienie różnych parametrów tych zadań:

ZADANIE 1. Uczeń oblicza wartość zakupionych towarów



Tomek zaprosił na przyjęcie 5 kolegów. W sklepie kupił 1 kg czekoladowych cukierków, pół kg winogron i po 1 jogurcie dla każdego kolegi.

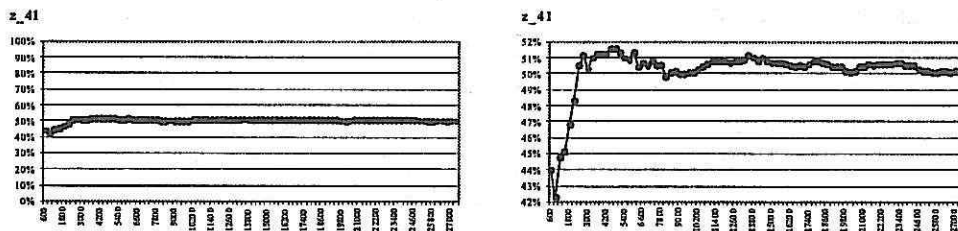
Ile chłopiec zapłacił w sklepie?

Tabela 1.

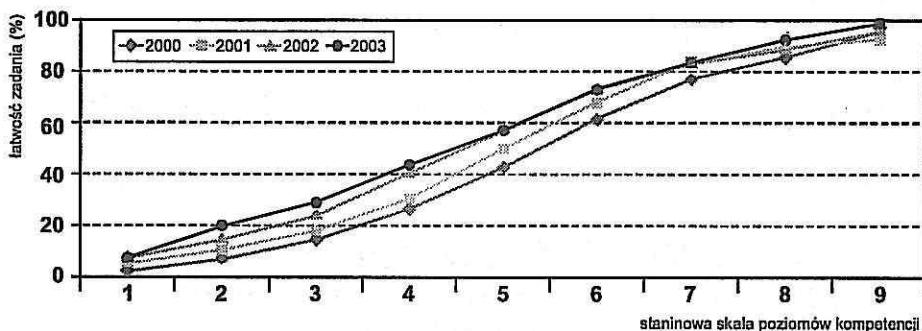
Łatwość zadania, rozkład typowych dystraktorów oraz frakcje opuszczeń w latach 2000–2003

kod		rok 2000 zadanie 30	rok 2001 zadanie 41	rok 2002 zadanie 44	rok 2003 zadanie 42
K	uczeń udziela odpowiedzi poprawnej: $8 + 2 + 4,50 = 14,50$	47%	50%	56%	59%
L	uczeń nie zauważa, że Tomek kupuje pół kg winogron i wykonuje dodawanie: $8 + 4 + 4,40 = 16,50$	6%	6%	6%	
M	uczeń dodatkowo nie łączy informacji „zaprosił 5 kolegów” i kupił „po jednym jogurcie dla każdego kolegi” i wykonuje dodawanie: $8 + 4 + 0,90 = 12,90$	42% (nie wyróżniono kodu M)	12%	34%	5%
N	uczeń udziela innej błędnej odpowiedzi, na przykład.: $2 + 450 + 1 = 4,60$; $5 \times 2 + 5 \times 0,90 + 5 \times 8 = 54,50$		28%		33%
O	uczeń nie udziela żadnej odpowiedzi	5%	4%	4%	3%

Wykres 1.
Kształtowanie się empirycznej normy tej umiejętności
na przykładzie roku 2001



Wykres 2.
Łatwość zadania w poszczególnych staniach w latach 2000–2003



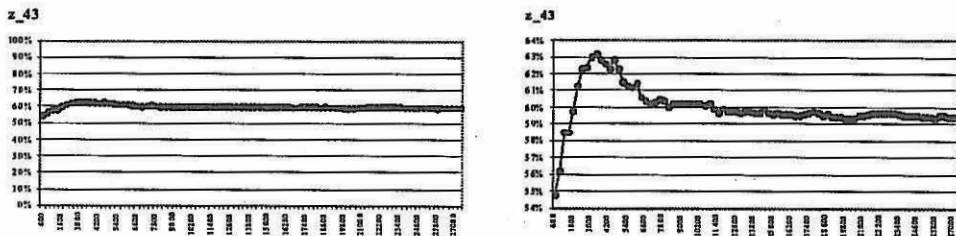
ZADANIE 2. Uczeń mnoży liczbę trzycyfrową przez jednocyfrową

$$149 \cdot 6 =$$

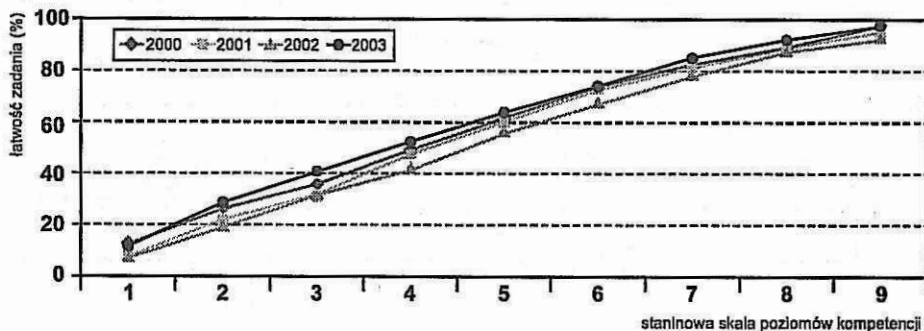
Tabela 2.
Łatwość zadania, rozkład typowych dystraktorów oraz frakcje opuszczeń
w latach 2000–2003

kod		rok 2000 zadanie 30	rok 2001 zadanie 41	rok 2002 zadanie 44	rok 2003 zadanie 42
K	uczeń udziela odpowiedzi poprawnej: 894	61%	60%	57%	65%
L	najczęściej pojawiający się niepoprawny wynik: 844	1%	1%	1%	1%
N	uczeń udziela innej, błędnej odpowiedzi	30%	31%	33%	29%
O	uczeń nie udziela żadnej odpowiedzi	8%	8%	9%	6%

Wykres 3.
Kształtowanie się empirycznej normy tej umiejętności
na przykładzie roku 2001



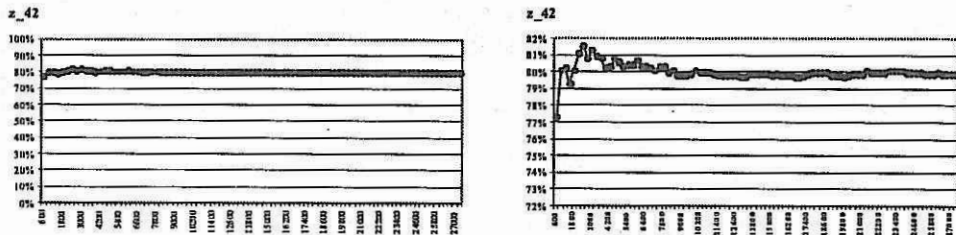
Wykres 4.
Łatwość zadania w poszczególnych staninach w latach 2000–2003



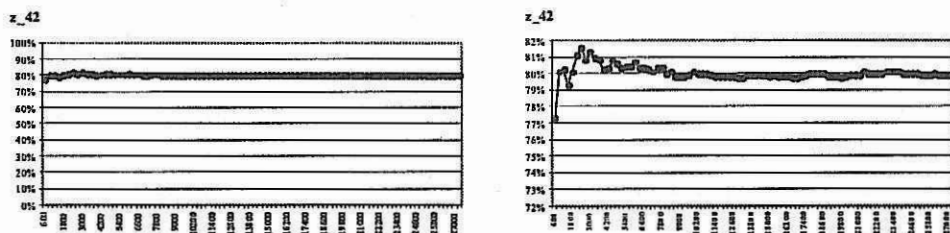
ZADANIE 3. Uczeń odejmuje sposobem pisemnym

$$\begin{array}{r} 384 \\ - 165 \\ \hline \dots \end{array}$$

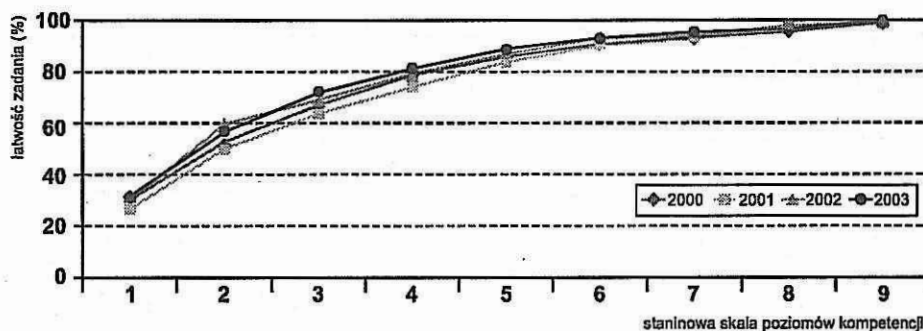
Tabela 3.
Łatwość zadania, rozkład typowych dystraktorów oraz frakcje opuszczeń
w latach 2000–2003



Wykres 5.
Kształtowanie się empirycznej normy tej umiejętności
na przykładzie roku 2001



Wykres 6.
Łatwość zadania w poszczególnych staninach w latach 2000–2003



ZADANIE 4. Uczeń oblicza wartość zakupionych towarów

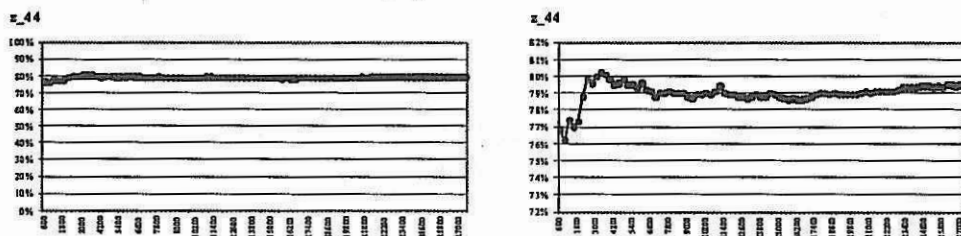
Babcia kupiła 2 kg jabłek po 4 zł i 3 kg mąki po 2 zł za kilogram. Który zapis pozwala wyliczyć, ile babcia zapłaciła za zakupy? (Zaznacz krzyżykiem.)

- $2 \cdot 2 + 3 \cdot 4$
 $2 \cdot 4 + 3 \cdot 2$
 $(4 + 2) \cdot (2 + 3)$

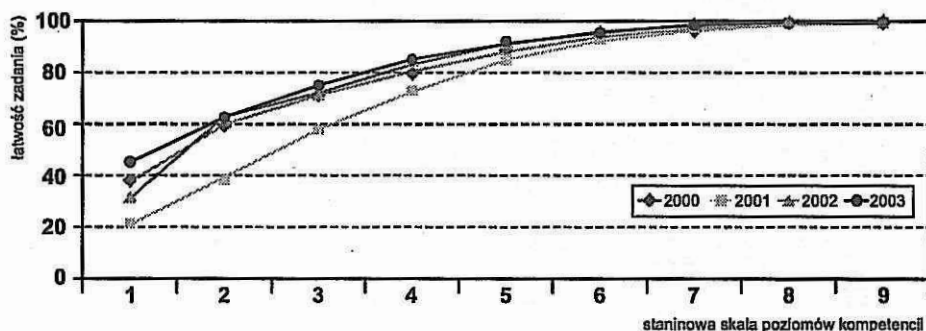
Tabela 4.
Łatwość zadania, rozkład typowych dystraktorów oraz frakcje opuszczeń
w latach 2000–2003

kod		rok 2000 zadanie 29	rok 2001 zadanie 44	rok 2002 zadanie 45	rok 2003 zadanie 43
K	uczeń udziela odpowiedzi poprawnej: $2 \times 4 + 3 \times 2$	85%	79%	86%	87%
L	uczeń udziela niepoprawnej odpowiedzi: $(4 + 2) \times (2 + 3)$	14%	8%	5%	4%
M	uczeń udziela niepoprawnej odpowiedzi: $2 \times 2 + 3 \times 4$		3%	3%	3%
N	uczeń udziela innej, błędnej odpowiedzi		4%	4%	5%
O	uczeń nie udziela żadnej odpowiedzi	1%	6%	2%	1%

Wykres 7.
Kształtowanie się empirycznej normy tej umiejętności
na przykładzie roku 2001



Wykres 8.
Łatwość zadania w poszczególnych staniach w latach 2000–2003



Hipoteza III

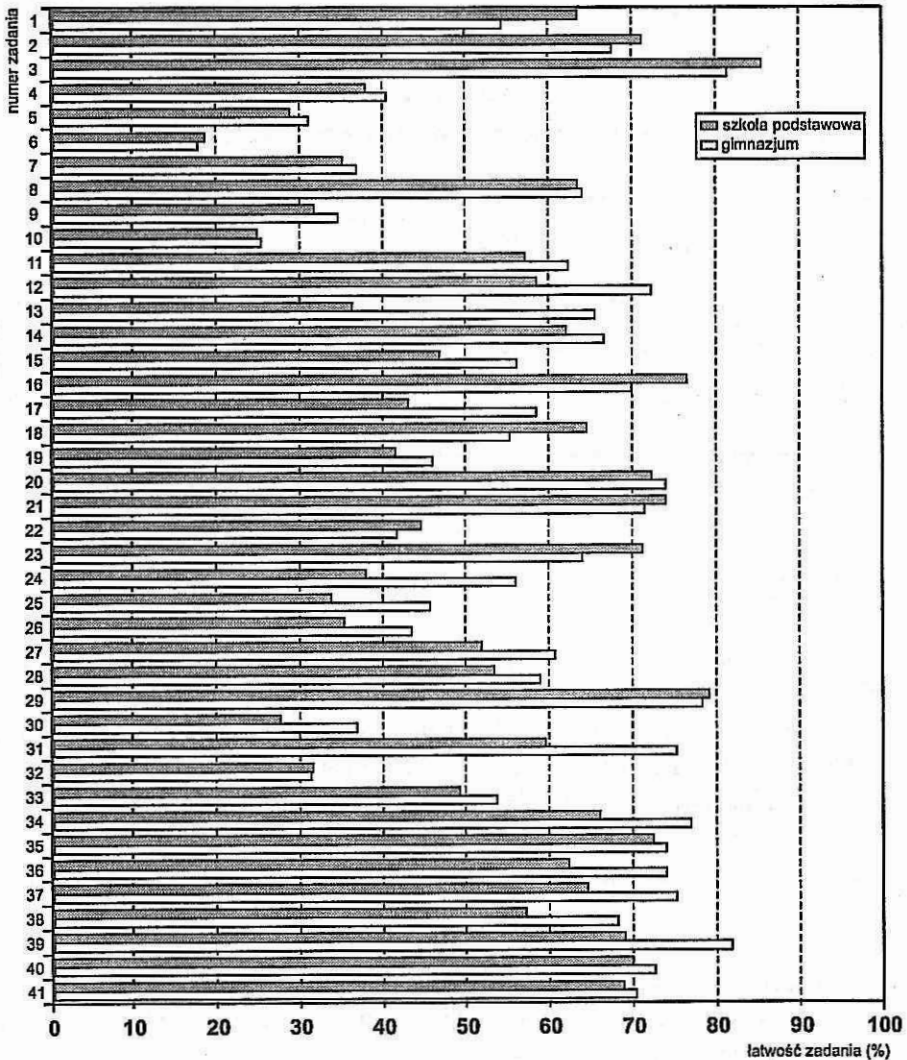
Badając osiągnięcia uczniów, próbujemy odpowiedzieć również na następujące pytania:

1. Jak będzie zmieniała się łatwość zadania, które w niezmienionej formie postawimy uczniom na różnych poziomach kształcenia?
2. Czy zadania łatwe pozostaną łatwymi, a trudne – trudnymi, niezależnie od poziomu kształcenia?
3. Czy zostanie zachowana hipotetycznie naturalna relacja, że im starszy jest uczeń, tym łatwiejsze stają się zadania?

Odpowiedzi na te pytania stanowią punkt wyjścia do budowania banków zadań, a w przyszłości także do konstruowania wewnętrznych testów jako narzędzi oceny osiągnięć uczniów z uwzględnieniem wielopoziomowości w nauczaniu. Poszukajmy ich w zestawieniach łatwości zadań trzech testów.

1. Testu eliminacyjnego konkursu „Moja przedsiębiorczość na drodze do Unii Europejskiej”, który rozwiązywali zarówno uczniowie klas VI szkół podstawowych, jak i I i II gimnazjum, powiatów wałbrzyskiego i świdnickiego w maju 2003 r.

Wykres 9.
Porównanie wyników testu eliminacyjnego
szkół podstawowych i gimnazjów

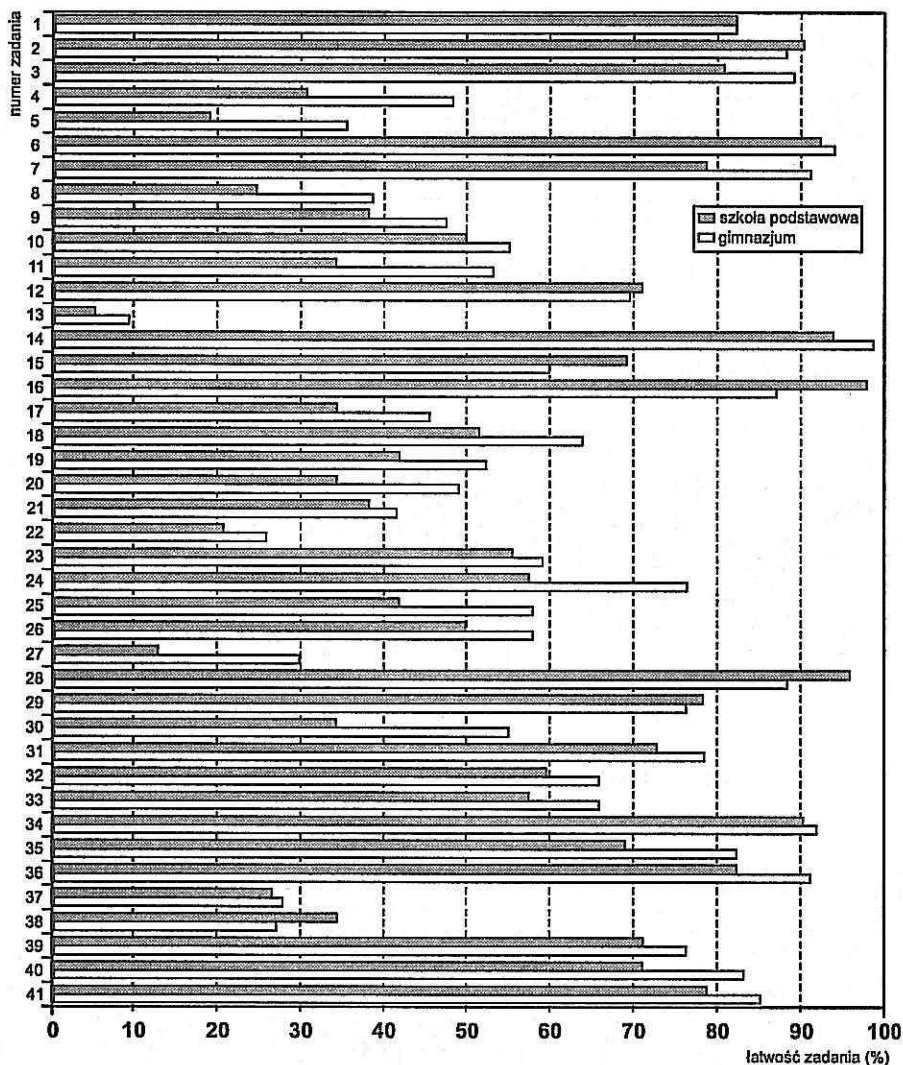


Liczba uczniów: szkoły podstawowe – 71, gimnazja – 169

Średni wynik: szkoły podstawowe – 23,85 – 58%; gimnazja – 25,88 – 63%

2. Testu finałowego tegoż konkursu.

Wykres 10.
Porównanie wyników testu eliminacyjnego
szkół podstawowych i gimnazjów

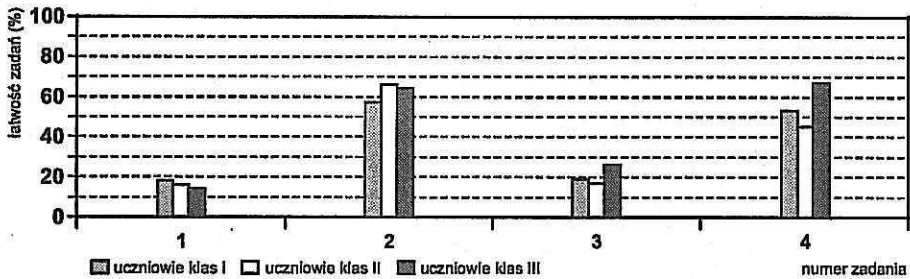


Liczba uczniów: szkoły podstawowe – 52, gimnazja – 103

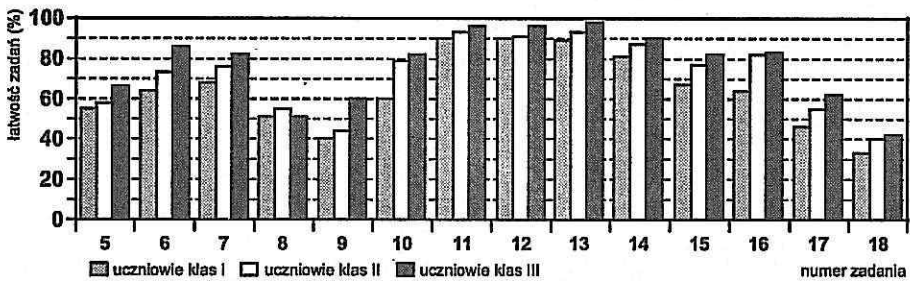
Średni wynik: szkoły podstawowe – 30,33 – 58%; gimnazja – 33,81 – 65%

3. Testu predyspozycji językowych, który dano do rozwiązania uczniom klas I, II, i III gimnazjum na początku roku szkolnego 2001/2002.

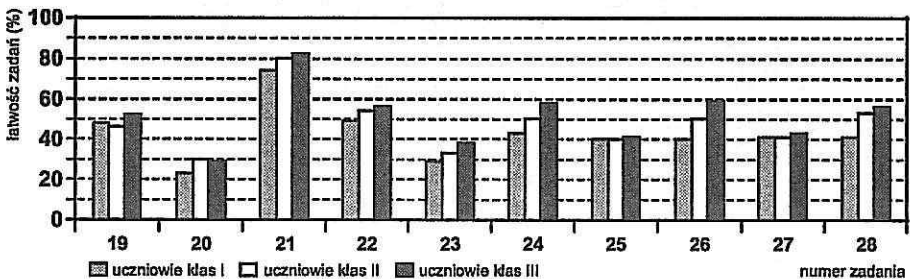
Uczeń wykazuje się wycuciem stylistycznym i umiejętnością stosowania poprawnych struktur zdaniowych



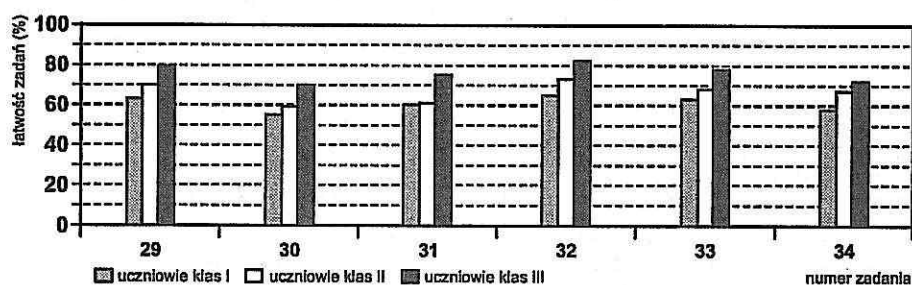
Uczeń wykazuje się zasobem słownictwa i związków frazeologicznych



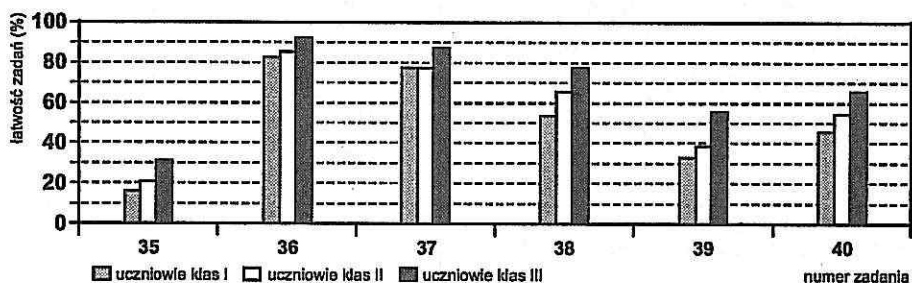
Uczeń rozróżnia poprawne formy gramatyczne na podstawie języka polskiego, dostrzega m.in. analogie



Uczeń rozróżnia poprawne formy gramatyczne na podstawie języka polskiego (język łaciński)



Uczeń wyciąga wnioski i buduje reguły rządzące językiem obcym (fikcyjnym)



Każde z zestawień potwierdza hipotezę, że zadania łatwe pozostają łatwymi, a trudne – trudnymi na kolejnych poziomach kształcenia, a wartość dodana łatwości tych zadań tylko nieznacznie wzrasta wraz z przechodzeniem ucznia na wyższe poziomy kształcenia.

Hipoteza o niewielkim wzroście łatwości zadań posłużyła do skonstruowania testu z wykorzystaniem zadań, których pomiaru łatwości na dużej populacji dokonano podczas badania wykonanego wcześniej.

Do zbadania kompetencji uczniów na zakończenie klasy I gimnazjum w maju 2003 r. przygotowano test, który niemal w całości skonstruowano z zadań o znanych wcześniej parametrach łatwości. 13 zadań pochodziło z testu, którym w styczniu 2001 r. badano kompetencje uczniów klas VI szkoły podstawowej oraz II klasy gimnazjum. Przewidywano, że wskaźniki łatwości dla uczniów klas I gimnazjum powinny być zawarte pomiędzy odpowiednimi wskaźnikami łatwości użytych zadań.

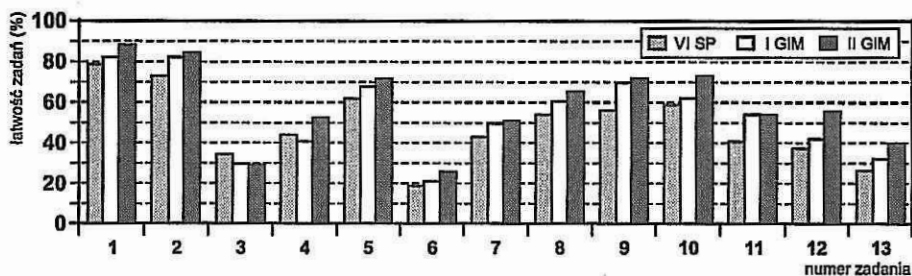
Oto zestawienie wskaźników łatwości zadań testu:

	styczeń 2001	maj 2003	styczeń 2001
	6 SP	I GIM	II GIM
1. Ruch obrotowy (wirowy) Ziemi powoduje, że: A. są pory roku B. zmienia się klimat Ziemi C. jest dzień i noc D. wieją wiatry	79%	82%	88%
2. Ruch obiegowy Ziemi powoduje, że: A. są wschody i zachody Słońca B. jest wiosna, lato, jesień, zima C. przemieszczają się chmury na niebie D. wieją wiatry	73%	82%	84%
3. Długość dnia i nocy na Ziemi ma związek z: A. grubością warstwy chmur B. ruchem obrotowym Ziemi C. grawitacją na Ziemi D. ruchem obiegowym Ziemi	25%	30%	29%
4. Gdy Słońce góruje w zenicie na Zwrrotniku Raka (23° 27' szerokości geograficznej północnej) to w Polsce jest: A. lato B. wiosna C. zima D. jesień	44%	41%	53%
5. Cień masztu pada w kierunku północno-wschodnim. Z którego kierunku świeci słońce? A. północno-wschodniego B. południowo-zachodniego C. południowo-wschodniego D. północno-zachodniego	62%	68%	72%
6. Maszt anteny telewizyjnej koło Warszawy rzuca cień na ziemię. Jest on najdłuższy o godzinie 12⁰⁰: A. 21 marca B. 21 września C. 22 grudnia D. 23 czerwca	19%	21%	26%
9. Zrównanie długości dnia i nocy następuje A. 4 razy w roku B. raz w roku C. 2 razy w roku D. 3 razy w roku	43%	50%	51%
10. Doba na Marsie jest o 40 minut dłuższa niż na Ziemi. Z tego wynika, że: A. Ziemia jest większa od Marsa B. Ziemia obraca się wokół swojej osi szybciej niż Mars C. Mars obraca się wokół swojej osi szybciej niż Ziemia D. Mars jest większy od Ziemi	54%	61%	65%

Europa jest najmniejszym z 5 stałych, zamieszkałych kontynentów, ale równocześnie najgęściej zaludniona. Rozciąga się od Atlantyku aż po Ural i od Morza Arktycznego po Morze Śródziemne. Na obszarze 10,5 mln km² stanowiących 10% stałych lądów planety żyje około 720 mln mieszkańców, a więc prawie tyle, ilu na obu kontynentach amerykańskich o czterokrotnie większej powierzchni. Takie zagęszczenie jest skutkiem dwóch stuleci bezprecedensowego rozwoju gospodarczego.

(tekst „Encyklopedia dla młodzieży” – Larousse 1994)

	styczeń 2001	maj 2003	styczeń 2001
	6 SP	I GIM	II GIM
16. Jaka jest powierzchnia stałych lądów Ziemi? A. 105 000 km ² B. 940500 km ² C. 94,5 mln km ² D. 105 mln km ²	56%	70%	72%
17. Jaka jest powierzchnia obu kontynentów amerykańskich? A. 14,5 mln km ² B. 42 mln km ² C. 4 mln km ² D. 420 mln km ²	59%	62%	73%
18. Który jest najmniejszym zamieszkałym kontynentem? A. Antarktyda B. Azja C. Europa D. Afryka	41%	54%	54%
19. Podaj zagęszczenie kontynentu europejskiego, czyli liczbę ludności przypadającą na 1 km² powierzchni. A. 0,14 B. 720 C. 7 D. 68	37%	42%	56%
20. Gęstość zaludnienia obu Ameryk jest: A. taka jak w Europie B. mniejsza niż w Europie C. większa niż w Europie D. brak danych by to określić	26%	32%	40%



Przewidywania w pełni potwierdziły naszą hipotezę o możliwości konstruowania nowych testów na podstawie znajomości wskaźników pomiarowych wcześniej użytych zadań. Zadania łatwe pozostały łatwe, a trudne – trudne. Dla uczniów klasy I gimnazjum były one łatwiejsze niż dla kolegów z klasy VI szkoły podstawowej, ale równocześnie nieco trudniejsze niż dla kolegów z klasy II gimnazjum.