

**Michał Hnatiuk**

Narodowy Uniwersytet „Politechnika Lwowska”, Lwów, Ukraina

Uczelnia Społeczno-Medyczna w Warszawie

ORCID: 0000-0002-7939-1010

Testy nauczycielskie. Teoretyczne i praktyczne aspekty zadań testowych

W artykule odzwierciedlono wybrane fragmenty dorobku naukowego w dziedzinie testów nauczycielskich. Opisano teoretyczne podstawy jakościowej i ilościowej analizy zadań testowych. Zwrócono uwagę na sposoby weryfikacji poprawności dydaktycznej, rzetelności oraz stosowności zadań testowych. Zaprezentowano także sposoby obliczenia zasadniczych parametrów ilościowej analizy testu, w szczególności: frakcji opuszczeń zadań, trudności i łatwości zadań oraz mocy różnicującej zadania testowe.

Uwzględniając praktyczne aspekty testu nauczycielskiego, opisano ogólne zasady konstrukcji zadań testowych. Zaprezentowano przykładowe testy oraz tabelę surowych wyników badań wiedzy i umiejętności uczniów. Na przykładzie wyników testowania obliczono podstawowe parametry ilościowej analizy zadań testowych. Uzyskane wskaźniki zinterpretowano, a na ich podstawie zweryfikowano poprawność dydaktyczną testu.

Słowa kluczowe: pedagogika, jakościowa i ilościowa analiza zadań testowych, ogólne zasady konstrukcji zadań testowych, testy nauczycielskie

Teacher's tests. Theoretical and practical aspects of test tasks

The article reflects selected fragments of scientific achievements in the field of teacher's tests. The theoretical basis of the qualitative and quantitative analysis of test tasks is described. Attention is paid to the methods of verifying the didactic correctness, reliability and appropriateness of the test tasks. The methods of calculating the basic parameters of the quantitative test analysis, in particular: the fraction of missing tasks, difficulty and ease of tasks, and the power to differentiate test tasks were also presented.

Taking into account the practical aspects of the teacher's test, the general principles of constructing test tasks are described. Sample tests and the table of raw results of research on

students' knowledge and skills are presented. The basic parameters of the quantitative analysis of test tasks were calculated on the basis of the test results. The obtained indicators were interpreted, and on their basis, the didactic correctness of the test was verified.

Keywords: pedagogy, qualitative and quantitative analysis of test tasks, teacher's tests, general principles of test task construction

Wprowadzenie

W praktyce pedagogicznej osiągnięcia edukacyjne wychowanków są sprawdzane i oceniane za pośrednictwem różnych metod. Do tradycyjnych zaliczane są: ustne, pisemne oraz oparte na obserwacji metody oceniania. Jednak, wraz z rozwojem dydaktyki coraz częściej wiedzę i umiejętności uczniów weryfikuje się za pośrednictwem metody testów nauczycielskich – testowania.

W systemie oświaty jednego państwa metoda testów w procesie dydaktycznym przybiera mniejsze, w systemie oświaty drugiego – większe znaczenie. W niniejszym artykule zdecydowano ujawnić role testów nauczycielskich w dwóch sąsiednich, a zarazem odmiennych systemach edukacji – polskim i ukraińskim. W związku z tym, zgromadzony materiał w głównej mierze oparto o dorobek naukowy polskich i ukraińskich badaczy w tym obszarze dydaktyki.

Według Iryny E. Bulakh i Maryny R. Mrugi „*testowanie* jako metoda – to proces pomiaru wskaźników ilościowych za pomocą testu” (Bulakh i Mruga, 2006, s. 9). W myśl tego twierdzenia możemy uznać, iż testowanie to metoda diagnostyki pedagogicznej pozwalająca za pośrednictwem testu zweryfikować poziom osiągnięć edukacyjnych uczniów w stosunku do wymogów programu nauczania.

Ze względu na wzrastającą rolę testów w procesie dydaktycznym w pracach naukowych ukraińskich badaczy (m.in. Liaszenko i in., 2012; Bykow i in., 2008) prowadzono szeroki dyskurs na temat stosunkowo nowej metody pomiaru pedagogicznego. Przedstawiono szereg definicji określających podstawowe pojęcia. W odniesieniu do *testu* jako narzędzia pomiaru pedagogicznego zaznaczono, że jest to „zestaw zadań, które najczęściej wymagają jednoznacznej odpowiedzi, skonstruowane według określonych zasad i procedur, przewidujące uprzednią weryfikację eksperymentalną i spełniające takie parametry efektywności, jak trafność i rzetelność” (Kuchar i Sergienko, 2010, s. 45). Szczególną uwagę zwrócono na praktyczne aspekty danej metody pomiaru. Opisano etapy konstruowania zadań testowych, zasady doboru treści, formy (formaty) zadań testowych, sposoby prezentacji wyników testów oraz wiele innych zagadnień dotyczących technologii konstruowania zadań testowych.

Ważnym aspektem danej problematyki jest to, że w dzisiejszych czasach testy nauczycielskie przeprowadza się w oparciu o technologie informacyjno-komunikacyjne. W Ukrainie w szkolnictwie ogólnokształcącym oraz wyższym powszechnie stosowane są technologie zadań testowych oparte na formularzach Google. Uwzględniając potrzeby edukacyjne uczniów oraz szerokie możliwości wymiany doświadczeń wśród pracowników pedagogicznych, na ukraińskich stronach internetowych prezentowane są różnorodne projekty edukacyjne, w tym także testy on-line. W ramach jednego z projektów powstała internetowa strona edukacyjna pt. *Na lekcji* («На Урок») na której zaprezentowano m.in. interaktywne zadania (testy z różnych przedmiotów o zróżnicowanym poziomie edukacyjnym) kontrolujące wiedzę i umiejętności uczniów oraz angażujące ich w aktywną pracę w klasie i w domu (Projekt edukacyjny *Na urok*).

Warto zaznaczyć, że pomiar wyników kształcenia za pośrednictwem testów nauczycielskich jest tematyką wielostronną. Z uwagi na jej obszerny zakres postanowiono ograniczyć się do ważnych, a zarazem pomocnych dla procesu dydaktycznego zagadnień – ogólnych zasad konstrukcji zadań testowych oraz ich jakościowej i ilościowej analizy. Postanowiono także zaprezentować przykład konstrukcji zadań testowych, sposób analizy testu nauczycielskiego oraz pozyskania informacji o rzetelności danego narzędzia.

Materiał badań i dyskusja

Teoretyczne podstawy pomiaru wyników kształcenia

Teoretyczne podstawy pomiaru wyników kształcenia opisano w różnych źródłach literaturowych, w różnych językach (m.in. Ciechanowicz, 1990; Ioannou-Georgiou i Pavlou, 2003; Kaniwicz, 2012; Liaszenko i in., 2012; Magnusson, 1981; Niemierko, 1975).

W tej części materiału postanowiono zaprezentować polski dorobek naukowy oraz opisać najważniejsze parametry, według których przeprowadza się jakościową i ilościową analizę testu nauczycielskiego.

Jakościowa analiza zadań testowych przewiduje sprawdzenie testu pod kątem *poprawności dydaktycznej* – czy zadania testowe są zgodne ze stanem wiedzy naukowej (poprawność rzeczowa), czy treść zadań jest dopasowana do norm językowych (poprawność redakcyjna) oraz czy ich stopień trudności dopasowany jest do wymogów programowych (odpowiednia trudność); *rzetelności* – czy pomiar osiągnięć uczniów będzie dokładny; oraz *stosowności* – czy będą mierzone tylko te umiejętności uczniów, które zamierzamy sprawdzić. Jakościowa analiza zadań ma dostarczyć odpowiedź na pytania, jakiej jakości produkt jest proponowany uczniom i czy ta jakość spełnia ich oczekiwania (Niemierko, 1999).

Nauczycielowi, w samodzielnie ułożonym teście, nie rekomenduje się przeprowadzania samodzielnej weryfikacji zadań testowych pod kątem poprawności dydaktycznej. Warto to zadanie powierzyć innemu nauczycielowi (lub grupie nauczycieli) danego przedmiotu edukacyjnego – tzw. grupie osób kompetentnych. Ich zadaniem będzie sprawdzić, czy zadania testowe są zgodne ze stanem wiedzy naukowej oraz czy ich trudność jest dopasowana do wymogów programowych. Oprócz tego nauczyciel językoznawstwa może sprawdzić czy zadania testowe będą zrozumiałe dla uczniów i czy są one dostosowane do norm językowych. Jeśli recenzja specjalistów będzie pozytywna, to możemy stwierdzić, że pod kątem poprawności dydaktycznej zadania testowe ułożono poprawnie.

Sprawdzanie zadań testowych pod kątem rzetelności to swego rodzaju sprawdzanie poprawności pomiaru. W celu weryfikacji poprawności pomiaru test należy przeprowadzić dwukrotnie, nieznacznie zmieniając za drugim razem zadania testowe. Sprawdzenia testu pod kątem rzetelności zazwyczaj dokonuje sam nauczyciel, autor testu. Jeśli uzyskane wyniki testowania będą porównywalne, to tym samym potwierdzą one rzetelność proponowanego testu.

Sprawdzając zadania testowe pod kątem stosowności, należy przeanalizować, czy ułożone zadania mierzą te umiejętności uczniów, które zamierzano sprawdzić. Warto zaznaczyć, że w literaturze pedagogicznej stosowność zadania przedstawia się także jako „trafność doboru zadania do testu, czyli zgodność zadania z planem testu” (Ruchała, 2008, s. 146). Weryfikację tego parametru po raz kolejny warto powierzyć innemu nauczycielowi (lub grupie nauczycieli), czyli „opinii sędziów kompetentnych” (Walczak, 2001, s. 146). Jeśli osoby oceniające stwierdzą, że dane zadania testowe sprawdzają tylko te umiejętności wychowanków, które planowano sprawdzić (zadanie zgodne z planem testu), to analizę testu pod kątem stosowności możemy uznać za pozytywną, a test za poprawnie ułożony.

Ilościowa analiza zadań to swego rodzaju zbieranie informacji, danych statystycznych o *ilości opuszczonych zadań* (frakcja opuszczeń zadań) przez uczniów, o *trudności i/lub łatwości zadania* dla ogółu uczniów czy wybranej grupy oraz o *mocy różnicującej zadania* – zdolności zadań do szczegółowej weryfikacji wiedzy uczniów, a ściślej mówiąc – ich osiągnięć w zakresie programowym (Niemierko, 1999). Należy zaznaczyć, że moc różnicująca zadania jest uważana za główny wskaźnik rzetelności i stosowności zadania.

Ilościowa analiza zadań testowych to kolejny krok w weryfikacji poprawności dydaktycznej testu. Wobec tego, wyżej wymienionym parametrom ilościowej analizy zadań warto poświęcić więcej uwagi i szczegółowo opisać każdy z nich. Zacznę od parametru frakcja opuszczeń zadania.

Frakcja opuszczeń zadania (częstotliwość pomijania danego zadania) jest stosunkiem liczby uczniów, które opuścili zadanie, do liczby uczniów testowanych. Wyraża się wzorem I.1 (Niemierko, 1999):

$$(I.1) \quad f = \frac{n_0}{n}$$

f – frakcja opuszczeń zadania,
 n_0 – liczba uczniów, którzy opuścili zadanie,
 n – liczba uczniów testowanych.

„Najczęstszą przyczyną opuszczenia zadań – zaznacza Franciszek Ruchała – jest niezrozumienie treści zadania, brak czasu (uczeń nie zdążył dojść do danego zadania) oraz utrata motywacji (uczeń zrezygnował z rozwiązywania testu)” (Ruchała, 2008, s. 137). Gdy frakcja opuszczeń jednego z zadań nie przekracza wartości 0,15, uważa się, że zadanie ułożono dydaktycznie poprawnie i nie jest ono zbyt trudne dla uczniów (Ruchała, 2008).

Drugim parametrem ilościowej analizy testu nauczycielskiego jest łatwość i/ lub trudność zadania, wyrażana poprzez odpowiednie wskaźniki.

„Łatwość zadania jest stosunkiem liczby punktów uzyskanych za rozwiązanie zadania przez uczniów biorących udział w testowaniu do maksymalnej liczby punktów możliwej do uzyskania przez tę liczbę uczniów” (Ruchała, 2008, s. 136). Wskaźnik łatwości zadania (zarówno w testach otwartych, jak i zamkniętych) wyraża się wzorem, który w różnych źródłach literaturowych może przybierać nieco inną formę.

Według Wojciecha Walczaka (2001) poziom łatwości całego testu możemy określić za pomocą następującego wzoru:

$$(I.2) \quad A = \frac{\sum x_1}{T}$$

Gdzie:
 A – oznacza łatwość testu,
 $\sum x_1$ – to suma punktów uzyskanych przez wszystkich uczniów,
 T – to suma punktów możliwych do uzyskania w teście.

Należy zaznaczyć, że podany wzór zaproponowano stosować w sytuacji, gdy dysponujemy narzędziem oceniającym zadania w skali 0–1. „Jeśli – zaznacza Walczak – zadanie mierzy kilka umiejętności i jest punktowane np. od 0 do 3 punktów, to w tabeli zapisuje się je odpowiednio jako zadanie: 1a 0–1, 1b 0–1, 1c 0–1” (Walczak, 2001, s. 137). Innymi słowy, żeby ułatwić ogólną weryfikację badanej właściwości, ocenianie zadań punktowanych powyżej jednego punktu sprowadza się do oceniania poszczególnych elementów składowych tych zadań. Warto dodać, że w podanej formule wskaźnik T (suma punktów możliwych do uzyskania w teście) należy rozumieć jako iloczyn dwóch czynników: liczby uczniów biorących udział w testowaniu oraz liczby zadań testowych.

Bolesław Niemierko (1999) wskaźnik łatwości zadań testowych punktowanych dowolną liczbą punktów zaproponował wyrazić wzorem:

$$(I.3) \quad p_{\bar{x}} = \frac{\sum x}{nk}$$

$p_{\bar{x}}$ – łatwość zadania,
 $\sum x$ – suma punktów uzyskanych za rozwiązanie zadania przez n uczniów,
 n – liczba uczniów testowanych,
 k – maksymalna liczba punktów możliwych do uzyskania przez jednego ucznia za rozwiązanie zadania.

W praktyce dydaktycznej najczęściej stosowanym testem jest test zamknięty, układany z zadań wielokrotnego wyboru. Zazwyczaj, zadania wielokrotnego wyboru oceniane są w skali 0–1. Dla zadań tego typu wskaźnik łatwości zadania obliczany jest według wzoru I.4 (Niemierko, 1999):

$$(I.4) \quad p = \frac{n_p}{n}$$

p – łatwość zadań punktowanych 0–1,
 n_p – liczba uczniów, którzy rozwiązali zadanie,
 n – liczba uczniów testowanych.

W literaturze pedagogicznej „wskaźnik łatwości” występuje jako pozytywny wskaźnik osiągnięć uczniów, a „wskaźnik trudności” – jako negatywny wskaźnik ich osiągnięć. Jednak, nie oznacza to, że analiza zadań pod kątem ich trudności jest analizą mniej ważną. Co więcej, wskaźnik trudności, jako wskaźnik przeciwny do wskaźnika łatwości zadań testowych, możemy nazwać *wskaźnikiem weryfikującym*, wspierającym lub podważającym twierdzenie o poprawności dydaktycznej testu.

Uznaje się, że „trudność zadania punktowanego 0–1 jest stosunkiem liczby uczniów, którzy nie rozwiązali prawidłowo tego zadania (podali błędną odpowiedź, opuścili zadanie lub nie doszli do niego), do liczby uczniów biorących udział w testowaniu” (Niemierko, 1999, s. 153). Wskaźnik trudności zadania wyraża się wzorem I.5 (Niemierko, 1999):

$$(I.5) \quad q = 1 - p$$

q – trudność zadania punktowego 0–1,
 p – łatwość zadania punktowego 0–1.

Takie wskaźniki, jak „łatwość zadania” i „trudność zadania” (punktowane 0–1), mogą przybierać wartość w przedziale od 0 do 1. Ich interpretacja, zaprezentowana w różnych źródłach literaturowych, jest podobna, choć w niewielkim stopniu (o kilka setnych wskaźnika) może różnić się między sobą. W tabeli 1 zaprezentowano interpretacje wskaźnika łatwości zadań testowych według Niemierki.

Tabela 1.

Interpretacja wskaźników łatwości zadań

Wskaźnik łatwości	Interpretacja zadania
0,00–0,19	bardzo trudne
0,20–0,49	trudne
0,50–0,69	umiarkowanie trudne
0,70–0,89	łatwe
0,90–1,00	bardzo łatwe

Źródło: Niemierko, 1999.

Jeśli wskaźnik łatwości zadań przybiera wartości graniczne (0,00–0,19 lub 0,90–1,00), to oznacza to, że zadania są bardzo łatwe lub bardzo trudne. Te wartości danego wskaźnika wskazują także na bardzo niską poprawność dydaktyczną zadania. Jak wynika z tabeli 1, tylko zadania umiarkowanie trudne i łatwe (wskaźnik łatwości 0,50–0,89) świadczą o poprawności dydaktycznej testu. Prawdopodobnie zadania o takich wskaźnikach będą motywować uczniów do dalszej pracy.

Na pewien aspekt związany z trudnością testu lub poszczególnych zadań testowych zwraca uwagę Walczak, podkreślając, iż „generalna zasada rozłożenia trudności zadań w sprawdzianie mówi o tym, aby zachęcająco proste zadania umieścić w pierwszym fragmencie testu; zaś trudne, stopniowo trudniejsze i najtrudniejsze – w jego części środkowej i znów, proste w ostatniej części arkusza” (2001, s. 142).

Trzecim parametrem ilościowej analizy zadań jest *moc różnicująca zadania*. „Moc różnicująca zadania – zaznacza Ruchała – określa związek wyników tego zadania z wynikami innych zadań testu. Jest korelacją wyników zadania z wynikami testu, w którym mieści się to zadanie” (Ruchała, 2008, s. 137).

Moc różnicującą zadania testowe można obliczyć na kilka sposobów, stosując do tego odpowiedni wzór. Ponieważ każdy wzór na obliczenia danej wartości jest zbyt skomplikowany, a obliczanie wyników zajmuje sporo czasu, możemy posługiwać się wskaźnikiem zastępczym, mniej dokładnym. Wskaźnik zastępczy mocy różnicującej dla zadań punktowanych 0–1 oblicza się według wzoru (Niemierko, 1999; por. Ruchała, 2008; Walczak, 2001):

$$(I.6) \quad D_{50} = p_W - p_N$$

D_{50} – wskaźnik zastępczy mocy różnicującej
(połówkowy),
 p_W – łatwość zadania w lepszej połowie wyników testu,
 p_N – łatwość zadania w słabszej połowie wyników testu.

W analizie testów nauczycielskich wskaźnik ten jest istotny, ponieważ jego wartość liczbową świadczy o łatwości lub trudności zadań, a tym samym weryfikuje ich poprawność dydaktyczną. Bardzo niskie, zerowe lub ujemne wskaźniki sygnalizują, że ułożone zadania są bardzo trudne lub bardzo łatwe dla uczniów, a poprawność dydaktyczna testu jest niska. Tylko zadania umiarkowanie trudne i trudne dają odpowiednio dobre wskaźniki korelacyjne i potwierdzają, że test ułożono poprawnie.

Warto dodać, że na moc różnicującą zadania wpływa nie tylko łatwość i trudność zadań, lecz także długość testu sprawdzającego. Im krótszy test, tym wyższa moc różnicująca zadania. Jednak w praktyce krótkie zadania testowe nie są preferowane, ponieważ obiektywnie nie odzwierciedlają osiągnięć uczniów. Niemierko (1999) nazywa taki wskaźnik korelacyjny *pozorną mocą różnicującą zadania testowe*. Minimalną i zadowalającą moc różnicowania w zależności od długości testu przedstawiono w tabeli 2.

Tabela 2.

Minimalna i zadowalająca moc różnicująca zadania w testach o różnej długości

Liczba zadań w teście	Minimalna rzeczywista moc różnicująca zadania	Zadowalająca moc różnicująca zadania
100	0,10	0,22
25	0,20	0,42
10	0,32	0,60
5	0,45	0,75

Źródło: Niemierko, 1999.

Praktyczne aspekty konstrukcji i analizy zadań testowych

Wraz z rozwojem teoretycznych podstaw konstrukcji zadań testowych poszerzają się konteksty praktyczne. Z uwagi na ich obszerny zakres zdecydowano przedstawić tylko niektóre elementy wymiaru praktycznego: ogólne zasady konstrukcji zadań testowych oraz ich ilościową analizę.

Ogólne zasady konstrukcji zadań testowych opisano w licznych źródłach literaturowych. W tej części materiału zdecydowano się zaprezentować ukraiński dorobek naukowy (m.in. Bykow i in., 2008; Bulakh i Mruga, 2006; Kuchar i Sergienko, 2011; Liaszenko i in., 2012). W oparciu o te i in. źródła każdy nauczyciel może układać testy samodzielnie. Przystępując do konstruowania testu, warto pamiętać, że:

1. Każde zadanie testowe powinno być poświęcone ważnemu z punktu dydaktyki zagadnieniu. Nie warto marnować czasu na ocenianie znajomości błahych faktów.
2. Pytanie powinno być zadane tak jasno, aby można było na nie odpowiedzieć, nawet nie patrząc na opcje odpowiedzi.
3. Wszystkie dystraktory (nieprawidłowe odpowiedzi) muszą być jednorodne, czyli należeć do tej samej kategorii, co prawidłowa odpowiedź.
4. Wszystkie dystraktory muszą być wiarygodne i mniej więcej tej samej długości, co poprawna odpowiedź.
5. Treść zadania może być długa, a odpowiedzi powinny być krótkie.
6. W treści zadań nie należy używać kategoriycznych stwierdzeń (zawsze, nigdy, wszystkie) i nieokreślonych sformułowań (często, rzadko, oczywiście) (Bykow i in., 2008).

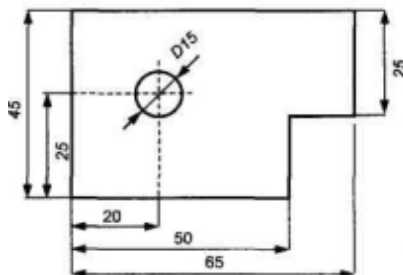
Uwzględniając powyższe zasady, układamy zadania testowe służące do weryfikacji wiedzy i umiejętności uczniów z danego przedmiotu edukacyjnego. Na przykładzie 1 zaprezentowano zadania sprawdzające osiągnięcia uczniów z rysunku technicznego w zakresie stosowania zasad wymiarowania. Każde zadanie testu jest oceniane w skali 1–0, czyli za prawidłową odpowiedź uczeń uzyskuje 1 punkt, a za nieprawidłową – 0.

Przykład 1

1. Odstęp między linią zewnętrznego konturu przedmiotu a najbliższą linią wymiarową powinien wynosić
 - a) 2–4 mm
 - b) 4–6 mm
 - c) 6–8 mm
 - d) 10 mm i więcej

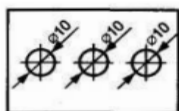
2. Ilość błędów na zwymiarowanym rysunku to¹

- a) jeden
- b) dwa
- c) trzy
- d) zero

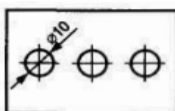


3. Rysunek, na którym zwymiarowano w sposób uproszczony otwory o jednakowej średnicy, to²

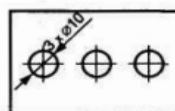
- 1) rys. 1
- 2) rys. 2
- 3) rys. 3
- 4) rys. 4



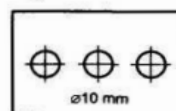
rys. 1



rys. 2



rys. 3

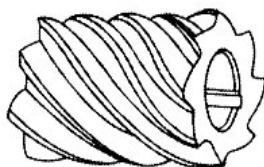


rys. 4

Przykładowe zadania testu nauczycielskiego z technologii mechanicznej sprawdzające wiadomości uczniów o narzędziach do kształtowania wyrobu zaprezentowano na przykładzie 2. Zasady oceniania odpowiedzi uczniów są takie same, jak w teście z rysunku technicznego (odpowiedź ucznia w każdym zadaniu jest oceniana jako prawidłowa lub błędna, w skali 1–0).

Przykład 2

1. Rysunek przedstawia



- a) frez pojedynczy walcowy
- b) frez zespołowy walcowy
- c) frez czołowy
- d) frez kątowy

Rysunek³

¹ Wojewódzki Urząd Pracy (2014, s. 2).

² Tamże, s. 3.

³ A. Górecki (1984, s. 252).

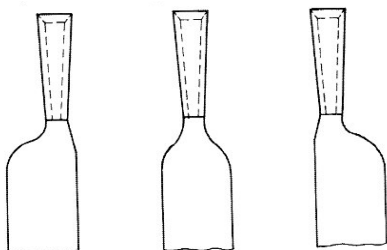
2. Poniższy rysunek przedstawia



Rysunek⁴

- a) wiertło kręte
- b) wiertło piórkowe
- c) wiertło sękownicze
- d) wiertło koronowe

3. Rysunek przedstawia rodzaj noży tokarskich. Są to noże



Rysunek⁵

- a) kształtowe
- b) oprawkowe
- c) odsadzone
- d) do toczenia gwintu

Ogólne (surowe) wyniki badań wiedzy i umiejętności podopiecznych z danego przedmiotu edukacyjnego (w zakresie danej tematyki przedmiotowej) należy zgromadzić w tabeli. W ramach przykładu zaprezentowano wyniki testowania z rysunku technicznego z tematyki *zasady wymiarowania* uzyskane w klasie pierwszej technikum (tabela 3). Uporządkowane wyniki osiągnięć edukacyjnych uczniów kolorystycznie podzielono na dwie połowy. W górnej połowie tabeli będą znajdować się najlepsze i dobre wyniki testowania, a w dolnej – przeciętne i najłabsze. Jeśli liczba uczniów biorących udział w testowaniu jest nieparzysta i uniemożliwia podzielenie wyników testowania na dwie połowy, to wynik środkowy pomijamy (wynik ucznia 4).

⁴ Wiertła piórkowe. Zaczepnięte 7 lutego 2015. Strona internetowa <http://www.dobrenarzedzia.waw.pl/?wiertla-piorkowe-9529,194>

⁵ A. Górecki (1984, s. 235).

Tabela 3.

Surowe wyniki badań wiedzy i umiejętności uczniów z rysunku technicznego w zakresie stosowania zasad wymiarowania

Zadanie/ uczeń	1	2	3	4	5	6	7	8	9	10	Suma pkt.
	Zasady wymiarowania										
9	1	1	1	1	0	0	1	1	0	1	7
5	1	1	1	1	0	0	1	1	0	0	6
6	0	1	1	1	0	0	1	1	0	0	5
8	1	0	1	1	1	0	1	0	0	0	5
2	–	1	0	1	0	1	0	1	0	0	4
4	1	1	0	1	0	1	0	0	0	0	4
10	0	1	0	1	0	1	0	0	0	1	4
7	0	1	0	1	0	1	0	0	0	0	3
3	0	1	0	1	0	1	0	0	0	0	3
11	0	0	1	0	0	1	0	0	0	0	2
1	0	1	0	1	–	0	0	0	0	0	2
f	0,09	0,00	0,00	0,00	0,09	0,00	0,00	0,00	0,00	0,00	
p	0,36	0,81	0,45	0,90	0,09	0,54	0,36	0,36	0,00	0,18	
p_w	0,60	0,80	0,80	1,00	0,20	0,20	0,80	0,80	0,00	0,20	
p_n	0,00	0,80	0,20	0,80	0,00	0,80	0,00	0,00	0,00	0,20	
D_{50}	0,60	0,00	0,60	0,20	0,20	-0,60	0,80	0,80	0,00	0,00	

Źródło: opracowanie własne.

Ilościową analizę zadań testowych przeprowadzamy w oparciu o uzyskane dane empiryczne. Na przykładzie wyników badań wiedzy i umiejętności uczniów z rysunku technicznego, zgromadzonych w tabeli 3, zweryfikujemy trzy parametry podstawowe: frakcję opuszczeń zadań, trudność (łatwość) zadania oraz moc różnicującą zadania.

1. Frakcję opuszczeń zadania f obliczamy według wzoru I.1.

Dla pierwszego zadania f_1 wynosi 0,09.

$$f_1 = \frac{n_o}{n} = \frac{1}{11} = 0,09$$

Dla drugiego f_2 równa się 0.

$$f_2 = \frac{n_o}{n} = \frac{0}{11} = 0,00$$

W podobny sposób obliczamy f_n dla pozostałych zadań. Uzyskane wyniki wpisujemy do tabeli.

Zgromadzone dane wskazują, że frakcja opuszczeń zadań nie przekracza wartości 0,15. Dokładne obliczenia pokazały, że waha się ona w granicach od 0,00 do 0,09. Bardzo niska frakcja opuszczeń zadań (mniejsza niż 0,15) świadczy o tym, że zadania ułożono poprawnie dydaktycznie.

2. Dla każdego zadania obliczamy łatwość zadania p . Stosujemy wzór I.4, ponieważ zadania testowe oceniano w skali 1–0.

Dla pierwszego zadania p_1 (łatwość zadania nr 1) równa się wartości 0,36.

$$p_1 = \frac{n_p}{n} = \frac{4}{11} = 0,36$$

W podobny sposób obliczamy wskaźnik p_n dla pozostałych zadań.

Dostrzegamy, że wskaźniki łatwości zadań wyrażane są wartościami liczbowymi: 0,00; 0,09; 0,18; 0,36; 0,45; 0,54; 0,81 oraz 0,90. Niektóre uzyskane wskaźniki przebiegają wartości graniczne (0,00; 0,09; 0,18 oraz 0,90), co sugeruje, że zadanie 4 jest zbyt łatwe, a zadania 5, 9 i 10 zbyt trudne dla niektórych uczniów. Pod kątem łatwości mogą one być ułożone dydaktycznie niepoprawnie. Pozostałe zadanie ocenia się jako trudne (0,36 i 0,45), umiarkowanie trudne (0,54) i łatwe (0,81).

Zgodnie z twierdzeniem naukowców

zadania, które są zbyt trudne, wywołują silne napięcie (stres), na którym uczniowie bardziej się koncentrują niż na pracy. Efekt ten wynika z faktu, że wielkość motywu osiągnąć można traktować w pewnym sensie jako różnicę pomiędzy dążeniem do sukcesu a lękiem przed porażką. Zbyt trudne zadania powodują, że od dążenia do sukcesu silniejszy staje się lęk przed niepowodzeniem, co uniemożliwia skuteczne działanie. Oznacza to, że optymalnym poziomem trudności zadania jest poziom średni. Wówczas jest on na tyle wysoki, że utrzymuje konieczny poziom napięcia, ale nie jest ono na tyle duże, aby dezorganizować działanie (Chomczyńska-Rubacha, 2008, s. 265).

Analiza wyników zadań testowych pod kątem łatwości wskazuje, że zbyt trudne zadania (zadania 5, 9, 10), podobnie jak zbyt łatwe zadania (zadanie 4) należałoby poprawić (zob. interpretacja wskaźnika p w tabeli 1).

3. Trzecim parametrem ilościowej analizy testu jest *moc różnicująca zadania*. Do obliczenia mocy różnicującej zadania stosujemy wzór I.6. Wymaga on obliczeń łatwości zadań w górnej (lepszey) i dolnej (słabszej) połowie wyników testu. Górną połowę oznaczono symbolem p_w , dolną – p_n .

Dla pierwszego zadania $p_{w1} = \frac{n_p}{n} = \frac{3}{5} = 0,60$, $p_{n1} = \frac{n_p}{n} = \frac{0}{5} = 0,00$, a połówkowy, zastępczy wskaźnik mocy różnicującej $D_{50} = p_{w1} - p_{n1} = 0,60 - 0,00 = 0,60$. W ten sposób obliczamy dla pozostałych zadań.

Bardzo niskie, zerowe lub ujemne wskaźniki świadczą o tym, że ułożone zadania testowe są bardzo trudne lub bardzo łatwe dla ucznia, a poprawność dydaktyczna jest zbyt niska. Z obliczeń wynika, że zadania 2, 9 i 10 posiadają bardzo niską, a zadanie 6 ujemną moc różnicującą. Pod kątem różnicowania uczniów te zadania mogą być ułożone dydaktycznie niepoprawnie. Szczególną uwagę należałoby zwrócić na zadania 9 i 10, które nie różnicują osiągnięć edukacyjnych, a w dodatku są zbyt trudne dla uczniów.

Warto zaznaczyć, że w danym przypadku powodów niskiej wartości wskaźnika D_{50} może być kilka. Po pierwsze, do analizy testu użyto wskaźnika zastępczego mocy różnicującej, który z reguły jest mniej dokładny. Po drugie, testowanie prowadzono w stosunkowo nielicznej grupie, gdzie poziom wiedzy i umiejętności uczniów jest zazwyczaj wyrównany. Po trzecie, liczba zadań testowych była niewielka (10 zadań), co spowodowało, że zróżnicowanie osiągnięć edukacyjnych uczniów okazało się nieznaczące. Te i inne czynniki w dużym stopniu przesądziły o niskiej mocy różnicującej zadań testowych. Podważyły również poprawność dydaktyczną niektórych zadań testu.

Wnioski

Na tle różnorodnych metod pomiaru dydaktycznego testowanie jest sprawliwą metodą oceniania. Stwarza ono dla wszystkich uczniów jednakowe warunki kontroli oraz eliminuje subiektywizm nauczyciela. W porównaniu do ustnej i pisemnej kontroli osiągnięć edukacyjnych uczniów testowanie zajmuje mniej czasu oraz generuje niższe wydatki, co sprawia, że z ekonomicznego punktu widzenia jest bardziej wydajne. Testy nauczycielskie, jako narzędzie pomiaru dydaktycznego uznawane jest za dokładniejsze narzędzie kontroli. W dodatku, może być narzędziem bardziej rozbudowanym, ponieważ może obejmować zadania z wielu tematów lekcyjnych (Kuchar i Sergienko, 2010).

Oprócz zalet testy nauczycielskie posiadają także pewne wady. Do najważniejszych niedostatków testowania zaliczono m.in. to, że testy negatywnie wpływają na kontrolę wiedzy uczniów, ponieważ zawężają funkcję kontroli w procesie

nauczania. Ograniczają one do minimum różnorodność danej formy weryfikacji osiągnięć edukacyjnych wychowanków. Poza tym, zadania testowe powodują trudności psychologiczne u wielu uczniów, z wyjątkiem tych, którzy mają teoretyczny typ myślenia, łatwo uczą się nowego sposobu działania oraz są w stanie przenieść to, czego się nauczyli do nowej sytuacji (Wychruszcz i in., 2018).

Pomimo niektórych wad testowanie oparte na technologii informacyjno-komunikacyjnej wkracza do grona podstawowych metod sprawdzania i oceniania wiedzy i umiejętności wychowanków. Zaistniała sytuacja pobudza wielu nauczycieli do konstruowania testów samodzielnie. Jednocześnie wymaga od nich poznania oraz przestrzegania ogólnych zasad konstrukcji zadań testowych. Innymi słowy, motywuje nauczycieli do zdobywania wiedzy z zakresu technologii testowania. Oprócz wiedzy ważna dla nauczyciela staje się także umiejętność przeprowadzenia analizy (jakościowej i ilościowej) zadań testowych. Dana analiza warunkuje poprawność dydaktyczną testu, a co za tym idzie – obiektywną kontrolę wiedzy i umiejętności podopiecznych.

Przeprowadzając analizę zadań testowych, „nauczyciel może też wnioskować o efektach swojej pracy i próbować odpowiedzieć na pytanie, czy wyniki testu są zgodne z jego oczekiwaniami. Może więc ocenić między innymi skuteczność stosowanych przez siebie metod nauczania, środków dydaktycznych” (Szaran, 2000, s. 45).

Dodam, że sprawdzone (pod kątem dydaktycznym) testy nauczycielskie oraz te zamieszczone na internetowych stronach edukacyjnych mogą być stosowane w praktyce pedagogicznej powszechnie (przez innych nauczycieli) i tym samym usprawniać i ułatwiać proces weryfikacji wiedzy i umiejętności uczniów.

Bibliografia

- Bulakh I.E. i Mruga M.R. (2006). *Stworzujemy jakisnyj test*. Kijów: Majster.
- Bykow W.Yu., Bogachkow Yu.M. i Żuk Yu.O. (2008). *Monitorynh równia nawczalnych dosiahneń z wykorzystanniam Internet-technologij*. Kijow: Pedahohiczna dumka.
- Chomczyńska-Rubacha M. (2008). Szkolne środowisko uczenia się. W: Z. Kwieciński i B. Śliwerski (red.), *Pedagogika 2. Podręcznik akademicki*. Warszawa: Wydawnictwo Naukowe PWN.
- Ciechanowicz A. (red.) (1990). *Kulturowa adaptacja testów*. Warszawa: Polskie Towarzystwo Psychologiczne.
- Górecki A. (1984). *Technologia ogólna: podstawy technologii mechanicznych*. Warszawa: Wydawnictwa Szkolne i Pedagogiczne.

- Ioannou-Georgiou S. i Pavlou P. (2003). *Assessing young learners*. Oxford: Oxford University Press.
- Kaniweć T.M. (2012). *Osnovy pedahohicznoho ociniuwannia*. Niżyn: PP Łysenko M.M.
- Kuchar L.O. i Sergienko W.P. (2010). *Konstrujuwannia testiw*. Łuck.
- Kuchar L.O. i Sergienko W.P. (2011). *Metodyczni rekomendaciji zi skladannia testowych zawdań*. Kijów: NPU.
- Liaszenko O.I., Łukina T.O., Bulakh I.E., Mruga M.R. (2012). *Metodyka i tehnolohiji ociniuwannia dijalności zahalnooświtnioho nawczalnoho zakładu*. Kijow: Pedahohiczna dumka.
- Magnusson D. (1981). *Wprowadzenie do teorii testów*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Niemierko B. (1999). *Pomiar wyników kształcenia*. Warszawa: Wydawnictwa Szkolne i Pedagogiczne.
- Niemierko B. (1975). *Testy osiągnięć szkolnych. Podstawowe pojęcia i techniki obliczeniowe*. Warszawa: Wydawnictwa Szkolne i Pedagogiczne.
- Projekt edukacyjny „Na urok”. Zaczepnięte 19 lipca 2020. Strona internetowa <https://naurok.com.ua/>
- Ruchała F. (2008). *Pomiar sprawdzający w kształceniu zawodowym*. Warszawa: CKE.
- Szaran T. (2000). *Pomiar dydaktyczny*. Warszawa: Wydawnictwa Szkolne i Pedagogiczne.
- Walczak W. (2001). *Jak oceniać ucznia?*. Łódź: Galaktyka.
- Wiertła piórkowe. Zaczepnięte 7 lutego 2015. Strona internetowa <http://www.dobrenarzedzia.waw.pl/?wiertla-piorkowe-9529,194>
- Wojewódzki Urząd Pracy (2014). *Projekt Modernizacja oferty kształcenia zawodowego w powiązaniu z potrzebami lokalnego/regionalnego rynku pracy*.
- Wychruszcz W.O., Kozłowskyj Yu.M. i Kowalczyk L.I. (2018). *Metodologija ta metodyka naukowoho doslidżennia*. Lwiw: Politechnika Lwiwska.