

OCENA EGZAMINU TESTOWEGO

parametry statystyczne

I. Wskaźnik trudności zadania T

Interpretacja

wartość wskaźnika T	stopień trudności zadania
0,000 – 0,250	bardzo trudne
0,251 – 0,400	trudne
0,401 – 0,750	umiarkowanie trudne
0,751 – 0,900	łatwe
0,901 – 1,000	bardzo łatwe

Optymalny układ zadań w teście dydaktycznym charakteryzuje się udziałem 27% zadań trudnych ($T < 0,400$), 46% zadań umiarkowanie trudnych ($0,401 < T < 0,750$) i 27% zadań łatwych i bardzo łatwych ($T > 0,750$).

W teście złożonym z zadań wielokrotnego wyboru z czterema lub pięcioma odpowiedziami, punktowanych 1 – 0, szansa zgadnięcia poprawnej odpowiedzi wynosi odpowiednio 0,250 i 0,200. Oznacza to, że w takich zadaniach praktycznie nie ma różnicy pomiędzy zgadywaniem, a znajomością odpowiedzi. Dlatego zadania o wskaźniku trudności T poniżej 0,250 raczej nie powinny być stosowane.

Tak niska wartość wskaźnika T może być skutkiem wadliwej konstrukcji zadania, niewłaściwego zrozumienia treści nauczania w trakcie procesu dydaktycznego, a także błędnego zaznaczenia poprawnej odpowiedzi przez autora testu.

Zadania, których wartość wskaźnika $T \leq 0,250$ muszą podlegać wnikliwej ocenie merytorycznej.

II. Moc różnicująca zadania testowego

W pomiarze dydaktycznym mocą różnicującą zadania, określa się związek tego zadania z wynikiem testu, w którym się ono mieści. Korelację wyniku zadania z wynikiem testu opisują dwa wskaźniki:

- wskaźnik mocy dyskryminacyjnej D

- wskaźnik korelacji punktowo-dwuseryjnej (RPBi)

Wskaźniki te przyjmują wartość w zakresie od -1 do +1.

Maksymalna wartość wskaźnika RPBi zależy od trudności zadania; wyniki najbardziej wiarygodne uzyskuje się przy wskaźniku trudności T około 0,500. Im większa jest różnica pomiędzy wyliczonym dla pytania wskaźnikiem trudności T, a wartością 0,500, tym mniej dokładna jest ocena trafności pytania przy pomocy współczynnika RPBi. Wartość wskaźnika mocy dyskryminacyjnej D jest tylko w nieznacznym stopniu zależna od trudności zadania.

Interpretacja

Zadania w których obydwa wskaźniki (RPBi i D) mają wartości ujemne, należy poddać wnikliwej analizie formalnej.

Jeżeli do wyboru są zadania o podobnej stosowności i poprawności dydaktycznej przy zbliżonych wartościach wskaźnika T, należy wybierać zadania o wyższych wskaźnikach mocy różnicującej D.

III. Rzetelność testu

Jest miarą wskazującą na zgodność, z jaką powtórzyłyby się wyniki w przypadku ponownego zastosowania testu. Określa ona również wewnętrzną zgodność testu, czyli stopień w jakim wszystkie zadania testu mierzą te same osiągnięcia zdających.

Wskaźnik rzetelności Kuder-Richardsona - KR20

Interpretacja

Im wyższy jest wskaźnik rzetelności, tym precyzyjniejszy jest instrument pomiarowy, czyli analizowany test. KR20 przyjmuje wartości w przedziale od 0 do 1,000, przy czym minimalna dopuszczalna wartość wynosi 0,700, a optymalna, świadcząca o bardzo wysokiej rzetelności testu $\geq 0,800$.

Średni wskaźnik mocy dyskryminacyjnej testu D_t

Jest średnią wartością wskaźników D zadań w teście. Zadania o ujemnej wartości D obniżają wartość D_t podobnie, jak zadania o niskiej wartości dodatniej.

IV. Wskaźnik trudności testu T

wartość wskaźnika T	stopień trudności testu
$<0,400$	test trudny
$\geq 0,400, <0,750$	test umiarkowanie trudny
$\geq 0,750$	test łatwy