

Test teoretyczny

Charakterystyka

Test teoretyczny polega na zadawaniu kandydatowi pytań lub wykonywaniu przez niego określonych zadań w zakresie wiedzy i umiejętności kognitywnych oraz na analizie rezultatów tych działań.

Testy teoretyczne przebiegają w warunkach kontrolowanych. Za ich pomocą można sprawdzić jednocześnie wiele osób. Wyniki testów teoretycznych łatwo odnieść do zewnętrznych standardów. Jednocześnie same podlegają standaryzacji. Oznacza to, że istnieje ujednolicony sposób posługiwania się nimi (obejmujący m.in. kolejność i treść pytań, które są zazwyczaj identyczne dla wszystkich kandydatów), a ich trafność (weryfikowane są te efekty uczenia się, które wybrano do sprawdzenia) i rzetelność (wyniki są powtarzalne niezależnie od czasu i miejsca) została sprawdzona na większych grupach osób. Testy zasadniczo powinny mierzyć jakość wiedzy, a nie jej ilość.

Test teoretyczny, zwany czasem testem wiedzy, jest jedną z najczęściej stosowanych metod weryfikacji. Test taki nie powinien być jedynie testem wiedzy zapamiętanej, ale przede wszystkim narzędziem pozwalającym ocenić umiejętności posługiwania się tą wiedzą. Pozwala na zastosowanie jasnych kryteriów oceny, a konstrukcja narzędzi i organizacja weryfikacji zapewniają bezstronność i rzetelność sprawdzania.

Dodatkowo specyfika teoretycznego testu pisemnego sprawia, że kandydat może pozostać anonimowy, a jego dane osobowe można zakodować w celu zwiększenia obiektywizmu oceny. Test pisemny powinien się składać z dużej liczby zadań obejmujących szeroki zakres wybranych zagadnień dotyczących kwalifikacji.

Testy ustne mogą (ale nie muszą) mieć wystandaryzowane pytania, identyczne dla wszystkich kandydatów. Ich zaletą jest możliwość zbadania głębi rozumienia danego zagadnienia przez kandydata.

Testy teoretyczne mogą przyjmować różne formy. Przedstawia je poniższa tabela.

Tabela 1. Rodzaje testów teoretycznych

Czynnik podziału	Rodzaje testów
Typ czynności wykonywanej przez osobę przystępującą do testu*	<ul style="list-style-type: none"> • Testy pisemne – wymagają od kandydatów wypowiedzi na piśmie. • Testy ustne – rozwiązywane przez kandydatów w formie słownej odpowiedzi.
Zasięg stosowania	<ul style="list-style-type: none"> • Testy standaryzowane (szerokiego użytku) – stosowane głównie w edukacji formalnej i pozaformalnej. Nad ich opracowaniem czuwają zespoły ekspertów, a ich stosowanie jest poprzedzone procedurą standaryzacji. Ten rodzaj testów powinien mieć instrukcję rozwiązywania dla kandydatów, karty odpowiedzi i podręcznik testowania dla asesorów, w którym umieszczone są dane techniczne testowania oraz normy dla poszczególnych ocen. • Testy niestandaryzowane – opracowywane na własny użytek przez asesora. Nie są poddawane procedurze standaryzacji. Wyników takich testów nie można

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

	porównywać pomiędzy poszczególnymi asesarami/ośrodkami.
Cel pomiaru	<ul style="list-style-type: none"> • Testy sprawdzające – ich wyniki pozwalają na ustalenie stopnia opanowania przez kandydatów efektów uczenia się. Na ich podstawie można stwierdzić, jakie czynności/umiejętności sprawiają osobom zdającym największe trudności i wysnuć wnioski na temat dalszego uczenia się. Szczególną odmianą są testy sprawdzające wielostopniowe, w których określone grupy zadań reprezentują kryteria weryfikacji przypisane poszczególnym ocenom. • Testy różnicujące – treść zadań testowych jest dobierana na podstawie aktualnych wyników innych osób. Wykorzystywane są przeważnie w selekcji kandydatów na określone stanowisko lub w badaniach naukowych.
* Ten czynnik jest podstawą podziału testów teoretycznych w Katalogu metod stosowanych w walidacji efektów uczenia się.	

Testy przeznaczone dla większej liczby zdających powinny być narzędziami standaryzowanymi. Innymi słowy, powinny pozwalać w obiektywny sposób zmierzyć poziom efektów uczenia się, a ich trafność i rzetelność została wcześniej sprawdzona i potwierdzona w badaniach na dużej liczbie osób.

Asesor powinien przestrzegać instrukcji dotyczących wypełniania narzędzi służących do oceny odpowiedzi.

Jeśli w testach stosuje się pytania otwarte, wówczas należy uważnie analizować odpowiedzi. Przy odpowiedziach ustnych zachodzi niebezpieczeństwo, że informacje zawarte w wypowiedziach kandydata będą zredukowane. Można temu zapobiec m.in. poprzez przygotowanie przez asesora oceny opisowej, w której znajdują się fragmenty odpowiedzi, które świadczą o osiągnięciu danych efektów.

Etapy walidacji, na których rekomenduje się stosowanie danej metody

Z raportów Cedefopu (agencji zajmującej się rozwojem i promocją edukacji zawodowej i uczenia się przez całe życie w Unii Europejskiej) wynika, że metodę testów teoretycznych stosuje się głównie na etapie weryfikacji. Tak dzieje się m.in. w Czechach, Estonii, Flandrii, Francji, Portugalii, Wielkiej Brytanii i w Republice Południowej Afryki.

Zakres efektów uczenia się, które można sprawdzać daną metodą

Wskazówki dotyczące efektów uczenia się, które można sprawdzać za pomocą testu teoretycznego, mają charakter orientacyjny. Zostały opracowane na podstawie analizy wymagań zawartych w Polskiej Ramie Kwalifikacji (w charakterystykach II stopnia typowych dla kwalifikacji o charakterze zawodowym), dotyczących efektów uczenia się.

Należy pamiętać, że wybór danej metody zawsze musi wynikać z analizy efektów uczenia się zdefiniowanych dla [danej kwalifikacji](#).

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

Wiedza

Za pomocą testu teoretycznego można sprawdzać wiedzę, zwłaszcza faktograficzną, ze wszystkich poziomów PRK.

Umiejętności

Metoda jest szczególnie przydatna do sprawdzania efektów uczenia się w zakresie:

- informacji („obliczenia, analiza, synteza i prognozowanie”) na poziomach 2–4 PRK (dotyczących wyszukiwania, porównywania, oceniania oraz analizowania lub przetwarzania informacji potrzebnych do wykonywania działań zawodowych); na poziomach 5–6 PRK (dotyczących dokonywania analizy prowadzonej działalności zawodowej z uwzględnieniem dostępnych danych ilościowych); na 7 poziomie PRK (dotyczącym prognozowania rozwoju sytuacji w dziedzinie działalności zawodowej) i na 8 poziomie PRK (dotyczącym tworzenia modeli rozwoju w dziedzinie działalności zawodowej i wnioskowania na ich podstawie);
- organizacji pracy („planowanie i korygowanie planów” oraz „korygowanie działań”) na poziomach 3 i 4 PRK;
- narzędzi i materiałów na poziomach 4–6 PRK;
- uczenia się i rozwoju zawodowego („wspieranie rozwoju innych osób”) na 7 poziomie PRK.

Kompetencje społeczne

Test teoretyczny można zastosować do sprawdzenia kompetencji społecznych w zakresie:

- przestrzegania reguł na poziomach 1–3 PRK (z wyjątkiem efektów uczenia się dotyczących rzetelnego informowania o rezultatach prac i dotrzymywania uzgodnień);
- odpowiedzialności na poziomach 2–5 PRK (dotyczącej uwzględniania różnego rodzaju skutków działań zawodowych).

[Wskazówki dotyczące efektów uczenia się, które można sprawdzić za pomocą metody Test teoretyczny](#)

Mocne i słabe strony testu teoretycznego

Mocne strony	Słabe strony
<ul style="list-style-type: none">• metoda powszechnie znana i oceniana jako trafna i rzetelna• tania i wygodna do zastosowania w porównaniu z innymi metodami• wyniki łatwo odnieść do zewnętrznych standardów• wszyscy zdający dostają takie same pytania• zachowana jest obiektywność oceny (jednakowe kryteria)• pozwala na różnorodność formalną i merytoryczną pytań; duża liczba pytań umożliwia objęcie wielu efektów uczenia się• łatwa do oceny• zapewnia powtarzalność wyników• metoda mało podatna na tzw. efekt asesora (wpływ subiektywnej oceny asesora na wyniki,	<ul style="list-style-type: none">• metoda nieodpowiednia do sprawdzenia wielu umiejętności (zwłaszcza praktycznych) i kompetencji społecznych• gotowych testów nie można dowolnie dostosowywać do indywidualnych potrzeb kandydatów• konstrukcja zadań może mieć wpływ na jakość testu (np. większość zadań może sprawdzać tylko odtwarzanie wiedzy a nie rozwiązywanie problemów)• sposób formułowania pytań może mieć wpływ na wyniki (interpretacja i klucz odpowiedzi)• istnieje ryzyko nadmiernego upraszczania i pytania o rzeczy błęde• ma ograniczoną liczbę pytań

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

zwłaszcza w testach pisemnych, można w dużym stopniu ograniczyć)

Ograniczenia w stosowaniu

Ograniczeniem w stosowaniu testu teoretycznego jest przede wszystkim rodzaj efektów uczenia się, które wybrano do sprawdzenia. Metoda ta jest przeznaczona głównie do weryfikowania wiedzy i umiejętności kognitywnych, natomiast zakres umiejętności praktycznych i kompetencji społecznych, które można zweryfikować za jej pomocą jest dość ograniczony.

Niepożądany wpływ na przebieg testu i końcowe wyniki kandydata mogą mieć różne czynniki. Trudności z wykazaniem efektów uczenia się za pomocą tej metody mogą mieć na przykład osoby o niskich umiejętnościach werbalnych i pisanie tekstów, a także takie, które wcześniej wypadły z systemu edukacji formalnej, mające trudności z rozwiązywaniem testów oraz osoby ze specjalnymi potrzebami.

Konieczne zasoby kadrowe i organizacyjno-materialne

Wymagane zasoby kadrowe i organizacyjno-materialne są uzależnione od celu testu i liczby osób sprawdzanych jednocześnie za pomocą tej metody.

Należy przewidzieć udział w walidacji osób przygotowujących testy i asesorów, a także koszty ich zatrudnienia lub przeszkolenia.

Aby móc zastosować metodę testu teoretycznego niezbędne są zasoby organizacyjno-materialne, obejmujące m.in.:

- przygotowanie testów, w tym ich standaryzację (jeśli uznano to za stosowne);
- materiały konieczne do przeprowadzenia testu (np. kartki papieru, stanowiska komputerowe);
- pomieszczenia, w których przeprowadzany jest test.

Koszty przygotowania i przeprowadzenia testu mogą się wahać od kilkuset złotych (np. druk materiałów egzaminacyjnych) do nawet kilkudziesięciu milionów – szczególnie kosztowne jest stworzenie testów pisemnych standaryzowanych na dużej próbie.

Institucje certyfikujące powinny liczyć się z koniecznością zaangażowania dodatkowych zasobów w przypadku kandydatów z niepełnosprawnością słuchu lub wzroku. Należy zadbać o odpowiednią organizację testu i przygotowanie dostosowanych materiałów (np. arkuszy pytań o dużej czcionce). Przeprowadzenie testu może wymagać dodatkowych umiejętności zarówno od asesorów, jak i od kandydatów (np. znajomość języka migowego, alfabetu Braille'a).

Możliwości łączenia testu teoretycznego z innymi metodami

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

W Polsce, zarówno w edukacji formalnej, jak i pozaformalnej, testy teoretyczne są często łączone z [obserwacją w warunkach symulowanych](#) (zadania praktyczne). Ponadto, test może uzupełniać bądź zastępować wywiad [swobodny i ustrukturyzowany](#).

Możliwe jest także uzupełnienie testu w formie odpowiedzi ustnej o obserwację w [warunkach rzeczywistych](#), [prezentację](#), [debate swobodną](#) oraz [analizę dowodów i deklaracji](#).

Przykładowe techniki

Technika to podejście w ramach danej metody wykorzystywane do zbierania i analizowania danych świadczących o posiadaniu przez daną osobę efektów uczenia się.

Odpowiedź ustna

Odpowiedź ustna przyjmuje najczęściej formę serii pytań zadawanych kandydatom. Umożliwia sprawdzenie przede wszystkim głębi rozumienia danego zagadnienia i umiejętności rozwiązywania problemów. Pozwala także na sprawdzenie wiedzy faktograficznej oraz umiejętności i kompetencji społecznych.

Odpowiedź ustna zazwyczaj nie jest ustrukturyzowana, a asesor na bieżąco dostosowuje pytania do odpowiedzi kandydata. Ta cecha upodabnia ją do [wywiadu swobodnego](#) (technika – rozmowa). Prowadzący wywiad posługuje się tzw. dyspozycjami do wywiadu, czyli ustalonym celem oraz luźno sformułowanymi tematami, które może omówić z kandydatem w dowolny sposób. Jednak wywiad swobodny nie jest z reguły ukierunkowany na sprawdzenie wiedzy, jak ma to miejsce w przypadku odpowiedzi ustnej.

Brak określonej struktury odpowiedzi ustnej sprawia, że nie można dokładnie przewidzieć przebiegu walidacji za pomocą tej techniki.

Rzetelność i trafność odpowiedzi ustnej zależy od jakości i formy pytań (np. zadawanie pytań problemowych pozwala na sprawdzenie umiejętności stosowania zdobytych dotychczas wiadomości w nowych sytuacjach oraz logicznego myślenia). Powinny one w jasny i precyzyjny sposób sprawdzać opanowanie wymaganych efektów uczenia się. Mogą dotyczyć rozumienia, uzasadniania, ustalania przyczyn i skutków. Forma pytań powinna być różnorodna, a odpowiedzi zwięzłe i logiczne.

Należy pamiętać, że gdy ocenie podlega wypowiedź kandydata, wszelkie zaburzenia w rozmowie (np. zmiana tematu), mogą mieć wpływ na ostateczną ocenę. Jest to szczególnie szkodliwe, gdy nowy wątek dotyczy efektów uczenia się, które nie były ujęte w scenariuszu walidacji. Asesor, który postępuje w ten sposób, nierzetelnie wykonuje swoje obowiązki. Jednak nawet w przypadku mniej istotnych zaburzeń, efekt takiego postępowania może np. zmienić atmosferę sprawdzania. W rezultacie kandydat może przestać odpowiadać tak szczegółowo lub szczerze. Rodzi to też stres u osoby podchodzącej do walidacji, co może skutkować gorszym wynikiem.

Można wyróżnić kilka typów odpowiedzi ustnych. Poniżej przedstawiono trzy z nich:

1. Viva voce

Jest to odpowiedź ustna na pytania zadawane przez komisję egzaminacyjną złożoną z zewnętrznego i wewnętrznego asesora. Technika viva voce jest stosowana głównie na poziomie uniwersyteckim (podyplomowym), po przedłożeniu pracy końcowej. Celem jest upewnienie się, że kandydat wie wystarczająco dużo na dany temat oraz że jest to jego własna praca.

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

2. Test ustny stanowiący uzupełnienie obserwacji w warunkach symulowanych (symulacji)

Jest to odpowiedź ustna na pytania zadawane przez asesora po wykonaniu przez kandydata zadania praktycznego. Stanowi część sesji podsumowującej. Zazwyczaj nie ma formalnej struktury, ale asesory zazwyczaj zadają zaplanowane wcześniej pytania. Możliwe jest także zadanie pytań ogólnych, na które odpowiadają wszyscy kandydaci w czasie wykonywania zadania praktycznego.

3. Rozumienie ze słuchu w teście językowym

Jest to test ukierunkowany na ocenę poziomu biegłości posługiwania się przez kandydata danym językiem. Test rozumienia ze słuchu jest często przeprowadzany z wykorzystaniem urządzeń odtwarzających pliki dźwiękowe. Powinny być one dopasowane do poziomu biegłości zdającego w zakresie rozumienia ze słuchu.

Ramka 1. Etapy odpowiedzi ustnej

Przebieg odpowiedzi ustnej zależy od jej typu, można jednak wyróżnić następujące etapy:

1. Pytanie wprowadzające, dające możliwość spontanicznej odpowiedzi na wybrany temat, przeanalizowania jakiejś kwestii.
2. Pytania pogłębiające, służące zaangażowaniu kandydata w rozmowę, np. „Skąd to wiedziałeś?”, „Jakiej metody użyłeś, aby dojść do tego wniosku?”.
3. Pytania naprowadzające, w których asesor może dać zdającemu pewne wskazówki, mogące naprowadzić go na właściwą odpowiedź, np. „Pamiętasz fakt xx?”, „Z czym jest powiązany?”.
4. Pytania problemowe, służące zmierzeniu głębi rozumienia zagadnień, np. „Czy możesz udowodnić, że dana metoda jest lepsza od innej?”.

Mocne strony odpowiedzi ustnej:

- gwarantuje natychmiastowy wynik (asesory od razu otrzymują odpowiedzi) i daje możliwość bezpośredniej obserwacji reakcji kandydata;
- daje możliwość sformułowania dodatkowych pytań, prześledzenia poprawności toku myślenia, zapytania kandydata, w jaki sposób doszedł do odpowiedzi;
- daje możliwość bezpośredniego poznania kandydata i ewentualnego ukierunkowania jego dalszej ścieżki edukacyjnej;
- dobrze uzupełnia zadania praktyczne;
- umożliwia ocenę umiejętności poprawnego, zwięzłego i logicznego wypowiedzania się kandydata;
- pozwala na poznanie cech osobowościowych kandydata;
- daje kandydatowi swobodę w formułowaniu odpowiedzi;
- uniemożliwia oszukiwanie.

Słabe strony odpowiedzi ustnej:

- w zależności od liczby kandydatów sprawdzanych jednocześnie, zastosowanie tej techniki może być czasochłonne i tym samym zwiększać koszty walidacji;
- wymagane są jasne kryteria oceniania, zmniejszające wpływ opinii asesora na wyniki kandydata;
- ocena jest narażona na wpływ subiektywnych czynników zewnętrznych i postronnych, co stwarza trudność w zapewnieniu obiektywnego, rzetelnego i trafnego sprawdzania;
- odpowiedź ustna może sprawiać trudność obcokrajowcom, którzy nie znają dobrze języka, w jakim ma być prowadzona weryfikacja;
- ocena odpowiedzi ustnej powinna być uzupełniona o raport z przebiegu weryfikacji, w celu zwiększenia jej rzetelności;
- odpowiedzi udzielane w czasie ustnego testu powinny być nagrywane lub dokumentowane, aby umożliwić procedurę odwoławczą lub rewizję oceny w przyszłości;
- nie zawsze da się przewidzieć długość odpowiedzi ustnej udzielanej przez kandydata;

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

- brak powtarzalności wyników;
- brak możliwości wykorzystania wyników do celów statystycznych.

Test pisemny

Test teoretyczny pisemny jest zbiorem zadań do rozwiązania, reprezentujących wybrany zakres poznawczej treści uczenia się. Wykonuje się go zazwyczaj podczas pojedynczej sesji.

Czynności osoby wykonującej test mogą obejmować:

- odtwarzanie wiadomości z pamięci,
- rozumienie,
- rozwiązywanie znanego problemu,
- rozwiązywanie nowego problemu,
- krytyczną ocenę sytuacji,
- syntezę wiedzy.

Narzędziem testów pisemnych jest zadanie testowe. To względnie niezależny, osobno punktowany element, którego rozwiązanie polega na wybraniu odpowiedzi spośród zaproponowanych lub udzieleniu własnej odpowiedzi.

Szczególnym rodzajem testu teoretycznego pisemnego jest studium przypadku. Ta technika może być stosowana podczas obserwacji w warunkach symulowanych (symulacji) oraz w teście teoretycznym pisemnym.

Ramka 2. Studium przypadku

„Studium przypadku” ma wiele definicji. Pojęcie to odnosi się m.in. do metod badawczych, stosowanych w naukach społecznych oraz do metod dydaktycznych. Z tego powodu może być różnie rozumiana przez poszczególne osoby.

W odniesieniu do walidacji studium przypadku jest techniką sprawdzania efektów uczenia się, która polega na wykonaniu zadania lub udzieleniu odpowiedzi na pytania na podstawie informacji zawartych w opisie konkretnej sytuacji (prawdziwej albo wymyślonej). Wynikiem studium przypadku może być m.in. diagnoza danego zdarzenia, propozycja rozwiązania problemów przedstawionych w opisie przypadku, prognoza dalszego rozwoju sytuacji czy propozycja działania.

Za pomocą studium przypadku można sprawdzić m.in. umiejętność analizy informacji, wyciągania wniosków, krytycznego myślenia oraz zastosowania wiedzy w praktyce. Kandydat może na przykład określić przyczyny danej sytuacji, wskazać główne problemy poruszone w opisie przypadku, wskazać zależności pomiędzy faktami, przewidzieć lub ocenić skutki opisanych działań/zdarzeń.

Przykłady zastosowania testu teoretycznego

Przykłady zagraniczne

Przykłady na zastosowanie testu teoretycznego są zawarte w [Bazie Dobrych Praktyk. Walidacja, zapewnianie jakości walidacji i certyfikowania](#).

Metoda ta jest stosowana m.in. w wypadku kwalifikacji:

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

- [opiekuna dziecięcego](#) w Estonii – technika testu pisemnego,
- [asystenta administracyjnego](#) we Włoszech - technika odpowiedzi ustnej.

Przykład polski

Poniższy przykład opracowano na podstawie poradnika dla autorów zadań i egzaminów testowych, przygotowanego przez Henryka Rebandela.

Nazwa instytucji	Warszawski Uniwersytet Medyczny Zakład Dydaktyki i Efektów Kształcenia Wydziału Nauki o Zdrowiu WUM Pracownia Ewaluacji i Doskonalenia Kształcenia Medycznego Biuro ds. Egzaminów Uczelnianych
Kraj pochodzenia	Polska
Kontekst funkcjonowania instytucji	Za kompleksową obsługę egzaminów testowych przeprowadzanych przez WUM odpowiedzialne jest bezpośrednio Biuro ds. Egzaminów Uczelnianych, utworzone 22 maja 2014 roku. Do głównych zadań Biura należy: <ul style="list-style-type: none"> • obsługa testów przeprowadzanych w wersji papierowej przez jednostki naukowo-dydaktyczne WUM, w szczególności: <ul style="list-style-type: none"> ◦ przygotowanie i wydruk kart egzaminacyjnych, ◦ analiza parametrów jakościowych pytań egzaminacyjnych, ◦ archiwizacja danych; • wdrożenie i obsługa uczelnianego systemu egzaminów elektronicznych, w szczególności: <ul style="list-style-type: none"> ◦ koordynacja pracy nad stworzeniem i rozwojem bazy pytań egzaminacyjnych, ◦ analiza parametrów jakościowych pytań egzaminacyjnych, ◦ przeprowadzanie egzaminów elektronicznych oraz nadzór nad ich prawidłowym przebiegiem, ◦ weryfikacja wyników egzaminów elektronicznych; • współpraca z Zakładem Dydaktyki i Efektów Kształcenia oraz jednostkami naukowo-dydaktycznymi WUM w zakresie przygotowywania i przeprowadzania testów.
Nazwa kwalifikacji	doktor nauk medycznych, doktor nauk farmaceutycznych, dyplom ukończenia studiów I stopnia na kierunku techniki dentystyczne i inne kwalifikacje
Krótki opis stosowania danej metody	Test teoretyczny
Możliwości samodzielnego zdobycia informacji o metodzie/z pomocą doradcy	Na początku danego okresu nauczania (zazwyczaj na początku roku akademickiego) studentom udostępniana jest informacja o wymaganiach egzaminacyjnych, formie, objętości i czasie trwania egzaminu, a także o kryterium progowym jego zaliczenia (graniczna liczba punktów).
Czynniki, jakie	Metodę testu wybrano ze względu na to, że pozwala na dogłębne sprawdzenie wiedzy

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

stały za wyborem danej metody w danych okolicznościach

studentów. Umożliwiają to m.in.: duża liczba zadań, identyczne zestawy zadań dla wszystkich zdających, krótki czas egzaminowania (istotny w wypadku grup o dużej liczebności), obiektywna skala ocen (przyznawania punktów).

Dobrze zaprojektowany test pozwala szybko ocenić stopień opanowania efektów uczenia się zakładanych dla danego przedmiotu.

Przebieg

Student przystępuje do testu po identyfikacji danych osobowych na podstawie dokumentu tożsamości. Wejście studenta na salę egzaminacyjną jest odnotowywane na liście obecności, co umożliwia ustalenie, czy po rozwiązaniu testu wszyscy zdający oddali karty egzaminacyjne.

Aby wykluczyć wszelkie pomyłki i oszustwa, karty testowe (egzaminacyjne) są personalizowane (imię, nazwisko, numer albumu) i podpisywane przez osoby zdające. Studenci składają podpisy również na książeczkach testowych. Umożliwia to zidentyfikowanie osób, które nie oddały książeczki testowej po zakończeniu egzaminu.

Książeczki testowe są traktowane jako chroniona prawem autorskim publikacja przeznaczona do użycia w ściśle określonym celu, tj. wyłącznie podczas testu. Ma to zapobiec nieautoryzowanemu ujawnianiu zbioru zadań.

Salę egzaminacyjną powinny być na tyle duże, by umożliwiały zachowanie odpowiednich odstępów pomiędzy wszystkimi osobami przystępującymi do testu. Studenci zajmują kolejne miejsca na sali wg listy, zazwyczaj uporządkowanej alfabetycznie, co zmniejsza prawdopodobieństwo, że obok siebie usiądą osoby „pomagające” sobie w udzielaniu odpowiedzi. Możliwość oszukiwania ograniczana jest również dzięki zapewnieniu odpowiedniej liczby osób pilnujących przebiegu testu, odebraniu kart osobom oszukującym oraz stosowaniu kilku wersji testu, aby sąsiadujące ze sobą osoby miały różne zestawy do rozwiązania.

Osoby przeprowadzające test mają obowiązek poinformować o obowiązujących zasadach przed rozdaniem kart egzaminacyjnych. Powinny przedstawić: czas trwania testu, sposób rozpoczęcia, godzinę i sposób zakończenia, sposób zgłaszania pytań, sposób postępowania w wypadku stwierdzenia błędów na kartach egzaminacyjnych, sposób postępowania w wypadku wcześniejszego rozwiązania testu, zasady korzystania z sanitariatu, zasady dotyczące posiadania telefonów komórkowych, skutki stwierdzenia zachowania niepożądanego.

Faktyczną godzinę rozpoczęcia i zakończenia testu ustala się po rozdaniu książeczek testowych wszystkim zdającym. Osoby przeprowadzające test powinny kilka razy w trakcie trwania egzaminu informować, ile czasu zostało do zakończenia pisania testu. Po upływie czasu wyznaczonego na rozwiązanie testu osoby zdające powinny pozostać na swoich miejscach, dopóki nie zostaną zebrane książeczki testowe wraz z kartami egzaminacyjnymi.

Przy ustalaniu czasu przeznaczonego na rozwiązanie testu należy wziąć pod uwagę wszystkie czynności związane z jego rozwiązywaniem, ale nie powinien on być zbyt długi, aby test dobrze różnicował lepiej i gorzej przygotowanych.

Ponadto test:

- może składać się z maksymalnie 100 pytań,
- zadanie testowe może zawierać maksymalnie 5 wariantów odpowiedzi (A, B, C, D, E).

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

	Można przygotować maksymalnie 4 wersje testu.
Wyniki procesu walidacji	<p>Test jest stosowany do weryfikacji efektów uczenia się.</p> <p>Test można zaliczyć albo nie zaliczyć. Nie przewiduje się potwierdzania części efektów uczenia się.</p>
Zasoby organizacyjne i materialne	<p>Do przeprowadzenia testu potrzebna jest sala dopasowana do liczby osób zdających, uwzględniająca konieczność zachowania między nimi odpowiednich odstępów.</p> <p>Metoda ta wymaga wyposażenia w postaci sprzętu komputerowego i oprogramowania do wprowadzania odpowiedzi i obliczenia wyników oraz badania parametrów statystycznych i psychometrycznych. Należy również pamiętać o kosztach związanych z wydrukowaniem odpowiedniej ilości książeczek egzaminacyjnych i spersonalizowanych kart egzaminacyjnych w kilku wersjach.</p> <p>Należy także zadbać o odpowiednią liczbę osób, które będą przeprowadzać test i pilnować jego przebiegu.</p> <p>Powinno się przygotować bank zadań z gotowymi, wystandaryzowanymi pozycjami testowymi. Do kolejnego testu można układać nowe zadania, jednak korzystniej jest wybrać część zadań o znanych parametrach z banku zadań, a tylko pewną część (np. 20%) powinny stanowić zadania nowe.</p> <p>Zalecane jest, by zadania testowe były układane przez przynajmniej kilku specjalistów z danej dziedziny – daje to gwarancję jakości i różnorodności stosowanych narzędzi pomiarowych.</p>
Ograniczenia w stosowaniu	<p>Test nie zawsze jest odpowiednią metodą do badania konkretnych umiejętności praktycznych czy rozwiązywania problemów w sytuacjach nietypowych.</p> <p>Na stworzenie dużego banku zadań potrzeba czasu i doświadczenia ze strony układających zadania testowe.</p>

Efekt asesora

Efekt asesora można zdefiniować jako błędy i różnice w ocenianiu efektów uczenia się w procesie walidacji, które są zależne od asesora i wpływają na wyniki weryfikacji efektów uczenia się.

Efekt asesora można podzielić na następujące kategorie:

Efekt asesora zależny od zastosowanej metody walidacji			Efekt asesora niezależny od zastosowanej metody walidacji (powszechne błędy oceniania)		
Uwarunkowania wewnętrzne efektu asesora	Uwarunkowania zewnętrzne efektu asesora	Wpływ etapu weryfikacji na efekt asesora	Błędy uwarunkowane cechami	Błędy atrybucji	Błędy związane z tematyką oraz formą oceniania

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

			asesora		
Są one specyficzne dla poszczególnych metod. Przykłady można znaleźć tutaj .	Należą do nich: osoba kandydata, czas, miejsce, kultura organizacyjna instytucji, skład komisji oceniającej. Przykłady można znaleźć tutaj .	Inne błędy mogą się pojawić przy przygotowaniu walidacji, w przebiegu weryfikacji oraz przy podjęciu decyzji o wyniku walidacji. Przykłady można znaleźć tutaj .	To czynniki i cechy osoby oceniającej obniżające trafność postrzegania, co w konsekwencji powoduje błędy w ocenianiu. Przykłady można znaleźć tutaj .	Asesor przypisuje większe znaczenie cechom ocenianego niż np. niekorzystnym uwarunkowaniom zewnętrznym. Przykłady można znaleźć tutaj .	Przykłady można znaleźć tutaj .

Minimalizacja efektu asesora ma prowadzić do zwiększenia rzetelności i obiektywności w ocenianiu efektów uczenia. Strategie minimalizowania efektu asesora można znaleźć [tutaj](#).

Błędy specyficzne dla testu teoretycznego

Uwarunkowania wewnętrzne efektu asesora

Ta metoda zakłada większy obiektywizm asesora niż inne.

Asesor badający wyniki testu zawierającego pytania zamknięte nie ma możliwości swobodnej interpretacji lub nadinterpretacji wyników, pod warunkiem precyzyjnego doboru pytań i klucza odpowiedzi.

Ryzyko występowania efektu asesora może być większe w przypadku testów zawierających pytania otwarte; anonimowość gwarantuje wyższy obiektywizm.

Uwarunkowania zewnętrzne efektu asesora

Testy są metodą obiektywnego sprawdzenia efektów uczenia się pod warunkiem prawidłowej konstrukcji merytorycznej i metodologicznej, co wymaga zaangażowania specjalisty w tym zakresie. To może prowadzić do rozbieżności między oczekiwaniami asesora i wymaganiami konstruktora testu oraz przejęcia przez jednego z nich dominującej roli.

Jednakże w dużej mierze adekwatny wynik otrzymuje się jedynie w obszarze wiedzy.

Wpływ etapu weryfikacji na efekt asesora

Przygotowanie walidacji

Wpływ na ocenę mogą mieć m.in.: niedostosowanie zestawu pytań do efektów uczenia się; testy z pytaniami zamkniętymi zawierają wizję twórcy testu na temat efektów uczenia się, która może być rozbieżna od praktyki; testy z pytaniami otwartymi wzmacniają subiektywizm oceny; konieczność dopasowania rodzaju testu (pytania zamknięte, otwarte, pytania mieszane) do potrzeb i możliwości kandydata.

Jeśli kandydat nie ma doświadczenia w rozwiązywaniu tekstów, istnieje duże ryzyko wystąpienia efektu asesora.

Weryfikacja

Wpływ na ocenę mogą mieć m.in.: presja czasu (nieuwzględnienie specyficznych potrzeb kandydatów, np. z niskimi umiejętnościami podstawowymi, z niepełnosprawnością); złe doświadczenia kandydatów z wcześniejszych okresów

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

edukacji.

Decyzja o wyniku walidacji

Testy są skuteczną i obiektywną formą sprawdzenia efektów uczenia się z obszaru wiedzy.

Wybór tej metody powinien być świadomą decyzją – nie odzwierciedli ona całokształtu efektów uczenia się i może stanowić tylko jeden z elementów walidacji.

Wnioski końcowe

Właściwie zaprojektowane testy pozwalają na wiarygodne sprawdzenie wiedzy, a przy odpowiednim doborze pytań – także umiejętności zastosowania wiedzy w praktyce (głównie w obszarze związanym z wykorzystaniem informacji, np. ich analizy, syntezy, obliczenia).

Konstruowanie testów wymaga wysokich kompetencji metodologicznych, co często pociąga za sobą współpracę ze specjalistą w tym zakresie.

Podjęcie decyzji o zastosowaniu tej metody powinno być poprzedzone pozyskaniem wiedzy o preferencjach kandydatów. Mogą to być osoby o niskiej sprawności pisania i czytania, z negatywnymi doświadczeniami z testami lub nie mające wprawy w ich rozwiązywaniu.

Strategia minimalizowania efektu asesora w odniesieniu do testów teoretycznych

Efekt asesora można minimalizować poprzez następujące działania:

- konsultowanie pytań do testu z praktykami w danej dziedzinie, doradcą walidacyjnym;
- udział eksperta w dziedzinie konstruowania testów;
- standaryzacja testu;
- badanie predyspozycji kandydata, np. związanych ze stopniem percepcji czy niepełnosprawności różnego rodzaju i dostosowanie formy przeprowadzenia testu do możliwości i dyspozycyjności kandydata (np. zastosowanie testów online);
- zastosowanie narzędzi statystycznych do oceny testu;
- uwzględnienie opcji anonimowości kandydata i asesora.

Zasady i narzędzia minimalizowania wpływu efektu asesora

Minimalizowanie ryzyka efektu asesora oznacza zapobiegania takim błędom i różnicom popełnianym przez asesora w ocenianiu efektów uczenia się w procesie walidacji, które wpływają na wyniki weryfikacji efektów uczenia się.

Ogólne zasady minimalizowania wpływu efektu asesora	Narzędzia rozwojowe dla asesorów, aby zmniejszyć prawdopodobieństwo wystąpienia efektu asesora
<p>Ryzyko występowania efektu asesora ma związek z wiedzą i umiejętnościami, jakie posiada asesor w związku z ocenianiem za pomocą różnorodnych metod oraz z postawą, jaką prezentuje, m.in. wobec norm etycznych, swojej pracy oraz samodoskonalenia.</p> <p>1. Dobór asesorów ze względu na posiadane przez nich efekty uczenia się oraz znajomość i doświadczenie w stosowaniu określonych metod</p>	<p>Metody rozwojowe pozwalają na minimalizowanie czynników i ryzyka występowania wypalenia zawodowego, a także efektu asesora. Należą do nich:</p> <ul style="list-style-type: none"> • superwizja <p>Superwizja to udzielanie wsparcia metodycznego i merytorycznego danej osobie przez superwizora (osobę z większym doświadczeniem, której wiedza i umiejętności są</p>

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

- walidacji.
2. Pozytywne oczekiwania asesora wobec kandydatów, bazujące na indywidualizacji i podmiotowości w ocenianiu. Uwrażliwienie na czynniki społeczne, ale nie poddawanie ich ocenie czy uwzględnianie w ocenie (np. niepełnosprawność kandydata, narodowość itp.).
 3. Ostrożne korzystanie z informacji zawartych w dokumentach kandydata, nieuleganie opiniom osób trzecich zawartych w dokumentach kandydata; dystans do opinii np. doradców walidacyjnych na temat kandydata, czy innych osób związanych z kandydatem.
 4. Współpraca z otoczeniem procesu walidacyjnego: doradcami walidacyjnymi, moderatorami, obserwatorami. Ustalanie przebiegu weryfikacji oraz określenie, jakie metody i w jakim zakresie będą wykorzystywane w danym przypadku.
 5. Ukierunkowanie i poinformowanie kandydata o głównym celu weryfikacji, sposobach oceny i etapach weryfikacji. Gotowość do składania dodatkowych wyjaśnień i objaśnień na wniosek kandydata, dotyczących technicznego i merytorycznego przebiegu walidacji. Upewnienie, że kandydat zrozumiał przekaz.
 6. Tam gdzie to możliwe, zapewnienie anonimowości kandydata i asesora, np. przy sprawdzaniu testów.
 7. Przygotowanie i przedstawienie informacji zwrotnej niezależnie od wyniku weryfikacji z dbałością o komunikowanie pozytywnych aspektów przebiegu procesu, szczególne uzasadnienie dla aspektów niedostatecznych (negatywnych).
 8. Oddzielenie „człowieka od problemu”, uwrażliwienie na przebieg procesu a nie na osobę kandydata.
 9. Unikanie weryfikacji przez pryzmat emocji („patrzenie przez różowe okulary” versus „patrzenie przez czarne okulary”).
 10. Systematyczna samokontrola, refleksja nad własnym działaniem; korzystanie z badania opinii kandydatów na temat przebiegu walidacji (wykorzystanie np. ankiety ewaluacyjnej); ciągłe doskonalenie.

potwierdzone, np. rekomendacją trenerską Polskiego Towarzystwa Psychologicznego III stopnia).

Do zadań superwizora/superwizorki należy w szczególności zbudowanie relacji kształcącej, stworzenie przestrzeni dla refleksji, rozwoju i zmiany, eksplorowanie obszarów, które do tej pory były niedostępne dla osób poddających się superwizji, pomoc przy dokonaniu zmiany w funkcjonowaniu w roli zawodowej.

Proces superwizji kończy się informacją zwrotną, w której superwizor doradza i/lub uczy konkretnych umiejętności i strategii interwencyjnych. Dzięki tej metodzie minimalizowane będzie ryzyko efektu asesora w odniesieniu do błędów uwarunkowanych cechami osoby oceniającej oraz w odniesieniu do błędów atrybucji.

- **coaching**

Coaching pozwala na przyspieszenie tempa rozwoju, osiągnięcia celów i polepszenia efektów działania. Stanowi również podstawę do zweryfikowania podejścia do wykonywanych obowiązków. Może prowadzić do podniesienia poczucia własnej wartości, co może wpłynąć na minimalizowanie efektu asesora w odniesieniu do błędów związanych z cechami osoby oceniającej.

- **mentoring**

Mentoring może być wykorzystana w początkowym okresie pracy asesora, kiedy zdobywa on doświadczenie. Skorzystanie ze wsparcia bardziej doświadczonego mentora, znajomego z wykonywanej funkcji, pozwoli na poznanie zagrożeń, błędów oceniania. Dzięki tej metodzie asesor będzie potrafił szybciej zauważyć popełniane przez siebie błędy w ocenianiu i wyeliminować błędy, które do nich prowadzą.

- **tutoring**

Tutoring to metoda rozwijania kreatywności i samodzielnego myślenia podczas cyklicznych spotkań tutora z podopiecznym. W odróżnieniu od mentoringu, tutoring to praca w obszarze wartości, talentów i celów. Zadaniem tutora jest wspieranie podopiecznego, aby ten pracował autentycznie, świadomie i spójnie, a nie dopasowywał się powierzchownie lub konformistycznie do jakiegoś zewnętrznego wzorca.

Tutoring bazuje na mocnych stronach podopiecznego, a jeżeli nie są w pełni uświadomione, ma za zadanie je uwypuklić. Istotna w tutoringowi jest również świadomość słabych stron, nad którymi należy pracować, aby je minimalizować lub wyeliminować.

Tutoring w minimalizowaniu efektu asesora może mieć zastosowanie na każdym etapie pełnienia tej funkcji. Możliwość rozwoju czy też samouświadomienia mocnych stron stanowi kluczowy bodziec do analizy zaangażowania w pracę. Refleksja, która towarzyszy tutoringowi może mieć znaczenie w uświadomieniu sobie nieetycznych zachowań w procesie walidacji.

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

- **trening intrapsychiczny**

Trening intrapsychiczny umożliwia poszerzenie i pogłębienie świadomości dotyczącej problemów emocjonalnych, ich przyczyn, jak i konsekwencji. Pozwala na lepsze zrozumienie siebie, efektywniejsze wykorzystywanie swoich predyspozycji, przezwyciężanie wewnętrznych ograniczeń i poznania swoich możliwości. Jego celem jest także zwiększenie samoświadomości i samoakceptacji. Pozwala na spojrzenie w głąb siebie, a w konsekwencji efektywniejsze tworzenie relacji z innymi ludźmi.

Efekt asesora niezależny od stosowanej metody walidacji - powszechne błędy oceniania

Błędy uwarunkowane cechami asesora

Należą do nich m.in.:

- mała złożoność struktur poznawczych – człowiek słabo różnicuje świat, przecenia podobieństwa pomiędzy sobą a innymi;
- niski poziom inteligencji poznawczej i emocjonalnej prowadzi do problemów ze stosowaniem wiedzy do poznawania innych oraz trudności w rozpoznawaniu emocji innych;
- wysoki poziom neurotyczności i nieakceptowanie samego siebie prowadzi m.in. do negatywnych stanów emocjonalnych, co ogranicza zasoby poznawcze wykorzystywane do oceny innych;
- konkretność umysłu, spostrzeganie innych na zasadzie "czarno-białej", mała tolerancja na wieloznaczność – schematyzm w myśleniu i działaniu, powielanie opinii;
- stany napięć i konfliktów wewnętrznych prowadzą do tzw. widzenia przez „czarne okulary” – dostrzegania przede wszystkim słabszych stron osób ocenianych;
- wysoki poziom autorytarności wyraża się m.in. w destruktywności i cynizmie, co przekłada się na uogólnioną wrogość w stosunku do ludzi oraz skłonność do deprecjacji innych;
- egocentryzm czyli koncentracja na własnej osobie nie sprzyja przejmowaniu perspektywy innego, uwzględniania złożonych uwarunkowań wpływających na zachowania innych;
- uległa bądź agresywna postawa życiowa: uległość sprzyja manipulowaniu przez innych; agresywna postawa wywołuje konflikty z otoczeniem i ogranicza zdolność do kompromisu;
- zależność od autorytetów i stereotypów prowadzi do niesamodzielności w ocenie innych, sprzyja manipulowaniu przez innych oraz wrogości wobec innych;
- machiawelizm jako cecha osobowości (siła nastawienia do manipulowania innymi) to tendencja do działania egoistycznego, instrumentalnego traktowania innych, u podstaw czego leży trafność spostrzegania, oceny innych;
- zaniżona lub zawyżona samoocena prowadzi do nieadekwatnej oceny siebie i innych;
- niechęć do wglądu we własną psychikę, brak motywacji do doskonalenia siebie ogranicza samowiedzę, prowadzi m.in. do schematyzmu i konserwatyzmu w ocenie innych;
- rutyna zawodowa prowadzi do stopniowego ograniczania wiedzy o sobie i innych.

Błędy atrybucji

Do podstawowych błędów atrybucji zalicza się:

- uleganie stereotypom: postrzeganie innych przez pryzmat utartych sądów, np. dotyczących płci, wykształcenia, wykonywanego zawodu, pochodzenia, religii;
- efekt halo: tendencja do automatycznego, pozytywnego (efekt aureoli) lub negatywnego (efekt Golema) przypisywania cech osobowościowych na podstawie pierwszego wrażenia. Na efekt halo duży wpływ ma wygląd zewnętrzny, którego uogólniona ocena może powodować obniżenie lub podwyższenie oceny w innych obszarach;
- błąd łagodności i surowości: błąd łagodności polega na przecenianiu zalet innych ludzi i niedocenianiu wad, występuje częściej w stosunku do osób, które oceniający (asesor) zna dłużej; błąd surowości objawia się

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

zaniżaniem oceny osób wcześniej nieznanymi (pojawia się rzadziej, niż błąd łagodności);

- błąd tendencji centralnej: polega na unikaniu ekstremalnych stopni charakterystycznych dla danej skali ocen, czyli na wykluczaniu ocen najwyższych i najniższych; błąd ten jest spowodowany jej najczęściej nadmierną ostrożnością i niepewnością oraz przekonaniem o braku wystarczających informacji o osobach ocenianych.

Błędy związane z tematyką oraz formą oceniania

Rozbieżności w ocenianiu mogą zależeć od czynników takich jak:

- forma weryfikacji efektów uczenia się: w zależności od tego, czy weryfikacja jest ustna, pisemna, czy mieszana, może być obciążona różnymi błędami (np. na weryfikację ustną mogą mieć wpływ błędy atrybucji, na weryfikację pisemną - np. styl opisu, poprawność językowa, charakter pisma);
- charakter efektów uczenia się: jeśli do oceny wiedzy, umiejętności lub kompetencji społecznych zastosuje się niewłaściwą (niedopasowaną do tego) metodę, ostateczny wynik będzie błędny; prace pisemne, zadania oraz wypowiedzi ustne w dziedzinach związanych z humanistyką ocenia się różnorodnie ze względu na niejednoznaczny charakter tematyki w dziedzinach humanistycznych (w naukach ścisłych najczęściej nie ma takiego problemu);
- efekt kontrastu: wystawienie wyższej oceny za poprawną odpowiedź po serii odpowiedzi słabych i odwrotnie – wystawianie niższej oceny po serii dobrych odpowiedzi; przy ocenianiu osób w zespole (np. debata) ocena tej samej osoby może różnić się w zależności od pozostałych uczestników - dodatni kontekst powoduje obniżenie oceny, zaś ujemny jej podwyższenie;
- efekt pierwszeństwa i świeżości: surowiej oceniane są błędy, które pojawiają się na początku, a nie na końcu pracy pisemnej czy wypowiedzi ustnej; w debatach: argumenty mówcy, który przemawia jako pierwszy, wywierają większy wpływ na słuchaczy niż argumenty dalszych mówców; w wywiadzie pierwsze poprawne/oczekiwane odpowiedzi prowadzą do wyższej oceny ogólnej i odwrotnie – niepoprawne/ niezgodne z oczekiwaniami odpowiedzi prowadzą do zaniżania oceny;
- reputacja miejsca: nazwa instytucji (miejsca), z której pochodzi kandydat może mieć wpływ na ocenę, np. „ośrodek zagraniczny”; firma o dużym prestiżu;
- patrzenie przez różowe i czarne okulary: uczucia i nastrój oceniającego (asesora) mają wpływ na ocenę innych osób i ich wytworów;
- projekcja zachowań/ kompetencji: asesor wyżej ocenia kompetencje, które sam posiada, ponieważ mimowolnie lubimy ludzi podobnych do siebie;
- skrajne reakcje oceniającego (asesora): pragnienie bycia bardzo profesjonalnym, bądź bardzo przyjacielskim powoduje popadanie w skrajności (od bycia zbyt pobłażliwym i nieformalnym do wysoko rozwiniętej surowości);
- błędy w metodach i narzędziach weryfikacji efektów uczenia się:
 - niewłaściwy dobór metody/ narzędzia do przedmiotu oceny,
 - różne predyspozycje kandydata do poszczególnych form weryfikacji;
- niekorzystne warunki towarzyszące weryfikacji powodujące dyskomfort fizyczny.

Bibliografia

American Educational Research Association, American Psychological Association i National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington: AERA. Pobrane z <http://www.apa.org/science/programs/testing/standards.aspx>

European Commission, DG Education and Culture i Cedefop (2010). *European inventory on validation of validation of non-formal and informal learning 2010. Thematic Report: Assessment methods*. Luxembourg: Publications Office.

European Commission, Cedefop i ICF International (2014). *European inventory on validation of validation of non-formal and informal learning 2014. Thematic report: validation methods*. Luxembourg: Publications Office.

Jablow, P. (2015). *5 things to know about the financial cost of testing*. Philadelphia Public School. The Notebook. Pobrane z <http://thenotebook.org/latest0/2015/11/24/5-things-to-know-about-the-financial-cost-of-testing>

Materiał pobrany ze strony walidacja.ibe.edu.pl/metody

Kubiszyn, T. i Borich, G. (2009). *Educational Testing and Measurement: Classroom Application and Practice*. John Wiley & Sons, Inc.

Komunikat dyrektora Centralnej Komisji Egzaminacyjnej z 1 września 2016 r. w sprawie szczegółowych sposobów dostosowania warunków i form przeprowadzania egzaminu potwierdzającego kwalifikacje w zawodzie do potrzeb i możliwości zdających ze specjalnymi potrzebami edukacyjnymi w roku szkolnym 2016/2017. Pobrane z https://cke.gov.pl/images/_KOMUNIKATY/Komunikat%20o%20dostosowaniach%20w%20roku%20szkolnym%202017-2018.pdf

Łobocki, M. (2000). *Metody i techniki badań edukacyjnych*. Kraków: Oficyna Wydawnicza „Impuls”.

Łoś, E. i Reszka, A. (2009). *Metody nauczania stosowane w kształtowaniu kompetencji kluczowych. Matematyka. Podręcznik metodyki operacyjnej*. Lublin: Innovatio Press Wydawnictwo Naukowe Wyższej Szkoły Ekonomii i Innowacji.

Solarczyk-Szwec, H., Machałowska, K. (2020). *Efekt asesora i minimalizowanie ryzyka efektu asesora w ZSK - ekspertyza*. Warszawa: Instytut Badań Edukacyjnych [materiał niepublikowany]

Strony internetowe:

Centralna Komisja Egzaminacyjna: <https://cke.gov.pl/>

Fundacja Vocational Competence Certificate: <http://vccsystem.eu/>

Rebandel, H. *Poradnik dla autorów zadań i egzaminów testowych*. Warszawski Uniwersytet Medyczny; <http://biurojakosci.wum.edu.pl/content/poradnik-dla-autorow-zadan-i-egzaminow-testowych-0>