

WebClass jako narzędzie do kontroli i oceny postępów w nauce

DOI: 10.47050/jows.2021.3.37-42

WOJCIECH MALEC,
MAGDALENA MALEC

Konieczność przejścia na zdalny tryb nauczania nakłoniła wielu pedagogów do poszukiwania narzędzi wykorzystujących nowoczesne technologie do monitorowania postępów uczniów oraz weryfikacji ich wiedzy i umiejętności. Jednym z nich jest autorski system zarządzania kształceniem (ang. *learning management system*) o nazwie WebClass (webclass.co), którego cztery podstawowe komponenty umożliwiają zarządzanie aktywnością użytkowników, komunikację, tworzenie materiałów dydaktycznych oraz kontrolę i ocenę postępów w nauce. Artykuł skupia się na module testowym platformy i omawia wykorzystanie go do tworzenia, przeprowadzania i analizy testów.

Testowanie wspomagane komputerowo (ang. *computer-assisted testing* lub *computer-based testing*) oraz przeprowadzane za pośrednictwem internetu (ang. *web-based testing*) różnią się od siebie trybem dostarczania treści (ang. *mode of delivery*) oraz językiem kodowania, którym w przypadku testowania w sieci jest HTML. W obu sytuacjach możemy jednak mówić generalnie o testowaniu komputerowym oraz o testach komputerowych. W literaturze szeroko pojętego pomiaru dydaktycznego można znaleźć wiele opracowań poruszających kwestię zalet i wad zarówno ogólnie pojętego testowania komputerowego, jak i ściślej testowania w sieci (np. Roever 2001; Dooley 2008; Brown 2016; Winke i Isbell 2017; Malec 2020).

Spośród zalet testowania w sieci, zwłaszcza w porównaniu z testami drukowanymi, często wymienia się na przykład łatwość dostępu do treści, większą autentyczność zadań czy też możliwość dostarczania natychmiastowej informacji zwrotnej. Zasadniczo zalety takiego sprawdzania wiedzy i umiejętności związane są z jego praktycznością wynikającą z zastosowania nowoczesnych technologii. Praktyczność (ang. *practicality*) oznacza wykonalność lub możliwość zrealizowania planu testu. Test praktyczny to taki, w przypadku którego wszelkie wymagane czynności można przeprowadzić z wykorzystaniem środków, jakie mamy do dyspozycji (por. Bachman i Palmer 1996). Według Greena (2014) strona praktyczna testowania jest podstawową i zarazem konieczną cechą użyteczności narzędzi pomiaru.

Szczególnie jeden aspekt praktyczności testowania z wykorzystaniem technologii wart jest podkreślenia. Mianowicie to, że testowanie komputerowe w sposób diametralny zmienia możliwości przeprowadzania analizy ilościowej wyników. Nie chodzi tu jedynie o automatyczne obliczanie wszelkiego rodzaju statystyk, ale także, a może nawet przede wszystkim, o samo gromadzenie danych liczbowych, które bez użycia technologii jest w wielu sytuacjach, zwłaszcza w przypadku sprawdzianów szkolnych, po prostu niepraktyczne. Istnieje wiele publikacji z dziedziny pomiaru dydaktycznego, w tym przeznaczonych szczególnie dla nauczycieli języków obcych (np. Brown 1996; Carr 2011), których spore fragmenty poświęcone zostały



mniej lub bardziej zaawansowanym metodom statystycznym. Jednakże trudno oprzeć się wrażeniu, że pomijana jest w literaturze kwestia samej wykonalności ręcznego tworzenia zbiorów danych liczbowych niezbędnych do przeprowadzenia stosownych analiz. Dla przykładu, w przypadku testu składającego się z 20 zadań oraz grupy 50 uczniów konieczne jest utworzenie zbioru danych zawierającego 1000 wyników. Trudno sobie wyobrazić, aby nauczyciele byli w stanie regularnie tworzyć tego rodzaju zbiory danych dla wszystkich sprawdzianów przeprowadzanych metodą tradycyjną. Testowanie komputerowe natomiast umożliwia automatyczne generowanie wymaganych zbiorów danych, a testowanie w sieci dodatkowo ułatwia dostęp do tych danych. Można więc powiedzieć, że technologia urealnienia możliwość dokonywania analizy ilościowej sprawdzianów językowych, a poprzez to istotnie podnosi ich praktyczność.

Inną ważną korzyścią płynącą z testowania komputerowego jest stuprocentowa dokładność punktowania, którą zapewniają zautomatyzowane algorytmy sprawdzające. Cecha ta pozytywnie wpływa na aspekt rzetelności pomiaru rozumiany jako wewnętrzna zgodność sędziów (Niemierko 1999), zarówno w przypadku zespołu (ang. *inter-rater reliability*), jak i pojedynczego sprawdzającego (ang. *intra-rater reliability*). Jak powszechnie wiadomo, w przypadku testów drukowanych sprawdzanych przez nauczycieli zapewnienie wysokiej dokładności punktowania jest często trudne do osiągnięcia, chociażby z tak błahych powodów jak zmęczenie czy niedopatrzenie.

Testowanie z wykorzystaniem nowoczesnych technologii nie jest bynajmniej pozbawione wad. Do najważniejszych należy zaliczyć konieczność posiadania odpowiedniego sprzętu komputerowego, możliwość wystąpienia różnego rodzaju awarii oraz niekompatybilności oprogramowania, ograniczenia systemów do testowania w zakresie tworzenia niektórych form zadań, a także zdecydowanie mniejsze możliwości weryfikacji niesamodzielnej pracy uczniów. Należy podkreślić, iż znaczenie tych potencjalnych wad i ograniczeń uzależnione jest od konkretnej sytuacji pomiarowej. Na przykład, kontrola niesamodzielności nie jest tak istotna w przypadku krótkich quizów zadanych jako praca domowa, jak w przypadku egzaminów końcowych. Innym czynnikiem wpływającym na ograniczenia związane z testowaniem komputerowym

jest wybór konkretnego systemu do testowania, który może albo ułatwiać, albo wręcz utrudniać wykonywanie czynności związanych z pomiarem.

Platforma WebClass

Platforma WebClass powstała z myślą o wykorzystaniu potencjału nowoczesnych technologii, a jednocześnie przez kilka lat ciągłej optymalizacji wprowadzane były poprawki eliminujące jej rozmaite niedoskonałości. Posiada ona dość obszerny moduł testowy, który może być wykorzystany do tworzenia, przeprowadzania i analizy statystycznej sprawdzianów różnego rodzaju (Malec 2018). Ogólnie rzecz ujmując, testy te mogą się składać z pozycji testowych (ang. *test items*) oraz zadań wymagających udzielenia rozszerzonych odpowiedzi¹.

W ciągu roku szkolno-akademickiego 2020/2021 z platformy korzystało 26 nauczycieli z Lublina, wśród których byli uczący w liceum oraz wykładowcy akademicy, a także 917 uczniów i studentów. Od września 2020 roku powstało około 900 testów i sprawdzianów, głównie leksykalno-gramatycznych oraz sprawdzających umiejętności czytania i słuchania ze zrozumieniem, ale także obejmujących inne obszary wiedzy i umiejętności, jak na przykład kulturoznawstwo, literatura amerykańska czy redakcja tekstów obcojęzycznych.

Testowanie na WebClass

Oto skrótowe przedstawienie testowania w systemie WebClass. Po utworzeniu testu nauczyciel przypisuje go do kont poszczególnych użytkowników lub całych grup (klas) uczniów, którzy po wykonaniu zadań mają możliwość sprawdzenia wyników oraz uzyskania informacji zwrotnej (w zależności od ustawień). Pozycje testowe punktowane są automatycznie na podstawie klucza, jednak w przypadku zadań otwartych nauczyciel może także przeprowadzić weryfikację punktacji — może na przykład zaakceptować odpowiedź, która zawiera jakiś nieistotny błąd. Należy tutaj podkreślić, że tego rodzaju decyzja skutkuje zaakceptowaniem odpowiedzi u wszystkich uczniów, co wyklucza problem pojawienia się różnic w punktacji między osobami udzielającymi identycznych odpowiedzi. Testy są również automatycznie oceniane na podstawie skali ocen uprzednio zdefiniowanej przez nauczyciela. Całościowe wyniki pozwalają obliczyć dwie średnie, arytmetyczna i ważona, a ocena końcowa może dodatkowo uwzględniać liczbę nieobecności na zajęciach oraz

1 Pozycje testowe to zadania testowe (ang. *test tasks*), które wymagają wskazania jednej z gotowych odpowiedzi lub udzielenia (względnie) krótkiej odpowiedzi (wg modelu Bachmana i Palmera 1996, 2010). Oba terminy (zadanie i pozycja testowa) są często używane zamiennie.

wyniki samooceny, podczas której uczniowie określają stopień opanowania celów kształcenia (sformułowanych przez nauczyciela zazwyczaj na początku semestru). Raporty dla poszczególnych sprawdzianów oraz całego semestru generowane są automatycznie i uaktualniane po każdym teście, a także po jakiegokolwiek zmianie dotyczącej punktacji zadań, np. po dodaniu nowych odpowiedzi do klucza. Raporty dla całych grup uczniów, zawierające wszystkie możliwe składowe końcowych ocen semestralnych, generowane są na koncie prowadzącego w formie zwartej tabeli, która dodatkowo umożliwia szybki dostęp do wielu szczegółów związanych z uzyskanymi ocenami (takich jak odpowiedzi i punktacja zadań wykonanych przez danego ucznia czy wyniki samooceny). Istnieje także możliwość ręcznego wprowadzania wyników testów przeprowadzanych poza platformą, np. drukowanych wersji sprawdzianów, w celu uwzględnienia ich przy obliczaniu końcowych wyników i ocen semestralnych. W pewnym sensie platforma odgrywa tu rolę zautomatyzowanego dziennika szkolnego.

Tworzenie testów

Proces testowania w systemie omawianym w niniejszym artykule rozpoczyna się od sporządzenia koncepcji i planu testu, po czym następuje budowanie poszczególnych zadań. Koncepcja lub projekt testu to w gruncie rzeczy zbiór różnego rodzaju decyzji, określanych mianem specyfikacji testu. Na tym etapie rola technologii, a w każdym razie platformy WebClass, sprowadza się do umożliwienia zapisania tych decyzji w postaci metryki testu oraz łatwego do niej dostępu. Oprócz specyfikacji z testem związane są różnego rodzaju ustalenia, np. słowa kluczowe oraz ustawienia dotyczące przeprowadzania pomiaru, takie jak ograniczenie dostępności testu do jakiejś konkretnej daty i godziny, limit czasu oraz próg zaliczeniowy.

Treść zadań testowych zapisywana jest na platformie w edytorze testów, wewnątrz elementów HTML posiadających atrybut *contenteditable* o wartości „true”, a nie w polach tekstowych, co ma miejsce w przypadku większości tego typu systemów internetowych. Takie rozwiązanie ma dobre i złe strony. Najważniejszą korzyścią wynikającą z zastosowania wspomnianego atrybutu jest możliwość edytowania treści w przeglądarce internetowej w sposób przypominający znane edytory tekstowe. Można tego dokonać za pośrednictwem poleceń języka JavaScript, których rola polega na wstawianiu do treści zadań odpowiednich elementów HTML oraz stylów formatowania, takich

jak pogrubienie, podkreślenie, kolor i wiele innych. (W przypadku pól tekstowych użytkownik ma jedynie możliwość edytowania samego tekstu, ale nie może zmieniać jego koloru, wielkości czcionek, wyrównania/justowania, ani też dodawać nowych elementów HTML, jak np. lista numerowana.) Dzięki możliwości niemalże dowolnego edytowania treści na platformie zadania testowe da się układać i modyfikować na wiele różnych sposobów. Dla przykładu, mogą one przybrać formę tabeli, w której użytkownik może dodatkowo umieścić elementy multimedialne, takie jak ilustracje, dźwięki i klipy wideo. W systemach, w których dodawanie elementów HTML jest zablokowane, zarówno położenie multimedii, jak i ogólny układ zadań są z góry narzucone przez konkretny szablon formularza testu. Tutaj natomiast użytkownik ma duże możliwości ingerencji w ostateczny wygląd testu, co pozwala nawet na tworzenie innowacyjnych form zadań poprzez (choćby nieznaczne) przekształcanie standardowych. Najistotniejszym minusem powyższego rozwiązania jest to, że w pewnych sytuacjach najlepszym sposobem na osiągnięcie zamierzonego celu jest modyfikacja kodu źródłowego edytowanych zadań, co z kolei wymaga przynajmniej podstawowej znajomości HTML. Teoretycznie istnieje także ryzyko dokonania takich modyfikacji, które będą skutkować nieprawidłowym wyświetleniem testu w przeglądarce internetowej. Jednakże w większości sytuacji ingerencja w kod źródłowy HTML absolutnie nie jest konieczna.

Istnieje możliwość tworzenia wielu rodzajów zadań, zarówno zamkniętych, jak i otwartych. Zamknięte obejmują zadania wyboru wielokrotnego (ang. *multiple choice*), pozycje alternatywne (ang. *binary choice*), zadania wielokrotnej odpowiedzi (ang. *multiple response*), tekst z lukami oraz wieloma opcjami do wyboru (ang. *multiple-choice cloze*), zadania na dobieranie (ang. *matching*), a także różne modyfikacje innych typów, np. zadania z luką, którym towarzyszy pula odpowiedzi do wyboru (ang. *banked gap-filling*). Do zadań otwartych typu *item* zaliczymy: zadania z luką (ang. *gap-filling*), tekst z lukami (*cloze*), tekst z lukami wraz z pierwszą literą lub połową każdego brakującego słowa (*C-test*), zestawy kilku zdań do uzupełnienia jedną odpowiedzią (ang. *gapped sentences*), transformacje (ang. *transformation*), transformacje z podanym wyrazem (ang. *use the word given*), pisanie zdań ilustrujących użycie słowa lub frazy (ang. *sentence writing*), poprawianie błędów (ang. *error correction*) oraz różne modyfikacje zadań z luką, jak np. krótka odpowiedź (ang. *short answer*), tłumaczenie (ang. *translation*) czy

słowotwórstwo (ang. *word-formation*). Dodatkowo system umożliwia tworzenie zadań rozszerzonej odpowiedzi, takich jak esej czy dyktando.

Sposoby konstruowania zadań

Istnieje kilka sposobów budowania testu na platformie. Pierwszy polega na układaniu pojedynczych pozycji testowych krok po kroku i „ręcznym” dodawaniu ich do konkretnej wiązki zadań. Teoretycznie możliwe jest łączenie pozycji różnego typu, jednak z wielu względów nie jest to najlepsze rozwiązanie, natomiast zadania rozszerzonej odpowiedzi nie mogą być łączone z pozycjami testowymi w obrębie jednej wiązki zadań. W przypadku rozszerzonych odpowiedzi pisemnych zadanie zawiera pole tekstowe, w którym uczeń wpisuje np. esej lub opowiadanie. Rzeczne pole może opcjonalnie zawierać tekst przeznaczony do modyfikacji, będący np. fragmentem tradycyjnego testu drukowanego. Opcja ta została dodana do edytora w ciągu ostatniego semestru, po sugestii nauczycieli korzystających z platformy. Warto zaznaczyć, że jedna z ciekawszych modyfikacji standardowych formatów zadań polega na dodaniu przycisków z symbolami fonetycznymi (nieznajdującymi się na klawiaturze), dzięki którym możliwe jest udzielanie odpowiedzi w postaci transkrypcji fonetycznej.

Drugim sposobem budowania zadań w omawianym systemie jest konwersja odpowiednio sformatowanego tekstu, co zasadniczo skraca czas tworzenia większych wiązek zadań testowych (Malec 2016). Konwerter tekstu (*text-to-items converter*) może być zastosowany wyłącznie do układania pozycji testowych, a więc do zadań typu *item*. Jedną z dodatkowych funkcji tego narzędzia jest możliwość automatycznego tworzenia zadań typu *cloze* oraz *C-test* (w kilku wariantach), na podstawie dowolnego tekstu wpisanego lub wklejonego przez użytkownika. Warto dodać, że konwerter działa też w „odwrotnym” kierunku, tzn. może być użyty do edycji istniejących już zadań (następuje wówczas konwersja zadań do fragmentu tekstu przeznaczonego do edycji).

Pozostałe dwa sposoby budowania formularza testu wymagają wcześniejszego utworzenia przynajmniej jednej puli pytań w banku zadań. Pierwszy z tych sposobów polega na importowaniu z niego wybranych pozycji do aktualnie tworzonego formularza. Mogą być one wskazane ręcznie lub wybrane losowo, opcjonalnie z zastosowaniem filtra określającego średnią wartość łatwości zadań (*item facility*) oraz mocy różnicującej (*item discrimination*) w przypadku testów różnicujących (*norm-referenced tests*) lub wskaźnika *B* (*B-index*) w przypadku testów sprawdzających (*criterion-referenced tests*).

Filtr ułatwia wybranie tych zadań, które spełniają żądane kryteria statystyczne. Ostatni sposób tworzenia formularza testu to zautomatyzowany generator testu, dzięki któremu utworzenie całego formularza wymaga tylko kilku kliknięć. Konkretniej, generator testu wymaga określenia (1) liczby wiązek zadań, (2) ich formatu oraz liczby w każdej wiązce, (3) przynajmniej jednej puli zadań w banku. Po uruchomieniu generatora następuje automatyczne budowanie formularza testu składającego się z pozycji losowo wybranych z banku zadań.

Stosowanie testów

Czynności związane z zastosowaniem testu obejmują przeprowadzenie pomiaru, sprawdzenie i punktowanie zadań oraz utworzenie raportów i udzielenie informacji zwrotnej. Testy mogą być prezentowane na ekranach uczniów w całości na jednej (przewijanej) stronie w przeglądarce bądź też każda pozycja pojedynczo (w tym przypadku system dodaje do formularza menu nawigacyjne). Dodatkowo kolejność wiązek oraz zawartych w nich zadań może być inna dla każdego ucznia, co wraz z aktywną opcją wyświetlania testu po jednym zadaniu znacząco utrudnia „ściągnięcie”, w sytuacji gdy uczniowie siedzący obok siebie widzą swoje ekrany, zwłaszcza jeśli czas na wykonanie wszystkich zadań jest ograniczony. Poza tym każdy test może być uruchomiony w trybie pełnoekranowym z opcją ciągłego sprawdzania, czy użytkownik nie opuszcza testu poprzez zmianę okna przeglądarki lub przejście do innego programu. W przypadku gdy ma to miejsce, praca zostaje zawieszona, a nauczyciel monitorujący sprawdzian zostaje o tym fakcie poinformowany i może pozwolić uczniowi na kontynuowanie testu. Warto też zauważyć, że podczas przeprowadzania pomiaru przeglądarka komunikuje się z serwerem co pięć sekund w celu zapisania w bazie danych aktualnie udzielonych odpowiedzi. Ma to istotne znaczenie, gdy zaistnieje konieczność przerwania pracy, na przykład z powodu jakiegś usterki technicznej. Dzięki zapisanym odpowiedziom uczniowie mogą zostać poproszeni w takiej sytuacji o kontynuowanie testu bez konieczności rozwiązywania zadań od początku.

Pozycje testowe są sprawdzane i punktowane automatycznie metodą, która polega na dokładnym porównaniu ich z kluczem odpowiedzi. W przypadku pozycji otwartych składa się on z trzech zbiorów odpowiedzi (ich liczba w każdym z tych zbiorów jest nieograniczona). Pierwszy z nich to odpowiedzi prawidłowe. Drugi to odpowiedzi, które zawierają nieistotne błędy (z punktu widzenia celu pomiaru), jak na

przykład drobne literówki. Trzeci zbiór to odpowiedzi, które zasługują na częściową akceptację. Wszystkie te zbiory odpowiedzi mogą być utworzone podczas budowania zadań, jednak drugi i trzeci najlepiej utworzyć po przeprowadzeniu testu, podczas weryfikacji automatycznej punktacji (o czym poniżej), gdyż na etapie układania zadań odgadnięcie konkretnych odpowiedzi, które powinny się tam znaleźć, może być niełatwe. Możliwe jest także automatyczne (ale tylko częściowe) zaakceptowanie odpowiedzi zawierających nie więcej niż określoną liczbę literówek. Może ona być inna dla każdego zadania. Ponadto algorytm sprawdzający może (opcjonalnie) całkowicie pomijać takie różnice między odpowiedzią udzieloną a prawidłową jak wielkie/małe litery, spacje oraz znaki interpunkcyjne.

Poprawne działanie algorytmu sprawdzającego jest w pewnym stopniu uzależnione od weryfikacji automatycznej punktacji. Przeprowadzana jest ona przez nauczyciela po zakończeniu testu z zastosowaniem narzędzia prezentującego każde zadanie testowe wraz z wszystkimi nieprawidłowymi odpowiedziami, których (automatyczna) punktacja może być łatwo zmieniona poprzez przeniesienie ich do jednego ze wspomnianych zbiorów w kluczu. W przypadku zadań zamkniętych nie ma oczywiście możliwości ingerowania w punktację. Z kolei dla zadań otwartych wymagających bardzo krótkich odpowiedzi, gdzie liczba potencjalnych odpowiedzi prawidłowych jest ograniczona do minimum, weryfikacja może być konieczna tylko w nielicznych przypadkach. Jest ona jednak dość istotna, gdyż ma na celu upewnienie się co do trafności punktowania, a co za tym idzie, do poprawności interpretacji wyników testu.

Jeśli chodzi o zadania rozszerzonej odpowiedzi, nie są one oceniane automatycznie, lecz za pomocą edytora HTML, w którym nauczyciel może zaznaczać błędy, wstawiać brakujące wyrazy, a także dodawać komentarze. Narzędzie do sprawdzania rozszerzonych odpowiedzi zawiera różne ułatwienia dla osoby sprawdzającej, jak na przykład wyszukiwanie słów i kolokacji w korpusie (BNC). Narzędzie to dodatkowo podlicza liczbę naniesionych poprawek oraz sugeruje ocenę w postaci wyniku procentowego. Należy podkreślić jednak, że zastosowany wzór na obliczanie tego wyniku jest w dużej mierze uzależniony od konkretnej sytuacji pomiarowej i nie jest uniwersalny.

Bardzo istotnym elementem procesu testowania są indywidualne raporty wyników (dla poszczególnych testów) oraz udzielanie informacji zwrotnej (*feedback*). Oprócz takich danych jak próg zaliczeniowy, ogólny wynik testu oraz ocena raporty zawierają wszystkie

zadania testowe, odpowiedzi udzielone przez ucznia, automatyczną ich korektę oraz punktację. Całość bardzo przypomina tradycyjny test drukowany poprawiony przez nauczyciela (Rys. 1). Raporty mogą dodatkowo zawierać komentarze na temat zadań i poszczególnych odpowiedzi. Mogą one być ogólne (prezentowane bez względu na udzieloną odpowiedź) lub szczegółowe (skierowane tylko do tych uczniów, którzy udzielili jakiejś konkretnej odpowiedzi).

Rys. 1. Informacja zwrotna zawierająca udzielone odpowiedzi, ich korektę, punktację oraz komentarze

Transformations – word given
Complete the second sentence so that it means the same as the first sentence, using the word given.

1. I respect him very much as he is the best surgeon I've ever met. **HAT**
I bow (take) my hat (off) to him as he is the best surgeon I've ever met. **0**
 Someone can **bow** her or his **head** (e.g. in shame).

2. This is, undoubtedly, the best road trip in my life. **SHADOW**
This is, without (a) shadow of a doubt, the best road trip in my life. **0.5**
 The indefinite article is necessary here.
ALSO: without the shadow of a doubt or beyond a shadow of a doubt or beyond the shadow of a doubt

3. Court officials ordered that he should be allowed to leave a Beijing prison earlier. **EARLY**
Court officials ordered his early imprisonment (at) (release from) a Beijing prison. **0**
 The opposite!

4. In view of the fact that they are now on holiday, we can't have them over for dinner. **SEEING**
Seeing that they are (now) now on holiday, we can't have them over for dinner. **+1**
ALSO: Seeing as they are

Analiza

Ostatnim komponentem systemu jest narzędzie do ewaluacji testów i pozycji testowych, które służy także do zapisywania pozycji w banku zadań. W innych systemach testowania wspomaganego komputerowo pozycje testowe mogą być układane bezpośrednio w banku zadań, czyli na samym początku cyklu ich tworzenia. Teoretycznie istnieje taka możliwość także na platformie WebClass. Jest to jednak nieco utrudnione, ponieważ narzędzie to raczej „zachęca” do zapisywania w banku tylko tych pozycji, których wskaźniki statystyczne są zadowalające, a to można stwierdzić dopiero po testowaniu próbnym.

Narzędzie do analizy testów prezentuje kilkanaście różnych wskaźników statystycznych dla testu jako całości. Wśród nich znajdują się współczynniki rzetelności pomiaru, zarówno te, które są właściwe dla pomiaru różnicującego, jak i te, które powinny być stosowane w pomiarze sprawdzającym (Brown i Hudson 2002; Bachman 2004). Warto podkreślić, że ostatnie z nich, z wyjątkiem *kappa kwadrat* (κ^2), szacowane są w ramach teorii uniwersalizacji (np. Brennan

2001). Jeśli chodzi o analizę pozycji testowych, jej wyniki prezentowane są w kontekście samych zadań, a nie w formie zestawienia tabelarycznego zawierającego jedynie numery odpowiednich pozycji. Taki sposób prezentacji pozwala szybko ocenić treść każdego zadania i zdecydować, czy i w jakim stopniu wskaźniki statystyczne potwierdzają jego poprawność konstrukcyjną. Wskaźniki te obejmują łatwość zadania (*IF*), wskaźnik zastępczy mocy różnicującej zadania (*ID*), korelację *r* Pearsona (dla zadań otwartych), korelację punktowo-dwuseryjną (dla zadań zamkniętych), wskaźnik progowy *B*, a także frakcję opuszczeń (ang. *omission rate*). W przypadku zadań wyboru wielokrotnego dodatkowo prezentowane są wyniki analizy dystraktorów, a wśród nich korelacja punktowo-dwuseryjna dla dystraktora (*PB_{DC}*), wskaźnik progowy *B* dla dystraktora (*B_{DC}*), a także wykresy liniowe pokazujące częstotliwości wybierania dystraktorów oraz prawidłowej odpowiedzi w pięciu grupach wyników.

Podsumowanie

Przeniesienie procesu dydaktycznego na zdalny tryb nauczania, wymuszone przez pandemię, spowodowało, że sprawdzanie postępów w nauce z wykorzystaniem nowoczesnych technologii nie było w ostatnim czasie jedynie wolnym wyborem, ale koniecznością. Warto jednak rozważyć korzyści wynikające z zachowania metod sprawdzania wiedzy i umiejętności wypracowanych podczas nauczania zdalnego, w tym także testowania komputerowego, po powrocie do stacjonarnego trybu pracy. Będzie to korzystne dla młodego pokolenia (szybka informacja zwrotna, możliwość sprawnego wykonania zadań z użyciem dobrze znanych urządzeń) oraz dla pedagogów (obiektywizm i szybkość oceny, a także zdolność statystycznego opracowania danych). Omówiona w artykule platforma WebClass powstała kilka lat przed obowiązkowym nauczaniem zdalnym, a utworzono ją z przekonaniem, że sprawdzi się jako obiektywne i wygodne narzędzie do pomiaru wiedzy i umiejętności uczniów oraz studentów również w stacjonarnym i hybrydowym trybie pracy.

BIBLIOGRAFIA

- Bachman, L.F. (2004), *Statistical Analyses for Language Assessment*, Cambridge: Cambridge University Press.
- Bachman, L.F., Palmer, A.S. (1996), *Language Testing in Practice: Designing and Developing Useful Language Tests*, Oxford: Oxford University Press.
- Bachman, L.F., Palmer, A.S. (2010), *Language Assessment in Practice*, Oxford: Oxford University Press.
- Brennan, R.L. (2001), *Generalizability Theory*, New York, NY: Springer-Verlag.
- Brown, J.D. (1996), *Testing in Language Programs*, Upper Saddle River, NJ: Prentice Hall Regents.
- Brown, J.D. (2016), *Language testing and technology*, [w:] F. Farr, L. Murray (red.), *The Routledge Handbook of Language Learning and Technology*, London: Routledge, s. 141–159.
- Brown, J.D., Hudson, T. (2002), *Criterion-referenced Language Testing*, Cambridge: Cambridge University Press.
- Carr, N.T. (2011), *Designing and Analyzing Language Tests*, Oxford: Oxford University Press.
- Dooley, P. (2008), *Language testing and technology: Problems of transition to a new era*, „ReCALL”, nr 20(1), s. 21–34.
- Green, A. (2014), *Exploring Language Assessment and Testing: Language in Action*, New York, NY: Routledge.
- Malec, W. (2016), *Automating the Construction of Selected-Response Items with a Text-to-Items Converter*, „CBU International Conference Proceedings”, nr 4, s. 864–872.
- Malec, W. (2018), *Developing Web-Based Language Tests*, Lublin: Wydawnictwo KUL.
- Malec, W. (2020), *Computer-Based Testing: A Necessary Evil or a Sensible Choice?*, „The Modern Higher Education Review”, nr 5, s. 100–113.
- Niemierko, B. (1999), *Pomiar wyników kształcenia*, Warszawa: Wydawnictwa Szkolne i Pedagogiczne.
- Roever, C. (2001), *Web-Based Language Testing*, „Language Learning & Technology”, nr 5(2), s. 84–94.
- Winke, P.M., Isbell, D.R. (2017), *Computer-Assisted Language Assessment*, [w:] S.L. Thorne, S. May (red.), *Encyclopedia of Language and Education, Volume 9: Language, Education and Technology*, Cham, CH: Springer International Publishing, s. 313–325.

DR HAB. WOJCIECH MALEC Pracuje jako adiunkt w Instytucie Językoznawstwa KUL. Jego zainteresowania naukowe obejmują wykorzystanie nowoczesnych technologii w nauczaniu języków obcych. Jest autorem i twórcą platformy WebClass, której moduł testowy służy do tworzenia, przeprowadzania i analizy statystycznej testów za pośrednictwem internetu.

MAGDALENA MALEC Dyplomowana nauczycielka języka angielskiego w XXI LO im. św. Stanisława Kostki w Lublinie, egzaminator OKE oraz Cambridge English, uczestniczyła w międzynarodowym projekcie Macmillan UK dotyczącym nauczania zdalnego. Jest zwolenniczką i propagatorką wykorzystywania narzędzi TIK w procesie dydaktycznym. Brała udział w licznych projektach i pracach badawczych dotyczących zastosowania nowoczesnych technologii w nauczaniu. Od początku istnienia platformy WebClass jest aktywnie zaangażowana w jej powstawanie i rozwój.