

KOMPUTEROWA EWALUACJA WYNIKÓW KSZTAŁCENIA W ŚWIETLE BADAŃ OPINII STUDENTÓW

Jacek Jędrzyckowski, Jarosław Wagner
Uniwersytet Zielonogórski,
Katedra Mediów i Technologii Informacyjnych
65-762 Zielona Góra, Al. Wojska Polskiego 69, tel. (068) 3272868
J.Jedryczkowski@kmti.uz.zgora.pl, J.Wagner@kmti.uz.zgora.pl

Coraz bardziej powszechne zjawisko samokształcenia i kształcenia na odległość będące konsekwencją dokonujących się przemian społecznych, ekonomicznych kulturowych i cywilizacyjnych stawia nowe wyzwania także w dziedzinie ewaluacji wyników kształcenia. Wymusza stosowanie obiektywnych procedur do szacowania wiadomości i umiejętności, które pozwalają wyeliminować efekty arbitralnej, subiektywnej i nieświadomie stronniczej oceny [4,5].

Sieciowe mechanizmy komunikacyjne umożliwiają obecnie testowanie dużych grup studentów za pomocą elektronicznych formularzy. Opracowywane testy mogą być zbudowane są w oparciu o szereg zadań zamkniętych oraz zadania otwarte z wyłączeniem zadań rozszerzonej odpowiedzi, do których realizacji nie opracowano jeszcze wydajnych i wiarygodnych algorytmów [2].

Tego typu rozwiązania są stosowane przez pracowników z Katedry Mediów i Technologii Informacyjnych Uniwersytetu Zielonogórskiego już od 1998 roku. Na podstawie badań ich opinii wytypowane zostały najistotniejsze cechy testów elektronicznych. Do najczęściej wymienianych należą: oszczędność czasu poświęcanego na sprawdzanie testów (liczenie punktów, wystawianie ocen), możliwość przeprowadzenia natychmiastowych obliczeń statystycznych oraz ich wizualizacja, archiwizacja wyników, możliwość szybkiej zmiany całości lub fragmentu testu, łatwość powielania, wydruk wyników całej grupy w chwili zakończenia pracy [3].

Mając na uwadze skuteczność stosowanych metod i środków ewaluacji, kluczową kwestią staje się optymalizacja formy testu elektronicznego rozumiana w tym wypadku jako wybór jednego lub kilku rozwiązań technicznych, które zapewnią studentom najlepsze wyniki testowania w określonym czasie [1].

Badając efektywność stosowanych rozwiązań należy wyodrębnić czynniki mające bezpośredni wpływ na uzyskiwane wyniki. Należą do nich: przygotowanie merytoryczne

studentów, poziom trudności zadań, typy zadań, rozwiązania techniczne. Już wstępna analiza wymienionych czynników pozwala stwierdzić, że nie jest możliwe znalezienie prostej zależności pomiędzy opiniami na temat zastosowanych rozwiązań technicznych, a uzyskanymi wynikami. W tej sytuacji poszukiwania optymalnych rozwiązań oparte zostały na badaniach ilościowych subiektywnych opinii studentów, z elementem obiektywizującym w postaci uzyskanych wyników testu, których znaczenia nie można jednak przeceniać.

Badania rozpoczęto od sformułowania problemu badawczego ujętego w formie pytania: Które rozwiązania techniczne dostępne w testach elektronicznych mają wpływ na jakość i tempo pracy?

Uszczegółowieniem problemu badawczego są następujące pytania:

1. Czy dołączona do testu (stale dostępna) interaktywna instrukcja wyjaśnia wszystkie wątpliwości i ogranicza liczbę pytań zadawanych w czasie trwania testu?
2. Czy udzielanie odpowiedzi poprzez manipulację kursorem myszy w obrębie elektronicznego formularza ma wpływ na zwiększenie komfortu i tempa pracy?
3. Czy system interaktywnych komentarzy przypominających o konieczności wypełnienia odpowiednich pól, wskazujący pominięte odpowiedzi oraz ostrzegający o błędach technicznych ma wpływ na zwiększenie komfortu i tempa pracy?
4. Czy wielobarwny interfejs, grupujący obszary o identycznych funkcjach ma wpływ na zwiększenie komfortu i tempa pracy?
5. Czy możliwość natychmiastowej korekty dokonanego wyboru ma wpływ na zwiększenie komfortu i tempa pracy?
6. Czy integracja wszystkich zaproponowanych rozwiązań technicznych w obrębie jednego testu ma wpływ na zwiększenie komfortu i tempa pracy?

Na podstawie przyjętych problemów szczegółowych wysunięte zostały następujące hipotezy robocze:

1. Dołączona do testu interaktywna instrukcja ograniczy do minimum pytania w trakcie testowania.
2. Wybór odpowiedzi dokonywany kursorem myszy eliminuje błędy techniczne, skraca czas pracy i ma wpływ na podniesienie jej komfortu.
3. Interakcja w teście eliminuje błędy techniczne, skraca czas pracy i ma wpływ na podniesienie jej komfortu.
4. Struktura funkcjonalna zbudowana w oparciu o różnobarwne pola interfejsu ułatwia odbiór treści, ma wpływ na sprawność nawigacji, co wpływa na skrócenie czasu pracy i podniesienie jej komfortu.

5. Nieograniczone możliwości korekty dokonanych wyborów eliminują sytuację stresową towarzyszącą temu zjawisku w przypadku testów tradycyjnych, przez co zwiększają wydawnie tempo pracy.
6. Integracja wszystkich uzasadnionych rozwiązań technicznych w obrębie jednego testu ma wpływ na zwiększenie komfortu i tempa pracy.

Badania zostały przeprowadzone w oparciu o interaktywną ankietę elektroniczną zintegrowaną z testem egzaminacyjnym. Udzielenie odpowiedzi miało charakter obligatoryjny. Uzupełnieniem badań ankietowych były obserwacje prowadzone w trakcie egzaminu. Studentom pozwolono zadawać pytania dotyczące konstrukcji testu, formy udzielania odpowiedzi, itp.. Po zakończeniu egzaminu i wypełnieniu ankiet przeprowadzono ze studentami kilkuminutowy wywiad.

Przedstawiony do oceny test elektroniczny zawierający 20 zadań, w tym: 3 zadania krótkiej odpowiedzi, 5 zadań z luką, 10 zadań wyboru wielokrotnego oraz 2 zadania na dobieranie. Odpowiedzi udzielano wykorzystując formularze pozwalające wpisać odpowiedź tylko w wyznaczonym miejscu i w określonej formie. Wybór opcji wykluczających się był niemożliwy, albo skutkowało natychmiastowym wyświetleniem ostrzegawczych komunikatów. Studenci byli przyzwyczajeni do tej formy ewaluacji, bowiem każdy z nich, co najmniej sześciokrotnie rozwiązywał tego typu test.

Z badań przeprowadzonych na 116 respondentach w tym 83 kobietach oraz 33 mężczyznach uzyskano wyniki dotyczące poszczególnych elementów elektronicznego testu. Wyniki opracowano w ujęciu całościowym, jednak dla pełniejszej analizy dodatkowo wyodrębniono opinie kobiet i mężczyzn, a także uwzględniono strukturę osiągniętych ocen.

Średnia ocen z egzaminu dla kobiet wyniosła 3,7, a dla mężczyzn 3,6. Analizując oceny, dokonano podziału na trzy grupy: I grupa – wypowiedzi studentów, którzy uzyskali oceny z przedziału 2-3, II grupa oceny 3,5 – 4, III grupa – oceny 4,5 – 5 (tab. 1).

	I GRUPA <i>oceny: (2-3)</i>	II GRUPA <i>oceny: (3,5-4)</i>	III GRUPA <i>oceny: (4,5-5)</i>
Liczba kobiet	20	36	27
Liczba mężczyzn	14	6	13
Razem:	34	42	40

Tab. 1. Respondentów podzielono na trzy grupy w zależności od otrzymanych w wyniku testowania ocen.

Ocena instrukcji

W grupie I instrukcję pomijano lub czytano pobieżnie, starając się poświęcić możliwie najwięcej czasu na rozwiązanie zadań. W trakcie egzaminu często pytano o zagadnienia wyjaśnione w instrukcji (19 pytań). W II grupie instrukcja była czytana z uwagą (zadano 11 pytań). W grupie III Około 62% studentów pominęło tekst instrukcji. Zaglądano do niej dopiero w przypadku wystąpienia wątpliwości lub problemów. Padło 6 pytań z prośbą o pomoc. Studenci wyjaśniali, że na tyle dobrze znają tego typu testy oraz standardy interfejsu, iż mogą bezpośrednio przystąpić do pracy. Podział na opinie kobiet i mężczyzn dotyczących konstrukcji i jasności sformułowań zawartych w instrukcji nie wykazał istotnych różnic. Wysokie oceny w obu grupach deklarowało 79% respondentów. 1% kobiet nisko ocenił przydatność instrukcji. Oceny umiarkowane to: 20% (kobiety) i 21% (mężczyźni).

Ocena sposobów wyboru wariantu odpowiedzi i interakcji arkusza testowego

U osób sporadycznie korzystających z komputera można było zaobserwować brak sprawności manualnej szczególnie przy wyborze wariantu odpowiedzi w formularzach typu „pole kombi” wykorzystywanych w zadaniach wyboru wielokrotnego. (4 osoby- I grupa, 6 osób – II grupa, 4 osoby- III grupa).

Opinie z grup II i III były bardzo zbliżone (odpowiednio: 76% i 75% ocen wysokich oraz 22% i 25% ocen umiarkowanych). Znacznie niżej formę udzielania odpowiedzi ocenili studenci z grupy I (68% ocen wysokich i 26% umiarkowanych), którzy w wywiadach opowiadali się za tradycyjnymi formami sprawdzania wiadomości (6 osób). Mimo, iż opinie nie były wyrażane w sposób otwarty, można było z poszczególnych wypowiedzi wywnioskować, że proponowany sposób ewaluacji utrudnia zachowania nieuczciwe i z tego względu nie jest mile widziany.

Podział na opinie kobiet i mężczyzn nie wykazał istotnych różnic. Zdania były zbliżone. Wysokie oceny w obu grupach deklarowało odpowiednio 72% i 76% kobiet i mężczyzn. Kobiety nie przedstawiły żadnych opinii negatywnych. Umiarkowane oceny stanowiły 28%. Mężczyźni deklarowali 15% opinii umiarkowanych i 9% negatywnych.

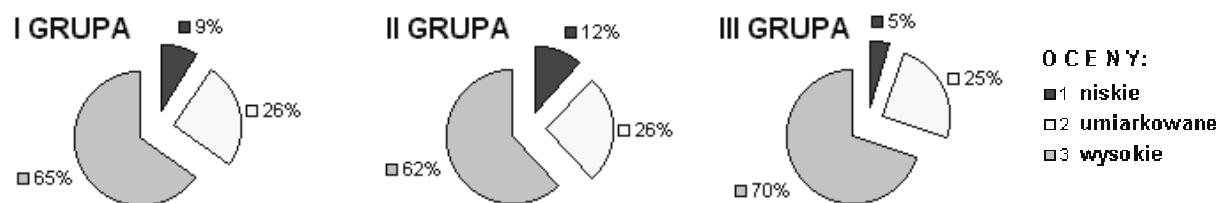
Bardzo wysoko oceniono interaktywny mechanizm komentarzy i powiadamiania o numerach pominiętych zadań.

Ocena kolorystyki interfejsu

Oceniany arkusz testu zawierał w górnej części ekranu ciemnoszary pasek narzędziowy z wielobarwnymi przyciskami funkcyjnymi i nawigacyjnymi. Poniżej na szarym tle umieszczono jasnozielone pola na dane osobowe studenta. Treść każdego zadania wpisana była w polu błękitnym, obok którego na niebieskim tle umieszczono jego numer. Propozycje

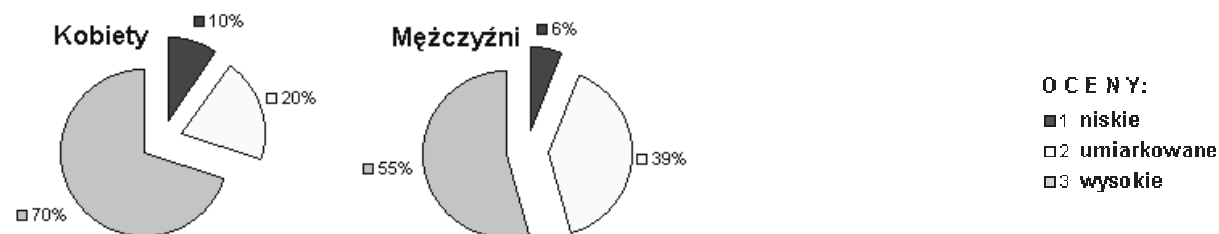
odpowiedzi lub pola, w których należało je wybrać lub wpisać były w kolorze białym. W zadaniach wyboru wielokrotnego poszczególne warianty odpowiedzi były oznakowane czarnymi literami na żółtym tle. Komunikaty ostrzegające o błędach technicznych wyświetlane były czerwoną czcionką na jasnoszarym tle. Uzasadnieniem zastosowania tak bogatej palety barw było znaczne uproszczenie i skrócenie treści instrukcji wyświetlanej na białym tle po naciśnięciu przycisku funkcyjnego.

W żadnej z grup nie zaobserwowano trudności z odczytaniem tekstu. Głównie mężczyźni z II grupy proponowali kolory bardziej stonowane. W III grupie studenci pracujący na co dzień z komputerem proponowali większe zbliżenie kolorystyki testu do standardowego interfejsu systemu Windows. Na podstawie załączonych wykresów (rys. 1) można stwierdzić, że niezależnie od przynależności do jednej z trzech wyodrębnionych grup opinie na temat kolorystyki interfejsu były bardzo zbliżone.



Rys. 1. Ocena kolorystyki testu w zależności od przygotowania merytorycznego do egzaminu.

Znaczne rozbieżności zaobserwowano porównując wypowiedzi kobiet i mężczyzn.



Rys. 2. Rozbieżność ocen kolorystyki testu w grupach kobiet i mężczyzn.

Analizując kształt powyższych wykresów (rys. 2) można stwierdzić, iż kobiety preferują wielobarwny interfejs w odróżnieniu od mężczyzn uznających barwy stonowane (najchętniej kolory szare i biel – podobnie jak w systemie operacyjnym Windows).

Możliwość natychmiastowej korekty dokonanego wyboru

W grupach: I, II i III wysoką ocenę mechanizmu natychmiastowej korekty dokonanego wyboru deklarowało odpowiednio: 96%, 92%, 91% respondentów. 90% kobiet i 97% mężczyzn wysoko oceniło prezentowany mechanizm. W wyniku przeprowadzonego wywiadu stwierdzono, iż 18% mężczyzn i 29% kobiet co najmniej raz zmieniło wybraną uprzednio wersję rozwiązania. Zaobserwowano 9 osób, które w sytuacji kończącego się czasu

losowały warianty odpowiedzi. Większość respondentów oceniała pozytywnie zaproponowany mechanizm. Pojawiły się jednak głosy sugerujące, iż takie rozwiązanie umożliwia zachowania nieuczciwe (3 respondentów).

Dobór i integracja rozwiązań technicznych oraz ich wpływ na jakość i tempo pracy

W grupach: I, II i III pozytywny wpływ zastosowanych rozwiązań technicznych w zintegrowanej postaci (interaktywna instrukcja, informacja o miejscu i liczbie opuszczonych zadań, mechanizm eliminujący błędy techniczne, kolorystyka i struktura interfejsu, forma wyboru odpowiedzi) na jakość i tempo pracy stwierdziło odpowiednio: 96%, 62% oraz 85% respondentów. Zgodnie z przyjętymi założeniami powyższe wyniki oraz wypowiedzi studentów wskazują, iż najistotniejszym czynnikiem wpływającym na tempo i jakość pracy jest przygotowanie merytoryczne. Jednak rozbieżności szczególnie między grupą I i dwoma pozostałymi sugerują, że respondenci z grupy I, poświęcający znaczną ilość czasu na analizę treści zadań, w pewnym stopniu rekompensowali tę stratę np. możliwością szybkiego zaznaczenia wybranego rozwiązania lub jego korekty. Potwierdzeniem powyższego stwierdzenia wydają się być odpowiednie wyniki kobiet i mężczyzn: 80% oraz 85%. Rozbieżność ta wynika najprawdopodobniej z faktu, iż mężczyźni stanowią około jednej trzeciej (28%) wszystkich respondentów, 42% spośród nich zakwalifikowało się do grupy I, gdzie 93% respondentów pozytywnie oceniło proponowane rozwiązania. W wyniku przeprowadzonego wywiadu zanotowano wypowiedzi dotyczące wpływu zastosowanych rozwiązań na tempo pracy. W ośmiu przypadkach respondenci zdecydowanie twierdzili, iż test elektroniczny wypełniali znacznie wolniej niż jego papierowy odpowiednik. Swoje opinie tłumaczyli koniecznością opanowania wielu czynności manualnych i obawą przed utratą efektów własnej pracy zawartej w wirtualnym formularzu. Osiemnaście osób nie stwierdziło wpływu konstrukcji testu i zawartych w nim mechanizmów na tempo pracy. Pozostali respondenci wypowiadali się bardzo pozytywnie na temat oszczędności czasu, twierdząc, że mieli go znacznie więcej niż w przypadku testowania tradycyjnego.

Na podstawie przeprowadzonych badań wysunięto następujące wnioski:

1. Nie potwierdziły się założenia dotyczące poprawy efektywności pracy z testem w wyniku załączenia interaktywnej instrukcji. Rozwiązaniem może być mechanizm wymuszający wcześniejsze zapoznanie się z instrukcją.
2. Hipoteza dotycząca usprawnienia pracy poprzez wybór części odpowiedzi za pomocą kursora myszki potwierdziła się tylko częściowo. U osób sporadycznie korzystających z komputera można było zaobserwować brak sprawności manualnej szczególnie przy wyborze wariantu odpowiedzi w formularzach typu „pole kombi”. Rozwiązaniem

może być stosowanie formularzy typu: „pole wyboru” lub „przycisk opcji”, które jednak wymagają większej powierzchni.

3. W całości potwierdziły się przewidywania dotyczące opinii studentów na temat wpływu interakcji w teście na eliminację błędów technicznych, skrócenie czasu pracy i podniesienie jej komfortu.
4. W opinii mężczyzn kolorystyka testu powinna być zbliżona do proponowanej w standardowym interfejsie systemu operacyjnego Windows. Żywe barwy są akceptowane w większej mierze przez kobiety. Przygotowując test elektroniczny lub ankietę można zaproponować kilka wariantów kolorystycznych. Na podstawie wywiadu i obserwacji stwierdzono, iż wyróżnione barwne hiperłącza i przyciski nawigacyjne miały wpływ na zwiększenie tempa pracy.
5. Bardzo wysokie oceny mechanizmu umożliwiającego anulowanie bez żadnych konsekwencji dokonanego uprzednio wariantu odpowiedzi wskazują, iż tego typu rozwiązanie ma niewątpliwy wpływ na podniesienie komfortu pracy. Jednak z punktu widzenia prawidłowości przebiegu procesu ewaluacji należy wziąć pod uwagę takie sytuacje, w których wybór odpowiedzi będzie możliwy tylko jednorazowo. Wyeliminuje to sytuacje, w których studenci bez większego zastanowienia losują warianty odpowiedzi, mając nadzieje na późniejszą korektę lub szczęśliwy traf. W minimalnym stopniu potwierdziły się oczekiwania dotyczące skrócenia czasu pracy. Wielu studentów pozostawiało zadania bez rozwiązania licząc na możliwość skontaktowania się z kolegami lub skorzystania z innych nieuczciwych rozwiązań.
6. Wysoka ocena wszystkich zawartych w teście elektronicznym mechanizmów świadczy o zasadności stosowania tego typu rozwiązań. Ułatwiają one pracę osobom niezorganizowanym i szczególnie wrażliwym na stres związany z tego rodzaju formą ewaluacji. Uproszczenie i organizacja procesu testowania wpływa na zwiększenie tempa pracy.

Przeprowadzone badania wykazały szereg prawidłowości, które zostaną wykorzystane przy konstruowaniu kolejnych testów elektronicznych. Celem przeprowadzonych badań nie było poszukiwanie nowych procedur testowania lub ich modyfikacja. Problem rzetelności, trafności i standaryzacji testów związany z ich zawartością merytoryczną i odpowiednim doбором zadań nie był analizowany. Autorzy skoncentrowali się wyłącznie na możliwości adaptacji istniejących rozwiązań do postaci elektronicznej. Zbadano wpływ dostępnych rozwiązań technicznych na efektywność i tempo pracy.

Otrzymane wyniki wskazują nowe obszary badań, np. nad dostosowaniem formy testu do indywidualnego systemu percepcji odbiorcy (stylu poznawczego). Analiza tych zagadnień pozwoli w przyszłości opracować szereg praktycznych rozwiązań, ściśle dostosowanych do wymogów ewaluacji wyników kształcenia dorosłych, ze szczególnym uwzględnieniem kształcenia na odległość.

Bibliografia:

1. KUŹNIAK I., Optymalizacja procesu kształcenia, Wydawnictwo Naukowe Uniwersytetu Adama Mickiewicza w Poznaniu, Poznań 1993.
2. NIEMIĘRKO B., Pomiar wyników kształcenia, WSiP, Warszawa 1999.
3. OSMAŃSKA-FURMANEK W., JĘDRYCZKOWSKI J., Przydatność prezentacji multimedialnych w kształceniu menedżerów w oparciu o holistyczną ewaluację osiągnięć studentów, [w:] Multimedia w biznesie, Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków 1999.
4. TUROS L., Andragogika ogólna, Wydawnictwo Akademickie „Żak”, Warszawa 1999.
5. ZIMBARDO P. G., Psychologia i życie, PWN, Warszawa 1999.