



ANNA STOLIŃSKA¹, MACIEJ BARTKOWSKI²

Analiza procesu percepcji danych przedstawianych na wykresach

Analysis of the Perception Process of Data Presented in Charts

¹ Doktor, Uniwersytet Pedagogiczny im. Komisji Edukacji Narodowej w Krakowie, Wydział Matematyczno-Fizyczno-Techniczny, Instytut Informatyki, Zakład Badań Edukacyjnych i Nowych Mediów, Polska

² Student, Uniwersytet Pedagogiczny im. Komisji Edukacji Narodowej w Krakowie, Wydział Matematyczno-Fizyczno-Techniczny, Instytut Informatyki, Zakład Badań Edukacyjnych i Nowych Mediów, Polska

Streszczenie

W artykule przedstawiono wyniki badań, w których wykorzystano technikę eyetrackingową w celu analizy procesu odczytywania informacji przedstawionych za pomocą wykresów. Badanie przeprowadzono wśród studentów Uniwersytetu Pedagogicznego w Krakowie. Rozważany problem dotyczy możliwości identyfikacji wzorców przetwarzania informacji za pomocą techniki śledzenia ruchów gałek ocznych.

Słowa kluczowe: percepcja wykresów, wzorce przetwarzania informacji, badania edukacyjne, okulografia

Abstract

This article presents the results of research, which used eye tracking technology to analyse the process of reading data from a graph. The survey was conducted among students of Pedagogical University of Cracow. The problem relates to the ability to identify patterns of information processing by means of eye tracking technology.

Keywords: graph perception, patterns of information processing, educational research, eye tracking

Wstęp

Graficzna reprezentacja danych wykorzystywana jest powszechnie zarówno w opracowaniach naukowych, jak i tekstach o charakterze publicystycznym, informacyjnym. Wykresy i diagramy stanowią istotne wzbogacenie publikacji w formie tekstowej, ponieważ nierzadko pozwalają dostrzec pewne zależności między poszczególnymi elementami, które nie są widoczne w innym zestawieniu, np. w formie tabeli.

Umiejętność przedstawiania, odczytywania i interpretowania informacji przedstawianych na wykresach odgrywa również dużą rolę w wielu zawodach, takich jak np. statystyk czy analityk giełdowy, lekarz czy naukowiec prowadzący badania w dowolnej dziedzinie nauki, stąd można uznać, że jest to niezwykle przydatna kompetencja. Uczniowie nabywają ją na lekcjach matematyki już w szkole podstawowej, jednak okazuje się, że stosunkowo wielu z nich ma problem, który matematyk i dydaktyk Turnau (1990, s. 73) ujmuje następująco: „jak pokazują powtarzane od czasu do czasu badania, bardzo wielu uczniów nie potrafi z wykresu odczytywać prawie żadnych własności funkcji, a niemal cała ich wiedza o wykresie sprowadza się do umiejętności rozpoznawania z niego nazwy funkcji”. Również wg statystyk Centralnej Komisji Egzaminacyjnej z 2015 r. (CKE, 2015, s. 17), zadanie maturalne z matematyki, które polegało na poprawnym odczytaniu zbioru wartości funkcji z wykresu, poprawnie rozwiązało jedynie 22% zdających. Należy zauważyć, że cała trudność zadania polegała bardziej na właściwym zinterpretowaniu wykresu niż na znajomości odpowiedniego działu matematyki. Trudności uczniów są spowodowane m.in. tym, że nie rozumieją oni zasad, wg których tworzone są wykresy. Podczas analizy wykresów powinno się skupiać uwagę przede wszystkim na ich kluczowych elementach, przy czym ważna jest również umiejętność dostrzegania pewnych reguł obowiązujących podczas ich komponowania. Według niektórych badaczy poprawne czytanie wykresów wymaga również pewnej wprawy, którą nabywa się poprzez ich częste oglądanie (Biecek, 2014, s. 108).

W niniejszym artykule podjęto próbę odpowiedzi na pytania, w jakim stopniu technika eyetrackingowa pozwala zidentyfikować wzorcowe schematy postępowania podczas odczytywania danych z wykresów. Celem podjętych badań jest też ustalenie (na podstawie analizy aktywności wzrokowej połączonej z badaniami sondażowymi) przyczyn trudności związanych z odczytywaniem informacji w formie graficznej. Odpowiedź na pytania badawcze przyczyniła się do sformułowania zaleceń dotyczących tworzenia wykresów i sposobów uczenia ich odczytywania.

Konstrukcja wykresów a trudności w ich odczytywaniu

Wykresy zamieszczone w zadaniach, np. z fizyki, często zawierają wiele nieznaczących merytorycznie detali, które mogą być przyczyną popełnienia błędów. Niejednokrotnie są to jednak takie elementy, która na wykresie muszą się znajdować, ponieważ jakby z natury rzeczy. Należą do nich m.in. punkt przecięcia dwóch linii albo odstające wartości. Uczniowie często sugerują się nimi i w efekcie udzielają błędnych, nieprzemyślanych odpowiedzi. Takie „pułapki” odkrył Madsen (2011, s. 1–13) w badaniach przeprowadzonych w 2011 r., które dotyczyły różnic w uwadze wizualnej osób poprawnie rozwiązujących zadania z fizyki i osób popełniających błędy w tych zadaniach. Ustalono, że uczestnicy

badani, którzy prawidłowo odczytywali dane z diagramów, więcej uwagi skupiali na istotnych elementach wykresu. Ci natomiast, którzy odpowiadali błędnie, więcej czasu poświęcili analizie obszarów, które przyciągały uwagę wizualną, natomiast nie wносиły istotnych treści (kluczowych informacji). Ponadto dostrzeżono, że luki w wiedzy z danej dziedziny ilustrowanej za pomocą wykresów sprawiają, że stają się one właśnie owymi „pułapkami”, podświadomie sterującymi uwagą badanych.

Te i inne badania (Simkin, Hastie, 1987, s. 454–465) dowodzą, że podczas projektowania wykresów należy stosować pewne reguły, które zredukują liczbę potencjalnych dystraktorów (GUS, 2014, s. 21–22):

- powinny charakteryzować się prostotą i zwięzłym opisem,
- tytuł musi jednoznacznie definiować tematykę wykresu,
- osie powinny być odpowiednio oznaczone,
- dane na wykresach powinny być pokazane w sposób jasny i precyzyjny,
- nie poleca się stosowania wykresów 3D,
- liczba elementów nie powinna być zbyt duża ze względu na czytelność,
- należy unikać zbyt bogatej kolorystyki.

Powyższe zalecenia są dość ogólne i zarazem najbardziej użyteczne. Badania nad aktywnością wzrokową mogą pomóc w opracowaniu dodatkowych reguł, co przełoży się na lepszą jakość wizualizacji danych.

Metodologia badań

Cele badań

Przedmiotem badań omówionych w niniejszym artykule jest analiza procesu percepcji danych przedstawianych na wykresach. Badania skupiają się na dwóch aspektach – po pierwsze, sprawdzają, w jaki sposób ludzie odczytują dane przedstawione na wykresach, po drugie, dokonuje się w nich weryfikacji (z zastosowaniem techniki śledzenia ruchów gałek ocznych) formułowanych przez badaczy hipotez i wniosków, dotyczących czynników powodujących trudność w odczytywaniu zwizualizowanych danych.

Analiza literatury pozwoliła na sformułowanie hipotezy: na poprawność odczytywania danych i zwiększenie trudności w tym zakresie wpływa nieodpowiednia kolorystyka, źle dobrany typ wykresu (np. zastosowanie wykresu radarowego lub efektu 3D), niewyróżnione etykiety lub ich brak, brak praktycznego doświadczenia w korzystaniu z różnego rodzaju wykresów oraz zbyt duża liczba elementów (dystraktorów).

Metody, techniki i narzędzia badawcze

Do zrealizowania celów oraz zweryfikowania powyższych hipotez postanowiono zastosować dwie metody badawcze: eksperyment z zastosowaniem techniki okulograficznej oraz sondaż diagnostyczny (technika: ankieta), będący komplementarnym uzupełnieniem pierwszej metody.

Do przeprowadzenia badań zastosowano eyetracker SMI iView X™ Hi-Speed 500/1250. Jest to urządzenie stacjonarne, zapisujące (w przypadku zrealizowanych badań) strumień danych z częstotliwością 500 Hz. Metoda pomiaru jest nieinwazyjna, natomiast ruch oka jest rejestrowany przez specjalną kamerę znajdującą się wewnątrz aparatury. W odległości 50 cm od urządzenia znajdował się monitor, na którym wyświetlane były zadania.

Na potrzeby badań skonstruowano kwestionariusz ankiety, której pytania odnosiły się m.in. do wykresów przedstawionych w badaniu okulograficznym.

Procedura badawcza

Badanie zostało przeprowadzone w Instytucie Informatyki Uniwersytetu Pedagogicznego w Krakowie. Badaniom poddano łącznie 35 studentów informatyki, matematyki, administracji, biologii oraz psychologii. Studenci kierunków ścisłych stanowili najliczniejszą grupę (80% wszystkich badanych). Średnia wieku uczestników wyniosła 23 lata. Ponad 50% uczestników zadeklarowało, że korzysta z wykresów więcej niż kilka razy w miesiącu, a tylko 10 studentów (31,3%) korzysta z nich rzadziej niż raz w miesiącu. Przed każdym badaniem informowano uczestnika o tym, jak będzie ono przebiegać. Po odpowiednim ustawieniu urządzenia (tj. optymalnym dopasowaniu go do badanej osoby) dokonywana była kalibracja w celu uzyskania rzetelnych i dokładnych wyników.

Na potrzeby badań przygotowano 9 slajdów, z których każdy zawierał wykres i jedno lub dwa pytania z nim związane. Wykresy sporządzono tak, by w niektórych sytuacjach wprowadzić dystraktory i czynniki utrudniające odczytywanie z nich danych, niektóre natomiast zostały wykonane poprawnie, wręcz wzorcowo.

Zadaniem badanego była dokładna analiza wykresu oraz udzielenie ustnej odpowiedzi na pytania zamieszczone bezpośrednio pod diagramem. Po zakończeniu badania, które trwało przez ok.10 minut, uczestnik był proszony o wypełnienie kwestionariusza ankiety w wersji elektronicznej, w którym zamieszczono wyświetlane podczas eksperymentu diagramy oraz pytania dotyczące takich kwestii, jak: częstotliwość korzystania z wykresów, umiejętność ich tworzenia oraz utrudnienia związane z odczytywaniem poszczególnych wykresów.

Wyniki

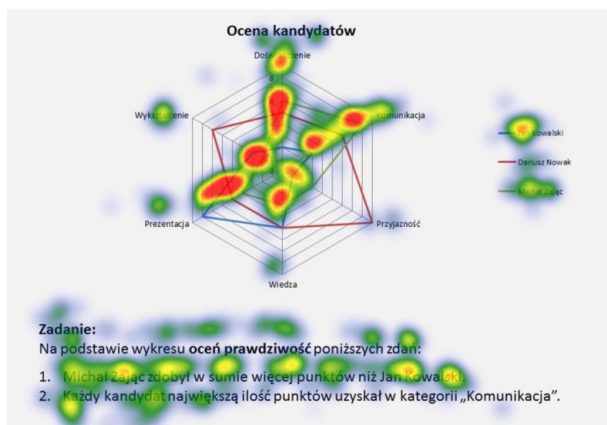
Wyniki badań eyetrackingowych zostały opracowane za pomocą aplikacji SMI BeGaze 3.7. W szczególności wzięto pod uwagę mapy cieplne, ścieżki wzroku, a także statystyki liczbowe na podstawie określonych obszarów zainteresowań (AOI). Należą do nich takie wskaźniki, jak: średni czas fiksacji, liczba fiksacji oraz czas patrzenia na wybrany obszar będący sumą czasów fiksacji i sakad. Ważną częścią były również ustne odpowiedzi badanych oraz wyniki uzyskane z ankiety. Wyniki 3 osób musiały zostać odrzucone ze względu na nieosiągnięcie założonego poziomu kalibracji.

Szczegółowe wyniki badań zostały przedstawione w pracy dyplomowej studenta informatyki Bartkowskiego (2017, s. 14–25). Poniżej zaprezentowano rezultaty uzyskane dla przykładowego (pierwszego) zadania.

W zadaniu został przedstawiony wykres radarowy z trzema seriami danych. Ponad 56% uczestników badań (18 osób) nigdy nie odczytywało danych z tego typu wykresu. Badany musiał udzielić odpowiedzi na dwa pytania. Pierwsze wymagało porównania dwóch serii celem ustalenia, który kandydat uzyskał największą liczbę punktów. Można je było rozwiązać na dwa sposoby: oszacować, która powierzchnia – ograniczona zieloną czy niebieską linią – jest większa lub zsumować liczbę punktów każdego kandydata. Poprawną odpowiedź na to pytanie udzieliły 23 osoby (71,88%). Drugie pytanie polegało na sprawdzeniu, czy maksymalna wartość w każdej serii danych była przypisana do tej samej, określonej kategorii. Tutaj bez wątplenia pomocne okazały się linie siatki. Wszyscy badani udzielili na to pytanie prawidłowej odpowiedzi.

Analizując mapy cieplne, czyli obszary, na których badani skupiali swoją uwagę, nie dało się zauważyć żadnych znaczących różnic. Każdy uczestnik, niezależnie od tego, czy udzielił poprawnej, czy błędnej odpowiedzi, koncentrował się na tych samych obszarach co inni. Analiza ścieżek wzroku również nie wykazała różnic. Średni czas patrzenia na wykres podczas rozwiązywania zadań wyniósł ok. 50 sekund.

Według badanych największą trudność podczas odczytywania danych z tego wykresu sprawiła duża liczba elementów (65,6%) oraz słabo widoczne etykiety/oznaczenia osi (50%). Utrudnieniem była również mała czcionka (15,6%), nieodpowiednia kolorystyka (9,4%). Dla 12,5% badanych nic nie utrudniało odczytywania danych.



Rysunek 1. Mapa cieplna odzwierciedlająca skupienie uwagi uczestnika badań podczas rozwiązywania zadania 1

Źródło: opracowanie własne.

Podsumowanie

Zarejestrowanie podczas eksperymentu ścieżek wzroku studentów przyczyniło się do wyszukania pewnego schematu, który zachodził wyłącznie wtedy, gdy badany udzielił właściwej odpowiedzi. Z uwagi na fakt, że prawidłowych odpowiedzi było zdecydowanie więcej niż niepoprawnych (w stosunku 7:3), wspomniany schemat można traktować jako pewnego rodzaju wzorzec, który prowadzi do bezbłędnej analizy danych zamieszczanych na wykresach.

Wzorzec ten wygląda następująco: należy 1) przeczytać tytuł wykresu, → 2) przeczytać całą treść zadania, włącznie z proponowanymi odpowiedziami, → 3) zapoznać się z wykresem: a) przeczytać legendę, b) zwrócić uwagę na rozmieszczenie elementów, c) zapoznać się z informacjami umieszczonymi na osiach, d) zapoznać się z etykietami, → 4) należy jeszcze raz przeanalizować treść zadania (dla przypomnienia): a) jeżeli zadanie składa się ono z więcej niż jednego pytania, to w pierwszej kolejności należy skupić się tylko na pojedynczym pytaniu; b) kolejne pytanie należy analizować dopiero po udzieleniu odpowiedzi lub wtedy, gdy odpowiedź na wcześniejsze pytanie nie jest znana, → 5) przystąpić do właściwej analizy wykresu: a) zwrócić uwagę na rodzaj wykresu i zinterpretować, co oznaczają jego składowe, b) znaleźć na nim elementy, które mają związek z pytaniem i ustalić odpowiadające im wartości, c) w razie potrzeby dokonać obliczeń (np. w pamięci), d) w razie potrzeby powrócić do treści pytania → 6) przed podaniem odpowiedzi można jeszcze raz powrócić do treści pytania.

Badanie potwierdziło hipotezy mówiące o tym, że nieodpowiednia kolorystyka, brak etykiet i zbyt duża liczba dystraktorów mają znaczący wpływ na poprawne odczytanie danych. Badanie zależności nie wykazało, by doświadczenie uczestnika badania w korzystaniu czy też tworzeniu wykresów miało wpływ na to, jakich udzielał odpowiedzi (współczynnik korelacji Spearmanar_s = -0,211; t_{kr} (α,30) = 2,0423). Nie wykazano również, by wykres radarowy sprawiał jakieś trudności w odczytywaniu danych, natomiast potwierdzona została hipoteza, iż wykresy sporządzone z wykorzystaniem efektu 3D zwiększały trudność związaną z odczytywaniem danych. Zdaniem uczestników badań zastosowanie tego efektu było największym utrudnieniem.

Badania pozwoliły wykazać, że eyetracking umożliwia szczegółową analizę aktywności wzrokowej – tego, jak człowiek odbiera informacje z otoczenia, a nawet wg jakiego schematu je przetwarza. Postępowanie zgodnie z tym schematem powinno naprowadzić odbiorcę na udzielenie właściwych odpowiedzi na pytania wymagające odczytywania wykresów i diagramów.

Konieczna wydaje się kontynuacja badań, które powinny być poszerzone o rzadziej wykorzystywane typy wykresów. Badaniu powinna również zostać poddana większa liczba osób, co umożliwi zastosowanie wnioskowania statystycznego w szerszym zakresie.

Literatura

- Bartkowski, M. (2017). *Technika eye trackingowa w badaniu wzorców przetwarzania informacji podczas analizy danych przedstawianych na wykresach*. Praca dyplomowa napisana pod kierunkiem dr Anny Stolińskiej. Kraków: Uniwersytet Pedagogiczny.
- Biecek, P. (2014). *Odkrywać! Ujawniać! Objaśnić! Zbiór esejów o sztuce prezentowania danych*. Warszawa: Fundacja Naukowa SmarterPoland.pl.
- CKE (2015). *Sprawozdanie z egzaminu maturalnego 2015. Matematyka*. Pobrane z: https://www.cke.edu.pl/images/_EGZAMIN_MATURALNY_OD_2015/Informacje_o_wynikach/2015/sprawozdanie/Sprawozdanie_matematyka_2015.pdf (17.04.2018).
- GUS (2014). *Graficzna prezentacja danych statystycznych. Wykresy, mapy, GIS*. Pobrane z: <https://stat.gov.pl/statystyka-regionalna/publikacje-regionalne/podreczniki-atlasy/podreczniki/graficzna-prezentacja-danych-statystycznych-wykresy-mapy-gis,2,1.html> (3.05.2018).
- Madsen, A., Larson, A., Loschky, L., Rebello, S. (2012). Differences in Visual Attention Between Those Who Correctly and Incorrectly Answer Physics Problems. *Physical Review Special Topics – Physics Education Research*, 8 (1), 1–13. DOI: <https://doi.org/10.1103/PhysRevSTPER.8.010122>.
- Simkin, D., Hastie, R. (1987). An Information-Processing Analysis of Graph Perception. *Journal of the American Statistical Association*, 82 (398), 454–465.
- Turnau, S. (1990). *Wykłady o nauczaniu matematyki*. Warszawa: PWN.