

ALFABETYZM STATYSTYCZNY A PUŁAPKI MYŚLENIA. ANALIZA LUKI KOMPETENCYJNEJ NA STYKU EDUKACJI ŚREDNIEJ I WYŻSZEJ

Ewa Wasilewska  <https://orcid.org/0000-0001-9065-5526>

Instytut Ekonomii i Finansów
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: ewa_wasilewska@sggw.edu.pl

Streszczenie: Celem artykułu jest ocena poziomu alfabetyzmu statystycznego uczniów szkół średnich i studentów oraz analiza wybranych komponentów tej kompetencji. Badanie przeprowadzono wśród 432 uczniów i studentów z wykorzystaniem autorskiego testu sytuacyjnego. Wyniki wskazują na istnienie luki kompetencyjnej pomiędzy efektami formalnej edukacji matematyczno-statystycznej a poziomem alfabetyzmu statystycznego ujawnianym podczas interpretacji danych i rozumowania probabilistycznego. Analiza ujawniła anomalię edukacyjną: uczniowie realizujący rozszerzony program matematyki wykazali wyższy poziom alfabetyzmu statystycznego niż studenci po akademickim kursie statystyki.

Słowa kluczowe: alfabetyzm statystyczny, luka kompetencyjna, edukacja statystyczna, heurystyki poznawcze, rozumowanie probabilistyczne, interpretacja danych

JEL classification: A21, A22, D91, C12, I21

WPROWADZENIE

We współczesnej gospodarce opartej na wiedzy i danych zdolność do rozumienia, interpretowania oraz krytycznej oceny informacji liczbowych staje się jedną z podstawowych kompetencji obywatelskich. W odpowiedzi na te wyzwania coraz większego znaczenia nabiera pojęcie alfabetyzmu statystycznego (statistical literacy), rozumianego jako zdolność do odczytywania, interpretowania, krytycznej oceny oraz komunikowania informacji statystycznych wykorzystywanych w życiu codziennym [Gal 2002]. W literaturze przedmiotu pojęcie to bywa zestawiane z pokrewnymi, lecz odrębnymi konstruktami – rozumowaniem statystycznym

<https://doi.org/10.22630/MIBE.2026.27.3.12>



(statistical reasoning) oraz myśleniem statystycznym (statistical thinking), przy czym nie ma pełnej zgodności co do precyzyjnych granic pomiędzy tymi pojęciami [Ben-Zvi i Garfield 2004]. Wysoki poziom alfabetyzmu statystycznego społeczeństwa stanowi istotny warunek podejmowania racjonalnych decyzji mikro- i makroekonomicznych. Jak podkreśla Oleński [2020], jakość statystyki publicznej oraz środowiska informacyjnego, w którym ona funkcjonuje, pełni funkcję istotnej bariery chroniącej przed dezinformacją i manipulowaniem danymi statystycznymi w przestrzeni publicznej. Rosnąca rola danych we współczesnym społeczeństwie, napędzana rozwojem technologii Big Data, dodatkowo wzmacnia potrzebę rozwijania kompetencji niezbędnych do krytycznej oceny i poprawnej interpretacji wyników badań [Szreder 2016].

Pomimo dynamicznego rozwoju narzędzi analitycznych, rzeczywista zdolność społeczeństwa do poprawnego i refleksyjnego rozumowania probabilistycznego wciąż pozostaje istotnym wyzwaniem. Przełomowe badania Amosa Tversky'ego i Daniela Kahnemana [1974], które dały początek współczesnej ekonomii behawioralnej, wykazały, że ludzie w warunkach niepewności często posługują się heurystykami poznawczymi – uproszczonymi regułami wnioskowania, które choć wymagają minimalnego nakładu wysiłku umysłowego i ułatwiają podejmowanie decyzji, mogą prowadzić do systematycznych błędów i złudzeń probabilistycznych. W realiach ekonomicznych brak świadomej kontroli nad tymi automatyzmami prowadzi do opisywanych w literaturze błędów poznawczych [Tversky i Kahneman 1974], których występowanie może utrudniać właściwą interpretację danych zarówno w życiu codziennym, jak i w praktyce gospodarczej.

Szczególnie krytycznym etapem na ścieżce edukacyjnej jest styk edukacji średniej i wyższej. Tradycyjny model dydaktyki nauk ilościowych kładzie silny nacisk na formalizm matematyczny, algorytmizację oraz pamięciowe opanowanie procedur obliczeniowych. Można jednak przypuszczać, że techniczna zdolność sprawnego posługiwania się aparatem matematycznym nie współwystępuje automatycznie z intuicją statystyczną i refleksyjną odpornością na zniekształcenia poznawcze. Zasadne staje się zatem zbadanie, w jakim stopniu formalna edukacja matematyczno-statystyczna sprzyja rozwojowi alfabetyzmu statystycznego oraz ogranicza podatność na błędy poznawcze ujawniające się podczas rozwiązywania problemów wymagających interpretacji danych i oceny niepewności.

PRZEGLĄD LITERATURY

Alfabetyzm statystyczny w procesie edukacji

Przywołana we Wprowadzeniu koncepcja Gala [2002] nosi pełną nazwę alfabetyzmu statystycznego dorosłych (Adult Statistical Literacy) i zakłada współwystępowanie dwóch komplementarnych komponentów. Pierwszy, poznawczy, obejmuje wiedzę statystyczną, umiejętności matematyczne oraz zdolność krytycznego myślenia. Drugi, dyspozycyjny, obejmuje postawy,

przekonania i motywację do podejmowania refleksyjnej oceny informacji. Rolę alfabetyzmu statystycznego w rozwoju kultury statystycznej oraz świadomym funkcjonowaniu społeczeństwa podkreślają również polscy autorzy, m.in. Domański [2016] i Oleński [2020].

Wraz z dynamicznym rozwojem technologii cyfrowych oraz rosnącą dostępnością dużych zbiorów danych, tradycyjne rozumienie alfabetyzmu statystycznego bywa obecnie rozszerzane o kompetencje określane jako data literacy, których rozwój Gould [2017] postuluje uwzględniać już na wczesnych etapach edukacji, między innymi poprzez wprowadzanie programów data science w szkołach średnich. Na gruncie polskim podobne stanowisko prezentują Jasiewicz i Stopnicka [2016], wskazując, że umiejętność krytycznej analizy danych staje się jednym z podstawowych elementów współczesnych kompetencji obywatelskich i zawodowych.

Jednocześnie liczne badania z zakresu edukacji statystycznej podkreślają, że opanowanie procedur obliczeniowych nie jest tożsame z rozwojem kompetencji związanych z interpretacją danych i rozumowaniem statystycznym. Batanero i współautorzy [2016] wskazują, że skuteczne kształcenie statystyczne wymaga rozwijania nie tylko wiedzy proceduralnej, lecz również rozumienia pojęć oraz umiejętności interpretacji wyników analiz. Podobne stanowisko prezentują Wild i Pfannkuch [1999], którzy podkreślają znaczenie kontekstu oraz procesu wnioskowania statystycznego. Ujęcie to znajduje również uzasadnienie w psychologii poznawczej, gdzie teoria dwuprocesowego przetwarzania informacji zakłada istnienie dwóch komplementarnych sposobów rozumowania: szybkiego, intuicyjnego i automatycznego (System 1) oraz wolniejszego, analitycznego i świadomie kontrolowanego (System 2). Dominacja przetwarzania charakterystycznego dla Systemu 1 sprzyja wykorzystywaniu heurystyk, które mogą prowadzić do błędów poznawczych podczas oceny danych i rozumowania probabilistycznego.

Nowsze badania empiryczne potwierdzają, że rozwój alfabetyzmu statystycznego na kolejnych etapach kształcenia nie ma charakteru jednorodnego. Kurnia, Lowrie i Patahuddin [2024] wykazali różnice między uczniami z różnych poziomów edukacyjnych, wskazując zarazem, że znaczna część badanych posługiwała się wiedzą statystyczną w sposób niekrytyczny. Z kolei Phadke, Beckman i Lock Morgan [2024] wykazali, że kontekst zadań może wpływać na wyniki pomiaru alfabetyzmu statystycznego studentów, co podkreśla znaczenie osadzania diagnozy tej kompetencji w realistycznych sytuacjach problemowych.

Przytoczone koncepcje wskazują, że alfabetyzm statystyczny wykracza poza sprawność wykonywania obliczeń i obejmuje zdolność poprawnego rozumowania w sytuacjach wymagających interpretacji danych oraz oceny niepewności. W praktyce oznacza to konieczność uwzględniania podczas kształcenia nie tylko aspektów rachunkowych, lecz również umiejętności krytycznej interpretacji i analizy informacji statystycznych.

Pułapki myślenia w rozumowaniu probabilistycznym – rola heurystyk poznawczych

Konsekwencją dominacji Systemu 1, opisanej w poprzednim podrozdziale, są systematyczne błędy poznawcze, które w niniejszym artykule określono jako pułapki myślenia [Kahneman 2011]. W literaturze przedmiotu szczególnie często opisywane są cztery rodzaje błędów poznawczych mających znaczenie dla rozumowania statystycznego. Jednym z nich jest niewrażliwość na wielkość próby, polegająca na traktowaniu niewielkich prób jako reprezentatywnych dla całej populacji, mimo że są one znacznie bardziej podatne na przypadkowe odchylenia od wartości oczekiwanych [Tversky i Kahneman 1971]. Istotnym zniekształceniem poznawczym jest również złudzenie gracza, przejawiające się błędnym przekonaniem, że losowe sekwencje zdarzeń powinny w krótkim okresie samoczynnie wyrównywać wcześniejsze odchylenia, co prowadzi do naruszenia zasady niezależności zdarzeń [Kahneman i Tversky 1972]. Do często analizowanych zniekształceń należy także niewrażliwość na zmienność danych, polegająca na koncentrowaniu uwagi wyłącznie na miarach położenia, z pominięciem informacji o rozproszeniu danych, co może prowadzić do błędnej oceny ryzyka i stabilności analizowanych zjawisk [Ben-Zvi i Garfield 2004]. Szczególne miejsce w literaturze zajmuje również błąd proporcji podstawowej (base rate fallacy), polegający na pomijaniu informacji o częstości występowania zjawiska na rzecz pojedynczych informacji szczegółowych, mimo że to właśnie proporcja podstawowa stanowi kluczowy element poprawnego rozumowania probabilistycznego [Bar-Hillel 1980]. Badania Gigerenzer'a i Hoffrage'a [1995] pokazują, że błąd ten występuje nawet u osób posiadających wysokie kwalifikacje zawodowe.

Opisane błędy poznawcze, będące konsekwencją działania heurystyk, stanowiły teoretyczną podstawę konstrukcji dystraktorów wykorzystanych w autorskim narzędziu badawczym. Celem badania nie był pomiar samych heurystyk, lecz ocena wybranych komponentów alfabetyzmu statystycznego w sytuacjach, w których heurystyki mogły prowadzić do wyboru niepoprawnych odpowiedzi.

CEL I METODYKA BADAŃ

Cel badań

Celem badania jest ocena poziomu alfabetyzmu statystycznego oraz identyfikacja ewentualnej luki kompetencyjnej między nim a poziomem formalnej edukacji matematyczno-statystycznej. Na potrzeby badania lukę tę zoperacjonalizowano jako brak oczekiwanego, jednokierunkowego wzrostu wyników testu wraz z kolejnymi poziomami formalnej edukacji matematyczno-statystycznej, ze szczególnym uwzględnieniem porównania studentów po ukończeniu akademickiego kursu statystyki z uczniami realizującymi matematykę w zakresie podstawowym i rozszerzonym. W ramach badania analizowano również

podatność respondentów na wybrane błędy poznawcze ujawniające się podczas rozumowania probabilistycznego. Szczególną uwagę poświęcono umiejętności poprawnej interpretacji informacji statystycznych w sytuacjach wymagających oceny ryzyka, niepewności oraz prawdopodobieństwa.

W nawiązaniu do omówionych wcześniej pułapek myślenia wynikających z działania heurystyk poznawczych analizowano w szczególności, czy wzrost poziomu formalnego wykształcenia matematyczno-statystycznego wiąże się z ograniczeniem podatności na wybrane błędy poznawcze opisywane w literaturze dotyczącej rozumowania probabilistycznego.

W związku z tak sformułowanym celem postawiono następujące pytania badawcze:

- PB1. Czy poziom alfabetyzmu statystycznego różni się pomiędzy uczniami szkół ponadpodstawowych realizującymi matematykę na poziomie podstawowym i rozszerzonym oraz studentami kierunków ekonomiczno-menedżerskich?
- PB2. Które obszary alfabetyzmu statystycznego sprawiają największe trudności poszczególnym grupom respondentów?
- PB3. Czy wraz ze wzrostem poziomu formalnej edukacji matematyczno-statystycznej zmniejsza się podatność na wybrane błędy rozumowania probabilistycznego?

Materiał badawczy i przebieg badania

Badanie zrealizowano w okresie maj–czerwiec 2026 roku metodą sondażu diagnostycznego z wykorzystaniem celowego doboru próby ($N = 432$ respondentów). Strukturę próby tworzą trzy grupy:

- Uczniowie (matematyka w zakresie podstawowym), $N = 123$: uczniowie klas trzecich (przedmaturalnych) dwóch publicznych liceów ogólnokształcących, realizujący matematykę w zakresie podstawowym.
- Uczniowie (matematyka w zakresie rozszerzonym), $N = 41$: uczniowie klas trzecich tych samych placówek, realizujący program rozszerzony.
- Studenci, $N = 268$: studenci I i II roku kierunków ekonomiczno-menedżerskich dwóch publicznych uczelni w Polsce centralnej, którzy pomyślnie ukończyli akademicki kurs statystyki.

Dobór uczniów klas przedmaturalnych był celowym zabiegiem metodologicznym: na tym etapie nie realizowali oni jeszcze szkolnego modułu statystyki i rachunku prawdopodobieństwa. Pozwoliło to na wyizolowanie ich naturalnej sprawności analitycznej, tworząc punkt odniesienia dla studentów posiadających sformalizowaną wiedzę akademicką. Warto zaznaczyć, że liczebność grupy realizującej program rozszerzony ($N = 41$) odzwierciedla rzeczywistą proporcję uczniów wybierających matematykę rozszerzoną w badanych placówkach. Nierównoliczność grup uwzględniono za pomocą testu t-Welcha.

Sesje badawcze przeprowadzono w kontrolowanych warunkach audytoryjnych techniką ankiety papierowej (PAPI), co wyeliminowało użycie

kalkulatorów, narzędzi AI czy pomocy zewnętrznych. Udział był dobrowolny i anonimowy. Aby zminimalizować presję czasową i ułatwić aktywację Systemu 2, czas odpowiedzi nie był ograniczany. Instrukcja celowo nie sugerowała matematycznego charakteru zadań, zapobiegając sztucznemu wywołaniu podwyższonej czujności poznawczej.

Z uwagi na specyfikę celów badawczych, w pracy zastosowano nielosowy dobór celowy, ukierunkowany na wyizolowanie grup respondentów o ściśle zdefiniowanych profilach edukacyjnych. Choć nielosowy charakter próby nakłada pewne ograniczenia w zakresie bezpośredniej generalizacji wyników na całą populację, zastosowanie aparatu statystyki indukcyjnej (testów chi-kwadrat (χ^2) oraz t-Welcha) posiada silne uzasadnienie metodologiczne. Testy istotności służą jako formalne kryterium oceny, czy zaobserwowane dysproporcje między grupami mają charakter systematyczny, czy też są dziełem przypadku wewnątrz badanej próby. Dodatkowo, w celu uniezależnienia wnioskowania od specyfiki rekrutacji próby, analizę uzupełniono o badanie wielkości efektu (współczynniki V Craméra), co pozwala na rzetelną ocenę rzeczywistej siły zidentyfikowanych zależności. W konsekwencji wyniki testów istotności należy interpretować jako opis relacji zaobserwowanych w obrębie badanej próby, nie zaś jako podstawę do bezpośredniej generalizacji na całą populację uczniów i studentów w Polsce.

Dla przeprowadzonych testów podano wartości statystyk testowych oraz odpowiadające im wartości p (p-value).

Konstrukcja i charakterystyka narzędzia badawczego

Narzędzie badawcze stanowił autorski test sytuacyjny obejmujący cztery zamknięte zadania jednokrotnego wyboru, osadzone w realistycznych kontekstach społeczno-gospodarczych. Test nie miał charakteru sprawdzianu wiedzy szkolnej, lecz został zaprojektowany jako narzędzie diagnostyczne służące ocenie wybranych komponentów alfabetyzmu statystycznego ujawniających się w sytuacjach wymagających interpretacji danych, oceny ryzyka oraz rozumowania probabilistycznego.

Konstrukcja narzędzia została oparta na klasycznych problemach opisywanych w literaturze dotyczącej alfabetyzmu statystycznego, rozumowania statystycznego oraz psychologii poznawczej [Gal 2002; Ben-Zvi i Garfield 2004; Kahneman 2011]. Poszczególne zadania odwoływały się do czterech opisanych w literaturze wzorców błędnego rozumowania: prawa małych liczb, złudzenia gracza, niewrażliwości na zmienność danych oraz błędu proporcji podstawowej, szczegółowo omówionych w części Przeglądu literatury poświęconej pułapkom myślenia w rozumowaniu probabilistycznym. Zmieniono kontekst sytuacyjny, osadzając zadania w realiach współczesnej gospodarki i życia społecznego, co miało zwiększyć ich komunikatywność oraz ograniczyć możliwość odwoływania się przez respondentów do zapamiętanych schematów szkolnych. Każde zadanie reprezentowało odmienny komponent alfabetyzmu statystycznego.

Poprawne rozwiązanie każdego zadania wymagało analitycznej interpretacji informacji statystycznych, charakterystycznej dla Systemu 2, natomiast odpowiedzi niepoprawne (dystraktory) skonstruowano tak, aby odpowiadały wzorcom intuicyjnego rozumowania wiążanym w literaturze z działaniem Systemu 1. Wybór dystraktora wskazywał zatem na odpowiedź zgodną z określonym wzorcem błędnego rozumowania, nie stanowił jednak bezpośredniego pomiaru zastosowania konkretnej heurystyki.

Przed rozpoczęciem badania właściwego narzędzie poddano pilotażowi z udziałem nauczyciela matematyki oraz pięciu uczniów szkoły ponadpodstawowej. Celem pilotażu była ocena zrozumiałości poleceń, jednoznaczności odpowiedzi oraz poprawności konstrukcji zadań. Zgłoszone uwagi posłużyły do doprecyzowania treści pytań i instrukcji. Ogólną strukturę logiczną narzędzia przedstawiono w Tabeli 1.

Tabela 1. Specyfikacja logiczno-strukturalna zadań testowych

Zadanie	Badany komponent alfabetyzmu statystycznego	Dominująca pułapka poznawcza (System 1)	Kompetencja wymagana do poprawnego rozwiązania (System 2)
Z.1: Influencerzy (Zasięgi a proporcja polubień)	Wpływ liczebności próby na zmienność wyników	Prawo małych liczb (heurystyka reprezentatywności)	Ocena wpływu wielkości próby na zmienność wyników
Z.2: Logistyka (Sekwencje tirów przy bramkach)	Rozumienie losowości i niezależności zdarzeń	Złudzenie gracza	Rozpoznanie niezależności kolejnych zdarzeń losowych
Z.3: Zarządzanie (Ocena efektywności pracowników)	Interpretacja zróżnicowania danych	Pomijanie zmienności (koncentracja wyłącznie na średniej)	Wykorzystanie miar rozproszenia do oceny stabilności wyników
Z.4: Systemy AI (Wykrywanie podróbek ofert)	Interpretacja prawdopodobieństwa warunkowego	Błąd proporcji podstawowej	Uwzględnienie częstości bazowej przy ocenie prawdopodobieństwa

Źródło: opracowanie własne

WYNIKI BADAŃ

Analiza wyników skrajnych – podatność i odporność na błędy poznawcze

Pierwszym etapem analizy empirycznej była ocena wyników skrajnych (0 lub 4 punkty). Wynik 0 punktów interpretowano jako pełną podatność na analizowane błędy poznawcze, natomiast wynik 4 punktów jako pełną odporność na te błędy, rozumianą jako zdolność do konsekwentnego odrzucania intuicyjnych, lecz błędnych odpowiedzi i stosowania poprawnego rozumowania analitycznego we wszystkich zadaniach testowych. Wyniki pośrednie (1–3 pkt) poddano odrębnej analizie w kolejnych częściach pracy.

Rozkład wyników skrajnych w poszczególnych grupach badawczych przedstawiono w Tabeli 2.

Tabela 2. Zestawienie wyników skrajnych (0 i 4 punkty) w zależności od profilu edukacyjnego badanych

Grupa badawcza	N	Pełna podatność na analizowane błędy poznawcze (0 pkt)	Pełna odporność na analizowane błędy poznawcze (4 pkt)
Uczniowie (matematyka w zakresie podstawowym)	123	15 (12,2%)	3 (2,4%)
Uczniowie (matematyka w zakresie rozszerzonym)	41	0 (0,0%)	23 (56,1%)
Studenci kierunków ekonomiczno-menedżerskich	268	48 (17,9%)	23 (8,6%)

Objaśnienia: procenty w nawiasach odnoszą się do całkowitej liczebności danej grupy.

Źródło: opracowanie własne na podstawie badań empirycznych

Analiza rozkładu odpowiedzi ujawnia głębokie zróżnicowanie badanych grup:

- Niemal powszechna odporność uczniów z programem rozszerzonym – w tej grupie nie odnotowano ani jednego przypadku całkowitej uległości wobec dystraktorów (0,0%), a ponad połowa respondentów (56,1%) rozwiązała bezbłędnie wszystkie zadania.
- Wyraźna podatność studentów kierunków ekonomiczno-menedżerskich – odsetek osób osiągających wynik zerowy (17,9%) okazał się najwyższy spośród badanych grup i ponad dwukrotnie przewyższał odsetek wyników maksymalnych

(8,6%), co czyni tę grupę najbardziej podatną na intuicyjne pułapki myślenia spośród wszystkich trzech.

- Uczniowie z programem podstawowym bliżej studentów niż rówieśników z rozszerzeniem – niewielki odsetek wyników maksymalnych (2,4%) współwystępował z relatywnie wysokim udziałem wyników zerowych (12,2%).

Istotność statystyczną zależności między profilem edukacyjnym respondentów a poziomem podatności i odporności na błędy poznawcze zweryfikowano testem chi-kwadrat (χ^2). Założenia reguły Cochra zostały spełnione (minimalna liczebność oczekiwana wyniosła 7,88), co pozwoliło na prawidłowe przeprowadzenie testu pomimo zerowej liczebności w jednej z komórek tabeli. Analiza wykazała istotny związek między profilem edukacyjnym respondentów a występowaniem wyników skrajnych ($\chi^2(2) = 38,66$; $p < 0,001$). W celu oceny siły tej zależności obliczono współczynnik V Craméra ($V = 0,59$), którego wartość – zgodnie z kryteriami interpretacji efektu zaproponowanymi przez Cohena [1988] – wskazuje na dużą siłę związku. Te same kryteria interpretacji zastosowano konsekwentnie również w dalszej części pracy podczas analizy poszczególnych zadań.

Wyniki te dostarczają pierwszej empirycznej przesłanki odnoszącej się do pytania badawczego PB3: wyższy poziom formalnej edukacji matematyczno-statystycznej nie przekłada się automatycznie na większą odporność na analizowane błędy poznawcze. Co więcej, sam kierunek tej zależności bywa odwrotny do intuicyjnie oczekiwanego – zjawisko to wymaga pogłębionej analizy przedstawionej w dalszej części artykułu.

Rozkład wyboru dystraktorów wskazuje, że część odpowiedzi była zgodna z intuicyjnymi wzorcami przetwarzania informacji charakterystycznymi dla Systemu 1, podczas gdy poprawne rozwiązanie wymagało analitycznego rozumowania charakterystycznego dla Systemu 2. Ze względu na konstrukcję narzędzia nie można jednak utożsamiać wyboru konkretnego dystraktora z bezpośrednim pomiarem zastosowania określonej heurystyki.

Ogólny poziom alfabetyzmu statystycznego – analiza wyników średnich

Kolejnym etapem analizy była ocena ogólnego poziomu analizowanej kompetencji w badanych grupach, wyrażonego średnią liczbą poprawnych odpowiedzi. Statystyki opisowe, przedstawione w Tabeli 3, ujawniają wyraźną hierarchię wyników: najwyższy poziom kompetencji w badanym zakresie i największą jednorodność osiągnęli uczniowie realizujący matematykę w zakresie rozszerzonym. Odmienny obraz uzyskano wśród studentów kierunków ekonomiczno-menedżerskich, którzy osiągnęli średni wynik zbliżony do uczniów realizujących matematykę w zakresie podstawowym, przy jednocześnie największym rozproszeniu wyników.

Tabela 3. Statystyki opisowe wyników testu w badanych grupach

Grupa badawcza	Średnia	Odchylenie standardowe	Współczynnik zmienności
Uczniowie (matematyka w zakresie podstawowym)	1,46	0,86	58,9%
Uczniowie (matematyka w zakresie rozszerzonym)	3,12	1,17	37,4%
Studenci kierunków ekonomiczno-menedżerskich	1,62	1,18	72,6%

Źródło: opracowanie własne na podstawie badań empirycznych

Test Shapiro–Wilka wykazał odstępstwa od rozkładu normalnego we wszystkich grupach: uczniów realizujących matematykę w zakresie podstawowym ($W = 0,864$; $p < 0,001$), uczniów realizujących matematykę w zakresie rozszerzonym ($W = 0,722$; $p < 0,001$) oraz studentów ($W = 0,900$; $p < 0,001$). Test Browna–Forsythe’a wskazał ponadto na brak homogeniczności wariancji ($F(2, 429) = 4,06$; $p = 0,018$). Ze względu na nierówne liczebności grup i zróżnicowanie wariancji do porównania średnich zastosowano dwustronny test t-Welcha dla prób niezależnych. Przy liczebnościach analizowanych grup test ten pozostaje dostatecznie odporny na odstępstwa od normalności; stabilność uzyskanych wniosków zweryfikowano dodatkowo testem U Manna–Whitneya. Ponieważ przeprowadzono trzy porównania, zastosowano poprawkę Bonferroniego, przyjmując skorygowany poziom istotności $\alpha = 0,0167$.

Porównania parami średnich wyników ujawniają wyraźny wzorec:

- Rozszerzony program wyprzedza podstawowy – różnica średnich okazała się istotna statystycznie ($t(55,29) = 8,38$; $p < 0,001$), na korzyść uczniów realizujących matematykę w zakresie rozszerzonym.
- Rozszerzony program wyprzedza też studia – stwierdzono istotną statystycznie różnicę ($t(53,22) = 7,68$; $p < 0,001$), również na korzyść uczniów z programem rozszerzonym.
- Studenci nie odróżniają się od uczniów z programem podstawowym – nie stwierdzono istotnej statystycznie różnicy średnich ($t(314,64) = 1,47$; $p = 0,141$). Wartość p przekroczyła zarówno klasyczny, jak i skorygowany poziom istotności, co oznacza brak podstaw do odrzucenia hipotezy o braku różnic między średnimi wynikami obu grup. W badanej próbie średni poziom kompetencji studentów po ukończeniu akademickiego kursu statystyki nie różnił się istotnie od poziomu reprezentowanego przez uczniów realizujących matematykę wyłącznie w zakresie podstawowym.

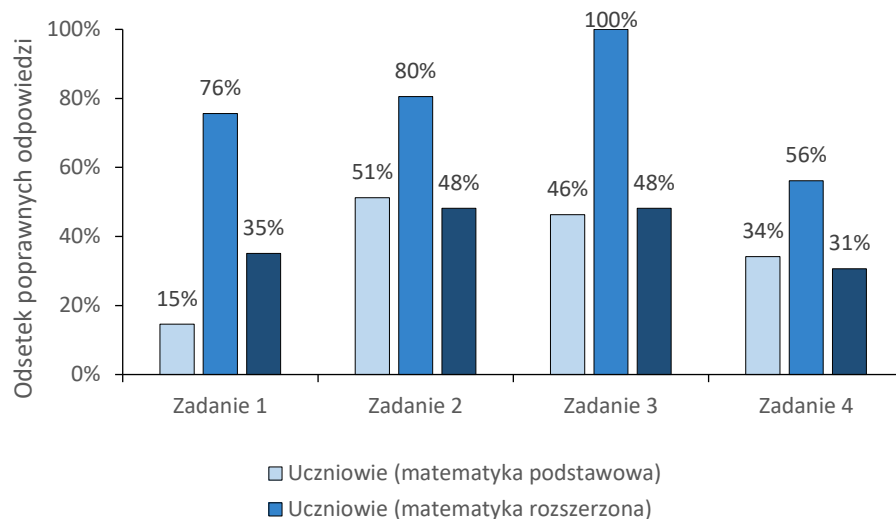
Kontrolne porównania testem U Manna–Whitneya potwierdziły wszystkie rozstrzygnięcia uzyskane testem Welcha: dla uczniów realizujących matematykę w zakresie rozszerzonym i podstawowym uzyskano $U = 802,5$; $p < 0,001$, dla uczniów realizujących matematykę w zakresie rozszerzonym i studentów $U = 8822,5$; $p < 0,001$, natomiast dla uczniów realizujących matematykę w zakresie podstawowym i studentów $U = 15705$; $p = 0,435$.

Uzyskane wyniki stanowią bezpośrednią odpowiedź na pytania badawcze PB1 i PB3. Analiza wykazała istotne zróżnicowanie ogólnego poziomu alfabetyzmu statystycznego pomiędzy badanymi grupami oraz wskazała, że wzrost poziomu formalnej edukacji matematyczno-statystycznej nie prowadzi automatycznie do wyższych rezultatów w tym zakresie. Świadczy o tym fakt, że studenci, pomimo ukończenia akademickiego kursu statystyki, nie uzyskali wyników lepszych od uczniów realizujących matematykę w zakresie rozszerzonym – to właśnie ta grupa osiągnęła najwyższy poziom analizowanej kompetencji. Anomalia ta wskazuje, że sam wzrost poziomu formalnej edukacji matematyczno-statystycznej nie gwarantuje ani wyższego poziomu alfabetyzmu statystycznego, ani większej odporności na analizowane błędy poznawcze. Możliwe przyczyny tego zjawiska omówiono w części poświęconej dyskusji wyników.

Analiza poszczególnych komponentów alfabetyzmu statystycznego

W kolejnym etapie analizy oceniono odsetek poprawnych odpowiedzi uzyskanych w poszczególnych zadaniach testowych. Pozwoliło to ustalić, które obszary alfabetyzmu statystycznego sprawiały respondentom największe trudności oraz w jakim stopniu wyniki różniły się pomiędzy badanymi grupami. Odsetek poprawnych odpowiedzi w podziale na profil edukacyjny badanych przedstawiono na Rysunku 1.

Rysunek 1. Odsetek poprawnych odpowiedzi w poszczególnych zadaniach testowych z podziałem na profil edukacyjny badanych



Źródło: opracowanie własne

Zależność pomiędzy profilem kształcenia a poprawnością odpowiedzi w poszczególnych zadaniach testowych zweryfikowano za pomocą testu chi-kwadrat (χ^2). Każde z analizowanych zadań zostało zaprojektowane w celu oceny innego komponentu, przedstawionego w części dotyczącej konstrukcji narzędzia badawczego. Ze względu na przeprowadzenie czterech niezależnych testów na tym samym zbiorze danych zastosowano poprawkę Bonferroniego, przyjmując skorygowany poziom istotności $\alpha = 0,0125$. Wszystkie uzyskane wartości p okazały się niższe od tego progu, co potwierdza istotne statystycznie zróżnicowanie poprawności odpowiedzi pomiędzy badanymi grupami.

Cztery zadania testowe ujawniają odmienne oblicza tej samej prawidłowości. Dla każdego zadania przedstawiono wartość statystyki χ^2 , odpowiadającą jej wartość p oraz współczynnik V Craméra jako miarę siły związku:

- Wpływ wielkości próby (Zadanie 1) – najwyższy odsetek poprawnych odpowiedzi uzyskali uczniowie realizujący matematykę w zakresie rozszerzonym (76%). Niższe wyniki odnotowano wśród studentów (35%) oraz uczniów realizujących matematykę w zakresie podstawowym (15%). Analiza wykazała istotny statystycznie związek między profilem edukacyjnym respondentów a poprawnością odpowiedzi ($\chi^2(2) = 52,87$; $p < 0,001$), a wartość współczynnika V Craméra ($V = 0,35$) wskazuje na umiarkowaną siłę tego związku.
- Rozumienie losowości (Zadanie 2) – 80% uczniów realizujących matematykę w zakresie rozszerzonym poprawnie rozpoznało niezależność zdarzeń, podczas

gdy wyniki pozostałych dwóch grup oscylowały wokół poziomu losowego zgadywania (51% i 48%). Rozkład poprawnych odpowiedzi pozostawał w istotnym statystycznie związku z profilem edukacyjnym respondentów ($\chi^2(2) = 14,97$; $p = 0,0006$). Wartość współczynnika V Craméra ($V = 0,19$) wskazuje jednak na słabą siłę tego związku.

- Interpretacja zmienności danych (Zadanie 3) – wszyscy uczniowie realizujący matematykę w zakresie rozszerzonym rozwiązyli zadanie poprawnie (100%), podczas gdy studenci oraz uczniowie realizujący matematykę w zakresie podstawowym uzyskali wyniki ponad dwukrotnie niższe (odpowiednio 48% i 46%). Zależność między profilem edukacyjnym respondentów a poprawnością odpowiedzi okazała się istotna statystycznie ($\chi^2(2) = 41,02$; $p < 0,001$), a wartość współczynnika V Craméra ($V = 0,31$) wskazuje na umiarkowaną siłę tego związku. Warto zauważyć, że jest to jedyne z czterech zadań, w którym wszyscy uczniowie realizujący program rozszerzony udzielili poprawnej odpowiedzi.
- Prawdopodobieństwo warunkowe (Zadanie 4) – komponent ten okazał się najtrudniejszy dla uczniów realizujących matematykę w zakresie rozszerzonym oraz dla studentów, którzy uzyskali tu swoje najniższe wyniki spośród wszystkich czterech zadań (odpowiednio 56% i 31%). Wyjątek stanowili uczniowie realizujący matematykę w zakresie podstawowym, dla których niższy wynik odnotowano jedynie w Zadaniu 1 (15%); w Zadaniu 4 uzyskali oni 34% poprawnych odpowiedzi. Poprawność odpowiedzi pozostawała w istotnym statystycznie związku z profilem edukacyjnym respondentów ($\chi^2(2) = 10,30$; $p = 0,0058$). Wartość współczynnika V Craméra ($V = 0,15$) wskazuje jednak na słabą siłę tego związku.

Analiza poszczególnych komponentów badanej kompetencji pozwoliła odpowiedzieć na pytanie badawcze PB2. Największe trudności respondentom sprawiało prawdopodobieństwo warunkowe – zwłaszcza uczniom z programem rozszerzonym oraz studentom – natomiast dla uczniów z programem podstawowym największym wyzwaniem okazało się rozumienie wpływu wielkości próby. We wszystkich analizowanych obszarach najwyższe odsetki poprawnych odpowiedzi uzyskali uczniowie realizujący matematykę w zakresie rozszerzonym. Wyniki te uszczegóławiają rezultaty przedstawione w części poświęconej porównaniu średnich wyników grup. Analiza siły zależności miarą V Craméra ujawniła, że za zróżnicowanie poziomu kompetencji pomiędzy badanymi grupami w największym stopniu odpowiadały komponenty związane z wielkością próby (Zadanie 1) oraz interpretacją zmienności danych (Zadanie 3), gdzie odnotowano umiarkowaną siłę związku.

Podsumowując, analiza poszczególnych komponentów potwierdziła, że wzrost poziomu formalnej edukacji matematyczno-statystycznej nie zawsze przekłada się na wyższy poziom badanej kompetencji. Obserwacja ta jest zgodna z przyjętym w niniejszym artykule rozumieniem luki kompetencyjnej jako rozbieżności pomiędzy efektami formalnej edukacji matematyczno-statystycznej

a poziomem alfabetyzmu statystycznego ujawnianym podczas interpretacji danych i rozumowania probabilistycznego.

WNIOSKI I DYSKUSJA

Synteza wyników

Przeprowadzone badanie wykazało trzy najważniejsze prawidłowości:

- Zróźnicowanie kompetencyjne między grupami – poziom alfabetyzmu statystycznego istotnie różnił się pomiędzy analizowanymi grupami, przy czym najwyższe wyniki uzyskali uczniowie realizujący matematykę w zakresie rozszerzonym.
- Zróźnicowanie trudności poszczególnych komponentów kompetencji – uczniom realizującym matematykę w zakresie rozszerzonym oraz studentom największą trudność sprawiało prawdopodobieństwo warunkowe, natomiast w grupie uczniów realizujących matematykę w zakresie podstawowym najniższy wynik odnotowano w zadaniu dotyczącym wpływu wielkości próby.
- Brak jednokierunkowego wzrostu kompetencji wraz z poziomem edukacji – studenci, mimo ukończenia akademickiego kursu statystyki, nie uzyskali statystycznie istotnej przewagi nad uczniami realizującymi matematykę w zakresie podstawowym. Najwyższy poziom badanej kompetencji stwierdzono natomiast wśród uczniów realizujących matematykę w zakresie rozszerzonym.

Dyskusja wyników

Przeprowadzone badanie ujawniło w analizowanej próbie swoistą anomalię edukacyjną: wyniki nie układały się zgodnie z oczekiwanym, jednokierunkowym wzrostem alfabetyzmu statystycznego w grupach reprezentujących kolejne etapy formalnej edukacji. Prawidłowość ta dotyczyła również odporności na analizowane błędy poznawcze. Wyniki nie wskazują zatem, aby osoby po ukończeniu kursu statystyki były w badanej próbie bardziej odporne na uleganie heurystykom. Podobny brak jednorodności rozwoju alfabetyzmu statystycznego pomiędzy kolejnymi poziomami kształcenia odnotowali również Kurnia, Lowrie i Patahuddin [2024], co wskazuje, że stwierdzona anomalia nie stanowi odosobnionego przypadku w literaturze przedmiotu.

Jednym z możliwych wyjaśnień uzyskanego układu wyników jest odmienny charakter kompetencji rozwijanych w toku edukacji szkolnej i akademickiej. Nauczanie statystyki na poziomie akademickim koncentruje się często na opanowaniu metod obliczeniowych i procedur statystycznych, podczas gdy zadania wykorzystane w niniejszym badaniu wymagały przede wszystkim interpretacji informacji, krytycznej oceny danych oraz świadomego wnioskowania w warunkach niepewności. Z kolei systematyczne rozwiązywanie w szkole średniej zadań wymagających uzasadniania i weryfikowania własnego rozumowania może sprzyjać gotowości do podważania pierwszych intuicji. Na stwierdzone różnice mogły wpływać również czynniki selekcyjne, w tym wcześniejsze kompetencje

matematyczne i motywacja uczniów wybierających zakres rozszerzony, których wpływu nie można oddzielić od wpływu sposobu kształcenia w badaniu przekrojowym. Interpretacja ta ma jednak charakter hipotetyczny i wymaga weryfikacji w dalszych badaniach.

Uzyskane rezultaty wpisują się w opisywaną w literaturze rozbieżność pomiędzy opanowaniem procedur statystycznych a umiejętnością ich praktycznego wykorzystania. Są również zgodne z koncepcją alfabetyzmu statystycznego przedstawioną przez Gala [2002] oraz Ben-Zviego i Garfielda [2004], zgodnie z którą poprawne rozumowanie statystyczne wykracza poza znajomość metod obliczeniowych i obejmuje interpretację informacji, krytyczną ocenę danych oraz podejmowanie decyzji w warunkach niepewności. Wyniki badania sugerują zatem, że kompetencje rachunkowe i interpretacyjne mogą rozwijać się częściowo niezależnie, co stanowi istotne wyzwanie dla dydaktyki statystyki.

Implikacje dydaktyczne

Uzyskane wyniki wskazują na potrzebę wzmacniania w kształceniu statystycznym umiejętności interpretowania danych i świadomego rozumowania w sytuacjach decyzyjnych. W szczególności zasadne wydają się:

- integrowanie treści statystycznych z podstawami ekonomii behawioralnej w celu uświadamiania studentom mechanizmów działania heurystyk poznawczych opisanych w literaturze [Tversky i Kahneman 1974];
- zwiększenie udziału zadań opartych na rzeczywistych danych i autentycznych problemach decyzyjnych oraz osadzonych w realistycznym kontekście gospodarczym;
- większe akcentowanie interpretacji wyników statystycznych i oceny wiarygodności wniosków, obok rozwijania umiejętności wykonywania obliczeń;
- wykorzystywanie zadań wymagających uzasadniania odpowiedzi i analizowania alternatywnych sposobów rozumowania, co może sprzyjać rozwijaniu alfabetyzmu statystycznego.

Ograniczenia badania

Interpretując uzyskane wyniki, należy uwzględnić trzy najważniejsze ograniczenia badania. Po pierwsze, zastosowano celowy dobór próby, co ogranicza możliwość uogólniania rezultatów na populację uczniów i studentów w Polsce oraz nie pozwala formułować wniosków o przyczynowym wpływie kolejnych etapów edukacji na poziom badanej kompetencji.

Po drugie, grupa uczniów realizujących matematykę w zakresie rozszerzonym była mniej liczna niż pozostałe grupy. Pomimo zastosowania metod statystycznych uwzględniających nierównoliczność prób wskazana jest ostrożność w interpretacji wielkości stwierdzonych różnic oraz ich weryfikacja na liczniejszych próbach.

Po trzecie, zastosowane narzędzie obejmowało cztery zadania odnoszące się do wybranych komponentów alfabetyzmu statystycznego. Wyniki dotyczą zatem

analizowanych obszarów kompetencji i nie wyczerpują całego zakresu tego złożonego konstruktu.

Kierunki dalszych badań

Naturalną kontynuacją badania będzie szczegółowa analiza rozkładu wyboru poszczególnych dystraktorów, pozwalająca zidentyfikować najczęstsze schematy błędnego rozumowania oraz lepiej rozpoznać mechanizmy prowadzące do nieprawidłowej interpretacji danych i oceny niepewności. Analiza ta zostanie przedstawiona w odrębnym opracowaniu poświęconym jakościowej interpretacji błędów poznawczych ujawnianych przez respondentów. Celowe wydaje się również rozszerzenie badań na studentów innych kierunków i osoby o zróżnicowanym doświadczeniu zawodowym, a także dalsze doskonalenie i walidacja zaproponowanego narzędzia badawczego.

Podsumowanie

Rozwój alfabetyzmu statystycznego wymaga nie tylko opanowania metod i narzędzi ilościowych, lecz także systematycznego kształtowania umiejętności interpretowania danych oraz poprawnego rozumowania w warunkach niepewności. Skuteczność kształcenia matematyczno-statystycznego warto zatem oceniać zarówno przez pryzmat poprawności wykonywanych obliczeń, jak i zdolności do świadomego i krytycznego wykorzystywania informacji statystycznych w procesie podejmowania decyzji.

BIBLIOGRAFIA

- Bar-Hillel M. (1980) The Base-Rate Fallacy in Probability Judgments. *Acta Psychologica*, 44(3), 211-233. [https://doi.org/10.1016/0001-6918\(80\)90046-3](https://doi.org/10.1016/0001-6918(80)90046-3)
- Batanero C., Chernoff E. J., Engel J., Lee H. S., Sánchez E. (2016) Research on Teaching and Learning Probability. ICME-13 Topical Surveys. Springer, Cham. <https://doi.org/10.1007/978-3-319-31625-3>
- Ben-Zvi D., Garfield J. (Red.) (2004) The Challenge of Developing Statistical Literacy, Reasoning and Thinking. Springer Science & Business Media. <https://doi.org/10.1007/1-4020-2278-6>
- Cohen J. (1988) Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Domański C. (2016) Wyzwania w zakresie edukacji statystycznej społeczeństwa. *Wiadomości Statystyczne. The Polish Statistician*, 9(664), 105-115. <https://doi.org/10.5604/01.3001.0014.1099>
- Gal I. (2002) Adults' Statistical Literacy: Meanings, Components, Responsibilities. *International Statistical Review*, 70(1), 1-25. <https://doi.org/10.1111/j.1751-5823.2002.tb00336.x>
- Gigerenzer G., Hoffrage U. (1995) How to Improve Bayesian Reasoning without Instruction: Frequency Formats. *Psychological Review*, 102(4), 684-704. <https://doi.org/10.1037/0033-295X.102.4.684>

- Gould R. (2017) Data Literacy is Statistical Literacy. *Statistics Education Research Journal Archives*, 16(1), 22-25. <https://doi.org/10.52041/serj.v16i1.209>
- Jasiewicz J., Stopnicka M. (2016) Data literacy: kompetencje w zakresie korzystania z danych w odniesieniu do przemian komunikacji naukowej. *Kultura Popularna*, 50(1), 36-49. <https://doi.org/10.5604/01.3001.0010.0042>
- Kahneman D. (2011) *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- Kahneman D., Tversky A. (1972) Subjective Probability: A Judgment of Representativeness. *Cognitive Psychology*, 3(3), 430-454. [https://doi.org/10.1016/0010-0285\(72\)90016-3](https://doi.org/10.1016/0010-0285(72)90016-3)
- Kurnia A. B., Lowrie T., Patahuddin S. M. (2024) The Development of High School Students' Statistical Literacy across Grade Level. *Mathematics Education Research Journal*, 36(Suppl 1), 7-35. <https://doi.org/10.1007/s13394-023-00449-x>
- Oleński J. (2020) Środowisko informacyjne statystyki publicznej. *Wiadomości Statystyczne. The Polish Statistician*, 65(4), 32-45. <https://doi.org/10.5604/01.3001.0014.1229>
- Phadke S., Beckman M., Lock Morgan K. (2024) Examining the Role of Context in Statistical Literacy Assessment. *Statistics Education Research Journal*, 23(1), Article 4. <https://doi.org/10.52041/serj.v23i1.529>
- Szreder M. (2016) O niektórych nowych wyzwaniach i oczekiwaniach wobec statystyki. *Wiadomości Statystyczne. The Polish Statistician*, 61(6), 1-9. <https://doi.org/10.5604/01.3001.0014.1000>
- Tversky A., Kahneman D. (1971) Belief in the Law of Small Numbers. *Psychological Bulletin*, 76(2), 105-110. <https://doi.org/10.1037/h0031322>
- Tversky A., Kahneman D. (1974) Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124-1131. <https://doi.org/10.1126/science.185.4157.1124>
- Wild C. J., Pfannkuch M. (1999) Statistical Thinking in Empirical Enquiry. *International Statistical Review*, 67(3), 223-248. <https://doi.org/10.1111/j.1751-5823.1999.tb00442.x>

STATISTICAL LITERACY AND THINKING TRAPS: AN ANALYSIS OF THE COMPETENCY GAP AT THE TRANSITION FROM SECONDARY TO HIGHER EDUCATION

Abstract: The aim of this paper is to assess the level of statistical literacy among secondary school and university students and to analyse selected components of this competence. The study was conducted on a sample of 432 secondary school and university students, using an original scenario-based test. The findings indicate the existence of a competency gap between the outcomes of formal mathematics and statistics education and the level of statistical literacy demonstrated in tasks involving data interpretation and probabilistic reasoning. The analysis revealed an educational anomaly: students completing the extended mathematics curriculum achieved a higher level of statistical literacy than university students who had completed an academic statistics course.

Keywords: statistical literacy, competency gap, statistics education, cognitive heuristics, probabilistic reasoning, data interpretation

JEL classification: A21, A22, D91, C12, I21