

Główne wyniki międzynarodowego badania umiejętności dorosłych PIAAC 2023

Autorzy raportu:

Michał Sitek
Artur Pokropek
Paweł Penszko
Katarzyna Chyl
Jacek Haman

Redakcja językowa i korekta: Monika Niewielska

Projekt graficzny i skład: Wojciech Maciejczyk

Zdjęcie na okładce: Shutterstock.com

Copyright © Instytut Badań Edukacyjnych, Warszawa 2024

Przedruk w całości lub w części wyłącznie za zgodą Instytutu Badań Edukacyjnych.

Cytowanie oraz wykorzystywanie danych jedynie z podaniem źródła.

Wzór cytowania:

Sitek, M., Pokropek, A., Penszko, P., Chyl, K., Haman, J. (2024). *Główne wyniki międzynarodowego badania umiejętności dorosłych PIAAC 2023*. Warszawa: Instytut Badań Edukacyjnych.

ISBN dla wersji drukowanej: 978-83-68313-40-6

ISBN dla wersji elektronicznej: 978-83-68313-39-0

DOI: 10.24131/978-83-68313-39-0

Publikacja sfinansowana ze środków Ministerstwa Edukacji Narodowej.

Wydawca: Instytut Badań Edukacyjnych
ul. Górczewska 8, 01-180 Warszawa
tel. 22 241 71 00; www.ibe.edu.pl



Spis treści

Wprowadzenie	4
1. Poziom i zróżnicowanie umiejętności dorosłych	8
Wstęp	8
Poziomy umiejętności.....	13
Podsumowanie.....	34
2. Zmiany wyników kluczowych umiejętności w ciągu ostatniej dekady	36
Wstęp	36
Zmiany w zakresie umiejętności w populacji dorosłych.....	39
Zmiany wyników PIAAC dla grup społeczno-demograficznych	45
Podsumowanie.....	52
3. Metodologia badania	53
Wstęp	53
Co mierzy PIAAC?	53
Populacja i próba w badaniu PIAAC w Polsce	61
Techniki i narzędzia badawcze	65
Skalowanie wyników.....	70
Bibliografia	72

Wprowadzenie

Czytaj Państwo raport z największego na świecie międzynarodowego badania kompetencji dorosłych. W 2023 roku 31 krajów OECD, w tym Polska, przeprowadziło nowy cykl badania, dostarczając wiedzy o aktualnych umiejętnościach dorosłych w zakresie rozumienia tekstu (*literacy*), rozumowania matematycznego (*numeracy*) i adaptacyjnego rozwiązywania problemów (*adaptive problem solving*). Metodologia badania umożliwia porównywanie wyników między uczestniczącymi krajami w danym okresie, a także na przestrzeni czasu. Badanie zaprojektowano tak, by umożliwić nie tylko pogłębione analizy zróżnicowania wyników wewnątrz krajów, uwarunkowań tych wyników związanych z płcią i innymi zmiennymi demograficznymi, ale też powiązań między umiejętnościami a korzystaniem z umiejętności, doświadczeniem zawodowym i pozycją na rynku pracy. Pozwala to na stawianie hipotez o przyczynach różnic i tworzenie mądrych polityk edukacyjnych, wspierających rozwój umiejętności i podgrupy dorosłych, którymi trzeba najpilniej się zająć.

Poprzednie badania międzynarodowe, z pierwszym cyklem PIAAC na czele, pokazały, że poziomy umiejętności dorosłych w zakresie czytania, matematyki i rozwiązywania problemów były silniej zróżnicowane wewnątrz krajów niż między nimi. Najważniejszymi czynnikami różnicującymi wyniki były: poziom wykształcenia, rodzaj wykonywanej pracy i status społeczno-kulturowy. Wyższe poziomy umiejętności mierzonych w PIAAC okazały się ważne dla sukcesu i produktywności w miejscu pracy, natomiast dorośli o niskich kompetencjach byli bardziej narażeni na bezrobocie, zaś ich miejsca pracy na likwidację lub automatyzację. Kraje z wyższymi średnimi poziomami umiejętności zazwyczaj miały lepsze wyniki gospodarcze i społeczne, i szybciej się rozwijały (OECD, 2019, Hanushek i in., 2017). W poprzednim cyklu PIAAC wyniki Polaków i Polek uplasowały się w okolicach średniej OECD dla umiejętności czytania, rozumowania matematycznego i rozwiązywania problemów, i wykazały jeden z największych skokowych wzrostów od czasów badania IALS z lat 90. (Rynko i in., 2013). Równie dynamicznie zmieniła się w tamtym czasie polska gospodarka i dojrzała młoda demokracja.

W ostatnich dziesięciu latach zaszły jednak kolejne przemiany społeczne i gospodarcze, które należy brać pod uwagę, rozpatrując kontekst nowego cyklu badania PIAAC. Zadajemy sobie nowe pytania o przyszłość rynku pracy, demokracji i nowych technologii. W czasie pomiędzy pierwszym i drugim cyklem PIAAC rozpoczął się największy po II wojnie światowej kryzys migracyjny, zrozumieliśmy lepiej zagrożenia związane ze zmianami klimatu, a także kampaniami dezinformacyjnymi i rosnącym populizmem. Zachodnie społeczeństwa systematycznie się starzeją. Czynnikiem zmieniającymi stosunki społeczne i gospodarcze w Europie była pandemia COVID-19 i pełnowymiarowa rosyjska inwazja na Ukrainę. W ostatnich latach rozwój sztucznej inteligencji i dużych modeli językowych stworzył nowe szanse, ale także zupełnie nowe zagrożenia.

W tym świecie przemian umiejętności badane w PIAAC wciąż stanowią podstawę dla sprawnego przetwarzania informacji i uczenia się. Bez adekwatnych umiejętności czytania,

rozumowania matematycznego i rozwiązywania nowych napotykanymi problemami trudno jest poruszać się we współczesności, która codziennie angażuje te zdolności. Z kolei wysoki poziom kompetencji, zakładający również umiejętność oceniania jakości informacji i poszukiwania wiarygodnych źródeł nie tylko jest ważny na rynku pracy, ale jest też niezbędnym warunkiem funkcjonowania demokracji.

To umiejętności związane z czytaniem wpływają na to, w jaki sposób dorośli zdobywają, wybierają i przetwarzają dostępne informacje, a więc także na to, jakie podejmują decyzje, zarówno te konsumenckie czy dotyczące zdrowia, jak i związane z prawami wyborczymi. Konteksty pomiaru umiejętności rozumienia tekstu w PIAAC są szerokie i obejmują wszelkie obszary życia codziennego, w których czytanie jest ważne. Kierując się tym, jakie treści są zrozumiałe i dostępne, dorośli wybierają media, z których czerpią wiedzę o świecie, i formułują opinie na temat złożonych problemów współczesności. W Polsce i innych krajach coraz częstszym źródłem wiedzy o świecie jest internet, co zostało w PIAAC uwzględnione zarówno poprzez włączenie do badania dużej liczby tekstów cyfrowych, jak i wykorzystanie tabletów zamiast papieru. Natomiast w sieci sprawnie poruszać się mogą tylko osoby o wysokich kompetencjach, które czytają ze zrozumieniem, krytycznie oceniają informacje i ich źródła, oraz poddają je refleksji. W kontekście formowania się opinii publicznej i tego, jak wpływa ona na wyniki wyborów i decyzje polityczne, duży odsetek dorosłych, którzy nie posiadają umiejętności niezbędnych do efektywnego uczestnictwa w złożonych procesach społecznych i politycznych, może mieć negatywne konsekwencje dla całego społeczeństwa. Indywidualne konsekwencje problemów z czytaniem mogą być również dotkliwe.

Podobnie z umiejętnością rozumowania matematycznego, w zakres której wchodzi m.in. rozumienie i interpretowanie danych liczbowych. Współcześnie osoby dorosłe coraz częściej stają przed koniecznością radzenia sobie z natłokiem informacji ilościowych, często zapośredniczonych przez nowe technologie. Takie informacje należy odpowiednio selekcjonować i interpretować, aby podejmować na ich podstawie jak najlepsze decyzje (Tout i in., 2021). Dane, choć wykorzystywane jako obiektywne źródło informacji, często są przedstawiane w sposób wybiórczy, aby poprzeć określone stanowisko. Umiejętność krytycznego czytania tabel, wykresów i poprawnego interpretowania wszelkich danych opartych na liczbach może ustrzec przed podejmowaniem niekorzystnych decyzji, np. dotyczących zdrowia (dawkowanie leku), finansów (decyzja o kredycie konsumenckim) czy czynnego udziału w wyborach (interpretacja sondaży przedwyborczych czy planów budżetowych). Może także ustrzec przed podawaniem dalej fałszywych informacji, a przez to – szerzeniem się dezinformacji. Nie ma wątpliwości, że pomimo łatwego dostępu do kalkulatorów i komputerów umiejętności rozumowania matematycznego leżą u podstaw sprawnego funkcjonowania w erze informacji.

Zarówno umiejętność rozumienia tekstu, jak i rozumowania matematycznego są oparte na bardziej ogólnych umiejętnościach metapoznawczych, takich jak adaptacyjne rozwiązywanie problemów w dynamicznej sytuacji, w której rozwiązanie nie jest od razu oczywiste. To

szerokie pojęcie, w którym zawiera się definiowanie problemów, ocena możliwych rozwiązań i monitorowanie postępów w realizacji celów (Greiff i in., 2021). Zdolność dorosłych do przystosowywania się do nowych okoliczności i uczenia się przez całe życie jest istotna w nowoczesnych społeczeństwach, które zmieniają się w coraz szybszym tempie. W nowym cyklu badania PIAAC włączono do pomiaru ten obszar kompetencji, zastępując nim mierzone w poprzednim cyklu rozwiązywanie problemów w środowiskach bogatych w technologię (*problem solving in technology-rich environment*). Choć zaawansowane technologie, również te oparte na sztucznej inteligencji, weszły już tak głęboko do naszej codzienności, że stały się niemal przezroczyste, umiejętność adaptacyjnego rozwiązywania problemów jest wciąż kluczowa w szybko zmieniającym się świecie.

Umiejętności mierzone w PIAAC to fundamenty do zdobywania innych kompetencji, zarówno poprzez uczenie się formalne, jak i nieformalne. Umiejętności rozumienia tekstu, rozumowania matematycznego i rozwiązywania problemów ułatwiają zdobywanie pogłębionej, specjalistycznej wiedzy we wszelkich dziedzinach i przystosowywanie się do zmieniających się warunków. To, jak bardzo w rozwiniętych gospodarkach pożądane są umiejętności oparte na przetwarzaniu danych i podejmowaniu na ich podstawie nierutynowych decyzji, pokazuje analiza trendów w ogłoszeniach o pracę (Deming, 2021). Osoby o wysokich kompetencjach mogą znakomicie poradzić sobie w gospodarce, która w coraz większym stopniu wykorzystuje automatyzację powtarzalnych czynności i oparta jest na nowych narzędziach cyfrowych, w tym sztucznej inteligencji. Z drugiej strony brak podstawowych umiejętności mierzonych przez PIAAC to ryzyko w obszarze bezpieczeństwa, zdrowia, finansów, polityki publicznej czy każdej innej sferze, w której przetwarzane są informacje oparte na liczbach i słowach. PIAAC daje możliwość identyfikacji grup najbardziej wrażliwych na wykluczenie i wykazuje, jak bardzo rozpowszechniony jest w społeczeństwie problem niewystarczających kompetencji podstawowych.

Pierwszy cykl badania PIAAC pokazał, że w Polsce wiele osób osiągnęło co najwyżej pierwszy poziom umiejętności rozumienia tekstu (18,8% osób badanych) i rozumowania matematycznego (23,5%). Analiza danych zebranych dzięki kwestionariuszom wykazała, że osoby, które uzyskiwały najniższe wyniki w PIAAC, to przede wszystkim te o niższym poziomie wykształcenia, starsze i posiadające w domu zaledwie kilka książek, a więc o niskim kapitale społeczno-kulturowym (Chłoń-Domińczak i in., 2015). Te dane były spójne z czynnikami ryzyka wykazanymi w analizie międzynarodowych danych PIAAC z pierwszego cyklu (Kyröläinen i Kuperman, 2021). Nakreśliły też obraz konsekwencji niskich kompetencji dla dorosłych. W Polsce osoby o najniższych kompetencjach wyjątkowo rzadko (16%) uczyły się formalnie lub nieformalnie, dzięki czemu mogłyby te kompetencje podnosić. Z kolei niski poziom wykształcenia i nieuczenie się były związane z bezrobociem, ryzykiem utraty pracy i większymi trudnościami w znalezieniu innego zatrudnienia (Chłoń-Domińczak i in., 2015). Problemy z czytaniem u dorosłych, również te wynikające z niepoddanej terapii dysleksji rozwojowej, mogą mieć poważne negatywne konsekwencje społeczne, ekonomiczne i zdrowotne (Vágvölgyi i in., 2016), a także emocjonalne,

takie jak niskie poczucie własnej wartości, lęk i długotrwały stres (Livingston, Siegel i Ribary, 2017). Dane z drugiego cyklu PIAAC pokazują, jak przez ostatnią dekadę kształtował się obraz umiejętności podstawowych w Polsce. Rosnąca liczba osób o niskich kompetencjach to sygnał, że większe grupy osób dorosłych zagrożonych problemami z czytaniem i matematyką powinny zostać objęte opieką przez politykę edukacyjną.

Badanie umiejętności czytania, rozumowania matematycznego i rozwiązywania problemów jest dużym wyzwaniem, zwłaszcza przy próbie stworzenia testów umożliwiających równoległy pomiar w kilkudziesięciu krajach (Chyl i in., 2024). Wyzwaniem jest wybór zadań, które będą możliwie niezależne kulturowo, a jest to warunkiem trafności pomiaru; nie zawsze udaje się to zrobić, co sugeruje polski przypadek badania IALS z 1994 r. (Gomez, 2000). Szczególnie istotną kwestią może być motywacja i zaangażowanie w wykonywanie zadań (Penk, Pöhlmann i Roppelt, 2014). Już w pierwszym cyklu badania Polska miała jeden z najniższych współczynników zaangażowania respondentów spośród wszystkich uczestniczących krajów (Goldhammer, Martens i Lüdtke, 2016). W badanie nie angażowały się przede wszystkim osoby o niższym poziomie wykształcenia i niższych umiejętnościach, szczególnie przy rozwiązywaniu trudniejszych zadań (Goldhammer, Martens i Lüdtke, 2017). Choć niskie zaangażowanie może być powiązane z niższym poziomem umiejętności, trudno rozdzielić wpływ obu tych czynników na końcowy wynik. Pamiętając o tych zastrzeżeniach, warto zauważyć, że PIAAC jest jedynym badaniem próbującym opisać rozkład kluczowych umiejętności dorosłych w Polsce na tle innych krajów. Nie dysponujemy też innymi badaniami krajowymi (które mogłyby być punktem odniesienia) do interpretacji różnic i zróżnicowania poziomu umiejętności Polaków. Tego rodzaju badania są szczególnie ważne w kontekście szybko zmieniających się wymagań rynku pracy i postępu technologicznego, które stawiają przed społeczeństwem nowe wyzwania i konieczność ciągłego doskonalenia kompetencji. Tę diagnozę społeczną warto rozszerzać w oparciu o inne narzędzia do pomiaru umiejętności rozumienia tekstu, rozumowania matematycznego i rozwiązywania problemów oraz badania krajowe, których w tej chwili brakuje.

1. Poziom i zróżnicowanie umiejętności dorosłych

Wstęp

Badanie Umiejętności Dorosłych PIAAC (Program Międzynarodowej Oceny Kompetencji Dorosłych) stanowi kompleksową analizę kluczowych kompetencji i ich uwarunkowań. Jego istotą jest pomiar trzech umiejętności. Umiejętność rozumienia tekstu (*literacy*) definiowana jest jako zdolność wyszukiwania, odczytywania znaczeń, oceny i refleksji nad pisanymi tekstami, umożliwiająca osiągnięcie zamierzonych celów, pogłębianie wiedzy i uczestnictwo w życiu społecznym. Rozumowanie matematyczne (*numeracy*) obejmuje wyszukiwanie, wykorzystywanie i krytyczne rozumienie treści matematycznych w różnych sytuacjach życiowych, natomiast adaptacyjne rozwiązywanie problemów (*adaptive problem solving, APS*) to zdolność osiągania własnych celów w dynamicznej sytuacji, wymagająca zaangażowania procesów poznawczych i metapoznawczych.

Badanie obejmuje kraje z różnych kontynentów, w tym Polskę, co pozwala na przeprowadzenie wielowymiarowego porównania międzynarodowego. W edycji 2022–2023 w badaniu PIAAC wzięło udział 31 krajów (lub ich części) z wielu kontynentów: Ameryki Północnej (Kanada, Stany Zjednoczone), Ameryki Łacińskiej (Chile), Europy (Anglia – Wielka Brytania, Austria, Belgia - Region Flamandzki, Chorwacja, Czechy, Dania, Estonia, Finlandia, Francja, Hiszpania, Irlandia, Litwa, Łotwa, Niderlandy, Niemcy, Norwegia, Polska, Portugalia, Słowacja, Szwajcaria, Szwecja, Węgry, Włochy), Azja (Japonia, Korea Południowa, Singapur) oraz z Bliskiego Wschodu (Izrael) i Oceanii (Nowa Zelandia).

Metodologię badania opisujemy szczegółowo w rozdziale 3. W tym rozdziale opiszemy najważniejsze wyniki drugiego cyklu badania. Omówimy średnie wyniki krajów oraz ich zróżnicowanie wewnątrz krajów, pokazując, jak wypadli polscy badani na tle dorosłych z innych krajów. Pokażemy także, w jaki sposób wyniki są powiązane z podstawowymi zmiennymi społeczno-demograficznymi: płcią, wiekiem i wykształceniem.

Porównanie średnich wyników

Głównym celem badania PIAAC jest porównanie poziomów i zróżnicowania badanych umiejętności między krajami. Ale zanim omówimy wyniki, warto podsumować krótko, jak je interpretować. Wyniki PIAAC są prezentowane na skali, której średnia wynosi 250 punktów, a odchylenie standardowe to 50 punktów. Jest to skala interwałowa – można porównywać wielkości różnic między wynikami dla różnych krajów czy grup badanych, ale nie można interpretować ich proporcji. Można np. powiedzieć, że różnica między 275 a 250 punktów jest taka sama jak między 250 i 225 punktów, ale nie ma sensu stwierdzenie, że 275 punktów to wynik o 10% lepszy niż 250 punktów. Skale dla rozumienia tekstu i rozumowania matematycznego

zostały ustalone w 2012 roku, na podstawie danych krajów uczestniczących wtedy w badaniu. Te skale są takie same jak w poprzedniej edycji badania, co umożliwia porównywanie zmian w czasie oraz włączanie nowych krajów do porównań. Skalę rozwiązywania problemów – nowej dziedziny w badaniu – wprowadzono w 2023 roku i zastąpiła ona mierzone 10 lat temu rozwiązywanie problemów w środowiskach bogatych w technologie (zob. Rynko, 2013).

Najwyższe średnie wyniki w rozumieniu tekstu uzyskali mieszkańcy Finlandii (296 punktów), Japonii (289), Szwecji (284) i Norwegii (281). Średnia OECD wyniosła 260 punktów¹. Podobnie wygląda ranking krajów w pomiarze umiejętności matematycznych. W zakresie rozwiązywania problemów najwyższe wyniki uzyskały Finlandia i Japonia (276 punktów), a następnie Szwecja (273) i Norwegia (271).

We wszystkich trzech pomiarach Polska uzyskała jeden z najsłabszych wyników. Średni wynik Polski (236) w rozumieniu tekstu jest statystycznie nieodróżnialny od wyniku Litwy i Portugalii, a w rozumowaniu matematycznym (239) od średniego wyniku Litwy. Należy jednak zachować ostrożność w interpretacji tych wyników ze względu na problemy z danymi uzyskanymi dla Polski. Jak podkreślamy w raporcie, Polska spełniła standardy techniczne badania, jednak wyniki polskich respondentów są niedoszacowane. To niedoszacowanie wynika z zaniżenia odsetka osób lepiej wykształconych w próbie względem danych pochodzących ze statystyki publicznej, ale przede wszystkim niższej motywacji polskich respondentów. Mimo zastosowanych metod analitycznych (np. wyłączenia z analiz części danych, zob. rozdział 3), nie udało się całkowicie wyeliminować wpływu tych czynników na wyniki.

Pamiętając o tych zastrzeżeniach, warto zwrócić uwagę na pozycję Polski na tle innych krajów regionu. Średnie wyniki w Polsce były niższe od wyników krajów Grupy Wyszehradzkiej i krajów bałtyckich: Estonia konsekwentnie osiąga najwyższe wyniki, znacznie przewyższając pozostałe kraje Grupy Wyszehradzkiej i kraje bałtyckie. Na drugim miejscu plasują się Czechy i Słowacja z podobnymi wynikami, a następnie Węgry i Litwa. W rozumieniu tekstu i rozumowaniu matematycznym wyniki Polski i Litwy są nieodróżnialne statystycznie. Różnica między Polską a Estonią wynosi ok. 45 punktów, a między Polską a Czechami i Słowacją – ok. 25 punktów, i są to różnice istotne statystycznie. Znacząco lepiej wypadli też dorośli z Niemiec i Szwecji: średnie wyniki w Niemczech były wyższe niż w Polsce we wszystkich trzech dziedzinach: o 37 punktów w rozumieniu tekstu, o 34 punkty w rozumowaniu matematycznym i o 31 punktów w rozwiązywaniu problemów. Szwecja również osiągnęła wyższe wyniki od Polski: o 49 punktów w rozumieniu tekstu, o 46 punktów w rozumowaniu matematycznym i o 45 punktów w rozwiązywaniu problemów.

Warto zwrócić uwagę, że uporządkowanie krajów w różnych dziedzinach jest bardzo podobne: kraje osiągające lepsze wyniki w jednej dziedzinie osiągają też lepsze wyniki w innych, a średnie

¹ Głównym punktem odniesienia są średnie krajów OECD z 2022 roku. Należy jednak pamiętać, że uwzględnia ona tylko część państw członkowskich, ponieważ nie wszystkie kraje OECD uczestniczyły w badaniu PIAAC.

z poszczególnych dziedzin są ze sobą mocno skorelowane. Trudno więc mówić, że konkretne kraje mają przewagi w jednych dziedzinach, a inne w drugich.

Tabela 1.1. Średnie wyniki rozumienia tekstu, rozumowania matematycznego i rozwiązywania problemów.

Rozumienie tekstu		Rozumowanie matematyczne		Rozwiązywanie problemów	
Finlandia	296 (0,6)	Finlandia	294 (1,2)	Finlandia	276 (1,0)
Japonia	289 (0,8)	Japonia	291 (0,8)	Japonia	276 (0,7)
Szwecja	284 (0,7)	Szwecja	285 (1,0)	Szwecja	273 (0,9)
Norwegia	281 (1,1)	Norwegia	285 (0,9)	Norwegia	271 (0,8)
Holandia	279 (2,0)	Holandia	284 (1,1)	Holandia	265 (0,9)
Estonia	276 (1,2)	Estonia	281 (0,6)	Dania	264 (0,7)
R. Flamandzki (Belgia)	275 (0,2)	R. Flamandzki (Belgia)	279 (0,9)	Estonia	263 (0,6)
Dania	273 (0,6)	Dania	279 (0,8)	R. Flamandzki (Belgia)	262 (0,8)
Anglia	272 (1,0)	Szwajcaria	276 (0,7)	Niemcy	261 (0,6)
Kanada	271 (2,2)	Singapur	274 (0,8)	Anglia	259 (0,9)
Szwajcaria	266 (1,4)	Niemcy	273 (0,7)	Kanada	259 (0,7)
Niemcy	266 (1,0)	Kanada	271 (0,9)	Szwajcaria	257 (0,7)
Irlandia	263 (1,1)	Anglia	268 (1,1)	Austria	253 (0,7)
Średnia OECD	260 (1,8)	Czechy	267 (1,1)	Singapur	252 (0,8)
Nowa Zelandia	260 (1,0)	Austria	267 (0,9)	Średnia OECD	251 (0,2)
Czechy	260 (0,8)	Łotwa	263 (0,9)	Czechy	250 (0,9)
Stany Zjednoczone	258 (1,0)	Średnia OECD	263 (0,2)	Nowa Zelandia	249 (2,1)
Francja	255 (0,8)	Słowacja	261 (1,2)	Irlandia	249 (0,9)
Singapur	255 (0,0)	Irlandia	260 (1,2)	Francja	248 (0,5)
Austria	254 (0,9)	Francja	257 (0,7)	Stany Zjednoczone	247 (1,4)
Słowacja	254 (0,9)	Nowa Zelandia	256 (2,1)	Słowacja	247 (1,1)
Chorwacja	254 (0,7)	Chorwacja	254 (1,6)	Łotwa	244 (0,8)
Korea Południowa	249 (0,9)	Węgry	254 (1,1)	Węgry	241 (0,9)
Węgry	248 (1,0)	Korea	253 (1,0)	Hiszpania	241 (0,8)
Łotwa	248 (1,0)	Hiszpania	250 (0,9)	Korea	238 (0,7)
Hiszpania	247 (1,0)	Stany Zjednoczone	249 (1,6)	Izrael	236 (1,0)
Włochy	245 (0,9)	Izrael	246 (1,2)	Chorwacja	235 (1,5)
Izrael	244 (1,5)	Litwa	246 (1,2)	Portugalia	233 (1,5)
Litwa	238 (0,9)	Włochy	244 (1,7)	Włochy	231 (1,3)
Polska	236 (1,7)	Polska	239 (1,1)	Litwa	230 (1,0)
Portugalia	235 (1,2)	Portugalia	238 (1,9)	Polska	226 (1,0)
Chile	218 (1,0)	Chile	214 (2,3)	Chile	218 (1,9)

Uwagi:

Kraje i gospodarki są uszeregowane w porządku malejącym według ich średniej. W nawiasach podano wartości błędów standardowych, które wskazują na poziom niepewności oszacowań średnich.

Badanie objęło dorosłych w wieku 16–65 lat; również tych, z którymi ze względu na barierę językową przeprowadzono jedynie ankietę zastępczą.

Polskie wyniki są najprawdopodobniej zaniżone i należy je interpretować z ostrożnością.

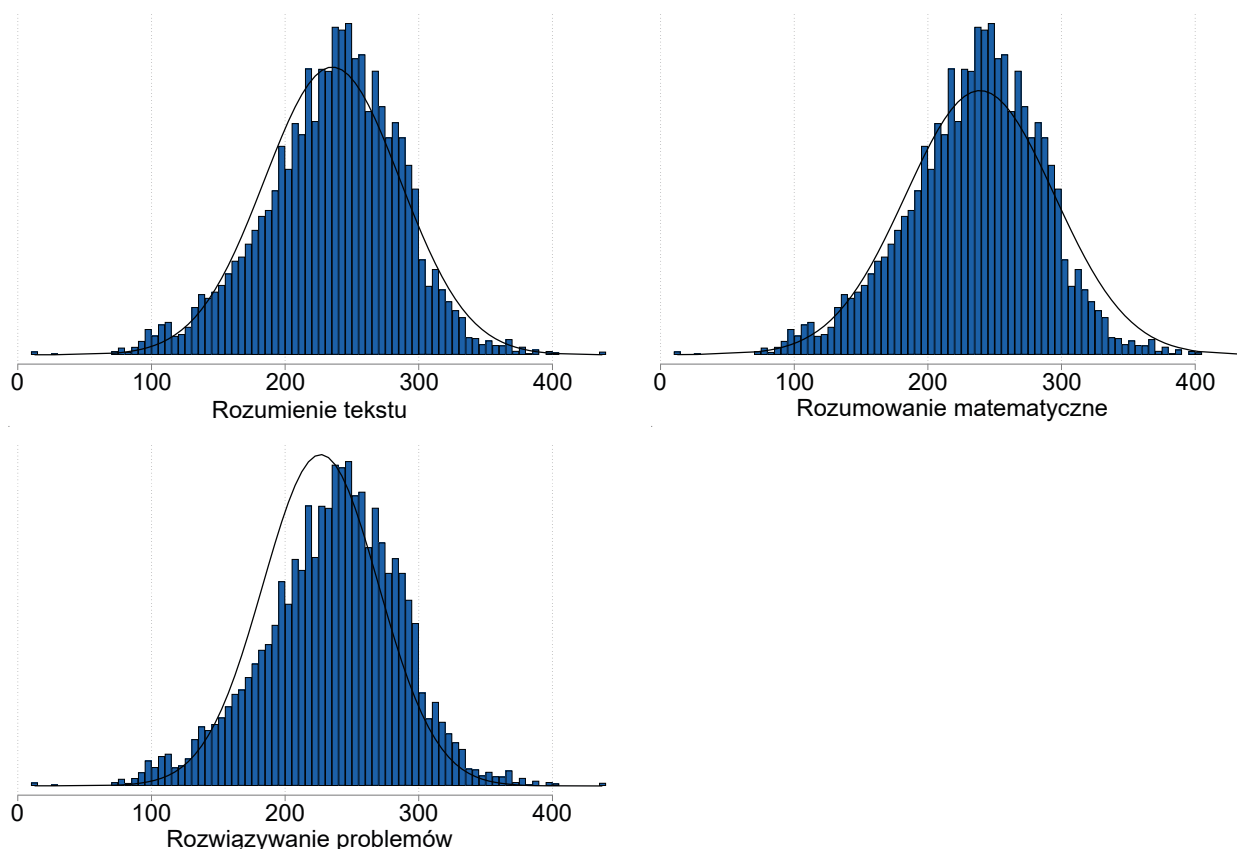
Źródło: opracowanie własne na podstawie: OECD, 2024, tabela A.2.1.

Jak różnią się dorośli pod względem poziomu umiejętności?

Dyskusje o wynikach porównań międzynarodowych często skupiają się na rankingach opartych na średnich wynikach w poszczególnych krajach. Tymczasem rzeczywistość jest bardziej złożona – w każdym kraju znajdziemy zarówno osoby świetnie radzące sobie z rozumieniem tekstu, jak i takie, które mają z tym poważne trudności. Różnice między umiejętnościami w poszczególnych krajach są większe niż różnice w średnich wynikach krajów. O ile więc średnie wyniki dają nam ogólny ogląd sytuacji, to dla zrozumienia tego, jak radzą sobie dorośli w rozumieniu tekstu czy rozumowaniu matematycznym, ważne jest przyjrzenie się różnym miarom zróżnicowania.

Zróżnicowanie wyników w Polsce najłatwiej zilustrować, pokazując częstości poszczególnych wyników (zob. rys. 1.1). Średnie dla rozumienia tekstu, rozumowania matematycznego i rozwiązywania problemów to odpowiednio 236, 239 i 230 punktów. Rozkład wyników jest zbliżony do rozkładu normalnego, czyli większość wyników koncentruje się wokół średniej, a wartości odbiegające od średniej są coraz mniej częste. W porównaniu z rozkładem normalnym, widoczne są niewielkie przesunięcia rozkładu wyników w jedną lub drugą stronę. W rozumieniu tekstu i rozumowaniu matematycznym jest więcej niższych wyników. W rozwiązywaniu problemów więcej jest wyników po prawej stronie rozkładu. W każdej z mierzonych dziedzin widoczna jest niewielka „nadwyżka” (w stosunku do rozkładu normalnego) respondentów z bardzo niskimi wynikami.

Rysunek 1.1. Rozkłady częstości wyników trzech dziedzin mierzonych w Polsce oraz rozkład normalny wyliczony dla średnich i odchyłeń standardowych tych wyników*.



* Na wykresie pokazano rozkład jednej wartości prawdopodobnej (plausible value), co daje przybliżony kształt rozkładu. W badaniu PIAAC każdemu respondentowi przypisuje się, po wyskalowaniu wyników, 10 wartości prawdopodobnych. Takie podejście pozwala uwzględnić niepewność pomiaru wynikającą z tego, że respondenci rozwiązują tylko część wszystkich zadań testowych i umożliwia precyzyjniejsze oszacowanie rozkładów wyników w populacji. Oszacowanie parametrów z wykorzystaniem wartości prawdopodobnych polega na przeprowadzeniu analiz osobno dla każdej wartości i wyliczenia średniej.

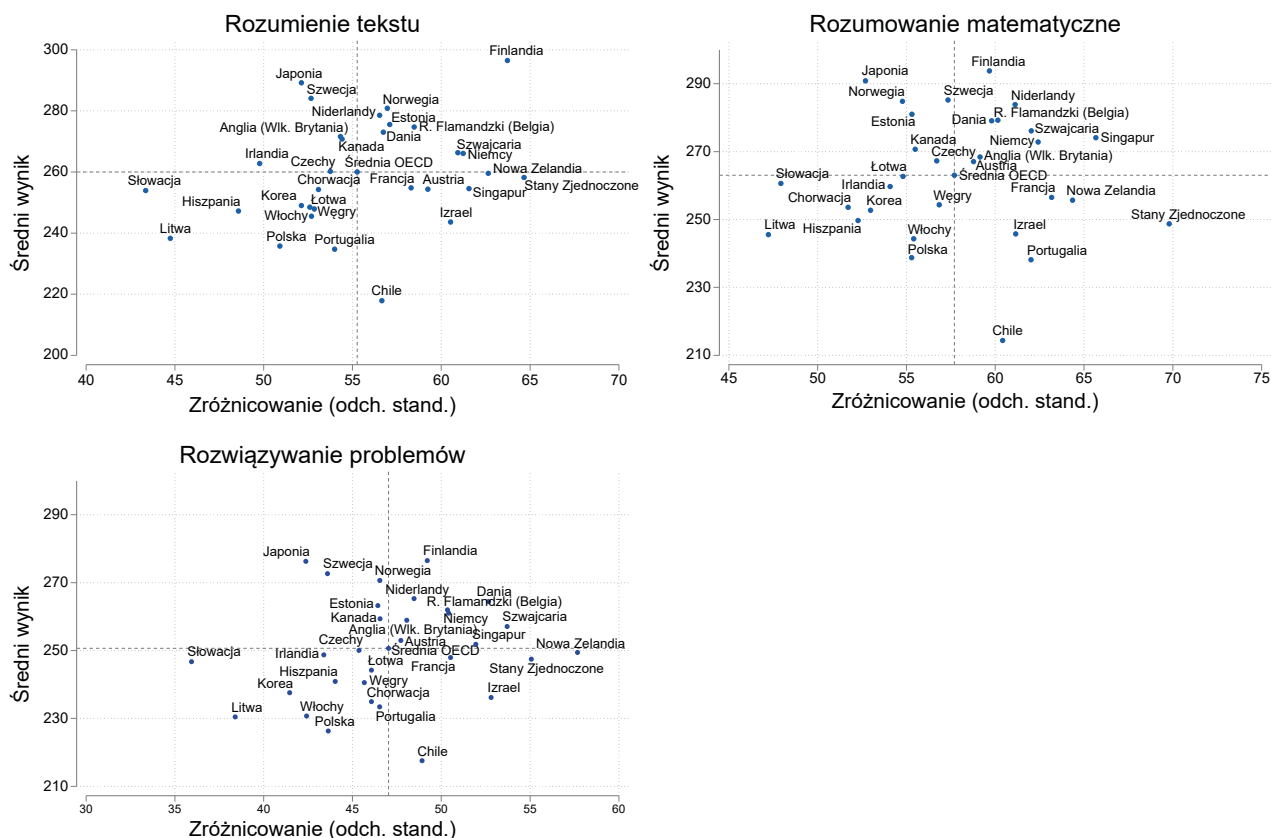
Źródło: opracowanie własne na podstawie danych PIAAC.

W niektórych krajach zróżnicowanie wyników jest większe niż w innych. Przydatną miarą tego zróżnicowania jest odchylenie standardowe – im jest ono większe, tym bardziej wyniki odbiegają od średniej. W Polsce było ono najwyższe w matematyce (odchylenie standardowe 55,3 punktu), następnie w rozumieniu tekstu (50,9 punktu), a najmniejsze w rozwiązywaniu problemów (43,6 punktu).

Jak zróżnicowanie wyników w Polsce wygląda na tle zróżnicowania wyników w innych krajach? Czy w krajach osiągających lepsze wyniki różnice są większe, czy mniejsze? Na pytanie te można odpowiedzieć, zestawiając średnie wyniki z ich odchyleniem standardowym. Zależności te podsumowuje poniższy wykres. Porównując średnie odchylenia standardowe, widać, że – podobnie jak w Polsce – zróżnicowanie jest na ogół najmniejsze w dziedzinie rozwiązywania problemów, a najwyższe w rozumowaniu matematycznym. Okazuje się, że zróżnicowanie wyników jest słabo powiązane ze średnim wynikiem. Są kraje osiągające wysokie, ale bardziej zróżnicowane wyniki, i takie, gdzie wyniki są mniej zróżnicowane. Polska jest w grupie krajów

o umiarkowanym zróżnicowaniu wyników, podobnie jak Czechy, Estonia, Holandia czy Szwecja. Największe zróżnicowanie wyników zaobserwowano w Austrii, Niemczech, Danii, Francji i Stanach Zjednoczonych. Pod względem najmniejszego zróżnicowania wyników wyróżniają się Litwa i Słowacja.

Rysunek 1.2. Średnie wyniki w zakresie rozumienia tekstu, rozumowania matematycznego oraz rozwiązywania problemów oraz ich zróżnicowanie (odchylenie standardowe) w krajach uczestniczących w badaniu.



Uwaga: Polskie wyniki są najprawdopodobniej zaniżone i należy je interpretować z ostrożnością.

Źródło: opracowanie własne na podstawie danych PIAAC.

Poziomy umiejętności

Do opisywania wyników w badaniu PIAAC używa się poziomów umiejętności. Ponieważ zarówno umiejętności badanych, jak i trudność zadań mierzono na tej samej skali punktowej, możemy przewidzieć, jak dobrze osoby o różnych umiejętnościach poradzą sobie z zadaniami o określonej trudności². Dzięki poziomom umiejętności i charakterystyce umiejętności mierzonych przez zadania na poszczególnych poziomach możemy dokładniej pokazać, co potrafią osoby na każdym poziomie, a także porównać kraje pod względem liczby osób

² Osoba na danym poziomie potrafi zwykle rozwiązać większość zadań z tego poziomu i z poziomów niższych, ale radzi sobie też z częścią trudniejszych zadań. Na przykład osoba przyporządkowana do poziomu 3. (z szacowanym poziomem umiejętności odpowiadającej środkowi tego poziomu) z prawdopodobieństwem 67% poprawnie rozwiązuje typowe zadanie o trudności odpowiadającej poziomowi 3., ale jeszcze lepiej radzi sobie z zadaniami z poziomów 1. i 2., a także potrafi rozwiązać niektóre zadania z poziomu 4. Próg 67% przyjęto arbitralnie – podobne rozwiązanie stosowane jest w wielu pomiarach edukacyjnych.

o najniższych i najwyższych umiejętnościach. W dziedzinie rozumienia tekstu i rozumowania matematycznego wyróżniono sześć poziomów (od: poniżej poziomu 1. do poziomu 5.). Dla adaptacyjnego rozwiązywania problemów skala jest krótsza i kończy się na poziomie 4.

Opis poszczególnych poziomów przedstawiono w poniższych tabelach.

Tabela 1.2 Opis tego, co potrafią dorośli na poszczególnych poziomach umiejętności rozumienia tekstu.

Poziom	Zakres punktów	Co potrafią dorośli na tym poziomie
Poziom 5.	376 punktów lub więcej	<p>Poziom 5. opisuje najwyższe umiejętności rozumienia tekstu. PIAAC nie pozwala jednak na ich bezpośredni pomiar, ze względu na ograniczony czas badania. Ograniczenia czasu wykluczały wykorzystanie długich, trudnych dokumentów czy zestawów wielu tekstów, a także skomplikowanych metod dostępu do nich, takich jak rozbudowane katalogi czy wyszukiwarki. Przyszłe badania, dążące do lepszego odzwierciedlenia pełnej skali kompetencji rozumienia tekstu, mogą obejmować też takie zadania, które rzeczywiście pokrywałyby poziom 5.</p> <p>Na podstawie cech najtrudniejszych zadań na poziomie 4. można jednak wywnioskować, co charakteryzuje umiejętność rozumienia tekstu na najwyższym, 5. poziomie. Dorośli na tym poziomie zapewne samodzielnie krytycznie oceniają stawiane przed nimi zadania, formułując własne cele w czytaniu na podstawie złożonych i niejawnych kryteriów. Potrafią wyszukiwać i zintegrować rozproszone informacje z wielu trudnych tekstów. Są w stanie wyciągać wnioski zarówno ze spójnych, jak i sprzecznych źródeł i punktów widzenia. Mogą oceniać wiarygodność argumentów, również takich, które pochodzą z nowych, nieznanych źródeł. Zadania na 5. poziomie powinny oceniać również umiejętność stosowania w praktyce abstrakcyjnych koncepcji, a także oceny wiarygodności źródeł i wybór informacji nie tylko adekwatnych tematycznie, ale także najbardziej godnych zaufania.</p>
Poziom 4.	od 326 do 376 punktów	<p>Dorośli na 4. poziomie umiejętności rozumienia tekstu są w stanie czytać wielostronicowe, złożone teksty i wykonywać wieloetapowe zadania obejmujące wyszukiwanie informacji, wnioskowanie na ich podstawie, ocenę i refleksję nad tekstami i źródłami. Potrafią wywnioskować, czego zadanie dotyczy, korzystając ze złożonych instrukcji niepodawanych wprost. Poprawne wykonanie zadań z tego poziomu często wymaga wnioskowania opartego na posiadanej wiedzy.</p> <p>Teksty i zadania na poziomie 4. mogą dotyczyć abstrakcyjnych lub nowych sytuacji. Często są długie i jest w nich wiele dodatkowych informacji, które mogą odwrócić uwagę badanych od ważnych informacji (dystraktory). Na tym poziomie dorośli są w stanie odpowiadać na złożone pytania, które tylko pośrednio pokrywają się z tym, jak informacje sformułowane są w tekście, lub są rozproszone w wielu jego częściach. Zadania mogą wymagać także oceny subtelnych argumentów lub perswazyjnego tonu i zależności między nimi, a także związków przyczynowo-skutkowych, które respondent powinien wziąć pod uwagę, odpowiadając na pytania. Odpowiedzi w zadaniach mogą zawierać ocenę i porządkowanie złożonych elementów.</p>

Poziom	Zakres punktów	Co potrafią dorośli na tym poziomie
Poziom 3.	od 276 do 326 punktów	Dorośli na 3. poziomie umiejętności potrafią interpretować i rekonstruować treści na podstawie dłuższych fragmentów tekstów, a także wykonywać wieloetapowe operacje, aby zidentyfikować i sformułować odpowiedź. Potrafią odnaleźć, zinterpretować lub ocenić jedną lub więcej informacji, często posługując się różnymi rodzajami wnioskowania. Jeśli zadanie tego wymaga, korzystają z różnych procesów, w tym wyszukiwania informacji, rozumowania i jej oceny. Dorośli na tym poziomie są w stanie porównywać i oceniać wiele różnych informacji, na podstawie ich istotności i wiarygodności. Na tym poziomie teksty są często złożone i długie, zarówno linearne, nieciągłe, jak i mieszane. Ważne informacje mogą być rozproszone na wielu stronach, czasem również w wielu źródłach, które mogą dostarczać sprzecznych danych. Rozumienie struktury tekstu ma kluczowe znaczenie dla rozwiązywania zadań na tym poziomie, zwłaszcza tych, które opierają się na złożonych tekstach cyfrowych i wymagają, aby po nich nawigować. Teksty mogą zawierać specjalistyczne i nieznanne słownictwo czy rodzaje argumentów. Dystraktory mogą być zawarte w tekstach, ale nie są prezentowane tak dobitnie jak kluczowe informacje. Zadania wymagają wyszukiwania jednej lub kilku informacji, ich interpretacji, a także oceny z wykorzystaniem różnego typu wnioskowania. Zadania na poziomie 3. wymagają również ignorowania nieistotnych informacji, aby poprawnie odpowiedzieć na pytanie. Najtrudniejsze zadania na tym poziomie mogą zawierać długie, złożone pytania dotyczące różnych kryteriów, bez jasnych wskazówek, co należy w nich zrobić.
Poziom 2.	od 226 do 276 punktów	Na poziomie 2. dorośli są w stanie wyszukiwać i poprawnie interpretować informacje w dłuższych tekstach, w których są też pojedyncze dystraktory. Potrafią poruszać się po prostszych, kilkustronicowych tekstach cyfrowych, aby znaleźć ważne dla pytania informacje, które mogą być zawarte w różnych częściach tekstu. Rozumieją parafrazy i wnioskuje, korzystając z pojedynczych lub sąsiadujących informacji. Odpowiadając na zadane pytanie, dorośli na poziomie 2. mogą uwzględnić więcej niż jedno kryterium lub wymaganie. Teksty na tym poziomie mogą zawierać wiele akapitów na jeden długiej stronie lub kilku krótszych, a także prostszych stronach internetowych. Teksty nieciągłe mogą zawierać prostą tabelę lub diagram. Wyszukanie kluczowej informacji może wymagać zastosowania narzędzi takich jak w dłuższych tekstach. Zadania i teksty mogą zawierać pojedyncze dystraktory. Niektóre zadania i teksty na tym poziomie mogą dotyczyć nowych, potencjalnie nieznanymi sytuacji. Wykonanie zadań wymaga niekiedy odnajdywania pośrednich powiązań między treścią tekstu i pytania, czasem opiera się też na długich instrukcjach. W niektórych zadaniach instrukcje zawierają bardzo oszczędne wskazówki dotyczące sposobu ich wykonania. Mogą wymagać wyciągania wniosków na podstawie jednej informacji, lub zebrania informacji w kilku cyklach.
Poziom 1.	od 176 do 226 punktów	Dorośli, którzy osiągnęli 1. poziom umiejętności, potrafią znaleźć informację na jednej stronie tekstu, znaleźć link na stronie internetowej lub wybrać odpowiedni tekst spośród kilku, o ile informacje są w nim jasno wskazane. Poruszają się po krótkich tekstach, korzystając z wypunktowanych list lub odrębnych sekcji w ramach jednej strony. Teksty na poziomie 1. mogą być linearne, nieciągłe lub mieszane, zarówno na papierze, jak i w cyfrowych dokumentach. Zazwyczaj jest to jedna strona z maksymalnie kilkuset słowami, a teksty raczej nie zawierają dystraktorów, lub jest ich bardzo niewiele. Teksty nieciągłe mogą mieć strukturę listy, np. na stronie z wynikami wyszukiwania, lub zawierać parę odrębnych sekcji, z ilustracjami lub prostymi diagramami. Zadania na poziomie 1. to proste pytania, które dostarczają pewnych wskazówek dotyczących ich wykonania, i są raczej jednoetapowe. Pytanie i informacja w tekście są sformułowane w podobny sposób. Niektóre zadania mogą wymagać analizy więcej niż jednej informacji.

Poziom	Zakres punktów	Co potrafią dorośli na tym poziomie
Poniżej poziomu 1.	poniżej 176 punktów	<p>Dorośli na poziomie poniżej 1. rozumieją zdania. W przypadku serii zdań o wzrastającej złożoności są w stanie określić, czy zdania mają sens, opierając się na ich wiarygodności (czy opisane zdarzenie mogłoby się wydarzyć) i wewnętrznej logice (czy zdanie jest sensowne). Potrafią czytać krótkie, proste fragmenty tekstów i w określonych miejscach tekstu rozpoznać, które słowo z dwóch pasuje do zdania i sprawia, że jest ono spójne z resztą tekstu (zadanie uzupełniania zdań). Potrafią również znaleźć pojedyncze informacje, np. słowa lub liczby, w bardzo krótkich tekstach, aby odpowiedzieć na proste, jasno sformułowane pytania.</p> <p>Teksty na poziomie 1. są krótkie, ciągłe, z niewieloma elementami takimi jak tytuły lub nagłówki. Nie zawierają dystraktorów ani elementów charakterystycznych dla dłuższych tekstów (np. spisów treści, linków, zakładek w tekstach cyfrowych). Zadania na poziomie 1. są proste i bardzo jasno określają, co i jak należy w nich zrobić. Wymagają rozumienia jedynie na poziomie zdania lub dwóch prostych, sąsiadujących zdań. Gdy tekst składa się z więcej niż jednego zdania, zadanie sprowadza się do pracy z informacjami takimi jak pojedyncze słowa lub frazy.</p>

Tabela 1.3. Opis tego, co potrafią dorośli na poszczególnych poziomach umiejętności matematycznych.

Poziom	Zakres punktów	Co potrafią dorośli na tym poziomie
Poziom 5.	376 punktów lub więcej	<p>Dorośli na poziomie 5. potrafią stosować różne strategie rozwiązywania problemów, aby przeanalizować, ocenić i poddać krytycznej refleksji złożone i formalne operacje matematyczne, również dynamiczne. Rozumieją pojęcia statystyczne i są w stanie krytycznie ocenić, czy dane mogą być wykorzystane do poparcia lub odrzucenia hipotezy. Potrafią wybrać najlepsze reprezentacje graficzne dla relacyjnych zestawów danych.</p>
Poziom 4.	od 326 do 376 punktów	<p>Dorośli na poziomie 4. potrafią stosować różne strategie rozwiązywania problemów, aby przeanalizować, ocenić i poddać krytycznej refleksji różnorodne informacje z dziedziny matematyki, często prezentowanych w nowych kontekstach. Takie informacje mogą być podawane nie wprost. Potrafią opracowywać i stosować różne strategie, aby rozwiązywać wieloetapowe problemy. Mogą one obejmować wnioskowanie o łączeniu informacji z różnych dziedzin matematyki lub stosowanie złożonych, formalnych procedur matematycznych.</p> <p>Na tym poziomie umiejętności dorośli potrafią:</p> <ul style="list-style-type: none"> • obliczać i interpretować proporcje; • opracować strategie porównywania dużych zbiorów danych; • odczytywać i interpretować dane z wieloma zmiennymi prezentowane na pojedynczym wykresie; • analizować złożone formuły algebraiczne, aby zrozumieć relacje między zmiennymi; • poddawać refleksji i rozumować matematycznie, aby oceniać prawdziwość wnioskowania statystycznego lub matematycznego, twierdzeń i argumentów, przy jednoczesnym uwzględnianiu istotnych warunków; • skonkretyzować postawiony problem, tak aby wynik był na odpowiednim poziomie szczegółowości w kontekście danej sytuacji.

Poziom	Zakres punktów	Co potrafią dorośli na tym poziomie
Poziom 3.	od 276 do 326 punktów	<p>Dorośli na poziomie 3. potrafią interpretować, stosować i oceniać rzeczywiste sytuacje z dziedziny matematyki. Oceniają, kiedy i jak zastosować daną informację i opracowują rozwiązanie problemu. Informacja może być mniej klarowna, osadzona w nie zawsze znanym kontekście, zaś język może być formalny i dość złożony. Dorośli rozwiązują zadania, w których procedury matematyczne wymagają stosowania dwóch lub więcej kroków, w których należy uwzględniać kilka warunków. Zadania mogą wymagać wykorzystywania, integracji lub manipulacji różnymi źródłami danych.</p> <p>Dorośli na tym poziomie potrafią:</p> <ul style="list-style-type: none"> oszacowywać lub wykonywać obliczenia z wykorzystaniem szerokiego zakresu liczb całkowitych, dziesiętnych, procentów, ułamków i miar, a także wnioskować o proporcjach; znaleźć brakującą liczbę w zestawie danych przy znanej średniej; rozpoznawać i wykorzystywać wzorce (wizualne i numeryczne) do szacowania wartości; stosować rozumowanie matematyczne, aby oceniać prawdziwość wniosków wyciąganych na podstawie danych, w tym również w odniesieniu do ograniczonego zestawu powiązanych warunków lub stwierdzeń; oceniać twierdzenia i zależności między nimi przy zastosowaniu różnych źródeł danych; rozpoznawać wzory matematyczne opisane w niestandardowym zapisie; wykorzystywać umiejętność wizualizacji przestrzennej, aby analizować bryły i przenosić je z reprezentacji trójwymiarowych do dwuwymiarowych.
Poziom 2.	od 226 do 276 punktów	<p>Dorośli na poziomie 2. potrafią interpretować, wykorzystywać i oceniać informacje matematyczne oraz proste twierdzenia w realistycznych kontekstach. Potrafią odczytywać i wykorzystywać informację prezentowaną w formie graficznej (wykresy kołowe, skumulowane wykresy słupkowe, wykresy liniowe), z formalną terminologią i dystraktorami. Są w stanie wykonywać wieloetapowe procesy matematyczne.</p> <p>Dorośli na tym poziomie potrafią:</p> <ul style="list-style-type: none"> korzystać z dynamicznych aplikacji do wykonywania prostych obliczeń; korzystać z danych prezentowanych w tabelach lub na interaktywnych wykresach, porządkować je; stosować proste rozumowanie proporcjonalne lub rozwiązywać problemy; formułować procesy i wyrażenia w celu matematycznego przedstawienia sytuacji, w tym zestawiania i łączenia informacji; stosować rozumowanie matematyczne podczas oceniania prawdziwości stwierdzeń; szacować lub wykonywać obliczenia obejmujące ułamki zwykłe, ułamki dziesiętne, czas, jednostki miar i mniej powszechnie używane procenty lub wykonywać typowe operacje (np. obliczanie średniej); zastępować i oceniać konteksty obejmujące autentyczne formuły algebraiczne; rozpoznawać wzory w dwuwymiarowych reprezentacjach geometrycznych.

Poziom	Zakres punktów	Co potrafią dorośli na tym poziomie
Poziom 1.	od 176 do 226 punktów	<p>Dorośli na poziomie 1. posługują się liczbami całkowitymi, dziesiętnymi, ułamkami zwykłymi i procentami. Potrafią posługiwać się informacjami matematycznymi prezentowanymi w złożonych, realistycznych kontekstach, w których informacja matematyczna jest prezentowana wprost i używana jest nieformalna terminologia, z niewielką ilością tekstu i dystraktorów. W celu znalezienia rozwiązania potrafią tworzyć proste strategie obejmujące jeden lub dwa kroki.</p> <p>Dorośli na tym poziomie potrafią:</p> <ul style="list-style-type: none"> • interpretować proste zależności przestrzenne, skale na mapach; • odczytywać informacje z tabel i prostych wykresów, uzupełnić prosty wykres słupkowy z liczbami całkowitymi; • rozpoznać największą liczbę w nieuporządkowanej liście, w tym części dziesiętne; • wykonywać podstawowe operacje arytmetyczne, w tym mnożenie i dzielenie liczb całkowitych, pieniędzy lub obliczać powszechnie znane procenty, np. 25% lub 50%.
Poniżej poziomu 1.	Poniżej 176 punktów	<p>Dorośli na poziomie poniżej 1. posługują się liczbami całkowitymi. Są w stanie stosować wiedzę matematyczną, aby rozwiązywać proste, jednoetapowe zagadnienia, w których informacje są przedstawione z wykorzystaniem ilustracji lub wyraźnej struktury. Takie zagadnienia są osadzone w realistycznych, znanych kontekstach, z niewielką ilością tekstu i dystraktorów. Stosowana matematyka jest nieformalna i przejrzysta.</p> <p>Dorośli na tym poziomie potrafią:</p> <ul style="list-style-type: none"> • liczyć do 20 przedmiotów o różnym poziomie organizacji (np. przedstawione losowo, w grupach lub uporządkowane); • układać wydarzenia według porządku chronologicznego; • porównywać nieuporządkowane listy liczb tak, aby znaleźć największą liczbę; • wskazać dane na wykresie; • dodawać i odejmować małe liczby całkowite.

Tabela 1.4 Opis tego, co potrafią dorośli na poszczególnych poziomach rozwiązywania problemów.

Poziom	Zakres punktów	Co potrafią dorośli na tym poziomie
Poziom 4.	326 punktów lub więcej	<p>Dorośli na tym poziomie potrafią rozwiązywać problemy w złożonych, bogatych w informacje kontekstach, integrując wiele źródeł i zależności między nimi, identyfikując i pomijając dystraktory oraz formułując odpowiednie wskazówki. Stosują wieloetapowe procedury, aby rozwiązywać złożone problemy. Dostosowują rozwiązanie do zmieniających się warunków, nawet jeśli są to zmiany subtelne, nieoczekiwane lub wymagające ponownego rozwiązania całego problemu. Rozróżniają istotne i nieistotne zmiany w procedurze, przewidują przyszłe zmiany sytuacji i biorą pod uwagę wiele kryteriów jednocześnie, aby ocenić, czy ich rozwiązanie będzie skuteczne.</p> <p>Dorośli na poziomie 4. angażują się w poniższe procesy poznawcze:</p> <ul style="list-style-type: none"> • tworzą złożone modele mentalne problemów, integrując informacje z wielu źródeł; • odkrywają relacje pomiędzy zadaniem i bodźcem, nawet jeśli połączenia te są trudne do wykrycia lub złożone; • opracowują strategie pozwalające na równoległe osiągnięcie kilku celów, stosują wieloetapowe rozwiązania oraz stale aktualizują modele umysłowe, poszukują strategii i nowych rozwiązań w trakcie rozwiązywania problemów. <p>Dorośli na poziomie 4. angażują się w następujące procesy metapoznawcze:</p> <ul style="list-style-type: none"> • monitorują i poddają refleksji proces rozwiązywania problemów, nawet jeśli środowisko jest złożone i zmienia się nieoczekiwanie; • nieustannie weryfikują i ponownie oceniają swój model mentalny, dostępne informacje i osiągnięte cele; • wykazują adekwatne i natychmiastowe reakcje na zmiany; • dostosowują się do częstych i nieprzewidywalnych zmian; • dostosowują swoje strategie do rozwiązywanych problemów.

Poziom	Zakres punktów	Co potrafią dorośli na tym poziomie
Poziom 3.	od 276 do 326 punktów	<p>Dorośli na poziomie 3. rozwiązują złożone, statyczne problemy lub problemy o umiarkowanej lub dużej dynamice. Rozwiązują problemy z wieloma założeniami lub równoległymi celami. Radzą sobie ze zmieniającymi się problemami wymagającymi przystosowywania się do nowych warunków, odróżniając zmiany istotne od nieistotnych. Planują i stosują wieloetapowe rozwiązania uwzględniające kilka ważnych zmiennych w tym samym czasie, biorąc pod uwagę to, jak mogą wpływać na siebie nawzajem. W przypadku szybko zmieniających się problemów, przewidują przyszły rozwój sytuacji w oparciu o informacje z przeszłości i odpowiednio dostosowują swoje zachowanie.</p> <p>Dorośli na poziomie 3. angażują się w następujące procesy poznawcze:</p> <ul style="list-style-type: none"> • generują modele mentalne dla umiarkowanej i bardzo złożonych problemów; • aktywnie poszukują rozwiązań poprzez ciągłą ocenę informacji dostarczanych w ramach stawianego problemu; • rozróżniają istotne i nieistotne informacje. <p>Dorośli na tym poziomie angażują się w następujące procesy metapoznawcze:</p> <ul style="list-style-type: none"> • monitorują własne zrozumienie problemu i zmian w nim zachodzących; • monitorują i oceniają postępy w osiągnięciu celu; • poszukują rozwiązań poprzez wyznaczanie celów częściowych i ocenę alternatywnych rozwiązań problemu; • poddają refleksji swoje podejście do rozwiązania problemu i, jeśli to konieczne, zmieniają swoją strategię.
Poziom 2.	od 226 do 276 punktów	<p>Dorośli na poziomie 2. potrafią znajdować i stosować kilkietapowe rozwiązania problemów, które wymagają uwzględnienia jednej lub kilku zmiennych. W szybko zmieniających się problemach identyfikują istotne informacje, jeśli taka była instrukcja zadania lub zmiany są wyraźne, zachodzą pojedynczo, są związane z jedną cechą i są raczej łatwo zauważalne. Na tym poziomie problemy do rozwiązania są prezentowane w zorganizowanych kontekstach, zawierają niewiele informacji istotnych dla sedna sprawy. Niewielkie zakłócenia w procedurze mogą być stosowane, o ile można je łatwo rozwiązać, zmieniając początkową strategię podejścia do problemu.</p> <p>Dorośli na poziomie 2. angażują się w następujące procesy poznawcze:</p> <ul style="list-style-type: none"> • tworzą modele mentalne dla prostych i średnio trudnych problemów, dostosowują je w miarę potrzeb; • odpowiednio reagują na zmiany, jeśli są przedstawiane w przejrzysty sposób; • dostosowują strategie do prostych lub umiarkowanie trudnych zmian w stawianym problemie. <p>Dorośli na tym poziomie angażują się w następujące procesy metapoznawcze:</p> <ul style="list-style-type: none"> • monitorują postępy w rozwiązywaniu problemu z jednym, jasno określonym celem; • poszukują najlepszych rozwiązań poprzez ocenę prostych lub umiarkowanie złożonych alternatyw; • poddają refleksji wybraną strategię, jeśli utkną w trakcie rozwiązywania problemu oraz wtedy, kiedy zostaną o to wprost poproszeni w zadaniu.

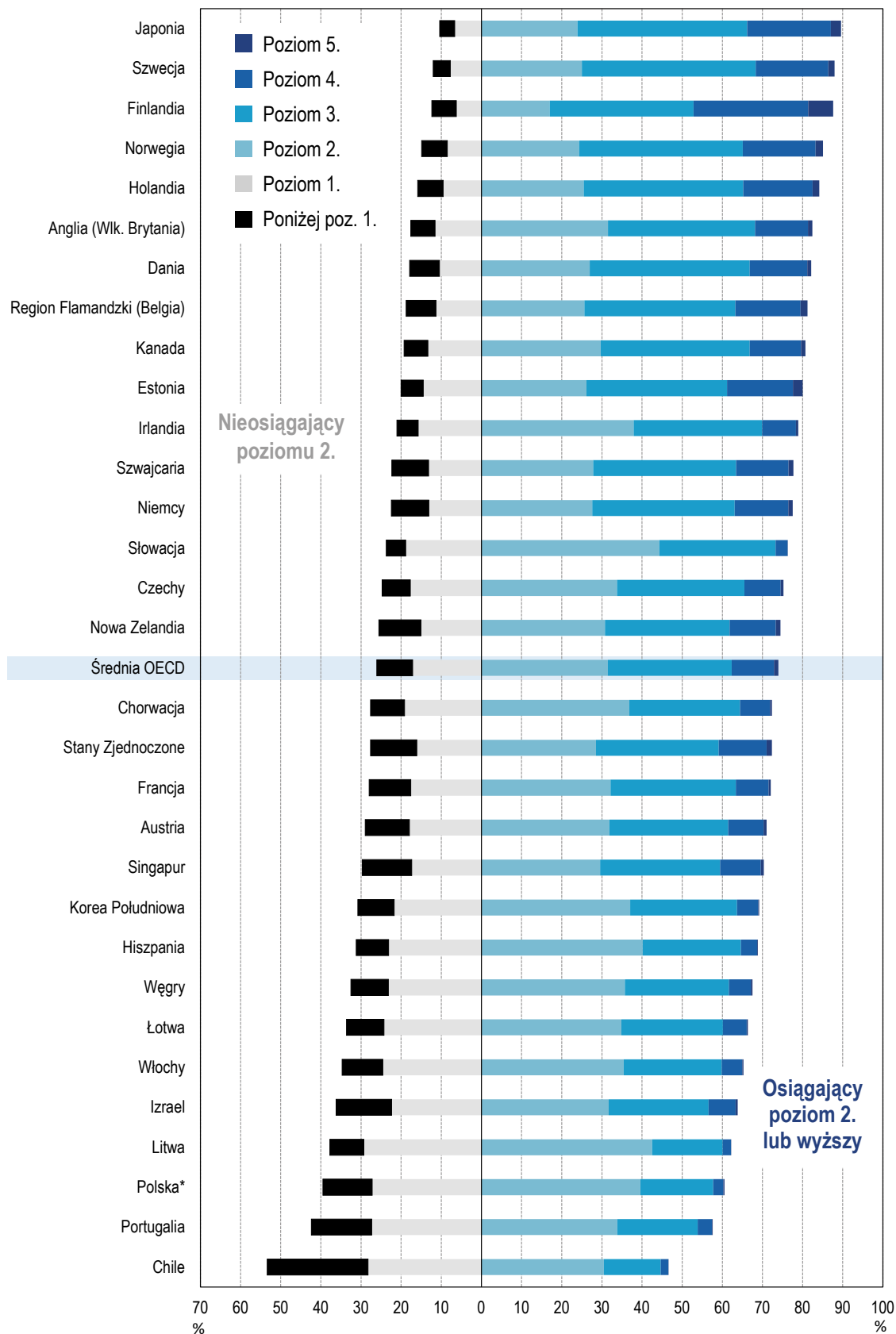
Poziom	Zakres punktów	Co potrafią dorośli na tym poziomie
Poziom 1.	od 176 do 226 punktów	<p>Dorośli na poziomie 1. rozwiązują proste problemy. Zawierają one niewielką liczbę elementów ważnych dla rozwiązania i niewiele dystraktorów. Rozwiązania na tym poziomie są proste i mają niewiele kroków. Konteksty dla problemów są znane, stosowane jest jedno lub dwa źródła informacji, cel zadania musi być pojedynczy i przejrzysty.</p> <p>Dorośli na poziomie 1. angażują się w następujące procesy poznawcze:</p> <ul style="list-style-type: none"> • tworzą modele mentalne dla prostych problemów o klarownej strukturze; • rozumieją związek między zadaniem a bodźcami, które są jasne i umieszczone w ustrukturyzowanym kontekście; • rozwiązują problemy, które się nie zmieniają, nie wymagają więc dostosowywania się do zmian.
Poniżej poziomu 1.	Poniżej 176 punktów	<p>Dorośli na poziomie poniżej 1. rozumieją proste, statyczne problemy w dobrze ustrukturyzowanym środowisku, bez ukrytych elementów lub dystraktorów. Problemy zazwyczaj wymagają jednoetapowego rozwiązania. Dorośli na tym poziomie umiejętności angażują podstawowe procesy poznawcze w celu rozwiązywania problemów, o ile otrzymują wsparcie i wyraźne podpowiedzi.</p>

W Polsce niepokojąco dużo badanych znalazło się na najniższych poziomach. Aż 39% badanych miało poważne trudności z rozumieniem tekstu, a 38% miało najniższe wyniki w matematyce (poziom 1. lub niższy). Co dziesiąta osoba (9%) nie radziła sobie nawet z najprostszymi zadaniami w obu dziedzinach. Są to znacznie gorsze wyniki niż w sąsiadujących z nami krajami. W Niemczech wyniki poniżej 1. poziomu miało ok. 17% osób w rozumieniu tekstu i 16% w rozumowaniu matematycznym, a w Czechach odpowiednio 19% i 18%. Z krajów naszego regionu najlepiej wypadła Estonia, gdzie tylko 8–9% dorosłych miało najniższe wyniki w obu obszarach. Warto zauważyć, że ranking krajów pod względem odsetka dorosłych z niskimi wynikami jest bardzo podobny do rankingu krajów pod względem średnich wyników.

Warto również spojrzeć na odsetki dorosłych z najlepszymi wynikami. W Polsce poziom 4. w rozumieniu tekstu miało zaledwie 2,5% badanych, a w matematyce 3,7%. Jeszcze mniej – tylko 0,3% w rozumieniu tekstu oraz 0,5% w matematyce – osiągało poziom 5. Były to wyniki znacznie słabsze niż w krajach sąsiednich: w Niemczech na poziomie 4. lub 5. znalazło się 19% w czytaniu i 18% w matematyce, a w Czechach odpowiednio 14% i 16%. Zdecydowanie wyróżniała się Estonia, gdzie wysokie wyniki uzyskało aż 29% badanych w czytaniu i 22% w matematyce.

W porównaniu z rozumieniem tekstu i umiejętnościami matematycznymi, dorośli Polacy wypadli jeszcze gorzej pod względem umiejętności rozwiązywania problemów. 48% z nich miało umiejętności na poziomie 1. lub niższym, z czego 8% osób nie przekroczyło progu poziomu 1. Bardzo niewielu badanych (5%) poradziło sobie z najtrudniejszymi zadaniami z poziomu 4.

Rysunek 1.3. Odsetki dorosłych na poszczególnych poziomach umiejętności rozumienia tekstu.

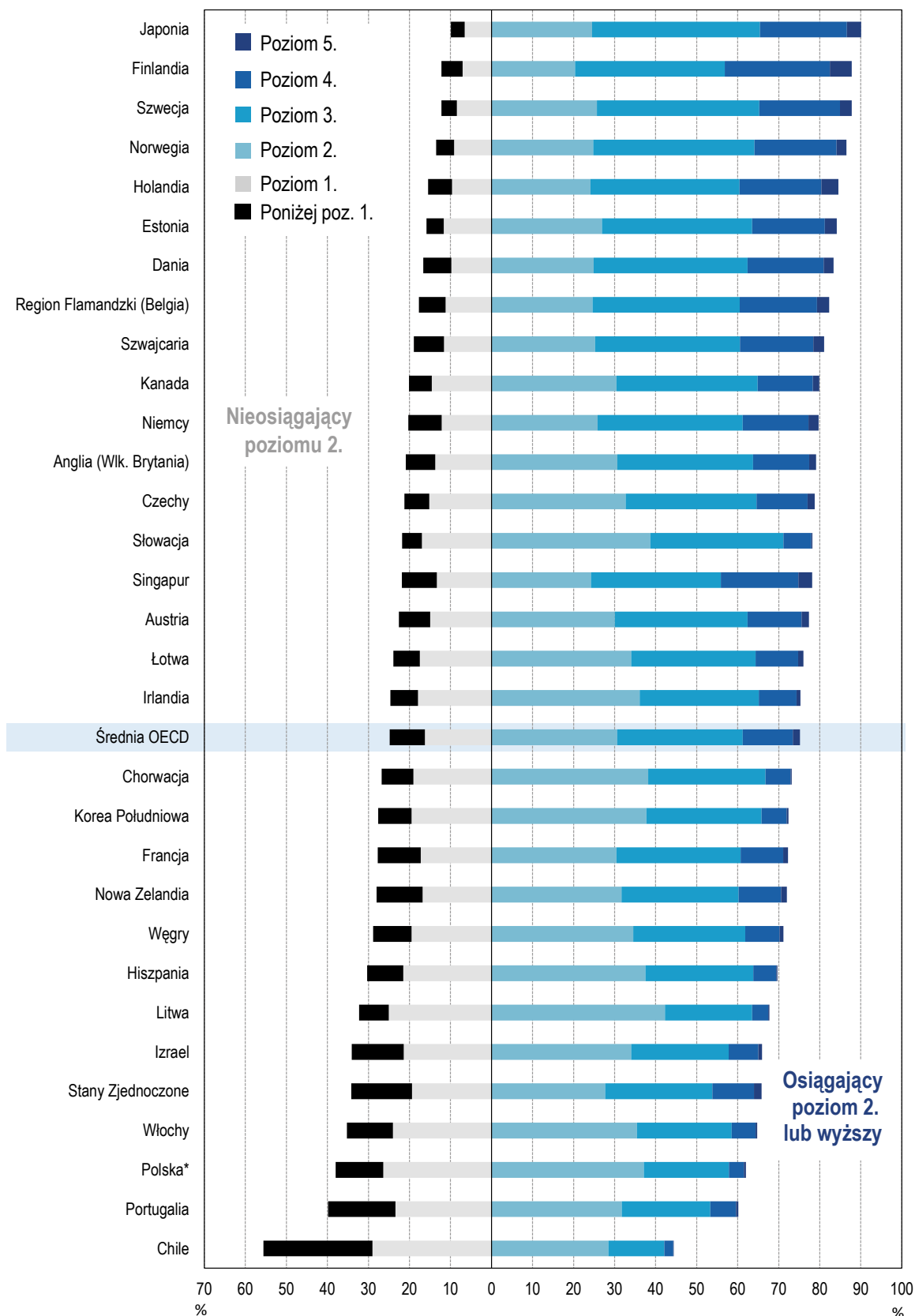


Uwaga: Kraje i gospodarki uporządkowano w kolejności rosnącej pod względem odsetka dorosłych poniżej 2. poziomu umiejętności.

*Polskie wyniki są najprawdopodobniej zaniżone i należy je interpretować z ostrożnością.

Źródło: opracowanie własne na podstawie: OECD, 2024.

Rysunek 1.4. Odsetki dorosłych na poszczególnych poziomach umiejętności rozumowania matematycznego.

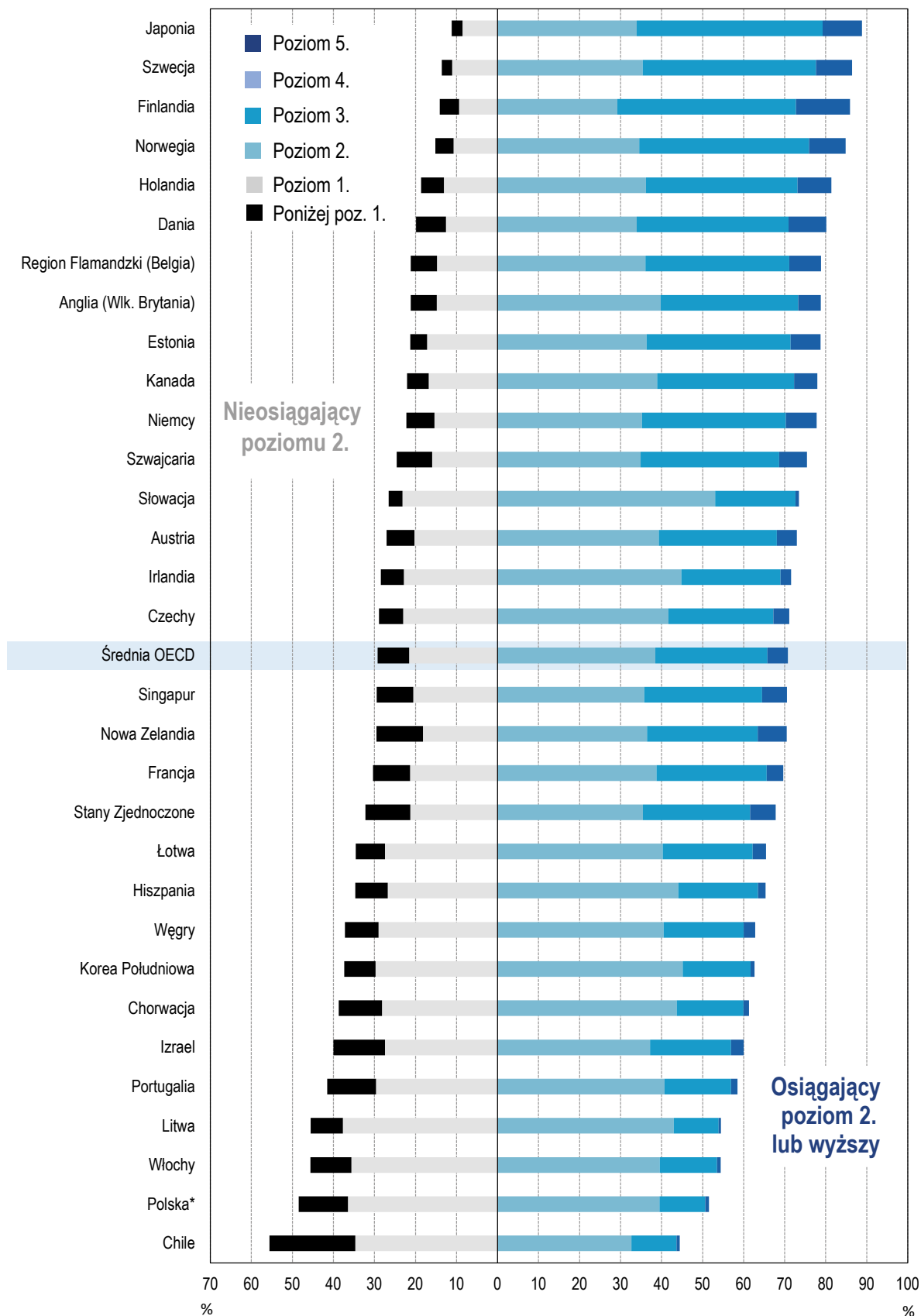


Uwaga: Kraje i gospodarki uporządkowano w kolejności rosnącej pod względem odsetka dorosłych poniżej 2. poziomu umiejętności.

*Polskie wyniki są najprawdopodobniej zaniżone i należy je interpretować z ostrożnością.

Źródło: opracowanie własne na podstawie: OECD, 2024.

Rysunek 1.5. Odsetki dorosłych na poszczególnych poziomach umiejętności rozwiązywania problemów.



Uwaga: Kraje i gospodarki uporządkowano w kolejności rosnącej pod względem odsetka dorosłych poniżej 2. poziomu umiejętności.

*Polskie wyniki są najprawdopodobniej zaniżone i należy je interpretować z ostrożnością.

Źródło: opracowanie własne na podstawie: OECD, 2024.

Wysokie odsetki osób z niskimi umiejętnościami są zaskakujące, jeśli weźmiemy pod uwagę wyniki poprzedniego badania oraz zmiany struktury wykształcenia i sytuacji społeczno-ekonomicznej, które zaszły w ostatniej dekadzie w Polsce. Częściowo można je wyjaśnić problemami metodologicznymi badania PIAAC w Polsce. Jak pokazujemy w rozdziale 3, w badaniu PIAAC niedoreprezentowane były, w porównaniu z innymi danymi dostępnymi w Polsce, osoby z wyższym wykształceniem. Alternatywne przeważenie danych poprawia polskie wyniki o kilka punktów. W przypadku odsetków osób z niskimi umiejętnościami, oszacowania tych osób są mniejsze o 2–3 punkty procentowe, co nie jest znaczącą zmianą. Drugim rodzajem problemów, który opisujemy szerzej w kolejnym rozdziale, porównując wyniki obu edycji badania, są problemy z motywacją respondentów i skalowaniem wyników.

Wskaźniki motywacji i zaangażowania respondentów, jakie możemy wyliczyć na podstawie danych PIAAC, to średni czas odpowiedzi oraz częstotliwość szybkich odpowiedzi, które mogą świadczyć o zgadywaniu (tzw. *rapid guessing*, np. czas odpowiedzi krótszy niż 5 sekund). Znaczenie ma także częstość pomijania pytań³. Krótsze odpowiedzi i częstość pominięć miały wpływ (jak pokazujemy w rozdziale 2) na niższe wyniki w Polsce. Badanie PIAAC miało złożony schemat badania i wyliczania wyników, ale możemy porównać te wskaźniki dla osób, które rozwiązywały poszczególne części testów. Takie porównanie wyraźnie pokazuje, że im krótszy czas poświęcony na zadania i wyższy odsetek pominięć, tym niższe były osiągnięte wyniki. Dla osób, którym przydzielono zadania z zakresu rozumienia tekstu, z pominięciem tych, które rozwiązywały tylko zadania z podstaw rozumienia tekstu (*reading components*), odsetek pominiętych zadań tłumaczy ok. 24% zróżnicowania wyników, a odsetek szybkich odpowiedzi ok. 10% zróżnicowania. Te dwie zmienne wyjaśniają łącznie ok. 28% zmienności wyników w tej grupie. W przypadku rozumowania matematycznego było to 30%, a rozwiązywania problemów – 24%.

Wśród respondentów rozwiązujących zadania z rozumienia tekstu, odsetek pominięć wyniósł 11%, średni czas na zadanie to 40 sekund, a odsetek odpowiedzi udzielonych w czasie krótszym niż 5 sekund – 16%. W grupie osób poniżej poziomu 2. tej umiejętności wskaźniki były znacząco gorsze: odsetek pominiętych odpowiedzi wzrósł do 18%, średni czas skrócił się do 28 sekund, a odsetek szybkich odpowiedzi zwiększył się do 24%. Podobne tendencje zaobserwowano w pomiarze umiejętności matematycznych. Średni czas odpowiedzi wyniósł 44 sekundy, podczas gdy w grupie osób poniżej 2. poziomu umiejętności były to średnio 33 sekundy. Odsetek osób poświęcających na zadanie mniej niż 5 sekund wyniósł odpowiednio 10% i 15%, a odsetek pominięć wśród osób o niskich umiejętnościach wyniósł 16% i 6%. Niemal identyczny wzorzec

³ W międzynarodowym skalowaniu wyników pominięcie zadania uznaje się za błędną odpowiedź. Wprowadzono jednak wyjątek: jeśli czas pracy nad zadaniem przekracza 5 sekund i podjęto minimalne wymagane działania, pominięcie traktuje się jako odpowiedź niepoprawną. Gdy te warunki nie są spełnione, zadanie oznacza się jako „niewidziane przez respondenta” (*not reached*). Nie jest to traktowane jako odpowiedź niepoprawna, ale jako brak danych. Większość pominiętych zadań trafiło do kategorii odpowiedzi niepoprawnych, ale niewielka część została uznana za brak danych, który wymaga oszacowania w modelu statystycznym.

występuje w zadaniach z rozwiązywania problemów. Osoby poniżej poziomu 2. umiejętności poświęcały mniej czasu na zadanie (średnia 34 vs. 40 sekund), częściej udzielały bardzo krótkich odpowiedzi (21% vs 15% zadań) oraz częściej pomijały zadania (16% vs 10%). Wyniki te jasno wskazują, że osoby bardziej zaangażowane, spędzające więcej czasu na rozwiązaniu, osiągały zdecydowanie lepsze rezultaty.

Niestety nie dysponujemy danymi jednostkowymi z innych krajów, które wniosłyby więcej informacji, pokazując relatywną sytuację w Polsce i innych krajach. Szerzej o tych problemach piszemy w kolejnym rozdziale, opisując zmiany w porównaniu z poprzednią edycją badania. Wszystko wskazuje na to, że na niski wynik dorosłych w Polsce wpłynął poziom zaangażowania respondentów. Wpływa on na nawigację po zadaniach testowych i ma znaczenie w sposobie wyliczania wyników.

Wiele wskazuje więc na to, że wyniki Polski są zaniżone, ale nie możemy precyzyjnie oszacować skali tego niedoszacowania. Podobne problemy mogą dotyczyć też innych krajów, ale Polska wyróżnia się pod tym względem, co uwzględniono w raporcie międzynarodowym z badania. Pamiętając o tych zastrzeżeniach, należy zauważyć, że ewentualne korekty nie doprowadziłyby do bardzo znaczących zmian pozycji Polski w rankingach.

Różnice w umiejętnościach mężczyzn i kobiet

Różnice w rozumieniu tekstu i umiejętnościach matematycznych między kobietami a mężczyznami kształtują się pod wpływem czynników społeczno-kulturowych, które ujawniają się już na wczesnych etapach edukacji. Badania pokazują, że stereotypy płciowe związane z umiejętnościami matematycznymi pojawiają się przed zaobserwowaniem rzeczywistych różnic w wynikach i wpływają na samoocenę oraz motywację uczniów (Cvencek, Meltzoff i Greenwald, 2011). Skutkuje to m.in. odmiennymi wyborami edukacyjnymi i zawodowymi – kobiety częściej wybierają kierunki humanistyczne i społeczne, podczas gdy mężczyźni dominują w dziedzinach związanych z matematyką i przedmiotami ścisłymi (STEM). Te wzorce utrwalają się w życiu zawodowym – mężczyźni częściej pracują w zawodach wymagających zaawansowanych umiejętności matematycznych, co pozwala im na ciągłe doskonalenie tych kompetencji, podczas gdy kobiety częściej rozwijają umiejętności językowe i społeczne. Dodatkowym czynnikiem są wyzwania związane z równoważeniem pracy i obowiązków rodzinnych, szczególnie w przypadku matek. Kobiety częściej wybierają elastyczne formy zatrudnienia i krótsze godziny pracy, co może ograniczać możliwości rozwoju kompetencji zawodowych.

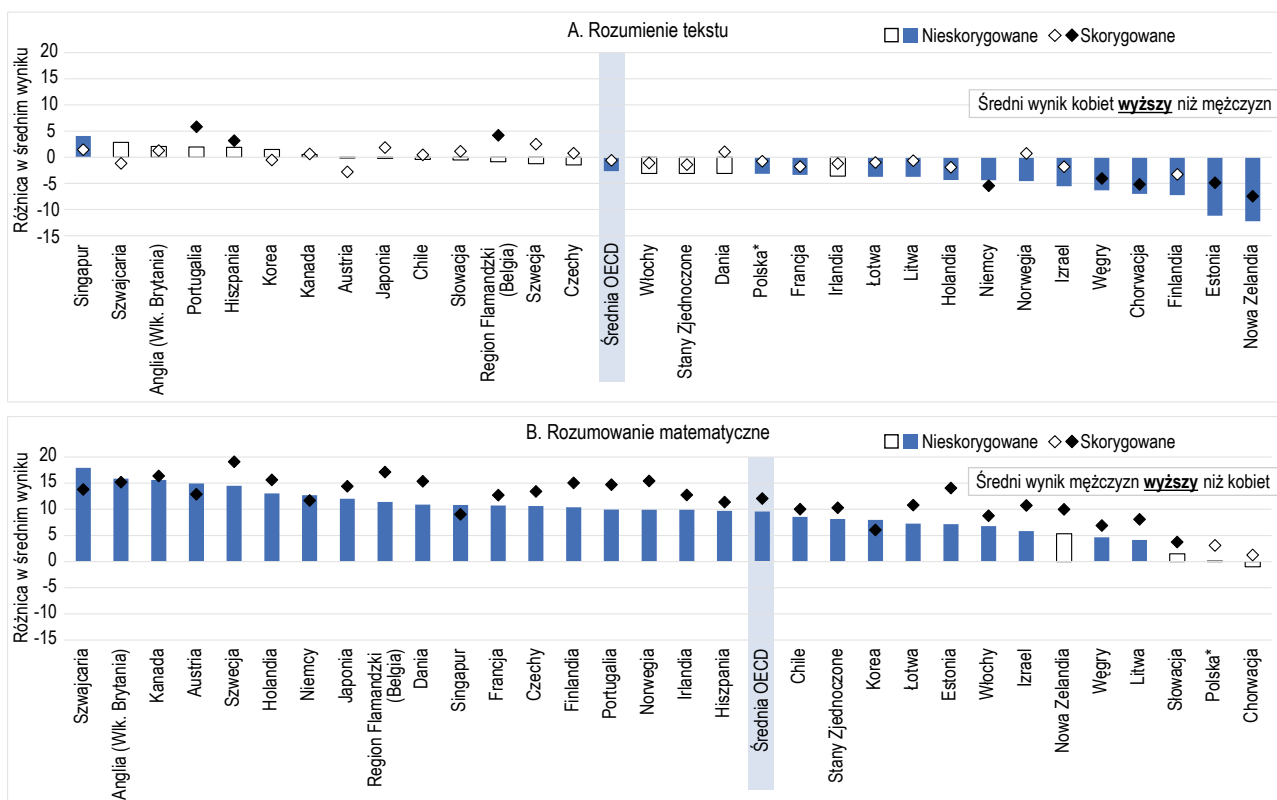
Badanie PIAAC dostarcza danych o różnicach w umiejętnościach, które są ważnym uzupełnieniem obrazu nierówności płciowych i ich uwarunkowań. W krajach OECD kobiety osiągają wyniki o 3 punkty lepsze niż mężczyźni w rozumieniu tekstu. Mężczyźni z kolei uzyskują o 10 punktów więcej w rozumowaniu matematycznym i o 3 punkty więcej w teście rozwiązywania problemów. Ta tendencja jest widoczna w wielu innych krajach europejskich. W wiodącej pod względem wyników Finlandii różnica na korzyść kobiet w rozumieniu

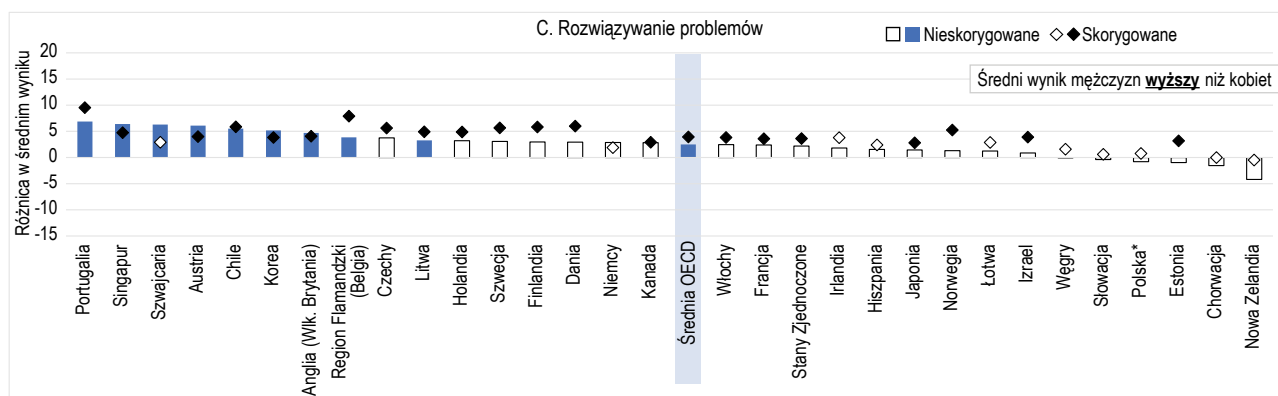
tekstu wynosi 11 punktów, a na korzyść mężczyzn w matematyce – 9 punktów. We Francji kobiety uzyskują o 9 punktów więcej w rozumieniu tekstu, a mężczyźni o 11 punktów więcej w matematyce. W Szwecji różnica na korzyść kobiet w rozumieniu tekstu wynosi 4 punkty, a na korzyść mężczyzn w matematyce – 13 punktów. W niektórych krajach europejskich różnice te są mniejsze lub są odwrócone.

W Polsce różnice między płciami są niewielkie i nieistotne statystycznie (3 punkty na korzyść kobiet w rozumieniu tekstu, 2 punkty na korzyść mężczyzn w matematyce i 1 punkt na korzyść mężczyzn w rozwiązywaniu problemów). Podobne są też odsetki kobiet i mężczyzn z niskimi umiejętnościami (poniżej 2. poziomu). Jedynie w rozumieniu tekstu jest nieco więcej mężczyzn niż kobiet z niskimi umiejętnościami (odpowiednio 41,2 i 37,8%, czyli o 3,4 pkt proc. różnicy).

Proste porównanie średnich wyników może być mylące, bo kobiety i mężczyźni różnią się pod względem wieku czy wykształcenia. Różnice skorygowane o takie zmienne jak wykształcenie, wiek, status migracyjny, język używany w domu i wykształcenie rodziców, pozwalają, poprzez kontrolę tych zmiennych, na bardziej precyzyjne oszacowanie rzeczywistych różnic w umiejętnościach kobiet i mężczyzn. Jak pokazują dane na rysunku 1.6, w niektórych krajach skorygowanie wyników prowadzi do zmniejszenia, a nawet odwrócenia różnic między płciami.

Rysunek 1.6. Skorygowane i nieskorygowane różnice między płciami w zakresie kluczowych umiejętności przetwarzania informacji (mężczyźni minus kobiety).





Uwagi: Różnice w wynikach między kobietami i mężczyznami przedstawiono w dwóch ujęciach: jako proste porównanie średnich wyników (różnice nieskorygowane) oraz z uwzględnieniem wpływu dodatkowych czynników, takich jak wykształcenie, wiek, status migracyjny, język używany w domu i wykształcenie rodziców (różnice skorygowane). Ciemniejsze kolory oznaczają różnice, które są statystycznie istotne na poziomie 5%.

Kraje i gospodarki są uszeregowane w porządku malejącym według nieskorygowanej różnicy biegłości między mężczyznami i kobietami.

**Polskie wyniki są najprawdopodobniej zaniżone i należy je interpretować z ostrożnością.*

Źródło: OECD, 2024, tabela A.2.7.

Różnice w umiejętnościach ze względu na wiek

Umiejętności kształtują się przez całe życie: są nabywane i rozwijane, ale mogą też zanikać (OECD, 2019). Na te procesy wpływa szereg czynników, w tym uwarunkowania genetyczne (Plomin i Deary, 2015) i procesy starzenia poznawczego (Desjardins i Warnke, 2012), warunki domowe, wsparcie rodziny, edukacja i udział w szkoleniach, korzystanie z poszczególnych umiejętności w trakcie wykonywania pracy (Heckman i Kautz, 2013), a także kultura (Baltes, 1997) oraz indywidualne przekonania, postawy i wartości. Ze względu na wzrost średniej długości życia i wydłużający się okres aktywności zawodowej, zrozumienie różnic umiejętności między grupami wiekowymi ma kluczowe znaczenie dla oceny i skuteczności działań na rzecz produktywności i integracji społecznej.

W drugim cyklu PIAAC w Polsce badano osoby urodzone w latach 1957–2007, czyli mające od 16 do 65 lat w czasie realizacji badania. Włączono więc zarówno osoby realizujące jeszcze obowiązek szkolny, studentów, jak i osoby aktywne na rynku pracy oraz te, które przeszły już na emeryturę. Różnice w umiejętnościach między różnymi grupami wiekowymi wynikają nie tylko z uwarunkowań biologicznych, ale także z efektów kohortowych (różne doświadczenia i polityka edukacyjna) oraz efektów okresowych (wpływy zmieniające się w czasie, np. kryzysy gospodarcze, pandemia COVID-19). Jednak badania przekrojowe, takie jak PIAAC, nie mogą w pełni oddzielić od siebie różnych efektów.

Rysunek 1.7 ilustruje średnią biegłość w rozumieniu tekstu, rozumowaniu matematycznym i rozwiązywaniu problemów w różnych grupach wiekowych. Porównanie średnich wyników w różnych kategoriach wieku pokazuje, że młodszy dorośli (16–34 lata) dominowali pod względem umiejętności we wszystkich testowanych obszarach. Grupa wiekowa 25–34 lata

osiągała najlepsze rezultaty w rozumieniu tekstu w 15 z 31 krajów, podczas gdy osoby w wieku 16–24 lat miały najwyższe wyniki w tej dziedzinie w 14 krajach. W rozumowaniu matematycznym rozkład był podobny: osoby w wieku 25–34 lat zajmowały pierwszą pozycję w 14 krajach, a grupa 16–24 lata – w 12 krajach. Z zadaniami mierzącymi umiejętność rozwiązywania problemów najlepiej poradzili sobie 25–34-latkowie w 16 krajach, natomiast 16–24-latkowie – w 14 krajach. Polska wyróżnia się pod tym względem: najmłodsza kategoria wieku (16–24 lata) uzyskała najwyższe wyniki we wszystkich trzech dziedzinach.

Rysunek 1.7. Umiejętność rozumienia tekstu, rozumowania matematycznego i rozwiązywania problemów w podziale na wiek.



Uwaga: Badanie objęło dorosłych w wieku 16–65 lat; również tych, z którymi ze względu na barierę językową przeprowadzono jedynie ankietę zastępczą.

Kraje uszeregowano w kolejności malejącej według średniego poziomu zaawansowania osób w wieku 25–35 lat.

**Polskie wyniki są najprawdopodobniej zaniżone i należy je interpretować z ostrożnością.*

Źródło: OECD, 2024 (tabela A.2.4).

Dla Polski porównanie średnich wyników trzech dziedzin pokazuje systematyczny spadek wyników wraz z wiekiem. Największą różnicę widać między grupami 45–54 i 55–65 lat, gdzie wyniki spadają o 14–19 punktów. Różnice między najmłodszymi i najstarszymi (16–24 vs 55–65) to różnice 27,3 punktu w matematyce, 25,4 punktu w rozumieniu tekstu oraz 23,5 punktu w rozwiązywaniu problemów. Jest to około połowy wartości odchylenia standardowego, które wynosiło odpowiednio: 55,3 punktu dla matematyki, 50,9 punktu dla rozumienia tekstu i 43,6 punktu dla rozwiązywania problemów, co sugeruje zbliżoną względną skalę zmian we wszystkich dziedzinach.

Porównanie średnich w podgrupach warto zestawić z innym, bardziej obrazowym, wskaźnikiem, jakim jest odsetek dorosłych z niskimi umiejętnościami. W Polsce najniższy odsetek słabych wyników występuje w grupie 16–24 lata (około 30–40%), a najwyższy w grupie 55–65 lat (50–62%). Podobnie jak w przypadku porównania średnich w podgrupach, największy wzrost odsetka osób z niskimi umiejętnościami widać, porównując dwie ostatnie kategorie: 45–54 lata i 55–65 lat, gdzie we wszystkich dziedzinach wzrost wynosi około 13–14 punktów procentowych.

Tabela 1.5. Średnie wyniki osób dorosłych oraz odsetek dorosłych w Polsce z niskimi wynikami rozumienia tekstu, rozumowania matematycznego i rozwiązywania problemów.

Badane umiejętności	Przedział wiekowy				
	16–24 lata	25–34 lata	35–44 lata	45–54 lata	55–65 lat
	Średnie wyniki				
Rozumienie tekstu	246,3 (2,2)	242,9 (2,0)	236,3 (1,8)	236,8 (2,0)	220,9 (2,0)
Rozumowanie matematyczne	249,6 (2,0)	246,0 (2,3)	239,6 (1,9)	241,1 (2,1)	222,3 (2,3)
Rozwiązywanie problemów	236,0 (2,0)	234,5 (1,6)	226,3 (1,8)	226,6 (1,9)	212,5 (1,7)
	Odsetki osób poniżej 2. poziomu umiejętności				
Rozumienie tekstu	32,1 (2,1)	33,8 (1,9)	37,8 (1,8)	38,7 (1,9)	51,9 (1,8)
Rozumowanie matematyczne	30,7 (2,1)	33,1 (2,3)	36,6 (1,9)	36,4 (1,9)	50,3 (1,9)
Rozwiązywanie problemów	39,6 (2,3)	40,5 (1,9)	47,7 (2,1)	48,7 (2,2)	61,8 (1,7)

Uwagi: Polskie wyniki są najprawdopodobniej zaniżone i należy je interpretować z ostrożnością.

Liczby w nawiasach oznaczają błąd standardowy.

Źródło: opracowanie własne na podstawie danych PIAAC.

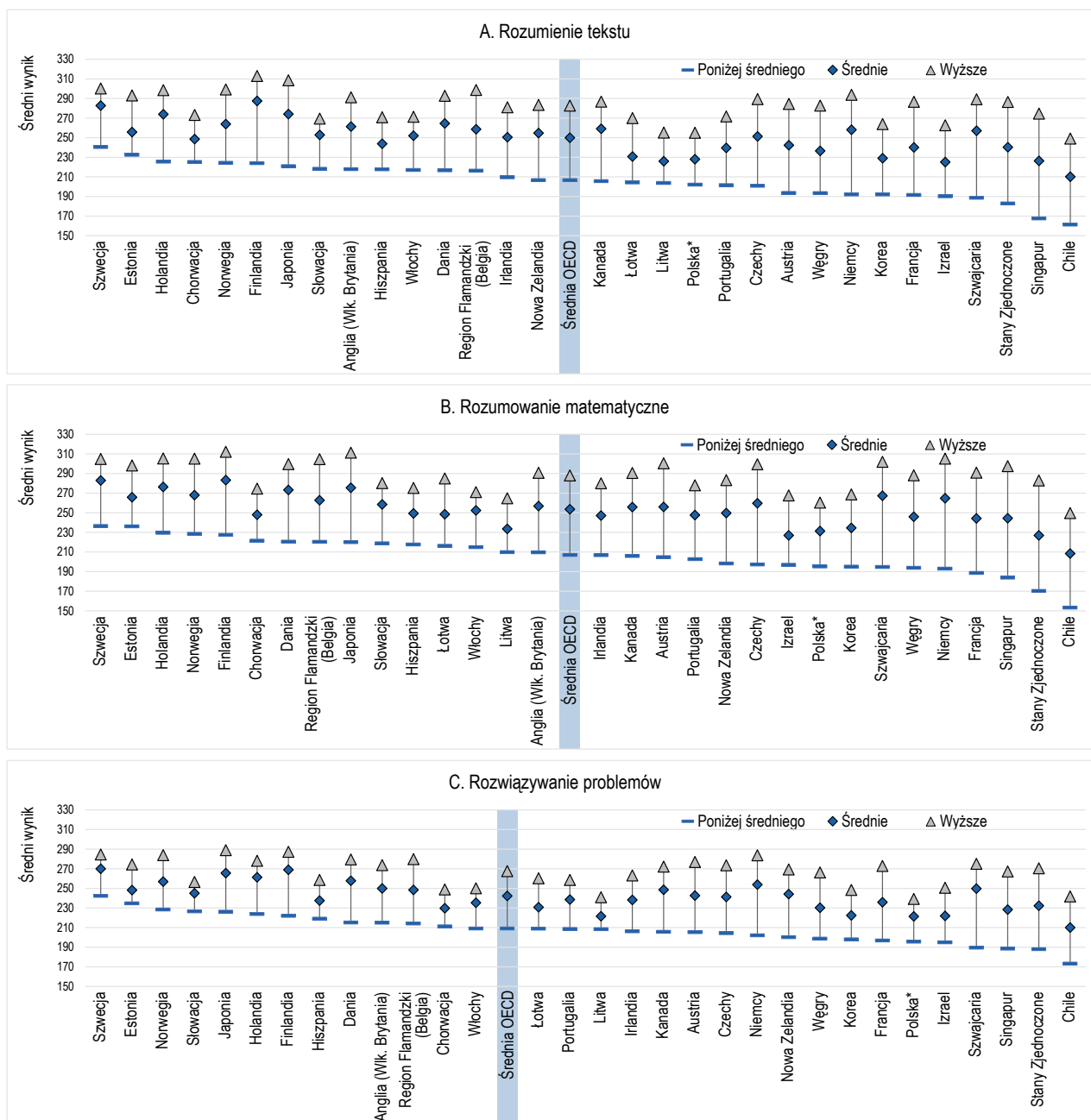
Różnice związane z poziomem wykształcenia

Umiejętności rozwijają się przez całe życie poprzez formalną edukację, udział w szkoleniach oraz nieformalne uczenie się. W każdym momencie życia umiejętności osób dorosłych odzwierciedlają ich wcześniejsze doświadczenia edukacyjne. Ten długi proces rozpoczyna się od edukacji przedszkolnej, kontynuowany jest przez formalne kształcenie szkolne i w różnych formach trwa w dorosłości. Ocena związku między edukacją, możliwościami uczenia się a biegłością w umiejętnościach dorosłych jest złożona, ponieważ rozwój umiejętności jest dynamiczny, a wcześniej nabyte kompetencje wpływają na późniejsze procesy uczenia się (Cunha i in., 2006). Zmiany te zachodzą w różnym tempie. O ile np. w krajach Europy Zachodniej wzrost popularności wyższego wykształcenia miał miejsce w latach 70. i 80., to w Polsce nastąpił dopiero od lat 90. Gdy obserwujemy różnice w poziomie umiejętności między starszymi a młodszymi osobami, musimy pamiętać, że nie wynikają one wyłącznie z wieku czy poziomu wykształcenia, ale także z tego, w jakich okresach i warunkach te osoby się kształciły. Na przykład wyższe wykształcenie było mniej powszechne wśród starszego pokolenia i stało się bardziej osiągalne dla młodszych pokoleń dzięki upowszechnieniu kształcenia w szkołach średnich i zwiększonej ofercie szkół wyższych. Respondenci badania PIAAC przechodzili przez system edukacji w bardzo różnych okresach historycznych: od głębokiego PRL-u, przez transformację ustrojową, reformy edukacji lat 90. i 2000., aż po czasy pandemii COVID-19. Różnice w wynikach mogą odzwierciedlać nie tylko czynniki indywidualne, ale także szersze zmiany społeczne i reformy edukacyjne. W badaniu przekrojowym, jakim jest PIAAC, nie możemy precyzyjnie rozdzielić, czy zaobserwowane różnice wynikają z wieku biologicznego (procesu starzenia się), przynależności do pokolenia (wspólnych doświadczeń osób urodzonych w tym samym czasie), czy też z wydarzeń wpływających na wszystkich jednocześnie.

Porównanie umiejętności dorosłych pod względem wykształcenia ograniczono do osób w wieku 25–65 lat, które w większości zakończyły już edukację. Takie ograniczenie pozwoliło na bardziej miarodajne porównania, eliminując zniekształcenia, które wiążą się z uwzględnieniem w porównaniach osób w trakcie nauki (np. studentów, którzy mają tylko wykształcenie średnie).

Rysunek 1.8. Średnia biegłość w kluczowych umiejętnościach przetwarzania informacji według poziomu wykształcenia.

Umiejętność rozumienia tekstu, rozumowania matematycznego i rozwiązywania problemów.



Uwaga: Badanie objęło dorosłych w wieku 25–65 lat; również osoby, z którymi ze względu na barierę językową przeprowadzono jedynie ankietę zastępczą, na podstawie której przeliczono podane informacje o latach nauki na poziomie wykształcenia. Kategoryzacja poziomów wykształcenia opiera się na Międzynarodowej Standardowej Klasyfikacji Kształcenia (ISCED) z 2011 roku. Klasyfikacja grupuje poziomy edukacji na: podstawowe i gimnazjalne (ISCED 1, 2 i 3 krótkie), średnie i policealne (ISCED 3 i 4) oraz wyższe (ISCED 5, 6, 7 i 8). Tam, gdzie to możliwe, kwalifikacje zagraniczne są porównywane z najbliższymi odpowiednikami w krajowym systemie edukacji.

*Polskie wyniki są najprawdopodobniej zaniżone i należy je interpretować z ostrożnością.

Źródło: opracowanie własne na podstawie danych PIAAC.

W Polsce dorośli 25+ z wykształceniem wyższym osiągnęli znacznie lepsze wyniki niż osoby z pozostałych kategorii, co potwierdza trend obserwowany w innych krajach. Warto zauważyć, że w niektórych krajach osoby z wykształceniem średnim uzyskiwały wyniki wyższe niż absolwenci

studiów wyższych w innych krajach. Przykładowo Finowie ze średnim wykształceniem uzyskiwali 288 punktów w rozumieniu tekstu, przewyższając osoby z wyższym wykształceniem z 19 spośród 31 krajów, w tym Polski. Interpretując wyniki, trzeba pamiętać, że udziały procentowe kategorii wykształcenia wpływają na statystyki. Kraje z większym odsetkiem osób z wyższym wykształceniem mogą mieć mniejsze zróżnicowanie wyników, stabilizując się wokół ogólnej średniej. W krajach z niskim udziałem osób słabo wykształconych możliwe są bardziej skrajne wyniki.

Przyjrzyjmy się dokładniej danym z Polski. Poza międzynarodowym podziałem na osoby z wykształceniem wyższym i poniżej średniego, warto uwzględnić też rozróżnienie między wykształceniem zasadniczym zawodowym a średnim (ogólnokształcącym, technicznym lub policealnym). Osoby z wyższym wykształceniem osiągają znacząco lepsze wyniki we wszystkich badanych umiejętnościach w porównaniu do pozostałych kategorii. Warto również zauważyć, że osoby z wykształceniem średnim uzyskują lepsze rezultaty niż absolwenci szkół zasadniczych zawodowych. Najwyraźniejsze różnice między tymi grupami widoczne są w rozumowaniu matematycznym – różnica wynosi 21,4 punktu. W przypadku rozwiązywania problemów i rozumienia tekstu różnice są mniejsze i wynoszą odpowiednio 13,2 i 17,1 punktu⁴.

Tabela 1.6. Rozkład procentowy osób w wieku 25–65 według poziomu osiągniętego wykształcenia oraz średnie wyniki rozumienia tekstu, rozumowania matematycznego oraz umiejętności rozwiązywania problemów w poszczególnych kategoriach wykształcenia w Polsce.

Wykształcenie	Udział procentowy	Rozumienie tekstu	Rozumowanie matematyczne	Rozwiązywanie problemów
gimnazjalne lub niższe	6,9 (0,4)	202,2 (3,3)	195,2 (3,7)	195,7 (3,0)
zasadnicze zawodowe	36,1 (0,6)	220,8 (1,9)	221,9 (1,7)	215,8 (1,4)
średnie lub policealne	27,9 (0,8)	237,9 (1,7)	243,3 (1,6)	229,0 (1,5)
wyższe	29,0 (0,8)	255,0 (1,9)	260,3 (2,1)	239,2 (1,8)

Polskie wyniki są najprawdopodobniej zaniżone i należy je interpretować z ostrożnością. Dodatkowe analizy pokazały też, że struktura osiągniętego wykształcenia w polskim zbiorze danych istotnie różni się od struktury wykształcenia w danych pochodzących ze spisu powszechnego oraz z innych danych, co wpływa na wartości oszacowań.

Liczby w nawiasie oznaczają błąd standardowy.

Źródło: opracowanie własne na podstawie danych PIAAC.

Porównanie średnich wskazuje na duże różnice ze względu na wiek i wykształcenie w Polsce. Ale za tymi wynikami kryje się bardzo duże zróżnicowanie wewnętrzne w ramach poszczególnych kategorii. Wiek, wykształcenie i płeć tłumaczą tylko około 10% różnic

⁴ Jak sygnalizujemy w rozdziale 3 (zob. Tabela 3.1, por. OECD 2024b), w danych PIAAC niedoreprezentowane są osoby z wykształceniem wyższym. Alternatywne przeważenie danych wykorzystujące rozkłady procentowe ze spisu powszechnego zmienia istotnie oszacowania udziałów kategorii wykształcenia, ale różnice w oszacowaniach wyników w ramach poszczególnych kategorii są bardzo podobne: różnice w punktowych oszacowaniach są minimalne i nie przekraczają 1–2 punktów.

w umiejętnościach osób dorosłych w Polsce. Choć są to ważne czynniki, to pozostałe 90% różnic wynika z innych przyczyn. Oznacza to, że nawet wśród osób o podobnym wieku i wykształceniu znajdziemy zarówno osoby świetnie radzące sobie z zadaniami, jak i te, które mają z nimi bardzo duże trudności.

Podsumowanie

Badanie PIAAC daje wyjątkowy wgląd w kompetencje dorosłych. Umiejętności te są kluczowe dla indywidualnego sukcesu, rozwoju społecznego i spójności społecznej. Osoby o wyższych kompetencjach są lepiej przygotowane do radzenia sobie w złożonej rzeczywistości współczesnego świata oraz do przyszłych wyzwań.

Badanie pokazuje duże zróżnicowanie poziomu umiejętności między krajami. Kraje skandynawskie (Finlandia, Szwecja i Norwegia) oraz Japonia uzyskały w badaniu najlepsze wyniki.

Polska we wszystkich trzech badanych dziedzinach (rozumienie tekstu, rozumowanie matematyczne i rozwiązywanie problemów) uzyskała jeden z najniższych wyników wśród krajów biorących udział w badaniu. Wiele wskazuje na to, że polskie wyniki badania PIAAC w Polsce są niedoszacowane. Niedoszacowanie wyników w Polsce wynika z wielu czynników. Niewielką część można przypisać niedoszacowaniu osób z wyższym wykształceniem. Drugim czynnikiem jest motywacja. Jak pokazują dane, spory odsetek szybkich odpowiedzi i braków odpowiedzi znacząco obniżył oszacowania poziomu umiejętności. Tego rodzaju problemy mogły też wystąpić w innych krajach, ale jak pokazujemy w kolejnym rozdziale, są one szczególnie duże w Polsce. Interpretując wyniki, trzeba brać pod uwagę te czynniki. Warto również podkreślić, że ewentualne korekty nie zmieniłyby radykalnie pozycji Polski w rankingu.

O ile na podstawie średnich wyników możemy przedstawić ranking krajów, to niewiele nam to mówi o faktycznych umiejętnościach dorosłych. Po pierwsze, różnice średnich między wieloma krajami są niewielkie. Co jednak ważniejsze, różnice w umiejętnościach w większym stopniu różnią się wewnątrz krajów niż między krajami. W Polsce aż 39% badanych miało poważne trudności z zadaniami testowymi mierzącymi rozumienie tekstu, a 38% z zadaniami mierzącymi praktyczne wykorzystanie umiejętności matematycznych. Co dziesiąta osoba nie radziła sobie nawet z najprostszymi zadaniami. To znacznie gorsze wyniki niż w Niemczech (16–17%), Czechach (18–19%) czy Estonii (8–9%). Niewielki był także odsetek osób z wysokimi umiejętnościami: tylko 2,5% badanych osiągnęło poziom 4. w rozumieniu tekstu, a 3,7% w matematyce. Jeszcze mniej osób osiągnęło poziom 5. Z krajów sąsiadujących z Polską znacznie lepiej wypadły pod tym względem Niemcy (18–19% osób na poziomie 4. lub 5.), Czechy (14–16%) i Estonia (22–29%).

Ważnym aspektem zróżnicowania wyników są też podstawowe charakterystyki społeczno-demograficzne. W Polsce różnice w wynikach między kobietami a mężczyznami są nieistotne statystycznie we wszystkich trzech badanych dziedzinach. W innych krajach europejskich

zazwyczaj kobiety osiągają lepsze wyniki w rozumieniu tekstu, a mężczyźni w matematyce, choć ogólnie różnice są niewielkie.

Porównując wyniki osób ze względu na wiek, okazuje się, że w Polsce najlepiej wypadli najmłodsi respondenci. Wyróżnia to Polskę na tle innych krajów, gdzie lepsze wyniki uzyskują zazwyczaj osoby w wieku 25–34 lat. W Polsce, podobnie jak w wielu innych krajach, zaobserwowano systematyczny spadek umiejętności wraz z wiekiem, szczególnie w grupie wiekowej 55–65 lat. Ten trend jest widoczny również w innych krajach, co jest związane z procesami starzenia się, ale także ze zmianami w systemach edukacji i różnicy w poziomie wykształcenia osób młodszych i starszych.

W Polsce, podobnie jak w innych krajach, osoby z wyższym wykształceniem osiągały znacznie lepsze wyniki niż osoby z pozostałych kategorii. W Polsce istotne różnice w wynikach zauważono również między osobami z wykształceniem zasadniczym zawodowym a średnim.

2. Zmiany wyników kluczowych umiejętności w ciągu ostatniej dekady

Wstęp

Badanie PIAAC zostało przeprowadzone w dwóch cyklach, oddzielonych od siebie około dziesięcioletnim odstępem. Dwadzieścia siedem krajów (lub ich części, które wzięły udział w badaniu) uczestniczyło w obu cyklach badania. Podczas gdy w drugim cyklu przeprowadzono jak dotąd tylko jedną rundę (w latach 2022/23), pierwszy cykl został zrealizowany w trzech rundach. Runda 1. odbyła się w latach 2011/12, runda 2. w latach 2014/15, a runda 3. w roku 2017. Ponieważ w poszczególnych rundach brały udział różne kraje, czas, jaki upłynął między zbiorami danych, nie jest taki sam dla wszystkich krajów. Większość z nich (21 z 27, w tym Polska) uczestniczyła w rundzie 1. pierwszego cyklu, 11 lat przed drugim cyklem. W tym rozdziale, dla uproszczenia, często mowa o zmianach „w ciągu ostatniej dekady”, choć odstęp między badaniami w różnych krajach mogą się różnić. Pięć krajów uczestniczyło w rundzie 2. pierwszego cyklu, osiem lat przed drugim cyklem. Węgry są jedynym krajem, który uczestniczył w rundzie 3. pierwszego cyklu, zaledwie sześć lat przed drugim cyklem. Schemat badania, wraz z datami realizacji poszczególnych rund i cykli oraz zestawieniem krajów biorących udział w każdej rundzie, przedstawiono na rysunku 2.1.

Rysunek 2.1 Harmonogram cykli badania PIAAC i różnice czasowe między rundami w poszczególnych krajach



Źródło: Opracowanie własne na podstawie danych PIAAC.

Jeżeli kraj posiada przynajmniej dwa punkty pomiarowe, pozwala to na analizę zmian w poziomie badanych umiejętności w czasie. Należy zaznaczyć, że badania te nie mają charakteru panelowego, co oznacza, że nie śledzono tych samych jednostek w kolejnych

punktach czasowych. Badania zostały przeprowadzone w dwóch odrębnych momentach na reprezentatywnych próbach populacyjnych, co umożliwia analizę porównawczą między grupami zebranymi w każdym z tych punktów czasowych. Należy jednak pamiętać, że do takich porównań trzeba podchodzić ostrożnie, biorąc pod uwagę ograniczenia narzucane zarówno przez zastosowane narzędzia badawcze, jak i przez zmiany warunków, w jakich przeprowadzono testy. I tak w obszarze rozwiązywania problemów wyniki z drugiego cyklu zostały uznane za nieporównywalne ze względu na różnice w mierzonych konstrukcjach.

Wyniki testów rozumienia tekstu (*literacy*) i rozumowania matematycznego (*numeracy*) zostały uznane za porównywalne i to one stanowią przedmiot porównań w niniejszym rozdziale. Zanim jednak przejdziemy do przedstawienia tych porównań, omówimy kilka istotnych czynników, na które należy zwrócić uwagę przy interpretacji uzyskanych wyników.

Instrumenty i metodologia pomiaru umiejętności rozumienia tekstu i rozumowania matematycznego w badaniu z 2023 roku są zbliżone do tych stosowanych w pierwszym cyklu, jednak zostały zaktualizowane, aby lepiej odzwierciedlać współczesne sposoby korzystania z informacji i ich przetwarzania. Wpływ potencjalnych różnic wynikających z tych aktualizacji został skontrolowany za pomocą metod statystycznych, które nie wykazały istotnych efektów z perspektywy konstrukcji skal pomiarowych. Pomimo braku znaczącego wpływu na statystyczną porównywalność wyników, warto jednak pamiętać o wprowadzonych zmianach, gdyż mogą one być istotne przy interpretacji wyników. Dotyczą one różnic w instrumentach pomiarowych, zmian metodologicznych w badaniu osób o niskich umiejętnościach, przejścia z komputera lub testu papierowego na tablet oraz zmiany nastawienia respondentów do udziału w badaniach.

Instrumenty pomiarowe różnią się między cyklami. W drugim cyklu badania umiejętności czytania ze zrozumieniem obejmują szerszy zakres tekstów cyfrowych, w tym teksty z wielu źródeł, z naciskiem na ocenę dokładności i trafności informacji. Z kolei ocena umiejętności matematycznych została dostosowana, aby uwzględnić konieczność interpretacji informacji matematycznych przedstawianych w formie dynamicznej (np. interaktywne strony internetowe) oraz strukturalnej (np. infografiki), a także krytycznej oceny tych informacji. Wprowadzone zmiany oznaczają, że pomimo dużej liczby zadań wspólnych dla obu cykli, co zapewnia silne powiązanie psychometryczne, struktury testu w obu cyklach nie są identyczne. W efekcie, chociaż wyniki można porównać, należy uwzględnić dodatkowy „błąd łączenia” (*linking error*), który odzwierciedla niepewność związaną ze zmianami w narzędziach pomiarowych.

Zmiany metodologiczne dotyczące pomiaru wyników osób o niskich umiejętnościach również mogą wpłynąć na wyniki. Pierwszy cykl badania obejmował test z podstaw czytania (*reading components*) zaprojektowany w celu uzyskania informacji o osobach dorosłych z bardzo niskim poziomem umiejętności czytania. Pomiarowi temu poddano jedynie dorosłych, którzy nie zaliczyli łatwego testu wstępnego, tak zwanego badania plasującego.

Oceniane były podstawowe umiejętności niezbędne do rozumienia tekstu pisanego (rozpoznawanie słów, rozumienie zdań i płynność czytania). Odpowiedzi na te zadania nie były brane pod uwagę przy oszacowaniu poziomu umiejętności czytania respondentów; poziom umiejętności był oparty na wynikach badania plasującego oraz na odpowiedziach udzielonych w kwestionariuszu. W drugim cyklu osoby dorosłe, które nie zaliczyły badania plasującego, były oceniane wyłącznie na podstawie zadań z podstaw czytania i matematyki (tzw. badania umiejętności podstawowych). Aby poprawić precyzję oszacowań na najniższym poziomie rozkładu umiejętności, w drugim cyklu uwzględniono wyniki z tych pomiarów przy oszacowaniu poziomu rozumienia tekstu i rozumowania matematycznego respondentów. Ta zmiana metodologiczna ma wpływ na oszacowanie umiejętności osób dorosłych, które po niezaliczeniu badania plasującego przystąpiły jedynie do zadań dotyczących podstaw czytania lub podstaw umiejętności matematycznych. Należy zachować ostrożność przy porównywaniu poziomu umiejętności w grupach, w których jest wielu takich respondentów, ponieważ różnice w obserwowanych wynikach umiejętności między pierwszym a drugim cyklem mogą być spowodowane zmianami w metodologii badania.

Kolejną różnicą, która mogła wpłynąć na wyniki, była zmiana narzędzia testowego z komputera i testu papierowego na tablet. W pierwszym cyklu badania były realizowane głównie na komputerach, jednak dla osób mających problemy z korzystaniem z komputera dostępna była również wersja papierowa. W Polsce wersję papierową ukończyło około 25% respondentów (Rynko i in. 2013). W drugim cyklu wprowadzono tablety dla wszystkich respondentów, aby w pełni wykorzystać zalety testowania komputerowego, takie jak automatyczne ocenianie oraz bardziej zaawansowane i interaktywne zadania. Wszyscy respondenci mieli dostęp do intuicyjnego interfejsu z instrukcjami obsługi tabletu (m.in. funkcji stukania, przeciągania i zaznaczania), co miało na celu zapewnienie jednolitych warunków testowania. Mimo że wyniki pilotażu potwierdziły porównywalność tej nowej formy zbierania danych, zmiana ta mogła wpłynąć na wyniki w grupach, które wcześniej korzystały z wersji papierowej.

Porównując wyniki między obiema edycjami badania, trzeba też wziąć pod uwagę, że zmieniło się nastawienie respondentów do udziału w badaniach, co może wpływać na uzyskane wyniki. Badania pokazują, że coraz mniej osób jest skłonnych brać udział w badaniach społecznych (Jabkowski, 2021), co potwierdzają też niższe niż 10 lat temu odsetki osób, które zgodziły się wziąć udział w badaniu. Badania pokazują też nasilające się w ostatnich latach problemy z motywacją respondentów do rzetelnego udzielania odpowiedzi (Aruguete et al., 2019; Steedle et al., 2019). Ten trend można przypisać narastającemu zjawisku zmęczenia ankietami, spadkowi zaufania do instytucji badawczych oraz większej niechęci do dzielenia się informacjami. Jest to szczególnie istotny problem w badaniach, w których wyniki nie mają wpływu na uczestników, jak PIAAC. W tego typu badaniach respondenci nie ponoszą konsekwencji swoich wyników, co może obniżyć ich zaangażowanie i sprawiać, że trudno jest odróżnić rzeczywiste umiejętności od poziomu motywacji do wykonania zadań (Pokropek, 2016). Badanie jest dobrowolne, więc choć badacze mogą zachęcać respondentów do aktywnego udziału, nie są w stanie zapewnić

pełnego zaangażowania w udzielanie odpowiedzi na pytania kwestionariusza ani w wykonanie zadań mierzących umiejętności.

Każdy z tych czynników – aktualizacja narzędzi pomiarowych, zmiany metodologiczne w ocenie niskich wyników, zmiana narzędzia testowego oraz problemy związane z motywacją – może wpływać na wyniki i powinien być brany pod uwagę przy interpretacji porównań między cyklami. Wyniki przedstawione w tym rozdziale pokazują, jak respondenci radzili sobie z zadaniami kognitywnymi i stanowią oszacowanie poziomu umiejętności dla poszczególnych krajów i grup. Jest to obecnie najbardziej rzetelne porównanie, jakim dysponujemy. Wyniki dostarczają wartościowych informacji na temat poziomu umiejętności w różnych krajach i grupach społecznych. Niemniej jednak interpretacja wyników powinna być dokonywana z ostrożnością. Mimo że porównania te są przydatne i pozwalają uchwycić ogólne trendy, wnioski muszą uwzględniać te ograniczenia i opierać się na krytycznej analizie, aby unikać nadinterpretacji.

Zmiany w zakresie umiejętności w populacji dorosłych

Zmiany średnich umiejętności

Rysunek 2.2 przedstawia zmianę średniej umiejętności rozumienia tekstu i rozumowania matematycznego między pierwszym a drugim cyklem badania umiejętności osób dorosłych. Biegłość w czytaniu i pisaniu znacznie wzrosła w Finlandii i Danii (odpowiednio o 15 i 9 punktów w ciągu 11 lat) i pozostała stabilna w 14 krajach. Jedenaście krajów doświadczyło znacznego spadku. W czterech z tych krajów spadek był szczególnie duży: 31 punktów w Polsce, 28 punktów na Litwie, 23 punkty w Korei i 21 punktów w Nowej Zelandii.

Tendencje w zakresie umiejętności rozumowania matematycznego były bardziej korzystne. Osiem krajów odnotowało znaczny wzrost tej umiejętności, przy czym największy wzrost zaobserwowano w Finlandii (17 punktów), Singapurze (17 punktów) i Estonii (9 punktów). Poziom wyników pozostał stabilny w 12 krajach i znacznie spadł w siedmiu. Największe spadki umiejętności matematycznych zaobserwowano na Litwie (22 punkty) i w Polsce (21 punktów).

Ogólnie rzecz biorąc, zmiany w umiejętnościach rozumienia tekstu i rozumowania matematycznego miały podobny charakter i przebiegały w tym samym kierunku, wzorce wzrostu lub spadku w jednej dziedzinie były zbliżone do tych obserwowanych w drugiej. Najbardziej godnym uwagi wyjątkiem od tego wzorca jest Singapur, gdzie znacznemu wzrostowi umiejętności rozumowania matematycznego towarzyszył stabilny poziom umiejętności rozumienia tekstu. Podobne tendencje obserwuje się w Kanadzie, Anglii (Zjednoczone Królestwo), Estonii, Holandii i Norwegii, choć zmiany w umiejętnościach matematycznych są stosunkowo mniejsze. Z kolei w Austrii, Czechach, we Francji i w Izraelu wyniki umiejętności testu rozumowania matematycznego pozostały stabilne, podczas gdy wynik testu umiejętności

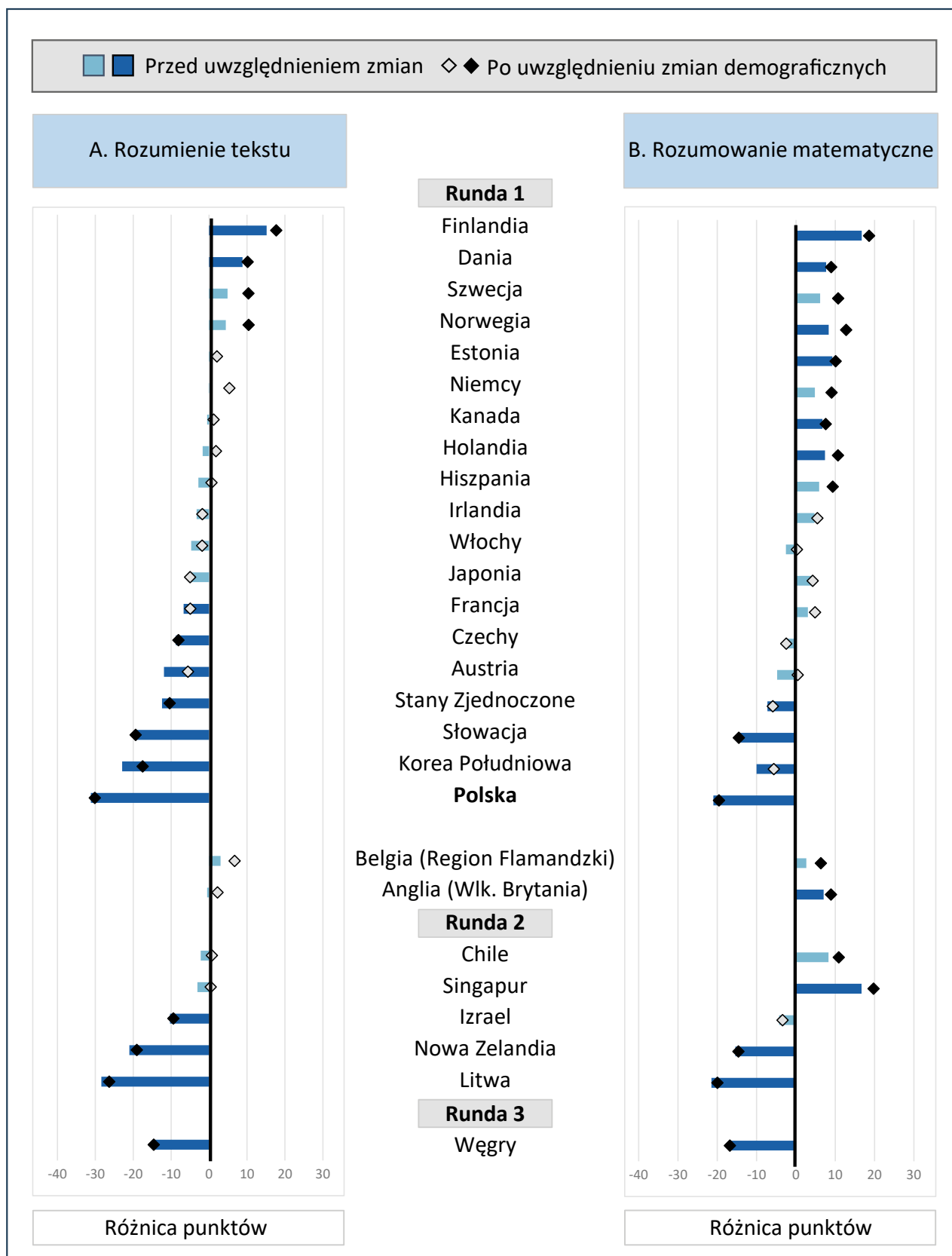
rozumienia tekstu spadł.

W większości krajów biorących udział w badaniu PIAAC struktura demograficzna dorosłej populacji zmieniła się w latach między cyklami. Większe przepływy migracyjne i starzenie się społeczeństw mogły przyczynić się do spadku umiejętności w poszczególnych krajach, ponieważ imigranci i osoby starsze często wykazują niższy poziom umiejętności. Aby wyodrębnić wpływ zmian demograficznych, dane z badania umiejętności dorosłych z 2023 r. zostały ponownie zważone, tak aby dopasować je do rozkładu demograficznego populacji zaobserwowanego w pierwszym cyklu badania (w zakresie wieku, płci i pochodzenia imigracyjnego). Procedura ważenia, analogiczna do poststratyfikacji, dostosowuje charakterystykę próby z 2023 roku do składu populacji w pierwszym cyklu. W pierwszym kroku próba z każdego cyklu oceny została podzielona na komórki, zdefiniowane na podstawie statusu imigracyjnego respondentów, płci oraz grupy wiekowej. W drugim kroku każda komórka została przeważona, aby suma wag w każdej z nich była stała w poszczególnych cyklach i równa sumie wag w próbie z pierwszego cyklu. Na tej podstawie obliczono średnie i rozkłady wyników testów umiejętności na przeważonych próbach, co pozwala oszacować hipotetyczny (kontrfaktyczny) rozkład umiejętności, jaki byłby obserwowany, gdyby skład próby w 2023 r. był identyczny z tym z pierwszego cyklu pod względem zmiennych uwzględnionych w tej procedurze ważenia. Warto podkreślić, że waga opiera się wyłącznie na danych z PIAAC, bez wykorzystania zewnętrznych danych, co wiąże się z konsekwencjami dotyczącymi pokrycia populacji, zastosowanych wag oraz błędu pomiarowego. Ostatecznym celem tej procedury było uzyskanie kontrfaktycznego rozkładu umiejętności, który uwzględnia zmiany demograficzne w badanej populacji.

Na rysunku 2.2 pokazano, że uwzględnianie zmian demograficznych ma wpływ na szacowany rozwój umiejętności. Wpływ ten jest jednak na ogół niewielki i nie dotyczy wszystkich krajów. Z analizy wynika, że gdyby profil demograficzny dorosłej populacji pozostał taki sam jak w pierwszym cyklu, średnie wyniki w zakresie umiejętności rozumienia tekstu wzrosłyby w Norwegii i Szwecji i pozostałyby na stałym poziomie, a nie spadły w Austrii i Francji. Innymi słowy, zmiany w strukturze demograficznej populacji są jednym z czynników, które odpowiadają za brak poprawy lub spadek umiejętności rozumienia tekstu w tych krajach. Podobnie, gdyby nie zmiany demograficzne, średnia biegłość w rozumowaniu matematycznym uległaby poprawie w Chile, Regionie Flamandzkim (Belgia), Niemczech, Hiszpanii i Szwecji, natomiast nie zmniejszyłaby się znacząco w Korei i Stanach Zjednoczonych.

Rysunek 2.2 Zmiana średniego poziomu umiejętności w zakresie rozumienia tekstu i rozumowania matematycznego między cyklami, skorygowana o zmiany demograficzne.

Różnica w średnich wynikach umiejętności między cyklami 2 i 1, po przeważeniu danych z cyklu 2. tak, aby odpowiadał rozkładowi wieku, pochodzenia imigracyjnego i płci w danych z cyklu 1.



Uwaga: Nieskorygowane różnice (przed uwzględnieniem zmian demograficznych) to różnice między średnimi wynikami w każdym cyklu. Skorygowane różnice (po uwzględnieniu zmian demograficznych) są oparte na metodzie analogicznej do poststratyfikacji, która przeważa próbki w cyklu 2., tak aby ich cechy demograficzne odpowiadały tym z cyklu 1. Uwzględnione cechy demograficzne to: wiek (w przedziałach dziesięcioletnich), płeć i pochodzenie imigracyjne. Skorygowane różnice przedstawiają hipotetyczny scenariusz, pokazujący, jak zmieniłyby się wyniki testów w populacji o profilu demograficznym odpowiadającym populacji z cyklu 1. w czasie.

Ciemniejsze kolory oznaczają różnice, które są statystycznie istotne na poziomie 5%.

Kraje uporządkowano w kolejności malejącej według nieskorygowanej zmiany średnich wyników w zakresie rozumienia tekstu.

Przyczyny zmian

Choć wiele krajów odnotowało spadki średnich wyników w badaniu PIAAC, Polska, wraz z Litwą, Koreą, Nową Zelandią, Słowacją i Stanami Zjednoczonymi, należy do tych, w których spadki te były największe. Sugeruje to, że polskim respondentom testy umiejętności poszły znacznie gorzej niż dekadę temu. Należy jednak zastanowić się nad przyczynami tych zmian. Jak wspomniano we wstępie do tego rozdziału, ważne jest, aby przy interpretacji wyników uwzględnić zmiany, które pojawiły się w ostatniej dekadzie. Analizy polskiego zespołu wskazują, że za znaczną część wyników może odpowiadać szczególnie duży spadek motywacji respondentów. Wątki te będą dokładniej analizowane w kolejnych publikacjach, ale warto już teraz zwrócić uwagę na pewne prawidłowości. Wyniki te silnie sugerują, że czynniki motywacyjne odgrywają kluczową rolę w porównaniach między cyklami.

Tabela 2.1 przedstawia rozkład punktacji oraz kategorii odpowiedzi w zadaniach łączących dla lat 2012 i 2023, z uwzględnieniem proporcji odpowiedzi poprawnych, niepoprawnych oraz braków odpowiedzi w dwóch głównych obszarach – rozumieniu tekstu i rozumowaniu matematycznym.

Tabela 2.1. Rozkład punktacji oraz kategorii odpowiedzi w zadaniach łączących (linkujących), czyli tych, których treść pozostała niezmienną między latami. Dane z Polski.

Punkty za zadanie	Odpowiedź	Rozumienie tekstu		Rozumowanie matematyczne	
		2012	2023	2012	2023
1	Odpowiedź poprawna	62,0%	54,4%	78,8%	57,7%
0	Odpowiedź niepoprawna	32,5%	31,4%	18,7%	32,7%
0	Brak odpowiedzi	5,6%	14,2%	2,5%	9,6%

Źródło: opracowanie własne na podstawie danych PIAAC.

W obszarze rozumienia tekstu (*literacy*) zauważamy, że udział odpowiedzi niepoprawnych pozostaje niemal niezmienny między cyklami, wynosząc odpowiednio 32,5% w 2012 roku i 31,4% w 2023 roku. Jednakże najbardziej wyrazistą zmianą jest trzykrotny wzrost odsetka braków odpowiedzi, który wzrósł z 5,6% w 2012 roku do 14,2% w 2023 roku. W kontekście oszacowań wyników testowych, gdzie brak odpowiedzi jest traktowany jako odpowiedź niepoprawna,

wzrost ten może istotnie wpływać na końcowe oszacowania poziomu umiejętności rozumienia tekstu, zwłaszcza w przypadku respondentów, którzy mogli zrezygnować z udzielania odpowiedzi z powodu braku motywacji lub trudności ze zrozumieniem treści zadania.

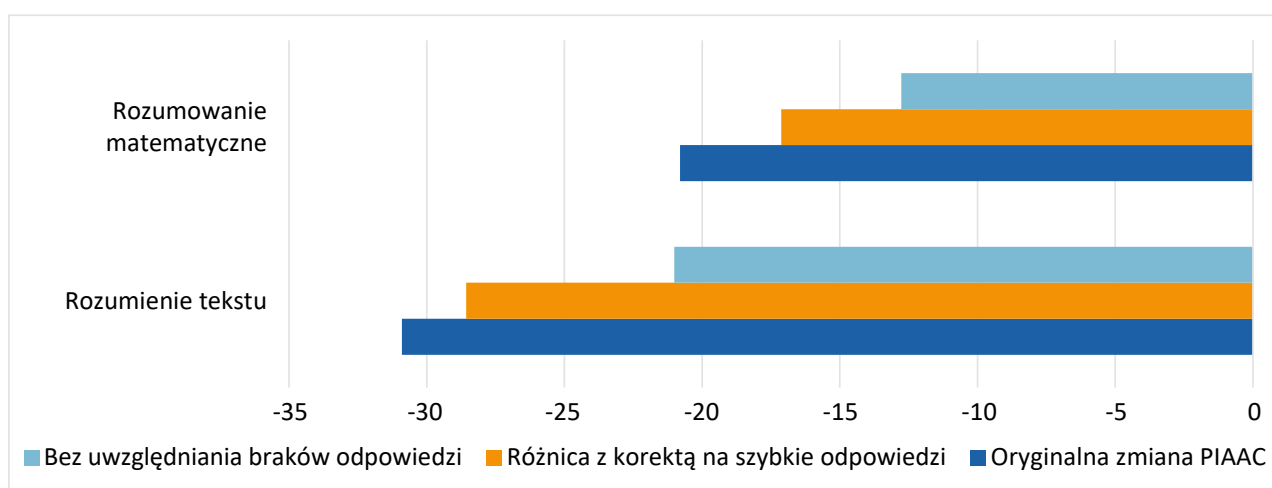
W obszarze rozumowania matematycznego (*numeracy*) odsetek braków odpowiedzi wzrósł aż czterokrotnie, z 2,5% w 2012 roku do 9,6% w 2023 roku, co wskazuje na znacznie częstsze przypadki, że respondenci zrezygnowali z udzielania odpowiedzi. Może to sugerować obniżenie poziomu zaangażowania i motywacji respondentów. Choć spadek odsetka odpowiedzi poprawnych (z 78,8% do 57,7%) oraz wzrost udziału odpowiedzi niepoprawnych (z 18,7% do 32,7%) wskazują na trudności respondentów w udzieleniu poprawnych odpowiedzi.

W kontekście problemu obniżonej motywacji warto zwrócić uwagę na klasyfikację szybkich odpowiedzi, które w literaturze są uznawane za potencjalne wskaźniki nieważności lub braku motywacji, określane mianem *rapid guessing*. Badania pokazują, że odpowiedzi udzielone w czasie poniżej 3 sekund lub poniżej 5 sekund na pytania w testach umiejętności, z dużym prawdopodobieństwem mogą być klasyfikowane jako *rapid guessing* (Goldhammer i inni, 2016). Odpowiedzi tego typu sugerują, że respondenci udzielają odpowiedzi bez pełnego zapoznania się z treścią zadania, co może wskazywać na brak zaangażowania lub niski poziom motywacji (oczywiście nie można w pewnych wypadkach wykluczyć niezwykle szybkiego czytania i odpowiadania, szczególnie na łatwe zadanie, ale jest to znacznie mniej prawdopodobne). W najnowszym cyklu zaobserwowaliśmy wzrost odsetka takich szybkich odpowiedzi (analizujemy wyłącznie zadania łączące, które w pierwszym cyklu były realizowane na komputerze). Przy 5-sekundowym progu w 2012 roku odsetek szybkich odpowiedzi dla umiejętności rozumienia tekstu wynosił 0,4%, natomiast w 2023 roku osiągnął poziom 6,2%. W rozumowaniu matematycznym odsetek odpowiedzi udzielanych w czasie poniżej 5 sekund wzrósł z 0,3% do 2,8%. Dodatkowo zauważalny jest spadek średniego czasu odpowiedzi na zadanie. W 2012 roku przeciętny czas spędzony na pojedynczym zadaniu wynosił 56,2 sekundy dla umiejętności rozumienia tekstu i 50,8 sekundy dla rozumowania matematycznego. W 2023 roku czasy te zmniejszyły się do odpowiednio 43,4 sekundy dla umiejętności rozumienia tekstu oraz 44,3 sekundy dla rozumowania matematycznego.

Trudno jest całkowicie wyizolować wpływ motywacji na wyniki, jednak istnieją metody, które mogą przynajmniej wskazać, czy motywacja ma znaczenie w interpretacji zmian wyników. Poniżej przedstawiamy wyniki dla Polski, które pokazują, jak zmieniłyby się spadek wyników, gdyby różnie traktować motywację respondentów. Na rysunku 2.3 przedstawiamy 3 wyniki (zarówno dla rozumienia tekstu jak i rozumowania matematycznego). Pierwszy to nieskorygowana zmiana w wynikach (tak jak na rysunku 2.2 w porównaniach międzykrajowych). W przypadku rozumienia tekstu spadek wynosi -30.9, a dla rozumowania matematycznego -20.8. To są wartości odzwierciedlające surową zmianę bez kontroli wpływu motywacji. W drugim modelu kontrolujemy szybkie odpowiadanie, wykorzystując modelowanie hybrydowe (Pokropek, 2016), które uwzględnia szybkie odpowiedzi potencjalnie wskazujące na zgadywanie. Przy takim

modelowaniu danych spadki są mniejsze (-28.57 dla rozumienia tekstu i -17.1 dla rozumowania matematycznego). Sugeruje to, że przynajmniej część spadku wyników można przypisać niskiej motywacji, która może prowadzić do szybkiego, nieprzemyślanego odpowiadania. W trzecim modelu traktujemy braki odpowiedzi jako brakujące dane, skupiając się wyłącznie na odpowiedziach, w których respondent podjął wysiłek zaznaczenia odpowiedzi (poprawnej lub nie). W takim podejściu obserwujemy dalsze zmniejszenie spadków wyników. Dla rozumienia tekstu spadek zmniejszył się do -21.0, a dla rozumowania matematycznego do -12.8.

Rysunek 2.3 Wpływ uwzględnienia motywacji na skalę spadków wyników w Polsce między cyklami badania PIAAC



Źródło: opracowanie własne na podstawie danych PIAAC.

Jak wynika z przedstawionych wyników, motywacja może mieć istotny wpływ na skalę spadku wyników. Nie wyjaśnia jednak całkowicie obniżenia rezultatów – nawet po zastosowaniu korekt uwzględniających motywację zmiany wyników pozostają istotnie negatywne, a spadek, zarówno z kontrolą, jak i bez kontroli motywacji, jest wyraźny. Te wyniki mają na celu przybliżenie potencjalnych przyczyn tak drastycznego spadku wyników Polski w punktach testowych PIAAC i sugerują, że tego spadku nie można interpretować jedynie z perspektywy rzeczywistych umiejętności respondentów, lecz również ich chęci do zaprezentowania tych umiejętności.

Zmiany wyników PIAAC dla grup społeczno-demograficznych

W tej sekcji skupiono się na tym, jak zmieniły się wyniki PIAAC między cyklami badania dla grup społeczno-demograficznych zdefiniowanych na podstawie poziomu wykształcenia, płci i wieku. W prezentacji wyników nacisk położono na umiejętność rozumienia tekstu. Wyniki w zakresie umiejętności matematycznych wykazują podobny wzorec i zostały jedynie pokrótce opisane, bardziej szczegółowe informacje dostępne są w dokumentacji raportu międzynarodowego.

Wykształcenie

W analizach w tym rozdziale poziom wykształcenia został ustalony na podstawie Międzynarodowej Standardowej Klasyfikacji Kształcenia (ISCED) z 2011 roku, która dzieli poziomy edukacji na trzy główne kategorie. Pierwsza kategoria to poziom **poniżej średniego** (obejmujący ISCED 1, 2 i krótkie programy na poziomie ISCED 3), który odpowiada w Polsce wykształceniu poniżej szkoły średniej i zasadniczej zawodowej. Druga kategoria to **wykształcenie średnie** (ISCED 3 i 4); w Polsce w tej kategorii mieści się również wykształcenie zasadnicze zawodowe. Trzecia kategoria to **wyższe** (ISCED 5, 6, 7 i 8), które obejmuje różne poziomy wykształcenia wyższego, od licencjackiego po doktorancki.

W większości krajów biorących udział w obu cyklach badania spadki wyników z rozumienia tekstu były największe wśród osób dorosłych o niskim wykształceniu (rys. 2.4). Średni poziom wyników dorosłych z wykształceniem poniżej szkoły średniej spadł między tymi dwoma cyklami we wszystkich krajach z wyjątkiem Kanady, Danii, Finlandii, Izraela, Hiszpanii i Szwecji. Spadki te wahały się od 13 punktów w Estonii do 46 punktów na Litwie. Wyniki dorosłych z wykształceniem średnim spadły w 19 krajach, nie zmieniły się w 7 krajach i wzrosły jedynie w Finlandii. Wśród dorosłych z wykształceniem wyższym, średni poziom wyników spadł w 13 krajach, a wzrósł tylko w Finlandii.

W Polsce wyniki wskazują na spadek umiejętności rozumienia tekstu we wszystkich grupach edukacyjnych w okresie od 2012 do 2023 roku. W grupie osób z wykształceniem poniżej średniego średni wynik zmniejszył się z 227,1 punktu w 2012 roku do 202,2 punktu w 2023 roku, co stanowi spadek o 24,9 punktu. Osoby z wykształceniem średnim odnotowały spadek z 254,5 punktu do 228,1 punktu, co oznacza różnicę wynoszącą 26,4 punktu. Natomiast w grupie osób z wykształceniem wyższym średni wynik zmniejszył się z 296,9 punktu w 2012 roku do 255 punktów w 2023 roku, co stanowi spadek o 41,9 punktu. Tak wysokie spadki, szczególnie wśród osób z wyższym wykształceniem, wydają się odzwierciedlać nie tylko zmiany w umiejętnościach, ale także możliwe obniżenie motywacji do rozwiązywania testów, problem wcześniej opisywany w tym rozdziale.

Zmiany w wynikach testu rozumowania matematycznego przebiegały według podobnego wzorca (szczegółowe wyniki dostępne są w raporcie międzynarodowym). W trzynastu

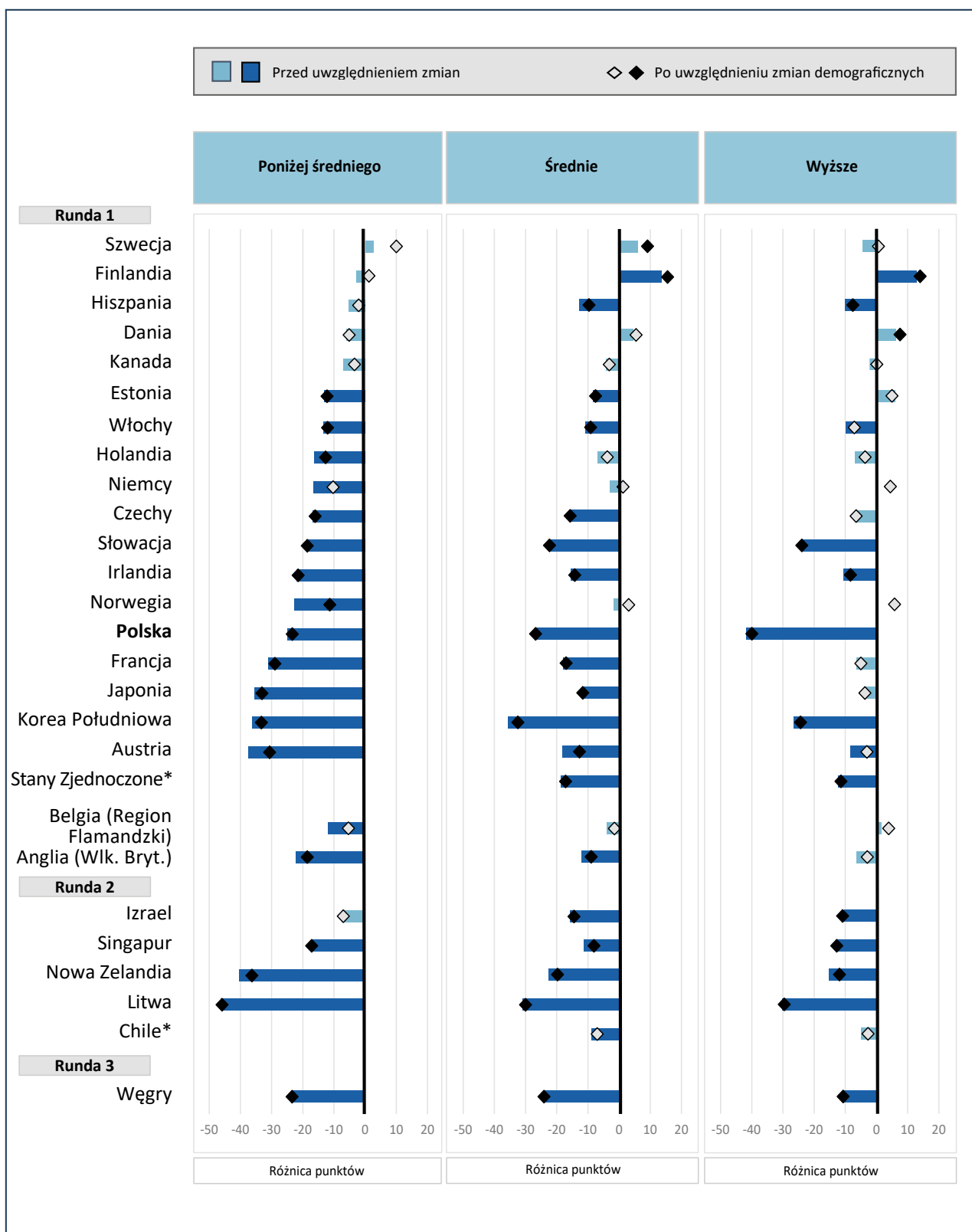
krajach odnotowano znaczący spadek średniego wyniku w tej grupie, podczas gdy jedynie Singapur wykazał wzrost. W przypadku dorosłych z wykształceniem wyższym średni wynik w teście rozumowania matematycznego wzrósł jedynie w Kanadzie, Estonii i Finlandii, natomiast spadł na Węgrzech, w Izraelu, we Włoszech, w Korei, na Litwie, w Nowej Zelandii, Polsce i na Słowacji. Szczególnie duże spadki odnotowano wśród osób z wykształceniem wyższym na Litwie, w Polsce i na Słowacji. W Polsce niekorygowane spadki wyników z rozumowania matematycznego dla trzech kategorii wykształcenia wyniosły odpowiednio: dla osób z wykształceniem poniżej średniego – 20,6 punktu, dla osób z wykształceniem średnim – 18,7 punktu, a dla osób z wykształceniem wyższym – 30 punktów.

Powszechny spadek wyników wśród osób z niskim wykształceniem pogłębił różnice edukacyjne w większości krajów uczestniczących w badaniu. W Austrii, Francji, Japonii, Nowej Zelandii i Norwegii wyraźnie wzrosły różnice w umiejętnościach między osobami z wykształceniem średnim i wyższym. W Singapurze, gdzie nierówności w rozumieniu tekstu pod względem wykształcenia są jedne z największych, różnica zmieniła się mniej, ponieważ umiejętności spadły zarówno w grupie z niskim, jak i wysokim wykształceniem. W Polsce natomiast różnica ta zmniejszyła się z powodu spadku wyników w rozumieniu tekstu wśród osób z wykształceniem wyższym.

Podobny obraz można zaobserwować dla rozumowania matematycznego. Różnica między dorosłymi z niskim i wysokim wykształceniem pogłębiła się w 11 krajach, przy czym duży wzrost (ponad 20 punktów) odnotowano w Austrii, Japonii i Nowej Zelandii. W Singapurze różnica w umiejętnościach matematycznych zmniejszyła się ze względu na poprawę wśród osób z niższym wykształceniem, podczas gdy w Izraelu i na Słowacji zmniejszyła się z powodu spadku wyników dorosłych z wyższym wykształceniem. W Polsce różnica ta nieznacznie się zmniejszyła, jednak nie jest to zmiana istotna statystycznie.

Rysunek 2.4. Zmiana w wynikach testu umiejętności rozumienia tekstu między cyklami, według poziomu wykształcenia.

Skorygowana i nieskorygowana różnica w średnich wynikach umiejętności czytania między cyklami, według poziomu wykształcenia (cykl 2. minus cykl 1.); dorośli w wieku 25–65 lat.



Uwaga: Nie uwzględniono osób dorosłych, z którymi w cyklu 2. przeprowadzono jedynie wywiad wstępny z powodu bariery językowej, aby maksymalnie zwiększyć porównywalność między cyklami. Nieskorygowane różnice (przed uwzględnieniem zmian demograficznych) to różnice między średnimi wynikami w każdym cyklu. Skorygowane różnice

(po uwzględnieniu zmian demograficznych) są oparte na metodzie analogicznej do poststratyfikacji, która przeważa próbki w cyklu 2., tak aby ich cechy demograficzne odpowiadały tym z cyklu 1. Uwzględnione cechy demograficzne to: wiek, płeć oraz pochodzenie imigracyjne.

* W Chile i Stanach Zjednoczonych ponad 20 procent podpróby dorosłych z wykształceniem poniżej poziomu średniego ukończyło jedynie test podstaw czytania i umiejętności matematycznych w drugim cyklu. Ponieważ metodologia oceny umiejętności osób, które przystąpiły jedynie do testu podstaw czytania, uległa zmianie między cyklami, zmiana poziomu umiejętności dla tej grupy nie jest raportowana.

Ciemniejsze kolory oznaczają różnice istotne statystycznie na poziomie 5%.

Kraje uporządkowano w kolejności malejącej, zgodnie z niekorygowaną zmianą poziomu umiejętności czytania wśród dorosłych z wykształceniem poniżej poziomu średniego.

Źródło: opracowanie własne na podstawie danych PIAAC.

Płeć

Pierwszy cykl PIAAC wykazał, że w większości krajów OECD różnice między kobietami a mężczyznami w wynikach badania umiejętności były stosunkowo niewielkie (OECD, 2019). W większości krajów nie odnotowano istotnych różnic w wynikach pomiaru umiejętności rozumienia tekstu między kobietami a mężczyznami, natomiast prawie wszystkie kraje odnotowały umiarkowane różnice w wynikach testów rozumowania matematycznego na korzyść mężczyzn.

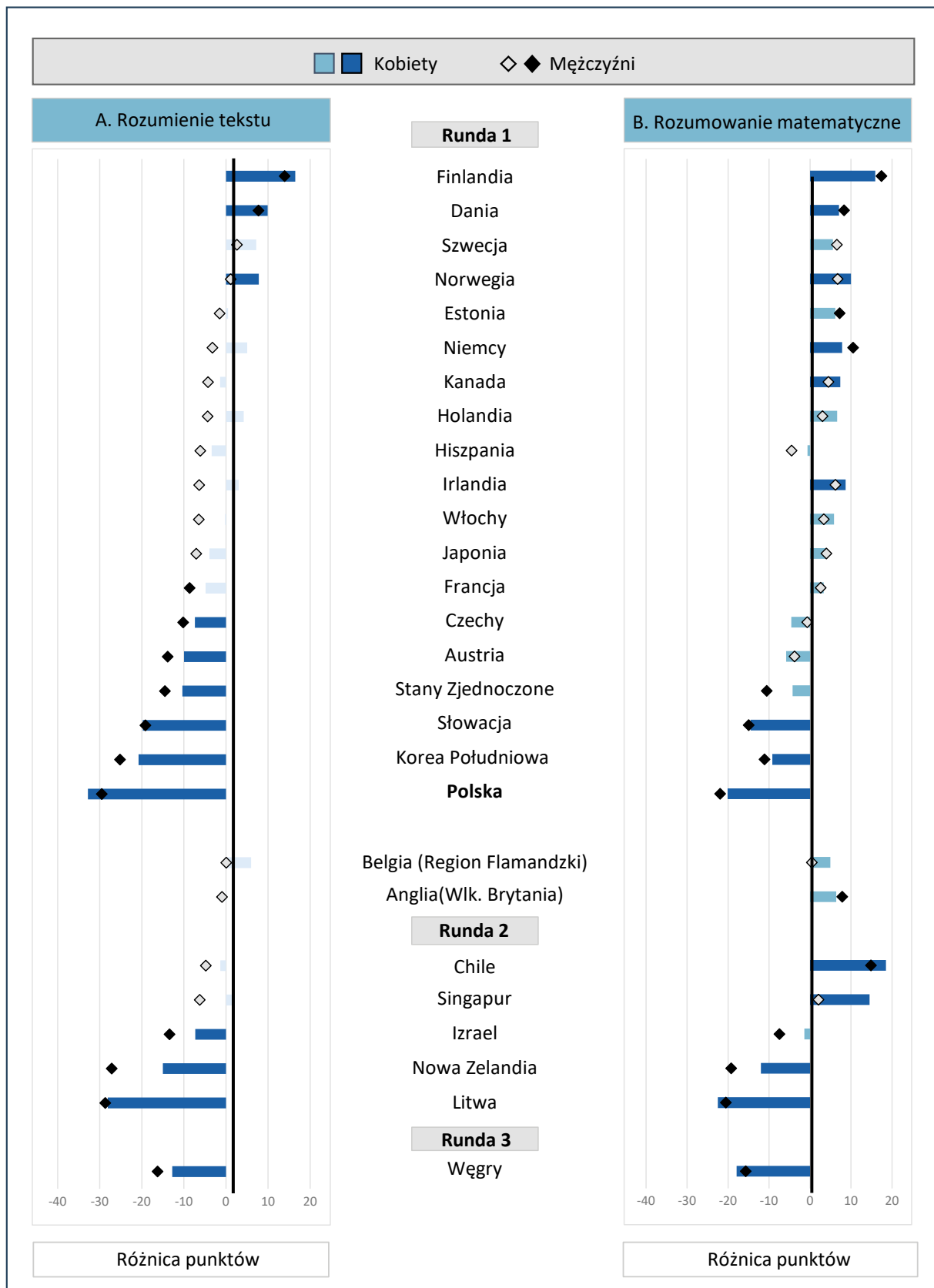
Rysunek 2.5 pokazuje, jak zmieniały się wyniki testów mierzących umiejętności mężczyzn i kobiet między cyklami badania. W jednej trzeciej krajów wyniki testów mierzących umiejętność rozumienia tekstu spadły u mężczyzn, podczas gdy u kobiet pozostały bez zmian lub spadły w mniejszym stopniu. Podobne prawidłowości obserwujemy w wynikach testu mierzącego rozumowanie matematyczne.

Silniejszy spadek wyników testów mierzących umiejętność rozumienia tekstu wśród mężczyzn doprowadził do zmniejszenia różnic między kobietami a mężczyznami w 10 krajach. W wielu krajach różnice między kobietami a mężczyznami uległy odwróceniu, a kobiety osiągały obecnie wyższe wyniki niż mężczyźni (zob. rozdział 1). Jeśli chodzi o umiejętność rozumowania matematycznego, różnica między płciami zmniejszyła się tylko w Chile, Izraelu i Stanach Zjednoczonych. Uwzględnianie różnic (takich jak poziom wykształcenia) w cechach społeczno-demograficznych między mężczyznami i kobietami ma niewielki wpływ na obserwowane zmiany w różnicach między płciami.

W Polsce w pierwszym cyklu różnica między płciami w wynikach badania umiejętności rozumienia tekstu wynosiła -6,4, co oznacza, że kobiety osiągały wyższe wyniki. W cyklu drugim różnica ta wynosiła -3,1, ta zmiana nie jest jednak istotna statystycznie. W zakresie umiejętności matematycznych różnica między płciami w pierwszym cyklu również nie była istotna statystycznie i wynosiła 1,9 na korzyść mężczyzn, a w cyklu drugim była bliska zeru. Polska pozostaje zatem krajem o najmniejszej różnicy między płciami w wynikach testów mierzących umiejętności zarówno matematyczne, jak i rozumienia tekstu spośród badanych krajów OECD.

Rysunek 2.5 Zmiana poziomu umiejętności rozumienia tekstu i rozumowania matematycznego między cyklami, według płci.

Różnica w średnich wynikach umiejętności rozumienia tekstu między cyklami, według płci (cykl 2. minus cykl 1.).



Uwaga: Badanie objęło dorosłych w wieku 16–65 lat. Na wykresie przedstawiono nieskorygowane różnice między średnimi wynikami w każdym cyklu. Ciemniejsze kolory oznaczają różnice istotne statystycznie na poziomie 5%.

Kraje uporządkowano w kolejności malejącej według nieskorygowanej zmiany w poziomie umiejętności rozumienia tekstu wśród mężczyzn.

Źródło: opracowanie własne na podstawie danych PIAAC.

Wiek

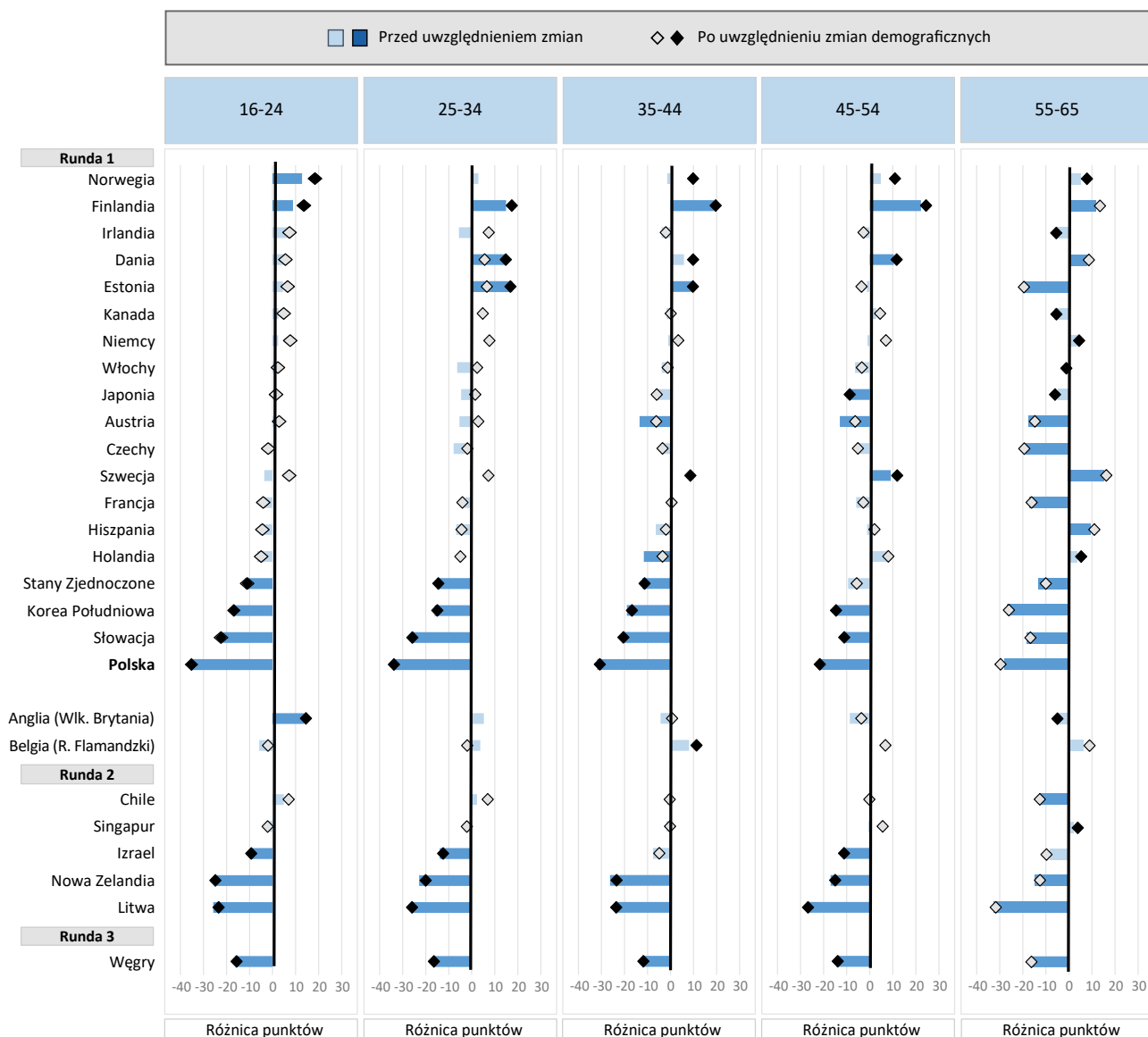
Rysunek 2.5 przedstawia różnice w wynikach testów z obydwu analizowanych dziedzin między osobami dorosłymi w tym samym wieku w czasie pierwszego i drugiego cyklu (tj. efekty kohortowe). Pokazuje on, jak zmieniały się wyniki testów w pięciu różnych grupach wiekowych w krajach uczestniczących w obu cyklach badania. Zmiana umiejętności w najmłodszej grupie wiekowej będzie w większym stopniu odzwierciedlać obecne i niedawne zmiany w edukacji, takie jak zwiększenie uczestnictwa lub poprawa jakości kształcenia w związku z niedawnymi reformami. Z kolei zmiany wyników w starszych grupach wiekowych mogą być efektem wpływów, które wystąpiły na długo przed okresem objętym badaniem – sięgając aż do lat 50. i 60. XX wieku. Ponadto różnice między osobami starszymi będą również odzwierciedlać niedawne zmiany w świecie pracy, które zapewniły lub ograniczyły możliwości utrzymania i rozwijania umiejętności.

Średnia wyników w zakresie rozumienia tekstu wśród młodych dorosłych w wieku od 16 do 24 lat znacznie się poprawiła w Anglii (Zjednoczone Królestwo), Finlandii i Norwegii między tymi dwoma cyklami. Z kolei w ośmiu krajach osoby w wieku 16–24 lat miały w drugim cyklu niższy średni wynik niż ich rówieśnicy w pierwszym cyklu, przy czym szczególnie wyraźne spadki odnotowano na Litwie, w Nowej Zelandii, Polsce i na Słowacji. Wśród osób w wieku 25–34 lata trzy kraje odnotowały poprawę wyników, a osiem krajów znaczny spadek. Tendencja w przypadku osób starszych była bardziej negatywna: w 12 krajach osoby w wieku 55–65 lat miały niższy wynik w dziedzinie rozumienia tekstu w drugim cyklu niż ta sama grupa wiekowa w poprzednim cyklu. Tylko w Finlandii, Danii, Hiszpanii i Szwecji odnotowano poprawę wyników wśród osób starszych. Polska jest jednym z niewielu krajów, w którym największe różnice odnotowujemy w grupie najmłodszej.

W zakresie umiejętności matematycznych poprawę wśród młodych dorosłych w wieku 16–24 lat odnotowano w Chile, Danii, Anglii (Zjednoczone Królestwo), Estonii, Finlandii, Irlandii, Japonii, Niderlandach, Norwegii i Singapurze. W większości tych krajów (wyjątkiem są Irlandia i Japonia), a także w Kanadzie umiejętność rozumowania matematycznego poprawiła się również wśród osób w wieku 25–34 lat. Biegłość w liczeniu wśród dorosłych w wieku 55–65 lat wzrosła w Finlandii, w Regionie Flamandzkim (Belgia), Niemczech, Holandii, Norwegii, Singapurze, Hiszpanii i Szwecji. Z kolei na Węgrzech, Litwie, w Nowej Zelandii, Polsce i na Słowacji odnotowano spadek umiejętności matematycznych we wszystkich grupach wiekowych (z wyjątkiem osób w wieku 55–65 lat w Nowej Zelandii). W Izraelu, Włoszech, Korei i Stanach Zjednoczonych wyniki spadły wśród osób w wieku 25–34 lat. Stany Zjednoczone odnotowały spadki wśród osób w wieku 55–65 lat.

Rysunek 2.6. Zmiana w wynikach testu mierzącego umiejętność rozumienia tekstu między cyklami, według wieku.

Różnica w średnich wynikach badania umiejętności rozumienia tekstu między cyklami 2 i 1 w poszczególnych grupach wiekowych.



Uwaga: Badanie objęło dorosłych w wieku 16–65 lat; nie uwzględniono osób, z którymi w cyklu 2. przeprowadzono jedynie wywiad wstępny z powodu bariery językowej, aby maksymalnie zwiększyć porównywalność między cyklami. Nieskorygowane różnice (przed uwzględnieniem zmian demograficznych) to różnice między średnimi wynikami w każdym cyklu. Skorygowane różnice (po uwzględnieniu zmian demograficznych) są oparte na metodzie analogicznej do poststratyfikacji, która przeważa próbki w cyklu 2., tak aby ich cechy demograficzne odpowiadały tym z cyklu 1. Uwzględnione cechy demograficzne to: płeć i pochodzenie imigracyjne.

Ciemniejsze kolory oznaczają różnice istotne statystycznie na poziomie 5%.

Kraje uporządkowano w kolejności malejącej, zgodnie z nieskorygowaną zmianą poziomu rozumienia tekstu wśród osób w wieku 16–24 lat.

Źródło: opracowanie własne na podstawie danych PIAAC.

Podsumowanie

Porównania międzynarodowe pokazują, że wyniki osiągnięte w teście rozumienia tekstu oraz rozumowania matematycznego dorosłej populacji spadły w wielu krajach między dwoma cyklami. Jedenaście krajów doświadczyło znacznego spadku wyników testów z rozumienia tekstu. W czterech z tych krajów spadek był szczególnie duży: 31 punktów w Polsce, 28 punktów na Litwie,

23 punkty w Korei i 21 punktów w Nowej Zelandii. Poziom wyników w teście rozumowania matematycznego pozostał stabilny w 12 krajach i znacznie spadł w siedmiu. Największe spadki umiejętności matematycznych zaobserwowano na Litwie (22 punkty) i w Polsce (21 punktów).

Dodatkowe informacje, jakie przyniosło badanie PIAAC, to zmiany w różnych grupach respondentów. W większości krajów biorących udział w obu cyklach badania spadki wyników z rozumienia tekstu były największe wśród osób dorosłych o niskim wykształceniu. W Polsce największe spadki odnotowujemy u osób z wyższym wykształceniem. W większości krajów nie odnotowano istotnych różnic w wynikach pomiaru umiejętności rozumienia tekstu między kobietami a mężczyznami, natomiast prawie wszystkie kraje odnotowały umiarkowane różnice w wynikach testów rozumowania matematycznego na korzyść mężczyzn. Polska pozostaje zatem krajem o najmniejszej różnicy między płciami w wynikach testów mierzących umiejętności zarówno matematyczne, jak i rozumienia tekstu spośród badanych krajów OECD. Zmiany w grupach wiekowych pokazują różnorodne tendencje. W większości krajów największe spadki odnotowuje się w grupach najstarszych. Polska jest jednym z niewielu krajów, w którym największe różnice odnotowujemy w grupie najmłodszej.

Warto pamiętać, iż wyniki testów nie muszą bezpośrednio odzwierciedlać poziomu umiejętności i przy interpretacji należy brać pod uwagę różnice pomiędzy cyklami. Dotyczą one zmian w instrumentach pomiarowych, zmian metodologicznych w pomiarze osób o niskich umiejętnościach, przejścia z komputera lub testu papierowego na tablet oraz zmiany nastawienia respondentów do udziału w badaniach. Przy czym wyniki cały czas pozostają miarą potencjału danych krajów i grup społecznych. Badania pokazują znaczący spadek wyników w większości krajów, co jest niepokojącym trendem. Przedstawione w raporcie pierwsze wyniki badania wskazują na potrzebę szerszej analizy czynników wpływających na wyniki testów, aby lepiej zrozumieć, co leży u podstaw tego spadku.

3. Metodologia badania

Wstęp

Kluczowym elementem każdego badania porównawczego jest precyzyjne określenie założeń, zapewnienie wspólnych standardów i ich dopełnienie przez kraje uczestniczące w badaniu. W badaniach porównawczych ważne jest przede wszystkim kompleksowe definiowanie ocenianych zdolności, porównywalność wyników między grupami i krajami oraz wiarygodność naukowa. W przypadku badania PIAAC założenia badania są sformułowane w osobnym dokumencie, który wyjaśnia cele badania, ale również definiuje dokładne dziedziny i wymiary mierzonych umiejętności, co przekłada się na rzetelność konstruowanych testów i trafność interpretowanych rezultatów. Podsumowujemy je w pierwszej części rozdziału. W drugiej części przedstawiamy główne informacje dotyczące realizacji badania. W projekcie PIAAC obowiązywały precyzyjne standardy techniczne obejmujące różne etapy i procesy badania⁵. Projekt zakładał systematyczne i wieloetapowe sposoby zapewniania jakości w kluczowych dla jakości badania obszarach: reprezentatywności doboru próby, tłumaczenia narzędzi, standaryzacji procedur zbierania danych poprzez szkolenie ankierów oraz kodowania i przetwarzania danych. W tej części raportu opisujemy najważniejsze aspekty tego procesu.

Przygotowanie i realizacja badania była długim procesem. Narzędzia przetestowano w badaniu próbnym, które zorganizował każdy z krajów biorących udział w badaniu, a poszczególne działania były szczegółowo uzgadniane z instytucjami odpowiedzialnymi za nie w konsorcjum międzynarodowym. Niezależni eksperci z Grupy Ekspertów Technicznych (*Technical Advisory Group*) w maju 2024 roku dokonali całościowej oceny jakości danych i przestrzegania standardów technicznych przez poszczególne kraje. Ocena obejmowała analizę reprezentatywności danych, podatności na błędy wynikające z braku odpowiedzi oraz porównywalności wyników między krajami. Ekspertzy stwierdzili, że wszystkie kraje spełniły standardy PIAAC, a jakość danych jest wystarczająco wysoka do ich publikacji i raportowania.

Co mierzy PIAAC?

Pomiar umiejętności PIAAC został przygotowany na podstawie precyzyjnych założeń teoretycznych OECD (OECD, 2021). Założenia badania, przygotowane i opublikowane na początku projektu, wyjaśniają cele badania, definiują dziedziny i wymiary mierzonych umiejętności. Te definicje były podstawą do stworzenia testów mierzących umiejętności i są ważne przy interpretowaniu wyników.

W 2023 pomiar był przeprowadzony wyłącznie z wykorzystaniem tabletów, inaczej niż w poprzedniej edycji, przeprowadzanej z użyciem komputerów lub papierowych arkuszy. Pomimo statystycznej równoważności między narzędziami papierowymi i cyfrowymi, tradycyjne

⁵ Standardy są dostępne na stronach OECD: [PIAAC CY2 Technical Standards and Guidelines.pdf](#)

papierowe arkusze ograniczały różnorodność zadań. Całkowite przejście na narzędzia cyfrowe było więc znaczącą zmianą, umożliwiającą przygotowanie bardziej interesujących i atrakcyjnych dla badanych zadań testowych, które lepiej odpowiadają współczesnym realiom.

Międzynarodowe badania umiejętności dorosłych, mimo wspólnego celu, przeszły ewolucję, która odzwierciedlała zmieniające się potrzeby i priorytety edukacyjne oraz rozwój metodologiczny. Pierwszymi tego typu inicjatywami były badania IALS oraz ALL, które stopniowo rozszerzały zakres analiz kompetencji dorosłych. Międzynarodowe Badanie Umiejętności Czytania i Pisania (IALS), przeprowadzone w 22 krajach w latach 1994–1998, było pierwszym porównawczym badaniem kompetencji osób dorosłych w wieku 16–65 lat. IALS koncentrowało się na pomiarze podstawowych umiejętności, takich jak rozumienie tekstów i wykonywanie obliczeń. Kontynuacją IALS było Badanie Umiejętności Czytania i Pisania oraz Umiejętności Życiowych (ALL). Przeprowadzono je w latach 2003–2008 w 11 krajach. Do badania wprowadzono pomiar rozumowania matematycznego oraz umiejętności rozwiązywania problemów. Pierwszy cykl Programu Międzynarodowej Oceny Kompetencji Dorosłych (PIAAC) przeprowadzono w 3 rundach w latach 2011–2017. Uczestniczyło w nim 39 krajów. Polska brała udział w pierwszej rundzie badania, przeprowadzonej w latach 2011–2012 (zob. Rynko, 2013). Kluczową innowacją wprowadzoną w pierwszym cyklu PIAAC było zastosowanie rozwiązań cyfrowych – badanie przeprowadzono z wykorzystaniem komputerów, co umożliwiło kompleksową ocenę umiejętności wykorzystania technologii informacyjno-komunikacyjnych. Drugi cykl badania PIAAC nawiązuje do założeń poprzedniej edycji, ale zaktualizowano je i uzupełniono, aby osiągnąć lepsze dopasowanie definicji mierzonych umiejętności do nowych sposobów korzystania z tych umiejętności. Założenia pomiaru w obu cyklach badania są jednak powiązane koncepcyjnie i praktycznie: część zadań była użyta w pierwszym cyklu badania (*linking items*, zadania łączące), co umożliwia porównywanie wyników obu cykli na tych samych skalach. Niemniej jednak ze zmianą warunków testowania wiążą się pewne problemy dotyczące porównywalności. Mimo że wyniki testów zostały uznane za psychometrycznie równoważne, należy pamiętać, że samo ich rozwiązywanie odbywało się w różnych warunkach, co mogło wpłynąć na doświadczenie uczestników oraz ich podejście i motywację do wykonywanych zadań.

Umiejętność rozumienia tekstu (*literacy*) jest zdefiniowana w PIAAC jako „wyszukiwanie, odczytanie znaczeń, ocenę i refleksję nad pisanymi tekstami, by osiągnąć zamierzone cele, pogłębić własną wiedzę i potencjał i uczestniczyć w życiu społecznym” (OECD, 2021). Umiejętność czytania, której kluczową częścią jest rozumienie tekstu, jest niezbędna w pracy, w kontaktach z innymi ludźmi i w wielu codziennych sprawach. Ważną jej częścią jest umiejętność przeglądania tekstu i wyszukiwania informacji. Osoby sprawnie czytające potrafią znaleźć w tekście konkretne treści, również korzystając z odpowiednich jego elementów, np. spisu treści lub nagłówków. Potrafią też zdecydować, które fragmenty tekstu są ważniejsze, i dostosowują odpowiednio sposób czytania, np. czytają niektóre fragmenty bardziej pobieżnie niż inne. W PIAAC taką umiejętność nazywa się „wyszukiwaniem informacji” (*accessing*). Ale czytamy po to, żeby zrozumieć tekst (*understanding*), co podkreśla kolejny element definicji tej

umiejętności. Odczytywanie znaczenia obejmuje zarówno dosłowne odczytanie przekazu, jak i łączenie dotychczasowej wiedzy z treścią tekstu i wyciąganie własnych wniosków. Kiedy mamy do czynienia z kilkoma tekstami na ten sam temat, musimy też umieć powiązać informacje z różnych źródeł, a w przypadku sprzecznych informacji – ocenić, które są bardziej wiarygodne. Czytając tekst, powinniśmy zadawać sobie pytania: Czy ten tekst jest dla mnie przydatny? Czy informacje są wiarygodne? Czy autor jest obiektywny? Czy przedstawione argumenty są spójne? Pytania te odnoszą się do wymienionego w definicji rozumienia tekstu aspektu oceny tekstu i zawartych w nim informacji. Dotyczy ona źródła i zawartości tekstu oraz jego wiarygodności, a także jakości znajdujących się w nim informacji oraz ich przydatności. Definicja rozumienia tekstu uznaje także to, że, z perspektywy indywidualnej, czytanie to narzędzie do osiągnięcia własnych celów. Potrzebujemy tej umiejętności zarówno w prostych sprawach, jak znalezienie drogi, jak i w trudniejszych – jak załatwianie spraw urzędowych. Czytanie jest też ważne w kontaktach z ludźmi, rozrywce, działaniu w społeczności i w pracy.

W drugim cyklu badania PIAAC, podobnie jak w pierwszym, wyróżniono cztery główne konteksty wykorzystania umiejętności czytania: zawodowy (związany z czytaniem np. instrukcji obsługi, raportów, e-maili, procedur bezpieczeństwa, specyfikacji technicznych), osobisty (np. czytanie dla przyjemności, informacje o zdrowiu, podróżach, hobby, komunikacja z rodziną i przyjaciółmi przez e-maile i media społecznościowe), społeczny (np. ulotki informacyjne, dokumenty urzędowe, artykuły prasowe, informacje z internetu dotyczące spraw publicznych) oraz edukacyjny, związany z czytaniem na potrzeby uczenia się.

W ciągu dnia czytamy teksty w wielu różnych kontekstach, ale korzystamy także z różnych form tekstów, od prostych, liniowych, po bardziej złożone interaktywne treści w internecie. Do pomiaru umiejętności rozumienia tekstu dobrano zróżnicowane teksty, kierując się, podobnie jak to miało miejsce w pierwszym cyklu badania, konkretnymi kryteriami: typu tekstu (np. opis, narracja czy instrukcja), formatu (ciągły, nieciągły lub mieszany), organizacji (długość, ilość informacji i ewentualnych narzędzi nawigacji) oraz źródła (jeden tekst lub wiele tekstów). Choć teksty mogły zawierać dodatkowe elementy, takie jak zdjęcia, wykresy czy ilustracje, pomiar umiejętności rozumienia tekstu był oparty tylko na tekście pisanym.

W badaniu PIAAC większość zadań jest utajnionych, więc prezentowane tu zadania mają tylko charakter poglądowy. Na rysunku 3.1 przedstawiono drugą część przykładowego zadania, które pokazuje typowe elementy zadań mierzących umiejętność rozumienia tekstu. Zadanie oparte jest na fikcyjnym regulaminie przedszkola. W pierwszym pytaniu respondent jest proszony o znalezienie na liście zasad godziny przyprowadzania dzieci do szkoły. Zadanie jest łatwe i sprawdza umiejętność przejrzenia i zrozumienia tekstu. Druga część zadania pokazana na rysunku 3.1. polega na znalezieniu zasady, która jest tymczasowo zawieszona w związku z Dniem Pluszaka opisanym w nowej, pojawiającej się na ekranie 2. pytania wiadomości e-mail od nauczyciela. To zadanie jest trudniejsze, ponieważ wymaga porównania informacji z dwóch różnych tekstów (łączenia informacji z różnych źródeł) oraz

rozumienia, że pluszak to rodzaj zabawki, o której mowa w regulaminie. Zadanie to ilustruje, podkreślone w założeniach drugiego cyklu badania, łączenie i interpretowanie informacji z różnych źródeł.

Rysunek 3.1. Przykład zadania mierzące umiejętność rozumienia tekstu.

The screenshot shows a digital task interface for PIAAC. At the top, it says 'PIAAC' with navigation icons. The task is labeled 'Zadanie 580 - Pytanie 1 / 1'. The instruction asks the user to read an email and answer a question. The email is from Marta Jankowska, titled 'Dzień Pluszaka', and contains information about a preschool event. To the right of the email is a document titled 'REGULAMIN PRZEDSZKOLA' with a row of colorful crayons at the top. The document contains a welcome message and a list of rules for the preschool, such as arrival time, clothing, and bringing toys.

Źródło: OECD, 2024.

Umiejętność rozumowania matematycznego (*numeracy*) obejmuje „wyszukiwanie, wykorzystywanie i krytyczne rozumienie treści matematycznych, informacji i idei reprezentowanych na wiele sposobów w celu angażowania się i radzenia sobie z wymaganiami matematycznymi obecnymi w różnych sytuacjach w dorosłym życiu” (Tout i in., 2021). Definicję tę zmodyfikowano w porównaniu z pierwszym cyklem. Podkreślono w niej rolę krytycznego rozumowania, zaakcentowano znaczenie dostępu do treści za pomocą technologii cyfrowych i różnorodne sposoby prezentacji treści matematycznych (teksty, symbole, obrazy czy treści w formie dynamicznej). Współczesny świat, w którym informacje są powszechnie dostępne za pośrednictwem internetu i nowych technologii, wymaga od nas krytycznego i świadomego podejścia do treści matematycznych, prezentowanych w różnorodnych formatach. Przykładowo, wyszukiwanie informacji jest czymś więcej niż techniczną umiejętnością zidentyfikowania elementów tekstu i obejmuje umiejętność efektywnego korzystania z dostępnych danych.

Zadania zawierają treści z czterech obszarów matematycznych: arytmetyki, geometrii, relacji i zależności funkcyjnych oraz statystyki i rachunku prawdopodobieństwa. Zadania

charakteryzują się różnym stopniem trudności, od prostych obliczeń po złożone równania i wzory. Poprawne rozwiązanie niektórych zadań może też wymagać od respondentów zastosowania wiedzy i umiejętności z kilku obszarów jednocześnie. Na przykład zadanie dotyczące obliczania kosztu zakupu pewnej ilości produktu w określonej cenie może wymagać zarówno umiejętności z obszaru obliczeń (wykonanie mnożenia), jak i rozumienia jednostek miar i ich przeliczania (np. przeliczenie kilogramów na gramy). Zadania PIAAC różnią się od typowych zadań szkolnych. Zgodnie z nurtem współczesnego myślenia o umiejętności rozumowania matematycznego zadania mierzą przede wszystkim umiejętność praktycznego zastosowania umiejętności matematycznych w realistycznych kontekstach codziennego życia. Zadania te są osadzone w trzech kontekstach odzwierciedlających różnorodność sytuacji, w których dorośli wykorzystują umiejętności matematyczne:

- osobistym – umiejętności matematyczne wykorzystywane w życiu codziennym, np. obliczanie budżetu domowego, planowanie podróży, czytanie etykiet produktów spożywczych, porównywanie cen,
- zawodowym – umiejętności matematyczne stosowane w pracy, np. obliczanie ilości potrzebnych materiałów, analiza danych sprzedażowych, sporządzanie raportów finansowych, interpretacja wykresów i diagramów,
- społecznym – np. interpretacja danych statystycznych, analiza wyników wyborów, ocena ryzyka inwestycji.

Rysunek 3.2 przedstawia przykładowe interaktywne zadanie mierzące umiejętność rozumowania matematycznego. W zadaniu do określenia liczby potrzebnych rolek tapety użyto kalkulatora. Pierwsza część wymaga samodzielnego obliczenia liczby rolek tapety oraz wprowadzenia parametrów ściany i rolki. Rozwiązując zadanie, można usuwać i zmieniać wprowadzone wartości. Druga część zadania polega na znalezieniu błędu i poprawieniu błędnie wpisanej szerokości lub długości rolki. Zadania sprawdzają umiejętności pracy z danymi przestrzennymi i liczbowymi. Ich trudność polega nie tylko na samych obliczeniach, ale przede wszystkim na konieczności dokładnego rozumienia różnych jednostek miary, przekładania informacji z diagramu na wartości liczbowe oraz umiejętności krytycznego myślenia przy weryfikacji wprowadzonych danych. Typowe popełniane przez rozwiązujących błędy to pomyłki w jednostkach (np. wpisanie centymetrów zamiast metrów), nieumiejętne przeliczanie skal oraz nieprawidłowe interpretowanie diagramu przestrzennego. Druga część zadania, polegająca na znalezieniu błędu, jest trudniejsza, bo wymaga nie tylko poprawnego wykonania obliczeń, ale także rozpoznania, która z wprowadzonych wartości jest nieprawidłowa. Test sprawdza umiejętności złożone, wykraczające poza proste kalkule: rozumienie kontekstu, przestrzenną wyobraźnię matematyczną oraz krytyczną analizę danych. Zadanie wykorzystano w badaniu próbnym PIAAC i jest ono umiarkowanie trudne.

Rysunek 3.2. Przykład zadania.

PIAAC

Zadanie 1 - Pytanie 1 / 2

Proszę spojrzeć na plan sypialni i kalkulator tapety. Proszę posłużyć się kalkulatorem tapety, aby odpowiedzieć na poniższe pytanie.

Ścianę za szafą i biurkiem (oznaczoną strzałką) należy pokryć tapetą od podłogi do sufitu.

Wprowadź właściwe informacje do kalkulatora tapety. Następnie naciśnij przycisk „Oblicz”, aby wyświetlić liczbę potrzebnych rolek.

Uwaga: rysunek nie jest w skali.

Kalkulator tapety obliczy, ile rolek tapety potrzebujesz.

Wykonaj kroki 1 i 2. Następnie naciśnij przycisk „Oblicz”. Jeśli chcesz zacząć od nowa, naciśnij przycisk „Wyczyść”.

1. Wprowadź informacje o tapecie.

Szerokość rolki (cm)

Długość rolki (m)

2. Wprowadź wymiary ściany

Szerokość (m)

Wysokość (m)

Oblicz
Wyczyść

Wysokość pokoju

2.50 metrów (m)

Rolka tapety

Szerokość: 52 centymetry (cm)
Długość: 10,05 metra (m)

PIAAC

Zadanie 1 - Pytanie 2 / 2

Proszę spojrzeć na ilustrację przedstawiającą wymiary ściany, która ma być pokryta tapetą. Proszę posłużyć się kalkulatorem tapety, aby odpowiedzieć na poniższe pytanie.

Liczba rolek, którą pokazuje kalkulator, nie jest poprawna. Przy wprowadzaniu wymiarów popełniono błąd.

Popraw w kalkulatorze wymiar, który został źle wprowadzony lub wymiary, które zostały źle wprowadzone.

Kalkulator tapety obliczy, ile rolek tapety potrzebujesz.

Wykonaj kroki 1 i 2. Następnie naciśnij przycisk „Oblicz”. Jeśli chcesz zacząć od nowa, naciśnij przycisk „Wyczyść”.

1. Wprowadź informacje o tapecie.

Szerokość rolki (cm)

Długość rolki (m)

2. Wprowadź wymiary ściany

Szerokość (m)

Wysokość (m)

Oblicz
Wyczyść

Potrzebujesz

369

liczba rolek

Źródło: OECD, 2024.

Żyjemy w świecie, który szybko się zmienia, dlatego dorośli muszą być elastyczni i otwarci na nowe wyzwania. Ważniejsze niż dawniej stają się takie wyzwania, jak zmiana wymagań zawodowych w pracy czy konieczność zmiany zawodu oraz zmieniające się sposoby funkcjonowania w życiu osobistym i społecznym. Wymaga to stałego uczenia się nowych rzeczy i radzenia sobie z niepewnością (Greiff i in., 2017). To było motywacją włączenia do PIAAC 2023 nowej dziedziny, która zastąpiła dziedzinę rozwiązywania problemów w środowiskach bogatych w technologię z pierwszego cyklu badania. Adaptacyjne rozwiązywanie problemów (*adaptive problem solving*, APS), które w tym raporcie określamy skrótowo *rozwiązywaniem problemów*, obejmuje „zdolność do osiągania własnych celów w dynamicznej sytuacji, w której metoda rozwiązania nie jest od razu dostępna. Wymaga to angażowania się w procesy poznawcze i metapoznawcze w celu zdefiniowania problemu, wyszukiwania informacji i zastosowania rozwiązania w różnych środowiskach i kontekstach informacyjnych” (Greiff i in., 2021).

Istotą adaptacyjnego rozwiązywania problemów jest elastyczność, dynamiczne podejście do ich rozwiązywania, umiejętność korzystania z różnorodnych zasobów oraz wykorzystania umiejętności metapoznawczych do oceny i refleksji nad własnym procesem myślowym. APS koncentruje się na zdolności adaptacji do zmieniających się okoliczności oraz na elastycznym modyfikowaniu strategii w trakcie rozwiązywania problemów. W rozwiązywaniu tego rodzaju problemów ważna jest umiejętność wyboru i efektywnego wykorzystywania zasobów, zarówno fizycznych, jak i cyfrowych. Założenia tej części pomiaru podkreślają też znaczenie oceny i refleksji nad własnym myśleniem (tzw. metapoznanie), np. sprawdzanie, na ile skuteczne są nasze działania i czy zbliżamy się do rozwiązania problemu, rozpoznanie, co utrudnia nam rozwiązanie problemu i jakie błędy popełniamy, a także modyfikacja strategii rozwiązania problemu, jeśli dotychczasowe metody okazują się nieskuteczne.

Rysunek 3.3 przedstawia przykładowe zadanie z tej dziedziny. Pierwsza część wymaga użycia interaktywnej mapy w celu znalezienia najszybszej trasy do osiągnięcia trzech celów w ograniczonym czasie: przewiezienia dziecka do szkoły, zrobienia zakupów spożywczych i powrotu do domu w określonym czasie, który zmienia się w zależności od wybranych opcji. To zadanie sprawdza, jak radzimy sobie z rozwiązywaniem problemów w określonych ramach: wymaga przeanalizowania różnych możliwych tras i czasów przejazdu. Druga część jest bardziej wymagająca, ponieważ pojawiają się nowe okoliczności, które wymuszają zmianę pierwotnego planu: awaria wodociągu blokuje jedną z tras, zmuszając uczestników do modyfikacji planu przy zachowaniu tych samych ograniczeń. Zadanie nie jest trudne. Wymaga umiejętności planowania, systematyczności w sprawdzaniu tras i zwrócenia uwagi na ograniczenia. O niskim poziomie trudności decyduje ograniczona liczba informacji, predefiniowane trasy i jasne ograniczenia czasowe. Główne wyzwanie polega na szybkim i logicznym dostosowaniu planu do zmieniających się okoliczności.

Rysunek 3.3. Przykładowe zadanie mierzące umiejętność rozwiązywania problemów.

PIAAC
?
◀ ▶

Zadanie 120 - Pytanie 1/2

Zapoznaj się z mapą i notatką poniżej. Wskaż miejsca na mapie, aby odpowiedzieć na poniższe pytanie.

Jest godzina 8:00. Musisz wykonać zadania wymienione w notatce poniżej.


Zaplanuj najszybszą trasę, aby zrealizować te zadania. Pamiętaj o ograniczeniach czasu.

Po zakończeniu, naciśnij strzałkę **DALEJ**, aby kontynuować. Jeśli chcesz zacząć od nowa, naciśnij przycisk **RESETUJ**. Całkowity czas przejazdu wyświetlony na dole mapy będzie się aktualizował w trakcie planowania trasy

- Odwieź dziecko do szkoły przed 8:30.
- Zrób zakupy (20 minut).
- Wróć do domu przed spotkaniem o 9:30

PIAAC
?
◀ ▶

8:00



RESETUJ

PIAAC
?
◀ ▶

Zadanie 120 - Pytanie 2/2

Zapoznaj się z mapą i notatką poniżej. Wskaż miejsca na mapie, aby odpowiedzieć na poniższe pytanie.

Zaplanowałeś pójście do sklepu A.

Jest 8:30. Odwiozłeś już dziecko do szkoły. Dostajesz alert z wiadomością, że wybrany przez Ciebie sklep został zamknięty z powodu awarii wodociągowej i zalania.

Dostosuj swoją trasę, aby zrealizować pozostałe zadania. Pamiętaj o ograniczeniach czasu.

Po zakończeniu, naciśnij strzałkę **DALEJ**, aby kontynuować.

- Odwieź dziecko do szkoły przed 8:30.
- Zrób zakupy (20 minut).
- Wróć do domu przed spotkaniem o 9:30

PIAAC
?
◀ ▶

8:30



RESETUJ

Źródło: OECD, 2024.

Ocena umiejętności rozumienia tekstu i rozumowania matematycznego w 2023 r. obejmowała też dodatkowe zestawy zadań stworzonych z myślą o osobach z umiejętnościami na bardzo niskim poziomie, które określamy jako badanie umiejętności podstawowych z zakresu czytania (*reading components*) oraz matematyki (*numeracy components*). Pierwszy cykl PIAAC wykazał, że znaczący odsetek dorosłych osiągnął niski poziom kompetencji w rozumieniu tekstu i rozumowaniu matematycznym, co stało się przesłanką do włączenia wyników z tego dodatkowego narzędzia do wyliczenia wyników. Badanie umiejętności podstawowych z zakresu czytania mierzy podstawowe umiejętności rozumienia znaczenia zdań i płynności czytania: umiejętności dekodowania, rozumienia zdań i krótkich fragmentów tekstów. Badanie umiejętności podstawowych z zakresu matematyki mierzy intuicje rozumienia liczb, ilości i relacji między liczbami, np. określenie ilości obiektów i porównywanie ich wielkości. Wykorzystanie tych zadań pomaga precyzyjniej oszacować umiejętności osób z niskimi umiejętnościami poniżej poziomu 1.

Populacja i próba w badaniu PIAAC w Polsce

Sposób doboru próby, poziom jej realizacji oraz zastosowane procedury ważenia danych to kluczowe elementy decydujące o wiarygodności wyników badania. Badanie PIAAC przeprowadzono na *reprezentatywnej próbie losowej*, a jego wyniki są uogólniane na populację, z której próba ta została wylosowana. Uogólnianie wyników uzyskanych w próbie losowej na populację obciążone jest zawsze błędami o charakterze *losowym* lub *niełosowym* (*systematycznym*). Błędy losowe, wynikające z tego, że nie zbieramy danych od wszystkich osób należących do danej populacji, minimalizuje się przez dobór odpowiedniej próby, a metody statystyczne pozwalają oszacować ich prawdopodobną wielkość. Błędy systematyczne, związane z niepełną reprezentatywnością próby, ogranicza się, stosując wysokiej jakości operat losowania, dbając o maksymalizację realizacji próby oraz odpowiednie ważenie danych, ale ich kierunek i wielkość można jedynie hipotetycznie oszacować poprzez analizy porównawcze z zewnętrznymi kryteriami.

Populacja docelowa

Docelową populacją badaną w badaniu PIAAC są stali mieszkańcy Polski w wieku od 16 do 65 lat⁶, bez względu na narodowość, obywatelstwo lub język, z wyłączeniem osób pozostających pod opieką lub nadzorem instytucjonalnym (np. przebywających w szpitalach, domach pomocy społecznej, więzieniach, koszarach). Zgodnie ze standardami technicznymi badania PIAAC, osoby z barierami uniemożliwiającymi udział w badaniu (nieznajomość języka, niepełnosprawności, problemy z pisaniem/czytaniem) są zaliczane do populacji docelowej, lecz nie uwzględnia się ich przy obliczaniu poziomu realizacji próby.

⁶ Ściśle rzecz biorąc, osoby, które ukończyły 16 lat i nie ukończyły jeszcze 66 lat w dniu przypadającym w połowie okresu realizacji terenowej badania PIAAC (w Polsce 1 stycznia 2023). Oznacza to, że najmłodszy uczestnicy badania mogli wypełniać testy PIAAC na kilka, kilkanaście tygodni przed 16. urodzinami (jeśli ankietę dotarł do nich na początku okresu realizacji badania), zaś najstarsi kilka, kilkanaście tygodni po 66. urodzinach (jeśli ankietę dotarł do nich pod koniec okresu realizacji terenowej).

Próba losowa: operat, sposób doboru; wielkość próby, warstwowanie i wiązkanie

Próbę do badania PIAAC wylosowano z operatu PESEL metodą dwustopniową: najpierw losowano miejscowości, następnie w nich samych badanych⁷. Próba zakładała 9800 osób (ostatecznie 9804), z minimalną liczbą 7 osób w każdej miejscowości. Ze statystyk PESEL wynikało, że jedna wiązka 7 osób (minimalna liczba osób wylosowanych w danej miejscowości) przypada na 16 660,7 osoby w wieku 16–65 lat.

Pierwszy etap losowania polegał na losowaniu miejscowości. Prawdopodobieństwo wylosowania danej miejscowości w pierwszym etapie losowania odpowiadało ilorazowi liczby jej mieszkańców w wieku 16–65 lat i wartości 16660,7.

Miejscowości, dla których wartość tego ilorazu była większa od 1 (miasta mające co najmniej 16661 mieszkańców w wieku 16–65 lat, co odpowiada w przybliżeniu 25 880 mieszkańców ogółem), włączane były do próby bez losowania („jednostki pewnościowe”). Liczba osób przewidzianych do wylosowania z każdej z tych miejscowości w drugim etapie losowania odpowiadała ilorazowi liczby mieszkańców w wieku 16–65 lat przez 2380,1 (zaokrąglonemu do najbliższej liczby całkowitej). Liczba takich miejscowości w próbie wynosiła 193. Miejscowości o mniejszej liczbie mieszkańców losowane były w procedurze losowania systematycznego, z listy uporządkowanej ze względu na (kolejno): podział miasto – wieś, województwo, liczba mieszkańców („warstwy *implicite*”). Prawdopodobieństwo wylosowania danej miejscowości było proporcjonalne do liczby mieszkańców danej miejscowości. W każdej wylosowanej miejscowości z tej grupy przewidywano wylosowanie w drugim etapie 7 osób. Do próby wylosowano łącznie 835 miejscowości z tej grupy.

W drugim etapie w każdej wylosowanej miejscowości losowano osoby. W każdej z wylosowanych miejscowości losowanie osób zostało przeprowadzone w procedurze losowania systematycznego, z listy uporządkowanej ze względu na (kolejno): płeć i wiek („warstwy *implicite*”). W miejscowościach „pewnościowych” losowano osoby w liczbie od 7 do 108⁸. W pozostałych wylosowanych miejscowościach losowano po 7 osób.

Dzięki przyjętemu schematowi losowania prawdopodobieństwo trafienia do próby dla każdej osoby z operatu losowania było w przybliżeniu takie samo⁹, natomiast zastosowanie losowania systematycznego z „warstwami *implicite*” zapewniło zgodność rozkładu w próbie kluczowych cech (typ i wielkość miejscowości, województwo, płeć, grupa wiekowa) z rozkładami w operacie losowania. Oba etapy losowania przeprowadzone zostały przez polski zespół badania PIAAC, na podstawie

⁷ „Miejscowość” na użytek losowania definiowana była jako jednostka stosowana w bazie PESEL. Jako osobne „miejscowości” traktowane były także dzielnice kilku największych miast.

⁸ Największą „miejscowością pewnościową” był Gdańsk: w przypadku większych miast jednostkami losowania („miejscowościami”) były ich dzielnice.

⁹ Prawdopodobieństwa te były identyczne (1/2380,1) dla mieszkańców miejscowości mniejszych, natomiast dla mieszkańców „miejscowości pewnościowych” były one nieznacznie zróżnicowane ze względu na konieczność zaokrąglania do wartości całkowitych liczby osób przypadających do wylosowania w danej lokalizacji.

danych z systemu PESEL udostępnionych przez Centralny Ośrodek Informatyki. Drugi etap losowania przeprowadzany był „na ślepo” – jego wynikiem był wykaz wylosowanych kolejnych numerów na uporządkowanych listach mieszkańców miejscowości. Niezbędne dane osobowe (imię, nazwisko, płeć, wiek, pełen adres) wylosowanych osób dołączył Centralny Ośrodek Informatyki.

Próba zrealizowana

Spośród 9804 osób wylosowanych do próby w Polsce badanie zostało przeprowadzone z 5013 osobami (51,1% oryginalnej próby), przy czym wyniki testów kompetencji uzyskano dla 4908 osób (50,1%). Właściwy „poziom realizacji próby” jest jednak wyższy (57%), gdyż od wylosowanych 9804 przypadków należy odliczyć te, które nie należą do populacji badanej (w szczególności osoby, które trwale opuściły Polskę, a także zmarłych w okresie między losowaniem próby a realizacją badania), oraz osoby, z którymi przeprowadzenie badania było niemożliwe ze względu na niewystarczającą znajomość polskiego, nieumiejętność pisania lub czytania, niepełnosprawność fizyczną lub intelektualną. Głównymi przyczynami niezrealizowania badania były odmowa (wyrażona osobiście przez wylosowanego bądź za pośrednictwem innej osoby) – ponad 42% niezrealizowanej próby – oraz niez uzyskanie kontaktu z badanym pomimo wielokrotnych prób – ponad 25%.

Poziom realizacji w innych krajach był bardzo zróżnicowany. W nielicznych krajach był on wysoki (powyżej 60%), poziom ten osiągnęły Korea (73%), Słowacja (70%), Singapur (62%), Izrael (61%) oraz Hiszpania (61%). Polska znalazła się w grupie krajów z relatywnie wysokim poziomem realizacji. Kraje z wartością wskaźnika poniżej 40% to Czechy (40%), Portugalia (39%), Austria (39%), Anglia (38%), Chorwacja (36%), Region Flamandzki (35%), Szwecja (31%), Szwajcaria (30%) i Dania (27%).

Ważenie danych

Uogólnianie wyników wymaga nadania respondentom odpowiednio dostosowanej „wagi” odpowiadającej udziałowi w populacji. W pierwszym etapie (ważenie poststratyfikacyjne) każdej osobie przypisuje się wagę odwrotną do prawdopodobieństwa wylosowania, by wyrównać obciążenia reprezentatywności. W drugim etapie (ważenie postrealizacyjne) wagi są korygowane, aby zminimalizować efekty nielosowych odmów udziału. W badaniu PIAAC ważenie przeprowadzono z uwzględnieniem następujących zmiennych (kategorii): płci, wieku, województwa, wielkości miejscowości, poziomu bezrobocia i wynagrodzeń oraz liczby miast w powiecie. Procedura ta zapewnia dostosowywanie wag w próbie tak, aby ich rozkłady były zgodne z rozkładami brzegowymi tych zmiennych w populacji.

Obciążenia wyników związane z niepełną realizacją próby

Ważenie danych pozwala ograniczyć obciążenie wyników ze względu na niepełną realizację próby, jednakże nie może go całkowicie wyeliminować. Aby ocenić skalę pozostałego obciążenia, wyniki PIAAC zostały poddane skoordynowanym międzynarodowo analizom (*nonresponse bias analysis*), w których symulowano zmiany wyników przy uwzględnieniu innych niż zastosowane schematów ważenia, także uwzględniających kalibrację rozkładów potencjalnie istotnych zmiennych (w szczególności wykształcenia) do dostępnych alternatywnych danych o populacji. Wyniki tych analiz wskazują, że Polska należy do krajów, w których efekty obciążenia ze względu na niepełną realizację badania są umiarkowane¹⁰.

Przy ocenie obciążenia wyników dla Polski warto szczególną uwagę zwrócić na kwestię wykształcenia. Wykształcenie nie było zmienną uwzględnianą w ważeniu danych (ze względu na brak informacji o wykształceniu osób nieuczestniczących w badaniu, a także brak spójnych danych dotyczących rozkładu osiągniętego wykształcenia). Wykształcenie jest silnie skorelowane z badanymi w PIAAC umiejętnościami i prawdopodobieństwem uczestniczenia w badaniu wylosowanej osoby, stąd „skrzywienie” próby ze względu na wykształcenie respondentów może znacząco przyczynić się do obciążenia ostatecznych wyników badania. Jednocześnie wykształcenie jest skorelowane z wiekiem oraz miejscem zamieszkania (również cechami skorelowanymi z prawdopodobieństwem uczestniczenia w badaniu, a wykorzystywanymi przy ważeniu danych), więc choć nie było bezpośrednio używane w ważeniu, procedura ważenia powinna przynajmniej przybliżyć rozkłady wykształcenia w próbie do rozkładów populacyjnych.

Dane o rozkładzie wykształcenia w populacji są jednak dostępne i choć ich wykorzystanie wiąże się z pewnymi problemami¹¹, można je porównać z danymi uzyskanymi w badaniu PIAAC (po uwzględnieniu ważenia). Jak pokazuje tabela 3.1, pomimo ważenia, dane z polskiej próby zaniżają odsetek osób lepiej wykształconych względem danych pochodzących ze statystyk publicznych. Warto przy tym zauważyć, że również w pierwszej edycji PIAAC z lat 2011/2012 ważne rozkłady wykształcenia odbiegały od danych GUS/EUROSTAT, tyle że w drugą stronę (większy był niż wedle GUS/EUROSTAT udział osób z wyższym wykształceniem).

¹⁰ Na 30 państw, dla których podane były wyniki analizy obciążenia ze względu na niepełną realizację próby, jako kraje o podobnej skali problemu jak w Polsce określono 12 innych państw, dla 6 krajów stwierdzono mniejsze nasilenie problemów, a dla 11 większe. W przypadku dwóch krajów (Włoch i Łotwy) ryzyko obciążeń uznano za wysokie.

¹¹ Inaczej niż dane o wieku, płci i miejscu zamieszkania, które pochodzą z rejestrów administracyjnych, dane o wykształceniu, podawane w statystyce publicznej, opierają się na deklaracjach (przy czym informatorem w spisie powszechnym nie zawsze była osoba, której informacja ta dotyczyła) i nie są poddawane innej weryfikacji. Ponadto dane o wykształceniu pochodzące z Narodowego Spisu Powszechnego Ludności i Mieszkań 2021 charakteryzowały się znaczącymi odsetkami braków odpowiedzi. Osobnym problemem jest publikowanie danych o wykształceniu jedynie w podziale na kategorie wiekowe nieodpowiadające ściśle definicji populacji badanej w PIAAC.

Tabela 3.1. Rozkład wykształcenia w grupie wiekowej 25–64 lata¹² wedle PIAAC oraz GUS/EUROSTAT.

	PIAAC 2022 Próba zrealizowana (dane nieważone)	PIAAC 2022 Dane ważone	GUS/ EUROSTAT 2022	PIAAC 2011 Dane ważone	GUS/ EUROSTAT 2011
Wykształcenie gimnazjalne, podstawowe lub niższe	7,1%	6,9%	6,5%	10,1%	11,1%
Wykształcenie zasadnicze zawodowe, średnie lub pomaturalne	65,7%	63,7%	59,2%	60,9%	65,6%
Wykształcenie wyższe	27,2%	29,4%	34,4%	29,0%	23,3%

Źródło: opracowanie własne.

Symulacje, w których do procedury ważenia włącza się wykształcenie¹³, wskazują jednak, że efekt niedoszacowania wykształcenia na średnie wyniki w pomiarze umiejętności jest stosunkowo niewielki i nie przekracza kilku punktów: po skorygowaniu rozkładu wykształcenia, średnia wyniku w zakresie rozumowania matematycznego wzrasta o 3,17 punktu, zaś w zakresie rozumienia tekstu o 2,45 punktu. Nie oznacza to, że dokładnie taką „korektę” należy przyjąć, odczytując polskie wyniki PIAAC, ale wynik ten wskazuje zarówno na kierunek, jak i na skalę obciążenia.

Techniki i narzędzia badawcze

W przypadku każdego respondenta procedury przewidywały pięć następujących po sobie kolejno części badania PIAAC:

- 1. Opis miejsca badania** (ankieta obserwacyjna) – po udaniu się pod adres osoby wylosowanej do badania, ale przed kontaktem z nią lub z inną osobą mieszkającą pod tym adresem, ankieter wypełniał krótki kwestionariusz, nazywany ankietą obserwacyjną (ang. *curbside observation*). Ankieta ta zawierała trzy pytania dotyczące budynku znajdującego się pod danym adresem. Dane te dostarczały informacji o osobach, które były wylosowane do próby, ale nie wzięły udziału w badaniu, co było pomocne w ocenie reprezentatywności badania.

¹² Ograniczenie się do grupy wiekowej 25–64 związane jest z kategoryzacją wieku dla rozkładów wykształcenia w dostępnych danych statystyk publicznych. Jednocześnie dla kategorii wiekowych młodszych niż 25 lat zarówno pomiar wykształcenia, jak i jego wartość predykcyjna jest znacznie niższa (w kategorii tej większość stanowią osoby wciąż się uczące).

¹³ W symulacjach przeprowadzono ważenie wieńcowe danych PIAAC z uwzględnieniem – poza zmiennymi uwzględnionymi w oryginalnej procedurze ważenia – pochodzące z danych statystyki publicznej rozkłady wykształcenia (4 kategorie: poniżej zasadniczego, zasadnicze zawodowe, średnie lub pomaturalne, wyższe) w grupach wiekowych (cztery kategorie: 25–34, 35–44, 45–54, 55–65 lat oraz kategoria 16–24 bez podziału na grupy wykształcenia).

- 2. Inicjalizacja** – po nawiązaniu kontaktu z respondentem i uzyskaniu jego zgody ankieter zadawał kilka pytań weryfikujących, czy rozmawia z właściwą, tj. wylosowaną do badania osobą.
- 3. Wywiad kwestionariuszowy** ankietera z respondentem realizowany metodą **ankiety CAPI** (wywiadu wspomaganego komputerowo). Kwestionariusz zawierał jedenaście bloków pytań dotyczących następujących zagadnień: cech społeczno-demograficznych respondenta (np. kraju urodzenia, blok A), jego udziału w edukacji i szkoleniach (np. poziomu wykształcenia, blok B), aktualnej i przeszłej sytuacji zawodowej oraz cech obecnej lub ostatniej pracy (np. formy zatrudnienia, wysokości wynagrodzenia, bloki C, D i E), wykorzystania umiejętności w pracy i życiu codziennym (bloki F i G), środowiska pracy (np. wykonywania pracy w zespole, blok H), postaw i działań respondenta (np. poziomu zadowolenia z życia, blok I), jego rodziny i gospodarstwa domowego (np. liczby dzieci, blok J) oraz kompetencji społeczno-emocjonalnych (np. ugodowości, blok K). Przeciętny wywiad CAPI w Polsce trwał 35 minut.
- 4. Badanie umiejętności** – po zakończeniu wywiadu z respondentem ankieter przekazywał mu tablet wraz z rysikiem, prosząc o samodzielne rozwiązanie pojawiających się nim zadań. Odpowiedzi udzielane przez respondenta w zadaniach były zasadniczą podstawą do szacowania poziomu umiejętności. Całe badanie umiejętności trwało w Polsce przeciętnie 48 minut i składało się z następujących kolejno sekcji:
 - a. Samouczek obsługi tabletu** – w tej sekcji respondent mógł poznać i w razie potrzeby przećwiczyć sposób obsługi tabletu (np. wybierania odpowiedzi).
 - b. Badanie plasujące** (ang. *locator*) – w tej sekcji przy użyciu tabletu respondent rozwiązywał jedenaście zadań z dziedziny rozumienia tekstu i dziedziny rozumowania matematycznego, co pozwalało wstępnie rozpoznać poziom jego umiejętności i na tej podstawie dobrać przedstawiane mu w dalszej części zadania. W przypadku 6,7% polskich respondentów poziom umiejętności został zdiagnozowany jako niski i w związku z tym brali oni udział tylko w badaniu umiejętności podstawowych, z pominięciem zasadniczego badania umiejętności. W przypadku 15,3% polskich respondentów poziom umiejętności został wstępnie zdiagnozowany jako średni, w związku z czym zostali oni zakwalifikowani zarówno do badania umiejętności podstawowych, jak i do zasadniczego badania umiejętności. O udziale w badaniu umiejętności podstawowych respondentów z wysokimi umiejętnościami decydowało losowanie. W rezultacie 67,2% polskich respondentów pominięło badanie umiejętności podstawowych ze względu na wstępnie zdiagnozowane wysokie umiejętności, a 10,8% mimo to brało w nim udział.
 - c. Badanie umiejętności podstawowych** (ang. *components*) – w tej sekcji respondent wykonywał podstawowe zadania z zakresu czytania i liczenia. Wprowadzenie tej sekcji do procedury badania służyło lepszemu rozpoznaniu umiejętności respondentów o niskim poziomie umiejętności.
 - d. Zasadnicze badanie umiejętności** – w tej części respondent otrzymał i samodzielnie rozwiązywał zestaw zadań z dwóch spośród trzech dziedzin: rozumienia tekstu, rozumowania matematycznego lub rozwiązywania problemów. Badanie w dwóch

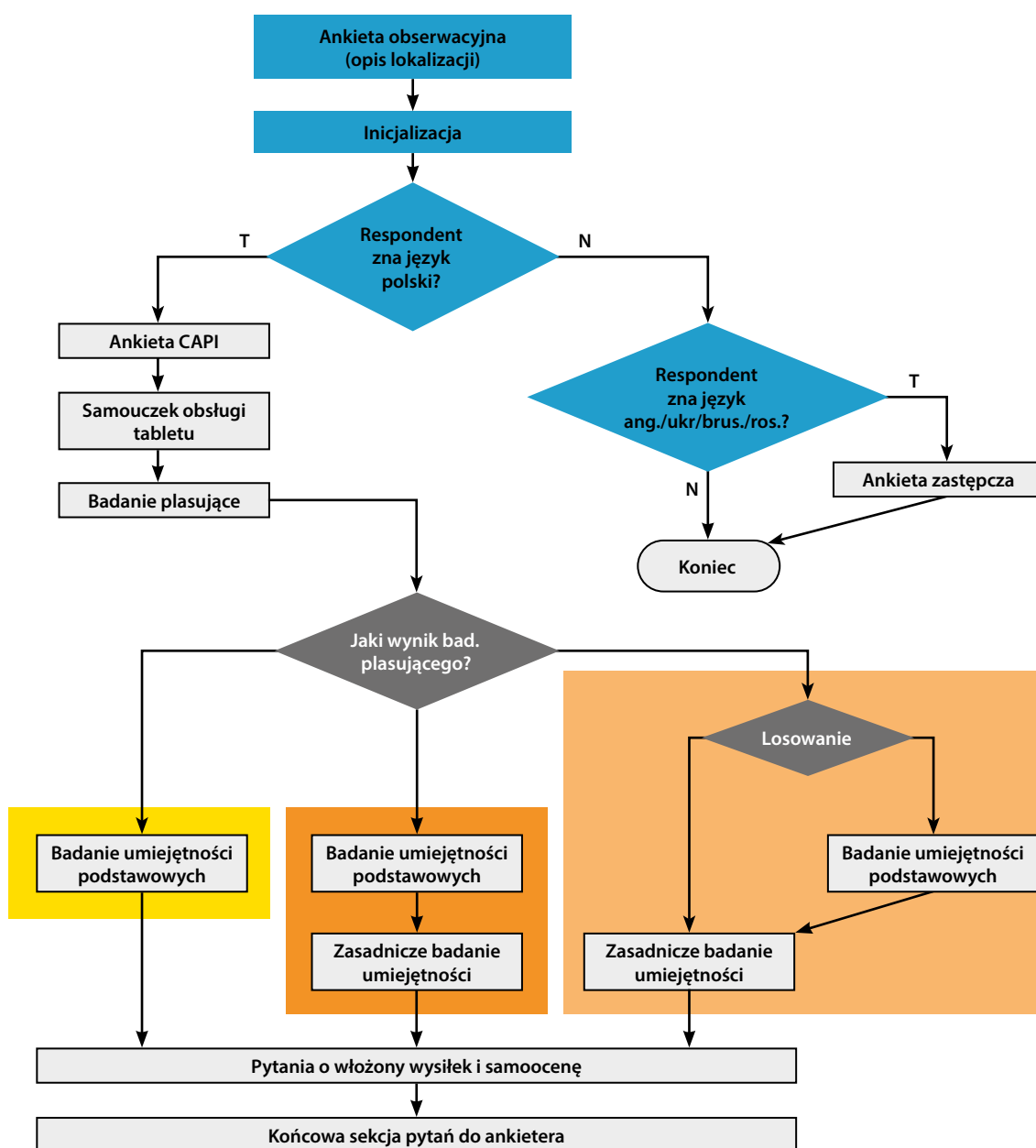
pierwszych z tych dziedzin miało charakter adaptacyjny, czyli dobór zadań przedstawianych respondentowi zależał od wyników rozwiązywania przez niego zadań wcześniejszych – co służyło bardziej precyzyjnemu pomiarowi jego umiejętności, a zarazem uniknięciu jego frustracji w zetknięciu ze zbyt łatwymi lub zbyt trudnymi zadaniami.

e. Samoocena – w tej sekcji respondent był proszony o opinię, na ile zadań poprawnie odpowiedział, i zadeklarowanie, jak duży wysiłek włożył w wykonywanie zadań.

5. Końcowa sekcja pytań do ankietera – ankieter po zakończeniu wywiadu był zobowiązany udzielić odpowiedzi na krótki zestaw pytań dotyczących okoliczności i przebiegu badania, np. tego, czy respondentowi ktoś pomagał lub przeszkadzał w badaniu umiejętności.

Przebieg badania podsumowuje poniższy schemat.

Rysunek 3.4. Schemat przebiegu badania z udziałem respondenta.



Źródło: Opracowanie własne.

Ankieta CAPI i badanie umiejętności były przeprowadzane w języku polskim. W przypadku, gdy osoba wylosowana nie znała języka polskiego, ale znała język angielski, rosyjski, ukraiński lub białoruski, procedury badania przewidywały wykonanie, zamiast ankiety CAPI, badania umiejętności i końcowej sekcji pytań do ankietera, tzw. ankiety zastępczej (ang. *doorstep interview*). Ankieta zastępcza wykorzystywała kwestionariusz dostępny w czterech wymienionych wyżej językach, zawierający sześć pytań o podstawowe cechy społeczno-demograficzne respondenta, który był wypełniany przez niego samodzielnie. Celem wprowadzenia ankiety zastępczej było uwzględnienie osób nieznających języka urzędowego w wazeniu danych i analizie reprezentatywności próby. W Polsce – w odróżnieniu od wielu innych krajów – znaczenie ankiety zastępczej było jednak marginalne ze względu na znikomą liczbę osób nieznających języka polskiego w próbie. Ankieta ta została wypełniona tylko przez jedną osobę.

Opisany wyżej schemat badania dotyczy badania zasadniczego w drugim cyklu badania PIAAC, zrealizowanego w Polsce na przełomie 2022 i 2023 roku. W rozdziale 2 znajdują się uwagi dotyczące różnic pod względem schematu badania między pierwszym a drugim cyklem badania PIAAC i ich możliwych konsekwencji dla porównywalności wyników w czasie.

Realizacja terenowa badania i procedury zapewniania jakości

Realizacja terenowa badania zasadniczego, czyli przeprowadzanie wywiadów i badań umiejętności z udziałem respondentów, trwała od początku września 2022 roku do końca kwietnia 2023 roku. Badanie zasadnicze było poprzedzone badaniem pilotażowym (od maja do września 2021 roku), które służyło przetestowaniu i udoskonaleniu stosowanych narzędzi i procedur, oraz szkoleniami koordynatorów sieci ankieterskiej i samych ankieterów.

Badania terenowe, podobnie jak wszystkie inne etapy prowadzenia badania PIAAC w Polsce, zrealizowano zgodnie z rozbudowanymi i szczegółowo opisanymi międzynarodowymi standardami technicznymi. Celem tych standardów było zapewnienie wysokiej jakości danych pochodzących z badania oraz ich porównywalności między wszystkimi uczestniczącymi krajami. Standardy określały:

- sposób organizacji pracy ankieterów,
- procedury nawiązywania kontaktu z osobą wylosowaną do badania (np. listy zapowiednie, wymóg sześciu prób nawiązania kontaktu o różnych porach dnia i w różne dni tygodnia),
- czynności i zachowania ankietera w kontakcie z respondentem oraz innymi osobami (przede wszystkim innymi domownikami),
- parametry wykorzystywanego sprzętu komputerowego,
- sposób rejestrowania przez ankietera informacji o przebiegu realizacji badania,
- przesyłanie przez ankietera na bieżąco danych wynikowych z prowadzonych działań i zrealizowanych wywiadów do centralnej bazy danych, co pozwoliło na bieżące monitorowanie przebiegu badania i dokonywanie korekt,
- inne aspekty związane ze zbieraniem danych w badaniu terenowym.

W badaniu PIAAC w Polsce funkcjonował złożony, wielowarstwowy system kontroli jakości prac terenowych. Kontrola ta była prowadzona przez:

- działy kontroli agencji badawczych, którym zlecono realizację prac terenowych (PBS Sp. z o.o. i Danae Sp. z o.o.), które prowadziły kontrolę terenową i nieterenową zgodnie ze standardami Programu Kontroli Jakości Pracy Ankieterów,
- agencję badawczą, której zlecono wykonanie dodatkowej kontroli terenowej, czyli ponownego nawiązania kontaktu z respondentem w celu potwierdzenia prawidłowości realizacji badania (EDBAD pracownia doradczo-badawcza MACIEJ MROCZEK w badaniu zasadniczym, Grupa BST sp. z o.o. w badaniu pilotażowym),
- zespół badania PIAAC w Instytucie Badań Edukacyjnych, który analizował dane pochodzące z systemu kontroli i informatycznego systemu monitoringu realizacji badania oraz porównywał je z danymi pochodzącymi z rejestru PESEL w celu weryfikacji,
- międzynarodowe konsorcjum koordynujące badania PIAAC i nadzorujące jakość jego realizacji w cyklu comiesięcznych sprawozdań i spotkań z zespołem krajowym.

W kontroli wykorzystywano również dokumentację fotograficzną wykonywaną obowiązkowo przez ankieterów i nagrania przeprowadzanych przez nich wywiadów. Rodzące się wątpliwości co do poprawności realizacji wywiadu rozstrzygano po uzyskaniu wyjaśnień od ankietera. Kontrolą zostało objęte co najmniej 10% adresów przydzielonych każdemu ankieterowi (co najmniej 30% dla 6 ankieterów, którzy przeprowadzali szczególnie dużą liczbę wywiadów) W przypadku stwierdzenia poważnego naruszania standardów realizacji przez danego ankietera, kontrolą obejmowano wszystkie przeprowadzone przez niego wywiady.

Tak rozbudowany system kontroli, duża liczba danych zbieranych komputerowo (np. o szczegółowym przebiegu realizacji każdego wywiadu) i możliwość porównania z danymi rejestru PESEL pozwalały na identyfikację przypadków naruszeń standardów badania, w tym tak rażących jak fałszowanie wywiadów, niewykonanie badania umiejętności czy niesamodzielne rozwiązywanie zadań przez respondenta. W przypadku uzasadnionego podejrzenia takiego naruszenia podejmowano działania korygujące. Z polskiego zbioru wynikowego usunięto dane, jeśli zachodziły wątpliwości co do ich rzetelności. W ten sposób odrzucono w całości 130 fałszywych lub nieprawidłowych wywiadów, a ponadto 30 niewykonanych w rzeczywistości badań umiejętności i 37 badań umiejętności, w przypadku których zachodziło podejrzenie udzielania pomocy przez ankietera lub osobę trzecią. W Polsce 8 ankieterów (spośród 120 zaangażowanych w badanie zasadnicze) odsunięto od prac terenowych z powodu zastrzeżeń do jakości ich pracy. Oczyszczone w ten sposób dane przekazano do międzynarodowego konsorcjum realizującego badanie kontrolne, które nie stwierdziło naruszeń w Polsce standardów technicznych niewyeliminowanych w procesie kontroli jakości.

W trakcie dodatkowej analizy danych międzynarodowych eksperci OECD zidentyfikowali ankieterów, wobec których powstały wątpliwości dotyczące jakości ich pracy związane ze stosunkowo częstymi przejawami niestarannego wykonywania badania umiejętności przez

ich respondentów (takimi jak krótki czas wykonywania badania umiejętności, bardzo szybkie przechodzenie przez kolejne zadania i niski wynik w badaniu plasującym mimo wyższego wykształcenia). W Polsce problem ten dotyczył 9 ankierów i 774 respondentów, z którymi przeprowadzali wywiady. Mimo braku stwierdzenia naruszeń standardów technicznych i braku zastrzeżeń do jakości danych z ankiety CAPI, OECD w porozumieniu z polskim zespołem badania PIAAC zdecydowało się dokonać w tym przypadku korekty ze względu na potencjalny wpływ jakości pracy ankierów na wiarygodność wyników dotyczących umiejętności. W związku z tym dane z badań umiejętności pochodzące od tych ankierów wyłączone z analiz służących oszacowaniu umiejętności.

Skalowanie wyników

W badaniu PIAAC zastosowano zaawansowane metody statystyczne, które zapewniają wiarygodność i porównywalność wyników. Metodologia ta czerpie z doświadczeń pierwszego cyklu badania oraz z innych międzynarodowych projektów badawczych, takich jak międzynarodowe badanie piętnastolatków PISA, i innych badań porównawczych, w których mierzone są umiejętności. Fundamentem procesu wyliczania wyników jest teoria odpowiedzi na pozycje testowe (IRT). To zaawansowane podejście statystyczne, które wykracza poza proste zliczanie poprawnych odpowiedzi. System bierze pod uwagę zarówno trudność pytań, jak i ich zdolność do różnicowania osób pod względem poziomu umiejętności. Metoda ta umożliwia też taki sposób konstrukcji testu, w którym badani rozwiązują różne zestawy zadań z bardzo dużej puli zadań pokrywających szeroki zakres umiejętności.

Jednym z głównych wyzwań w badaniu PIAAC jest zapewnienie porównywalności wyników między krajami i wersjami językowymi. Rozwiązaniem jest wykorzystanie tzw. pytań łączących (*linking items*) – zestawu zadań, które pozostają niezmiennie we wszystkich wersjach testu i kolejnych cyklach badania. Działają one jako punkt odniesienia, umożliwiając porównania w czasie i między krajami. Proces skalowania przebiega oddzielnie dla każdej dziedziny i ma charakter iteracyjny. Najpierw ustala się wspólne parametry dla pytań łączących, które następnie służą jako punkt odniesienia dla pozostałych zadań. W kolejnym kroku prowadzona jest szczegółowa analiza odpowiedzi z różnych krajów, pozwalająca wykryć pytania, których trudność może się różnić ze względu na kontekst kulturowy lub językowy. Porównuje się też moc różnicującą poszczególnych pytań, czyli to, na ile skutecznie pozwalają one rozróżnić osoby o różnym poziomie umiejętności. W takich przypadkach parametry są dostosowywane indywidualnie dla grup wykazujących odstępstwa. PIAAC wykorzystuje zaawansowaną metodę kalibracji wielogrupowej IRT. Proces rozpoczyna się od estymacji parametrów międzynarodowych dla pytań trendowych, po czym następuje identyfikacja i analiza interakcji między pytaniami a krajami. Procedura jest powtarzana przy stopniowo obniżanych progach wykrywania niezgodności, aż do uzyskania pełnej spójności skalowania. Istotnym elementem skalowania wyników jest również modelowanie regresji latentnej, które łączy wyniki z poszczególnych dziedzin z danymi ankietowymi respondentów. Takie podejście zwiększa precyzję oszacowań populacyjnych i pozwala uwzględnić

wzajemne powiązania między różnymi obszarami umiejętności.

W drugim cyklu badania PIAAC wprowadzono istotne zmiany w porównaniu do skalowania wyników z pierwszego cyklu. Dwie dziedziny obecne w pierwszym cyklu: rozumienie tekstu i rozumowanie matematyczne nadal opierały się na pytaniach łączących, co zapewniło możliwość komunikowania na tej samej skali co w pierwszym cyklu (czyli ustalonego wtedy rozkładu określonego średnią 250 i odchyleniem standardowym 50 dla obu tych dziedzin). W drugim cyklu w dziedzinach rozumienia tekstu i rozumowania matematycznego wprowadzono nowy hybrydowy schemat testowania wieloetapowego, umożliwiając lepsze dopasowanie zadań do poziomu umiejętności respondentów. Test składał się z trzech, niewidocznych dla respondenta etapów: najpierw wszyscy rozwiązywali te same zadania (badanie plasujące), potem dostawali zadania łatwe lub trudne (zależnie od wyników), a w ostatnim etapie poziom był jeszcze dokładniej dopasowany. Co czwarty uczestnik dostawał przez cały test zadania średnio trudne, co pomagało w kalibracji testu. Pomiar umiejętności z tych dwóch dziedzin uwzględniał (inaczej niż w pierwszym cyklu) także zadania mierzące podstawowe umiejętności rozumienia tekstu i rozumowania matematycznego (zob. rys. 3.4). Pomiar rozwiązywania problemów był przeprowadzony w prostszy sposób: zadania podzielono na pięć zestawów, z których każdy miał 12–14 zadań, a każdy respondent losowo otrzymywał dwa z tych zestawów. Skalowanie umiejętności z tej dziedziny wymagało oszacowania parametrów zadań od podstaw i było prostsze, bo nie wymagało powiązania wyników z poprzedniego cyklu.

Dodatkowym wyzwaniem było uwzględnienie tego, że w drugim cyklu badanie przeprowadzono tylko na tabletach, co wymagało weryfikacji, czy pytania z pierwszego cyklu funkcjonują w taki sam sposób w nowym środowisku. Dzięki tym innowacjom możliwe było zwiększenie dokładności wyników oraz lepsze dostosowanie procesu skalowania do zmieniających się potrzeb badania. Mimo to, jak podkreślamy w rozdziale 2, uwzględnienie wyników z zadań dotyczących podstawowych umiejętności czytania i matematyki w oszacowaniach umiejętności w drugim cyklu mogło wpłynąć na porównywalność wyników, zwłaszcza w grupach o dużym odsetku respondentów z bardzo niskimi umiejętnościami. Drugą ważną różnicą było wprowadzenie tabletów jako jedyne narzędzie testowe, co zastąpiło technikę mieszaną z pierwszego cyklu badania, w której część badanych korzystała z komputerów, a część rozwiązywała zadania w wersji papierowej. Zapewniło to jednolite warunki pomiaru i pozwoliło na wykorzystanie zalet testowania komputerowego, mogło jednak wpłynąć na wyniki osób, które korzystałyby z kwestionariuszy papierowych, gdyby była taka możliwość. Miało to szczególne znaczenie w Polsce, gdzie w pierwszym cyklu badania aż 36% respondentów rozwiązywało zadania testowe w papierowych zeszytach testowych (zob. Rynko, 2013).

Bibliografia

- Aruguete, M. S., Huynh, H., Browne, B. L., Jurs, B., Flint, E., & McCutcheon, L. E. (2019). How serious is the "carelessness" problem in surveys? An evaluation of four detection methods. *Journal of Research in Personality, 79*, 71–79. <https://doi.org/10.1016/j.jrp.2019.02.002>
- Baltes, P. B. (1997). On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as foundation of developmental theory. *American Psychologist, 52*(4), 366–380.
- Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin, 143*(1), 1–35.
- Chłoń-Domińczak, A., Grygiel, P., Holzer-Żelażewska, D., Humenny, G., Maliszewska, A., Palczyńska, M., Sitek, M., Świst, K. (2015). *Polacy z niskimi kompetencjami: aktywność zawodowa, uczenie się przez całe życie oraz wykluczenie społeczne – rekomendacje dla polityki*. Instytut Badań Edukacyjnych.
- Chyl, K., Dębska, A., Pokropek, A., Szczurbiński, M., Tanaś, Ł. L., & Sitek, M. (2024). Assessment of adults with low literacy skills: A review of methods. *Frontiers in Education, 9*, Article 1346073.
- Deming, D. (2021). The growing importance of decision-making on the job. *NBER Working Paper Series, No. 28733*. National Bureau of Economic Research, Cambridge, MA. <https://doi.org/10.3386/w28733>
- Desjardins, R. (2017). *Political Economy of Adult Learning Systems: Comparative Study of Strategies, Policies and Constraints*. Bloomsbury Academic.
- Desjardins, R., & Warnke, A. J. (2012). Ageing and Skills: A Review and Analysis of Skill Gain and Skill Loss Over the Lifespan and Over Time. *OECD Education Working Papers, 72*. OECD Publishing.
- Eccles, J. S. (2011). Gendered educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *International Journal of Behavioral Development, 35*(3), 195-201.
- Eurostat. (2022). Life expectancy statistics. Pobrane z https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Life_expectancy_statistics
- Goldhammer, F., Martens, T., & Lüdtke, O. (2017). Conditioning factors of test-taking engagement in PIAAC: An exploratory IRT modelling approach considering person and item characteristics. *Large-Scale Assessments in Education, 5*(1), 1–25.
- Goldhammer, F., Martens, T., Christoph, G., & Lüdtke, O. (2016). Test-taking engagement in PIAAC (*OECD Education Working Papers No. 133*). OECD Publishing. <https://doi.org/10.1787/5jlzfl6fhxs2-en>

- Gomez, S. V. (2000). The collective that didn't quite collect: Reflections on the IALS. *International Review of Education*, 46(5), 419–431.
- Greiff, S., et al. (2021). PIAAC Cycle 2 assessment framework: Adaptive problem solving. In *The Assessment Frameworks for Cycle 2 of the Programme for the International Assessment of Adult Competencies*. OECD Publishing. <https://doi.org/10.1787/3a14db8b-en>
- Hanushek, E., et al. (2017). Coping with change: International differences in the returns to skills. *Economics Letters*, 153, 15–19. <https://doi.org/10.1016/j.econlet.2017.01.007>
- Heckman, J. J., & Kautz, T. (2013). Fostering and Measuring Skills: Interventions that Improve Character and Cognition. *NBER Working Paper 19656*. National Bureau of Economic Research.
- Jabkowski, P. (2021). The decline in response rates in European comparative surveys: New evidence from ESS, EVS and SHARE. *Survey Research Methods*, 15(2), 143–159. <https://doi.org/10.18148/srm/2021.v15i2.7740>
- Kyröläinen, A. J., & Kuperman, V. (2021). Predictors of literacy in adulthood: Evidence from 33 countries. *PLOS ONE*, 16(3), e0243763.
- OECD (2019). *Skills Outlook 2019: Thriving in a Digital World*. OECD Publishing.
- OECD (2021). *Education at a Glance 2021: OECD Indicators*. OECD Publishing.
- OECD (2024a). *Do adults have the skills they need to thrive in a changing world? Survey of Adult Skills 2023*, OECD Publishing, <https://doi.org/10.1787/b263dc5d-en>
- OECD (2024b), *Survey of Adult Skills 2023 Reader's Companion*, OECD Publishing, <https://doi.org/10.1787/3639d1e2-en>
- OECD (2019). *OECD Skills Strategy 2019: Skills to Shape a Better Future*. OECD Publishing.
- OECD (2019). Overview: The future of work is in our hands. In *OECD Employment Outlook 2019: The Future of Work*. OECD Publishing. <https://doi.org/10.1787/e4718721-en>
- OECD (2019). *Skills Matter: Additional Results from the Survey of Adult Skills*. OECD Skills Studies, OECD Publishing, <https://doi.org/10.1787/1f029d8f-en>
- Penk, C., Pöhlmann, C., & Roppelt, A. (2014). The role of test-taking motivation for students' performance in low-stakes assessments: An investigation of school-track-specific differences. *Large-Scale Assessments in Education*, 2, Article 5. <https://doi.org/10.1186/s40536-014-0005-4>
- Plomin, R., & Deary, I. J. (2015). Genetics and intelligence differences: Five special findings. *Molecular Psychiatry*, 20(1), 98–108.

Pokropek, A. (2016). Grade of membership response time model for detecting guessing behaviors. *Journal of Educational and Behavioral Statistics*, 41(3), 300–325.

Rynko, M. (Red.), Burski, J., Chłoń-Domińczak, A., Palczyńska, M., Śpiewanowski, P. (2013). *Umiejętności Polaków – wyniki Międzynarodowego Badania Kompetencji Osób Dorosłych (PIAAC)*. Instytut Badań Edukacyjnych. <https://open.icm.edu.pl/handle/123456789/6089>

Steedle, J. T., van Rijn, P. W., & Drolenga, H. (2019). Measuring test-taking effort on low-stakes tests. *Applied Measurement in Education*, 32(2), 139–153.
<https://doi.org/10.1080/08957347.2019.1587682>

Tout, D., et al. (2021). PIAAC Cycle 2 assessment framework: Numeracy. In *The Assessment Frameworks for Cycle 2 of the Programme for the International Assessment of Adult Competencies*. OECD Publishing. <https://doi.org/10.1787/c4221062-en>

Vágvölgyi, R., Coldea, A., Dresler, T., Schrader, J., & Nuerk, H. C. (2016). A review about functional illiteracy: Definition, cognitive, linguistic, and numerical aspects. *Frontiers in Psychology*, 7, Article 1617. DOI: 10.3389/fpsyg.2016.01617