



INSTYTUT BADAŃ
EDUKACYJNYCH
Państwowy Instytut Badawczy

PROJEKT „PRZYJAZNA SZKOŁA”

RAPORT Z BADAŃ PILOTAŻOWYCH

**OCENA WŁAŚCIWOŚCI PSYCHOMETRYCZNYCH
TESTOWANYCH NARZĘDZI POMIAROWYCH**

CZĘŚĆ DRUGA

PAWEŁ GRYGIEL

SYLWIA OPOZDA-SUDER

MAJA GAJDA

3. Wielowymiarowa Skala Samooceny: Poczucie własnej wartości (WSS-PWW)

3.1. Analiza struktury czynnikowej

1. Jestem z siebie zadowolony/zadowolona.
2. Sądzę, że mam dużo dobrych cech.
3. Często myślę, że inni są lepsi ode mnie.
4. Mam powody do dumy z siebie.
5. Uważam, że jestem kimś wartościowym.
6. Myślę, że jestem naprawdę dobry/dobra w wielu rzeczach.

Jednoczynnikowy model CFA oszacowany dla 6-itemowej wersji skali poczucia własnej wartości uzyskał satysfakcjonujące dopasowanie we wszystkich analizowanych próbach. W próbie treningowej oraz w próbie pełnej wartości RMSEA nie przekraczały progu 0.05, przyjętego jako kryterium dobrego dopasowania. W próbie testowej oraz w próbie pewnej wartości RMSEA były nieco wyższe, ale nadal pozostawały poniżej progu 0.08, wskazując na dopasowanie akceptowalne. Równocześnie wartości CFI i TLI we wszystkich próbach przekraczały wartość 0.95, przyjętą jako próg dobrego dopasowania, a wartości SRMR były wyraźnie niższe od progu 0.08. Całościowo wyniki te wskazują, że jednoczynnikowy model 6-itemowej wersji skali dobrze odtwarzał strukturę danych (zob. Tabela 1).

Tabela 1. WSS-PWW: Wskaźniki dopasowania dla modelu CFA (model wyjściowy, 6-itemowy)

MODEL	χ^2	df	RMSEA [90% CI]	CFI	TLI	SRMR
Próba treningowa	17.63*	9	0.036 [0.008-0.061]	0.998	0.996	0.014
Próba testowa	33.03**	9	0.060 [0.039-0.083]	0.990	0.983	0.017
Próba pełna	40.39**	9	0.049 [0.034-0.065]	0.994	0.990	0.014
Próba pewna	41.04**	9	0.055 [0.039-0.073]	0.993	0.989	0.016

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; 90% CI = 90% przedział ufności dla RMSEA; CFI = porównawczy indeks dopasowania; TLI = indeks Tuckera–Lewisa; SRMR = standaryzowany pierwiastek średniokwadratowy reszt. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Analiza standaryzowanych ładunków czynnikowych potwierdziła, że wszystkie pozycje były istotnie związane z czynnikiem latentnym. W próbie pełnej bezwzględne wartości ładunków mieściły się w zakresie od 0.496 do 0.818. Najsilniejsze związki z czynnikiem

odnotowano dla pozycji 5, 4 oraz 1, natomiast relatywnie najniższy, choć nadal istotny, związek uzyskała pozycja 3. Ujemny znak ładunku tej pozycji wynikał z jej odwróconego kierunku treściowego. Podobny profil ładunków uzyskano także w pozostałych próbach (zob. Tabela 2).

Tabela 2. WSS-PWW: Standaryzowane ładunki czynnikowe dla modelu CFA (model wyjściowy, 6-itemowy)

Wskaźnik	Próba treningowa	Próba testowa	Próba pełna	Próba pewna
WSSPWW_1	0.818**	0.791**	0.804**	0.810**
WSSPWW_2	0.677**	0.723**	0.695**	0.665**
WSSPWW_3	-0.514**	-0.475**	-0.496**	-0.566**
WSSPWW_4	0.808**	0.814**	0.809**	0.815**
WSSPWW_5	0.820**	0.809**	0.818**	0.835**
WSSPWW_6	0.750**	0.768**	0.759**	0.774**

Uwaga. Wartości przedstawiają standaryzowane współczynniki STDYX. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Podsumowując, wyniki CFA wspierają jednoczynnikową strukturę 6-itemowej wersji skali WSS-PWW. Model osiągnął dobre lub akceptowalne parametry dopasowania we wszystkich analizowanych próbach, a wszystkie pozycje istotnie reprezentowały wspólny konstrukt. Nie stwierdzono zatem przesłanek przemawiających za redukcją liczby pozycji w tej skali.

3.2. Analiza niezmienności pomiarowej

Analizę niezmienności pomiarowej skali WSS-PWW przeprowadzono ze względu na płeć i klasę, porównując model konfiguralny z modelem skalarnym. W modelu konfiguralnym założono tę samą strukturę czynnikową w grupie chłopców i dziewcząt, natomiast w modelu skalarnym wprowadzono dodatkowe ograniczenia dotyczące ładunków czynnikowych i progów odpowiedzi, konieczne do porównywania średnich latentnych. Wyniki analizy przedstawiono w Tabeli 3 i 4.

Niezmienność ze względu na płeć

Model konfiguralny charakteryzował się dobrym dopasowaniem do danych, $\chi^2(18) = 48.07$, RMSEA = 0.048, CFI = 0.995. Po nałożeniu ograniczeń skalarności uzyskano model o parametrach $\chi^2(34) = 114.87$, RMSEA = 0.058, CFI = 0.987. Zmiany wskaźników dopasowania względem modelu mniej restrykcyjnego mieściły się w przyjętych granicach (Δ RMSEA = 0.010; Δ CFI = -0.008).

Tabela 3. WSS-PWW: Wskaźniki dopasowania modeli niezmienności pomiaru ze względu na płeć

MODEL	Test chi-kwadrat dopasowania modelu ¹					RMSEA	Δ RMSEA	CFI	Δ CFI
	χ^2	df	$\Delta \chi^2$	Δ df	p				
Konfiguralny	48.07**	18	—	—	—	0.048	—	0.995	—
Skalarny	114.87**	34	66.467	16	<.001	0.058	0.010	0.987	-0.008

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; CFI = porównawczy indeks dopasowania. Δ = różnica wartości χ^2 , RMSEA i CFI w odniesieniu do modelu mniej restrykcyjnego. ¹ Wartości $\Delta \chi^2$ i Δ df obliczono z wykorzystaniem procedury DIFFTEST w pakiecie Mplus. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Jednocześnie wynik procedury DIFFTEST wskazywał, że różnica χ^2 między modelami była istotna statystycznie ($\Delta \chi^2(16) = 66.467$, $p < .001$), co świadczy o formalnym pogorszeniu dopasowania po wprowadzeniu ograniczeń skalarności. Oznacza to, że przesłanki na rzecz pełnej niezmienności skalarnej ze względu na płeć nie są całkowicie jednoznaczne. Z jednej strony wartości Δ RMSEA i Δ CFI wspierają założenie porównywalności pomiaru między dziewczynkami i chłopcami, z drugiej jednak istotny wynik DIFFTEST nakazuje zachowanie ostrożności przy interpretowaniu porównań średnich latentnych.

Mając na uwadze powyższe zastrzeżenie, porównano średnie latentne między grupami. Uzyskany wynik sugeruje, że dziewczynki osiągały niższy poziom poczucia własnej wartości niż chłopcy, stanowiący grupę referencyjną ($\Delta M = -0.497$, $SE = 0.090$, $p < .001$). Różnicę tę należy interpretować z uwzględnieniem niejednoznacznych przesłanek pełnej niezmienności skalarnej.

Niezmienność ze względu na klasę

Model konfiguralny uzyskał dobre dopasowanie do danych: $\chi^2(18) = 50.29$, $p < .001$, RMSEA = 0.050, CFI = 0.994. Po wprowadzeniu ograniczeń skalarności dopasowanie modelu pozostało bardzo dobre: $\chi^2(34) = 61.95$, $p < .01$, RMSEA = 0.034, CFI = 0.995.

Porównanie obu modeli nie wskazywało na pogorszenie dopasowania po nałożeniu ograniczeń równości ładunków czynnikowych i progów odpowiedzi. Wartość RMSEA była niższa niż w modelu konfiguralnym ($\Delta\text{RMSEA} = -0.016$), a CFI nieznacznie wzrósł ($\Delta\text{CFI} = 0.001$). Taki układ zmian wskaźników przemawia za porównywalnością pomiaru w grupach uczniów klas 4–5 i 6–8. Wniosek ten potwierdza również nieistotny statystycznie wynik procedury DIFFTEST, $\Delta\chi^2(16) = 22.94$, $p = .115$. Łącznie uzyskane rezultaty dostarczają podstaw do przyjęcia pełnej niezmienności skalarnej pomiaru ze względu na klasę.

Po potwierdzeniu niezmienności skalarnej porównano średnie latentne. Uczniowie klas 6–8 uzyskali istotnie niższy poziom poczucia własnej wartości niż uczniowie klas 4–5, stanowiący grupę referencyjną ($\Delta M = -0.483$, $SE = 0.103$, $p < .001$).

Tabela 4. WSS-PWW: Wskaźniki dopasowania modeli niezmienności pomiaru ze względu na klasę

MODEL	Test chi-kwadrat dopasowania modelu ¹					RMSEA	Δ RMSEA	CFI	Δ CFI
	χ^2	df	$\Delta\chi^2$	Δ df	p				
Konfiguralny	50.29**	18	—	—	—	0.050	—	0.994	—
Skalarny	61.95**	34	22.940	16	.115	0.034	-0.016	0.995	0.001

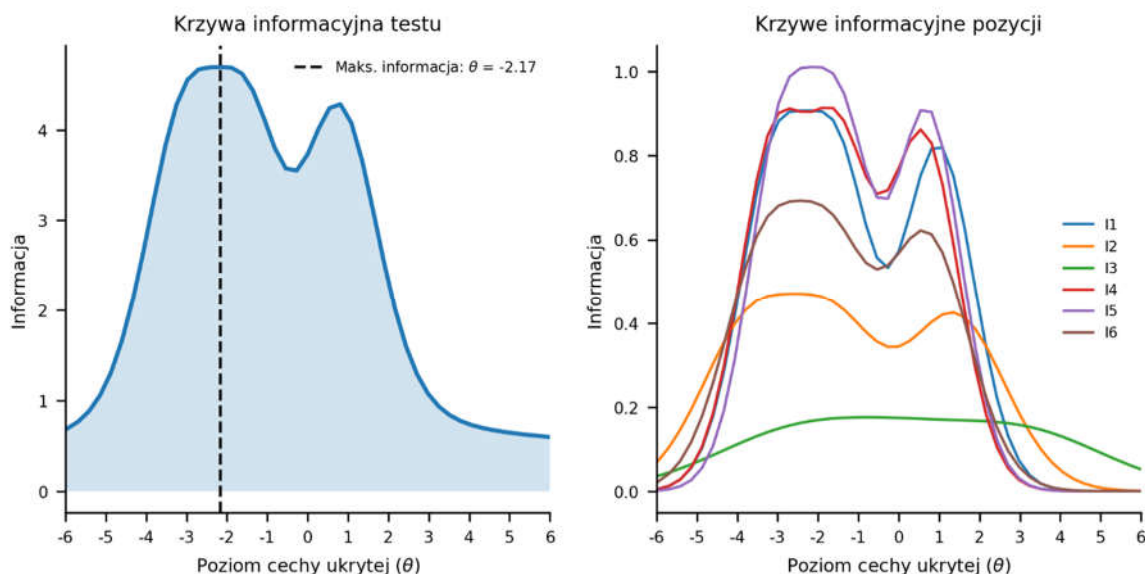
Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; *df* = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; CFI = porównawczy indeks dopasowania. Δ = różnica wartości χ^2 , RMSEA i CFI w odniesieniu do modelu mniej restrykcyjnego. ¹ Wartości $\Delta\chi^2$ i Δ df obliczono z wykorzystaniem procedury DIFFTEST w pakiecie Mplus. Istotne statystycznie: * $p < .05$. ** $p < .01$.

3.3. Krzywa informacyjna testu i pozycji testowych

Skala poczucia własnej wartości osiąga największą precyzję pomiaru w dolnym zakresie cechy, czyli przede wszystkim u osób z obniżonym poczuciem własnej wartości. Maksimum krzywej informacyjnej przypada około $\theta = -2.17$, gdzie rzetelność warunkowa wynosi około 0.79, a więc pozostaje bardzo blisko progu 0.80.

Jednocześnie skala zachowuje użyteczną dokładność w dość szerokim zakresie wyników. Rzetelność co najmniej 0.70 utrzymuje się od $\theta = -3.76$ do $\theta = 1.51$, co oznacza, że narzędzie dobrze różnicuje nie tylko osoby o bardzo niskim, ale także o przeciętnym i umiarkowanie podwyższonym poziomie poczucia własnej wartości. Największy wkład w informacyjność skali wnoszą pozycje I5, I4 i I1, natomiast najstabszy udział ma pozycja I3 (zob. Wykres 1).

Wykres 1. WSS-PWW: Krzywa informacyjna testu oraz krzywe informacyjne pozycji dla poczucia własnej wartości



Uwaga. Lewy panel przedstawia krzywą informacyjną całego testu, a prawy panel krzywe informacyjne poszczególnych pozycji. Oś X oznacza poziom cechy ukrytej θ , natomiast oś Y poziom informacji, czyli precyzję pomiaru na danym poziomie cechy. Wyższa wartość informacji oznacza większą dokładność pomiaru w danym zakresie θ . Linia przerywana w lewym panelu wskazuje poziom cechy, przy którym test osiąga maksymalną informację. Oznaczenia pozycji odnoszą się do kolejnych pozycji (itemów) wchodzących w skład analizowanego narzędzia.

3.4. Miary rzetelności

Rzetelność skali poczucia własnej wartości oszacowano z uwzględnieniem porządkowego charakteru odpowiedzi. Alfa porządkowa wyniosła $\alpha_{ord} = 0.871$, natomiast współczynnik omega całkowita McDonalda osiągnął wartość $\omega_t = 0.826$. Obie wartości wskazują na wysoką spójność wewnętrzną skali. Uzupełniając obliczono również klasyczną alfę Cronbacha, której wartość wyniosła $\alpha = 0.826$. Wartość AVE, czyli średniej wariancji wyjaśnionej, wyniosła 0.547.

3.5. Rozkład odpowiedzi na poszczególne twierdzenia skali

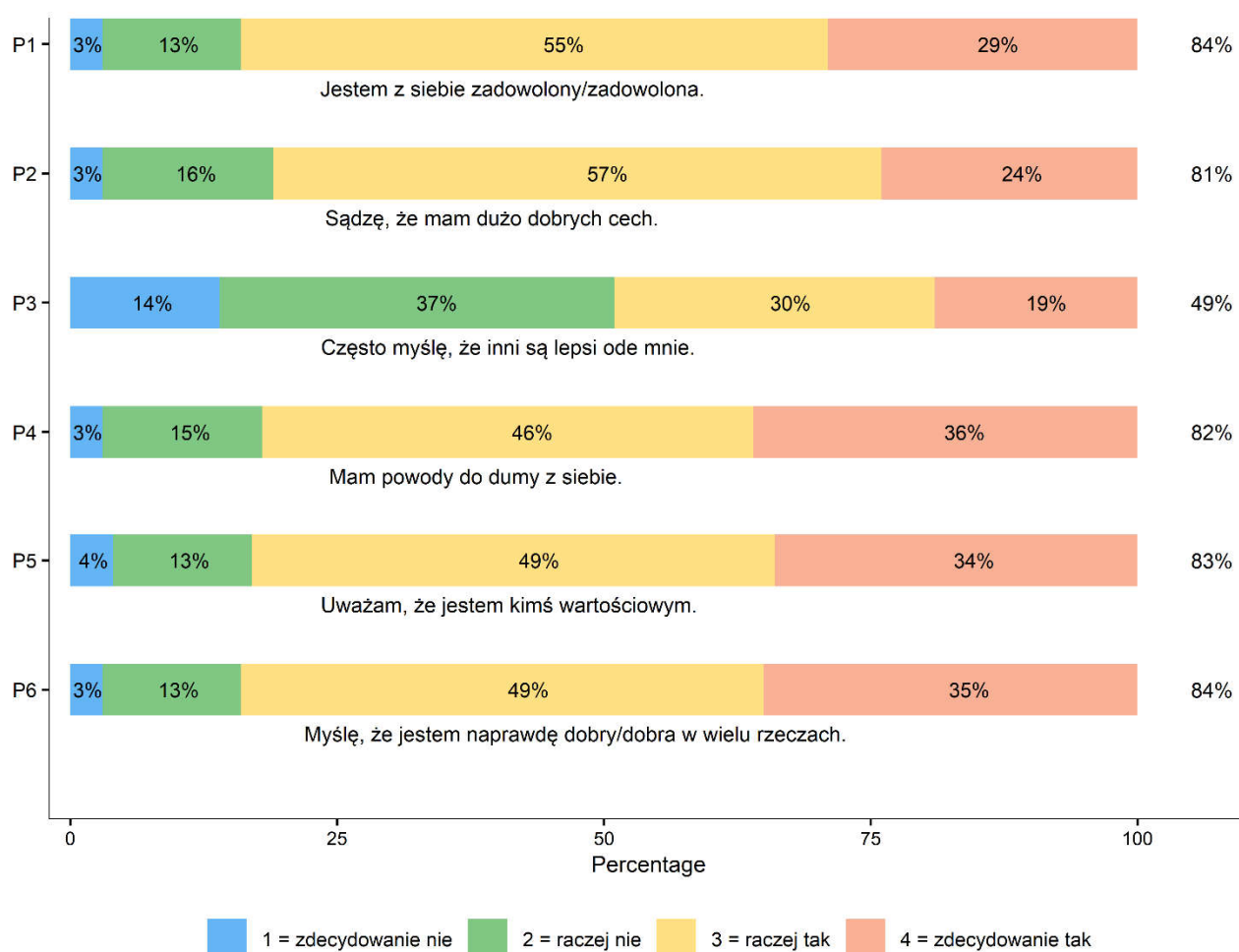
Rozkład odpowiedzi wskazuje, że uczniowie najczęściej oceniali własną wartość w sposób pozytywny. Najwyższy odsetek odpowiedzi pozytywnych uzyskano dla pozycji P1 („Jestem z siebie zadowolony/zadowolona”) oraz P6 („Myślę, że jestem naprawdę dobry/dobra w wielu rzeczach”) – w obu przypadkach łącznie 84% badanych raczej lub zdecydowanie zgodziło się z tymi stwierdzeniami. Bardzo zbliżony poziom aprobaty dotyczył także pozycji P5

(„Uważam, że jestem kimś wartościowym”) – 83% odpowiedzi pozytywnych, P4 („Mam powody do dumy z siebie”) – 82%, oraz P2 („Sądzę, że mam dużo dobrych cech”) – 81%.

Odmiennej kierunek interpretacji dotyczy pozycji P3 („Często myślę, że inni są lepsi ode mnie”), która była sformułowana negatywnie. Z tym twierdzeniem raczej lub zdecydowanie zgodziło się 49% uczniów, natomiast 51% wybrało odpowiedzi wskazujące na brak zgody. Wynik ten sugeruje, że mimo ogólnie korzystnego profilu samooceny znaczna część badanych doświadcza porównań społecznych obniżających ocenę siebie.

Podsumowując, uczniowie najczęściej deklarowali zadowolenie z siebie, poczucie własnej wartości oraz przekonanie o posiadaniu pozytywnych cech, choć odpowiedzi na pozycję odwróconą wskazują na pewne zróżnicowanie w zakresie niekorzystnych porównań z innymi (zob. Wykres 2).

Wykres 2. WSS-PWW: Procentowy rozkład odpowiedzi na poszczególne twierdzenia skali

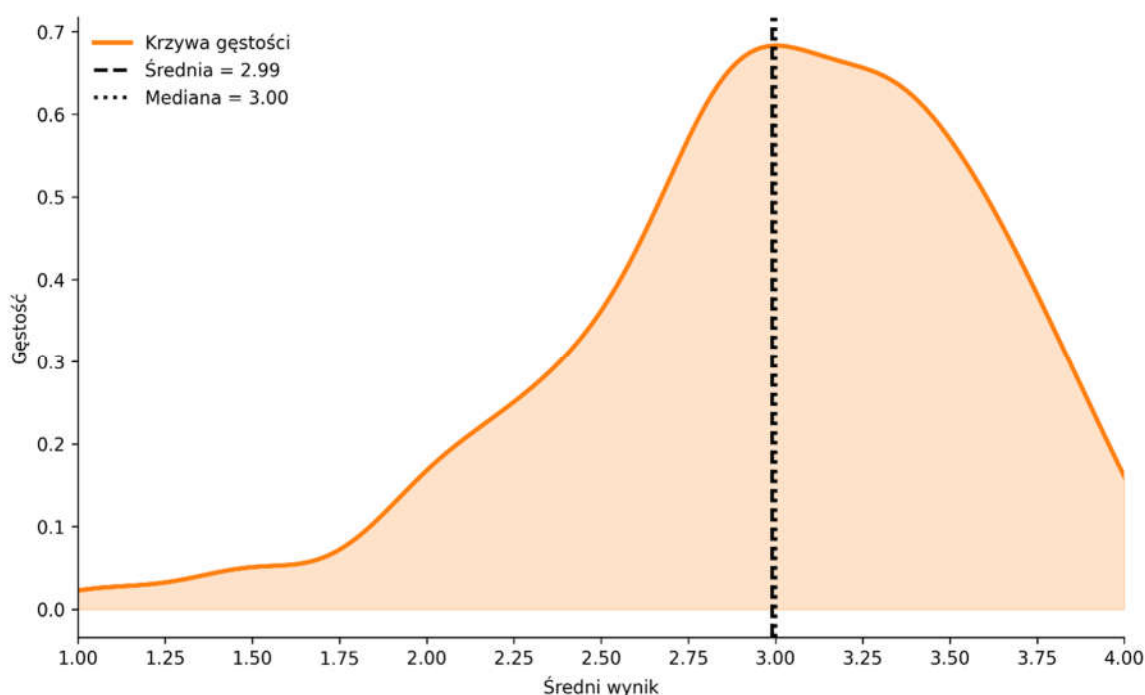


Uwaga. Wartości po prawej stronie wykresu przedstawiają łączny odsetek odpowiedzi wskazujących na zgodę z danym stwierdzeniem, czyli sumę kategorii „raczej się zgadzam” i „zdecydowanie się zgadzam”. Pozycja P3 była sformułowana negatywnie, dlatego zgodę z tym twierdzeniem należy interpretować odmiennie niż w przypadku pozostałych pozycji.

3.6. Charakterystyka rozkładu wyników (średnia punktów) skali

Uczniowie uzyskali wyniki wskazujące na raczej korzystne poczucie własnej wartości. Średnia wyniosła $M = 2.99$, odchylenie standardowe $SD = 0.59$, a mediana $Me = 3.00$. Wyniki obejmowały cały możliwy zakres, od 1 do 4, a punkt środkowy skali wynosił 2.50. Przy kierunku skali, w którym wyższe wartości oznaczają wyższe poczucie własnej wartości, średnia lokuje się powyżej tego punktu. Ujemna skośność rozkładu, wynosząca -0.69 , potwierdza, że średnie wyniki uczniów częściej lokowały się w górnym zakresie rozkładu. Z kolei lekko dodatnia kurtoza, równa 0.45 , wskazuje co najwyżej na niewielkie odstępstwo od rozkładu normalnego i nie sugeruje silnej koncentracji wyników skrajnych (zob. Wykres 3).

Wykres 3. WSS-PWW: Rozkład średnich wyników poczucia własnej wartości

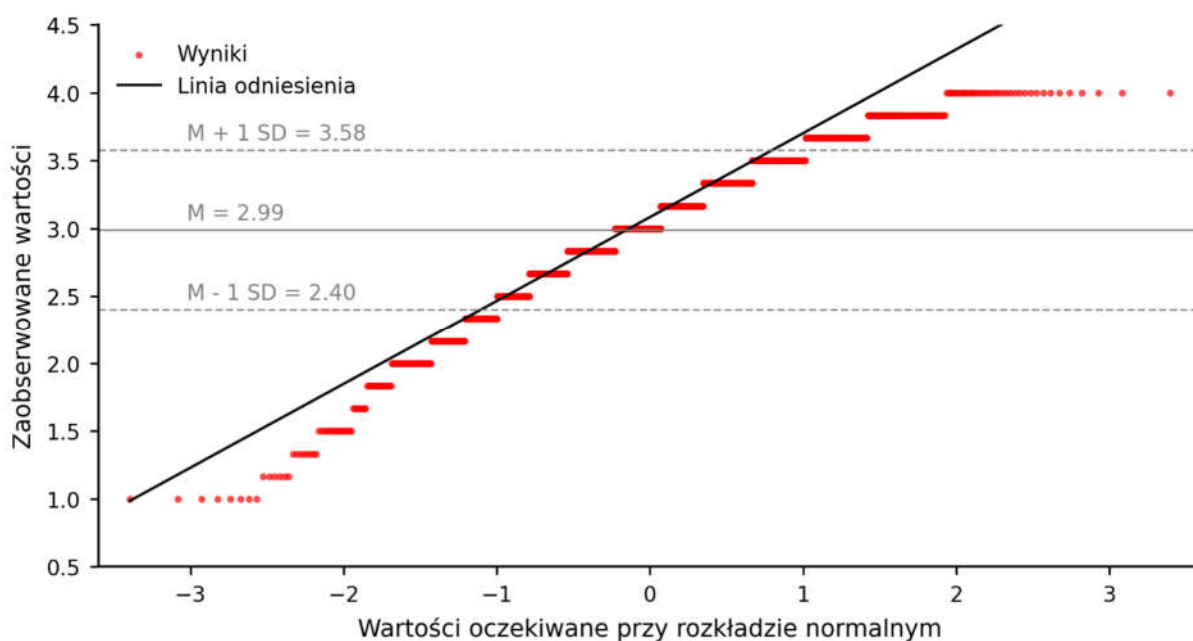


Uwaga. Na wykresie przedstawiono wygładzoną krzywą gęstości średnich wyników. Linia przerywana wyznacza średnią arytmetyczną (M), a linia kropkowana medianę (Me). Wysokość krzywej odzwierciedla względne zagęszczenie obserwacji w poszczególnych zakresach wartości.

Uzupełniając, wykres Q-Q (por. Wykres 4) wskazuje, że rozkład wyników tylko częściowo odpowiada rozkładowi normalnemu. W środkowej części rozkładu obserwacje układają się dość blisko linii odniesienia, co sugeruje względnie dobre dopasowanie do normalności dla większości wyników. Większe odchylenia pojawiają się natomiast na krańcach rozkładu, zwłaszcza w jego górnej części, gdzie widoczna jest koncentracja

obserwacji przy wyższych wartościach. Taki obraz jest spójny z ujemną skośnością i potwierdza, że uczniowie częściej uzyskiwali wyniki relatywnie wysokie niż niskie. Jednocześnie odchylenia od normalności nie mają charakteru skrajnego, lecz wskazują raczej na umiarkowane asymetryczne przesunięcie rozkładu w stronę wyższych poziomów poczucia własnej wartości.

Wykres 4. WSS-PWW: Wykres Q-Q średnich wyników poczucia własnej wartości względem rozkładu normalnego



Uwaga. Wykres przedstawia porównanie kwantyli empirycznych z kwantylami teoretycznymi rozkładu normalnego. Czerwone punkty oznaczają zaobserwowane wartości, a czarna linia pełni funkcję linii odniesienia dla rozkładu normalnego. Odchylenia punktów od tej linii wskazują na odstępstwa rozkładu empirycznego od normalności. Pozioma linia ciągła oznacza średnią arytmetyczną wyniku, oznaczoną jako M . Linie przerywane wyznaczają wartości o jedno odchylenie standardowe poniżej i powyżej średniej, czyli $M - 1 SD$ oraz $M + 1 SD$, gdzie SD oznacza odchylenie standardowe.

3.7. Wnioski

Podsumowując, wyniki analiz wskazują, że 6-itemowa skala WSS-PWW ma bardzo dobre właściwości psychometryczne. Analiza CFA potwierdziła jej jednoczynnikową strukturę, a wszystkie pozycje istotnie reprezentowały wspólny czynnik poczucia własnej wartości. Skala charakteryzowała się wysoką rzetelnością oraz korzystną precyzją pomiaru, szczególnie w zakresie obniżonego poziomu mierzonej cechy. Łącznie uzyskane rezultaty pozwalają traktować WSS-PWW jako spójny, rzetelny i trafny strukturalnie wskaźnik poczucia własnej wartości uczniów.

4. Wielowymiarowa Skala Samooceny: Kondycja fizyczna (witalność) (WSS-KF)

1. Myślę, że mam bardzo dobrą kondycję fizyczną.
2. Po wysiłku szybko odzyskuję siły.
3. ~~Na co dzień mam bardzo dużo energii.~~
4. Gdy biegam lub ćwiczę, męczę się szybciej niż inni.
5. Jestem aktywny/aktywna fizycznie i dużo się ruszam.
6. Myślę, że jestem wysportowany/wysportowana.
7. Sądzę, że bardzo dobrze radzę sobie z ćwiczeniami ruchowymi, np. bieganiem, skakaniem, graniem w piłkę.

4.1. Analiza struktury czynnikowej

W pierwszym kroku analiz oszacowano jednoczynnikowy model CFA dla pełnej, 7-itemowej wersji skali kondycji fizycznej/witalności. Model ten nie osiągnął akceptowalnego dopasowania w zakresie RMSEA w żadnej z analizowanych prób. Wartości wskaźnika mieściły się w przedziale od 0.091 do 0.102, przekraczając przyjęty próg 0.08. Jednocześnie wartości CFI i TLI przekraczały 0.95, a wartości SRMR pozostawały poniżej 0.08. Podstawowym problemem modelu wyjściowego okazała się więc zbyt wysoka wartość RMSEA (zob. Tabela 5).

Tabela 5. WSS-KF: Wskaźniki dopasowania dla modelu CFA (model wyjściowy, 7-itemowy)

MODEL	χ^2	df	RMSEA [90% CI]	CFI	TLI	SRMR
Próba treningowa	119.57**	14	0.102 [0.085-0.119]	0.987	0.981	0.032
Próba testowa	99.10**	14	0.091 [0.075-0.108]	0.989	0.984	0.028
Próba pełna	209.43**	14	0.098 [0.086-0.110]	0.988	0.981	0.028
Próba pewna	161.57**	14	0.095 [0.082-0.108]	0.987	0.981	0.028

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; 90% CI = 90% przedział ufności dla RMSEA; CFI = porównawczy indeks dopasowania; TLI = indeks Tuckera–Lewisa; SRMR = standaryzowany pierwiastek średniokwadratowy reszt. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Analiza ładunków czynnikowych pokazała, że wszystkie pozycje były istotnie związane z czynnikiem latentnym. Zgodnie z przyjętą zasadą redukcji skali zachowano jedną pozycję sformułowaną negatywnie, czyli pozycję 4. Spośród pozycji pozytywnie sformułowanych

najslabiej wspólny czynnik reprezentowała pozycja 3, która uzyskała najniższe ładunki w każdej z prób. Dlatego w pierwszym kroku redukcji usunięto właśnie tę pozycję (zob. Tabela 6).

Tabela 6. WSS-KF: Standaryzowane ładunki czynnikowe dla modelu CFA (model wyjściowy, 7-itemowy)

Wskaźnik	Próba treningowa	Próba testowa	Próba pełna	Próba pewna
WSSKF_1	0.920**	0.916**	0.916**	0.919**
WSSKF_2	0.671**	0.673**	0.671**	0.667**
WSSKF_3	0.557**	0.569**	0.563**	0.544**
WSSKF_4	-0.586**	-0.562**	-0.573**	-0.645**
WSSKF_5	0.842**	0.841**	0.841**	0.833**
WSSKF_6	0.901**	0.924**	0.912**	0.912**
WSSKF_7	0.827**	0.860**	0.843**	0.841**

Uwaga. Wartości przedstawiają standaryzowane współczynniki STDYX. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Po usunięciu pozycji 3 oszacowano model 6-itemowy. Dopasowanie modelu uległo częściowej poprawie, zwłaszcza w próbie treningowej, w której RMSEA spadło poniżej progu 0.08. W pozostałych próbach wartości RMSEA nadal jednak przekraczały poziom akceptowalnego dopasowania. Pozostałe wskaźniki prezentowały się korzystnie: CFI i TLI pozostawały powyżej 0.95, a SRMR było wyraźnie niższe od 0.08. Wyniki te sugerowały, że redukcja o jedną pozycję była niewystarczająca (zob. Tabela 7).

Tabela 7. WSS-KF: Wskaźniki dopasowania dla modelu CFA (model po redukcji, 6-itemowy)

MODEL	χ^2	df	RMSEA [90% CI]	CFI	TLI	SRMR
Próba treningowa	46.11**	9	0.075 [0.054-0.097]	0.995	0.992	0.020
Próba testowa	76.55**	9	0.101 [0.081-0.123]	0.991	0.986	0.023
Próba pełna	113.62**	9	0.089 [0.075-0.104]	0.993	0.989	0.020
Próba pewna	109.95**	9	0.098 [0.082-0.115]	0.991	0.985	0.022

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; 90% CI = 90% przedział ufności dla RMSEA; CFI = porównawczy indeks dopasowania; TLI = indeks Tuckera–Lewisa; SRMR = standaryzowany pierwiastek średniokwadratowy reszt. Istotne statystycznie: * $p < .05$. ** $p < .01$.

W kolejnym kroku przeanalizowano ładunki czynnikowe modelu 6-itemowego. Ponieważ pozycja 4 pełniła funkcję jedyne go wskaźnika o negatywnym kierunku treściowym, nie była usuwana mimo relatywnie niższego ładunku (zob. Tabela 8).

Tabela 8. WSS-KF: Standaryzowane ładunki czynnikowe dla modelu CFA (model po redukcji, 6-itemowy)

Wskaźnik	Próba treningowa	Próba testowa	Próba pełna	Próba pewna
WSSKF_1	0.923**	0.918**	0.920**	0.922**
WSSKF_2	0.650**	0.666**	0.656**	0.653**
WSSKF_4	-0.594**	-0.570**	-0.583**	-0.653**
WSSKF_5	0.844**	0.838**	0.841**	0.833**
WSSKF_6	0.904**	0.925**	0.914**	0.913**
WSSKF_7	0.824**	0.855**	0.839**	0.837**

Uwaga. Wartości przedstawiają standaryzowane współczynniki STDYX. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Następnie oszacowano model zmodyfikowany, w którym uwzględniono korelację reszt między pozycjami WSSKF_4 i WSSKF_2. Decyzja ta była oparta na indeksach modyfikacyjnych modelu 6-itemowego, wskazujących na lokalną współzmiennność tej pary pozycji niewyjaśnioną w pełni przez ogólny czynnik WSS-KF. Ponadto przyjęta modyfikacja miała również uzasadnienie treściowe. Pozycja 2 odnosi się do szybkiego odzyskiwania sił po wysiłku, natomiast pozycja 4 dotyczy szybszego męczenia się podczas biegania lub ćwiczeń w porównaniu z innymi osobami. Obie pozycje odnoszą się więc do zbliżonego obszaru funkcjonowania fizycznego: subiektywnie ocenianej wydolności organizmu, odporności na wysiłek oraz tempa regeneracji po aktywności fizycznej. Prawdopodobnie to właśnie wspólne odniesienie obu pozycji do doświadczenia wysiłku fizycznego i zmęczenia generuje dodatkową współzmiennność, której nie wyjaśnia w pełni jednoczynnikowy model WSS-KF.

Po wprowadzeniu korelacji reszt między WSSKF_4 i WSSKF_2 model uzyskał bardzo dobre lub co najmniej akceptowalne dopasowanie do danych we wszystkich analizowanych próbach. W próbie treningowej RMSEA wyniosło 0.048, w próbie pełnej 0.044, a w próbie pewnej 0.038, czyli mieściło się poniżej progu 0.05. W próbie testowej wartość RMSEA była nieco wyższa i wyniosła 0.061, co wskazuje na dopasowanie z pogranicza dobrego i akceptowalnego. Równocześnie pozostałe wskaźniki jednoznacznie przemawiały za dobrym dopasowaniem modelu: wartości CFI mieściły się w zakresie od 0.997 do 0.999, wartości TLI od 0.995 do 0.998, a SRMR było bardzo niskie i nie przekraczało 0.012. Oznacza to, że model

jednoczynnikowy z dodatkową korelacją reszt między pozycjami WSSKF_4 i WSSKF_2 dobrze odtwarzał strukturę zależności między pozycjami skali (zob. Tabela 9).

Tabela 9. WSS-KF: Wskaźniki dopasowania dla modelu CFA uwzględniającego korelację reszt WSSKF_4 z WSSKF_2 (model zmodyfikowany, 6-itemowy)

MODEL	χ^2	df	RMSEA [90% CI]	CFI	TLI	SRMR
Próba treningowa	21.57**	8	0.048 [0.024-0.073]	0.998	0.997	0.012
Próba testowa	29.53**	8	0.061 [0.038-0.085]	0.997	0.995	0.012
Próba pełna	31.17**	8	0.044 [0.029-0.061]	0.999	0.997	0.009
Próba pewna	21.27**	8	0.038 [0.019-0.057]	0.999	0.998	0.008

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; 90% CI = 90% przedział ufności dla RMSEA; CFI = porównawczy indeks dopasowania; TLI = indeks Tuckera–Lewisa; SRMR = standaryzowany pierwiastek średniokwadratowy reszt. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Analiza standaryzowanych ładunków czynnikowych wskazuje, że wszystkie pozycje były istotnie związane z czynnikiem samooceny kondycji fizycznej i witalności. W próbie pełnej wartości ładunków mieściły się w zakresie od -0.559 do 0.921. Najsilniej z czynnikiem związane były pozycje WSSKF_1 oraz WSSKF_6, natomiast relatywnie naj słabszy bezwzględny ładunek uzyskała pozycja WSSKF_4. Korelacja reszt między WSSKF_4 i WSSKF_2 była istotna statystycznie we wszystkich analizowanych próbach. W próbie pełnej wyniosła -0.194, a w pozostałych próbach mieściła się w zakresie od -0.167 do -0.231. Wartości te wskazują na stabilną, choć niezbyt silną lokalną zależność między tymi dwiema pozycjami. Ujemny znak korelacji jest spójny z przeciwnym kierunkiem powiązania pozycji WSSKF_4 z czynnikiem samooceny kondycji fizycznej i witalności (zob. Tabela 10).

Tabela 10. WSS-KF: Standaryzowane ładunki czynnikowe dla modelu CFA z korelacją reszt WSSKF_4 z WSSKF_2 (model zmodyfikowany, 6-itemowy)

Wskaźnik	Próba treningowa	Próba testowa	Próba pełna	Próba pewna
WSSKF_1	0.924**	0.920**	0.921**	0.923**
WSSKF_2	0.636**	0.647**	0.639**	0.626**
WSSKF_4	-0.576**	-0.547**	-0.559**	-0.625**
WSSKF_5	0.846**	0.841**	0.843**	0.836**
WSSKF_6	0.905**	0.926**	0.916**	0.916**
WSSKF_7	0.825**	0.858**	0.841**	0.840**
Kowariancje reszt: WSSKF_4 z WSSKF_2	-0.167**	-0.216**	-0.194**	-0.231**

Uwaga. Wartości przedstawiają standaryzowane współczynniki STDYX. Wiersze w sekcji „Kowariancje reszt” odnoszą się do kowariancji między błędami pomiaru wskazanych pozycji. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Podsumowując, wyjściowy model 7-itemowy wymagał modyfikacji przede wszystkim ze względu na zbyt wysokie wartości RMSEA, wskazujące na niewystarczające dopasowanie modelu do danych. Etapowe usunięcie pozycji 3, a następnie uwzględnienie korelacji reszt między pozycjami WSSKF_4 i WSSKF_2, pozwoliło uzyskać lepiej dopasowaną wersję skali. Jednocześnie niska wartość korelacji reszt nie podważa jednoczynnikowej interpretacji narzędzia, lecz wskazuje na obecność niewielkiej lokalnej zależności między tymi dwiema pozycjami. Zależność ta najprawdopodobniej wynika ze wspólnego odniesienia obu pozycji do subiektywnej oceny męczenia się i regeneracji po wysiłku. Końcowy model 6-itemowy, obejmujący pozycje 1, 2, 4, 5, 6 i 7, spełnia przyjęte kryteria dopasowania i zachowuje zgodność z założeniami konstrukcyjnymi skali.

4.2. Analiza niezmienności pomiarowej

Analizę niezmienności pomiarowej WSS-KF przeprowadzono ze względu na płeć i klasę, porównując model konfiguralny z modelem skalarnym. W obu modelach uwzględniono kowariancję reszt między pozycjami WSSKF_2 i WSSKF_4, zgodnie z przyjętym modelem pomiarowym dla tej skali. W modelu konfiguralnym założono tę samą strukturę czynnikową

w grupie chłopców i dziewcząt, natomiast w modelu skalarnym wprowadzono dodatkowe ograniczenia dotyczące ładunków czynnikowych i progów odpowiedzi, konieczne do porównywania średnich latentnych. Wyniki analizy przedstawiono w Tabeli 11 i 12.

Niezmienność ze względu na płeć

Model konfiguralny uzyskał akceptowalne dopasowanie do danych, $\chi^2(16) = 67.91$, $p < .01$, RMSEA = 0.067, CFI = 0.997. Po nałożeniu ograniczeń skalarności otrzymano model o następujących parametrach: $\chi^2(32) = 83.60$, $p < .01$, RMSEA = 0.047, CFI = 0.997. Wprowadzenie ograniczeń nie pogorszyło dopasowania modelu: wartość RMSEA zmniejszyła się względem modelu konfiguralnego ($\Delta\text{RMSEA} = -0.020$), a CFI pozostał na tym samym poziomie ($\Delta\text{CFI} = 0.000$).

Tabela 11. WSS-KF: Wskaźniki dopasowania modeli niezmienności pomiaru ze względu na płeć

MODEL	Test chi-kwadrat dopasowania modelu ¹					RMSEA	Δ RMSEA	CFI	Δ CFI
	χ^2	df	$\Delta \chi^2$	Δ df	p				
Konfiguralny	67.91**	16	—	—	—	0.067	—	0.997	—
Skalarny	83.60**	32	36.613	16	.002	0.047	-0.020	0.997	0.000

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; CFI = porównawczy indeks dopasowania. Δ = różnica wartości χ^2 , RMSEA i CFI w odniesieniu do modelu mniej restrykcyjnego. W obu modelach uwzględniono kowariancję reszt pozycji WSSKF_2 i WSSKF_4. ¹ Wartości $\Delta \chi^2$ i Δ df obliczono z wykorzystaniem procedury DIFFTEST w pakiecie Mplus. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Zgodnie z przyjętymi kryteriami zmian wskaźników dopasowania, uzyskane wartości ΔRMSEA i ΔCFI przemawiają za niezmiennością skalarną pomiaru ze względu na płeć. Jednocześnie wynik procedury DIFFTEST był istotny statystycznie, $\Delta \chi^2(16) = 36.613$, $p = .002$, co wskazuje, że formalnie model skalarny różnił się od modelu konfiguralnego. Ponieważ jednak pogorszenie dopasowania nie znalazło potwierdzenia w kryteriach opartych na zmianach RMSEA i CFI, a model skalarny zachował bardzo dobre wartości bezwzględnych wskaźników dopasowania, wynik DIFFTEST nie stanowi wystarczającej podstawy do odrzucenia porównywalności pomiaru między grupami.

Na tej podstawie porównano średnie latentne między grupami. Uzyskany wynik wskazuje, że dziewczynki osiągały niższy poziom samooceny kondycji fizycznej / witalności niż chłopcy, stanowiący grupę referencyjną ($\Delta M = -0.714$, $SE = 0.132$, $p < .001$). Interpretując tę

różnicę, należy pamiętać, że wniosek o porównywalności pomiaru opiera się przede wszystkim na kryteriach praktycznych, tj. zmianach RMSEA i CFI, przy jednocześnie istotnym wyniku procedury DIFFTEST.

Niezmienność ze względu na klasę

Model konfiguralny był bardzo dobrze dopasowany do danych: $\chi^2(16) = 30.39$, $p = .016$, RMSEA = 0.035, CFI = 0.999. Po wprowadzeniu ograniczeń skalarności uzyskano model, który również zachował bardzo dobre parametry dopasowania: $\chi^2(32) = 68.64$, $p < .001$, RMSEA = 0.040, CFI = 0.998.

Zmiany wartości praktycznych wskaźników dopasowania były niewielkie. Wzrost RMSEA wyniósł 0.005, natomiast CFI obniżył się jedynie o 0.001 (Δ RMSEA = 0.005; Δ CFI = -0.001). Obie wartości mieszczą się w przyjętych granicach, co przemawia za porównywalnością pomiaru samooceny kondycji fizycznej / witalności między uczniami klas 4–5 i 6–8.

Tabela 12. WSS-KF: Wskaźniki dopasowania modeli niezmienności pomiaru ze względu na klasę

MODEL	Test chi-kwadrat dopasowania modelu ¹					RMSEA	Δ RMSEA	CFI	Δ CFI
	χ^2	df	$\Delta \chi^2$	Δdf	p				
Konfiguralny	30.39*	16	—	—	—	0.035	—	0.999	—
Skalarny	68.64**	32	39.899	16	< .001	0.040	0.005	0.998	-0.001

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; CFI = porównawczy indeks dopasowania. Δ = różnica wartości χ^2 , RMSEA i CFI w odniesieniu do modelu mniej restrykcyjnego. ¹ Wartości $\Delta \chi^2$ i Δdf obliczono z wykorzystaniem procedury DIFFTEST w pakiecie Mplus. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Jednocześnie procedura DIFFTEST wykazała istotne statystycznie pogorszenie dopasowania po nałożeniu ograniczeń skalarności, $\Delta \chi^2(16) = 39.90$, $p < .001$. Oznacza to, że przesłanki dotyczące pełnej niezmienności skalarnej nie są całkowicie spójne: kryteria oparte na zmianach RMSEA i CFI wspierają przyjęcie porównywalności pomiaru, podczas gdy wynik formalnego testu różnicy modeli wskazuje na potrzebę zachowania ostrożności przy interpretowaniu porównań średnich latentnych.

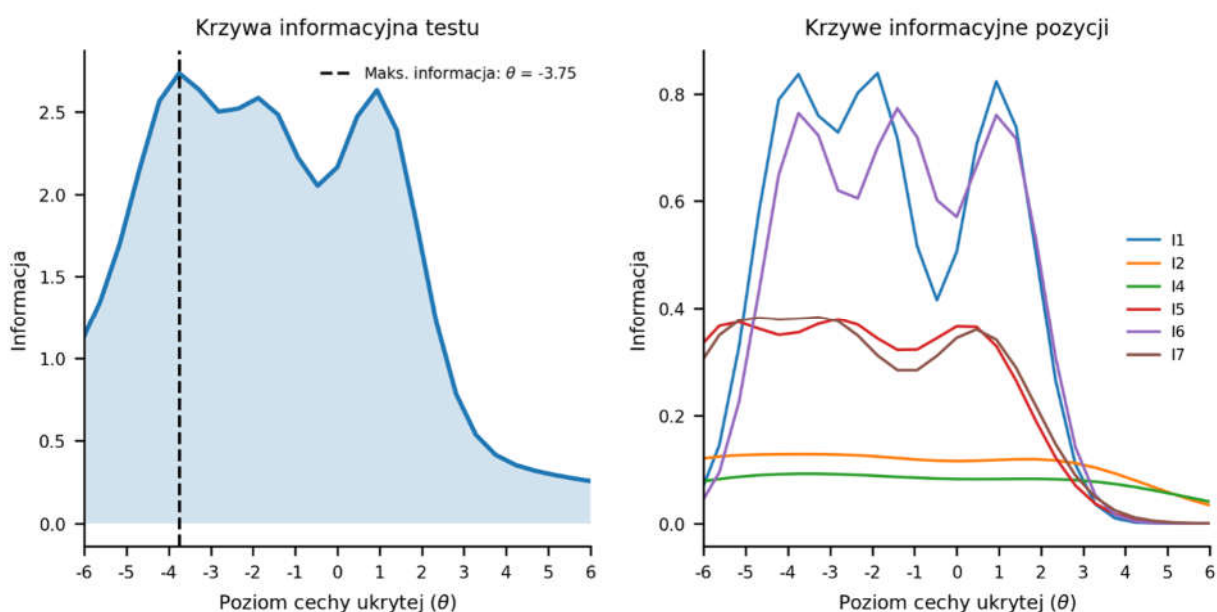
Z uwzględnieniem tego zastrzeżenia porównano średnie latentne. Uczniowie klas 6–8 uzyskali istotnie niższy poziom samooceny kondycji fizycznej / witalności niż uczniowie klas 4–5, stanowiący grupę referencyjną ($\Delta M = -0.629$, $SE = 0.131$, $p < .001$). Różnicę tę należy

interpretować biorąc pod uwagę niejednoznaczne przesłanki dotyczące pełnej niezmienności skalarnej.

4.3. Krzywa informacyjna testu i pozycji testowych

Skala kondycji fizycznej i witalności charakteryzuje się niższą informacyjnością niż większość pozostałych skal, co wskazuje na relatywnie słabszą precyzję pomiaru w ujęciu IRT. Najdokładniej mierzy ona bardzo niski poziom cechy ukrytej: maksimum informacji przypada około $\theta = -3.75$, a odpowiadająca mu rzetelność warunkowa wynosi około 0.73. Oznacza to, że skala przekracza próg użytecznej precyzji 0.70, ale nie osiąga poziomu 0.80, typowego dla bardziej precyzyjnego różnicowania indywidualnego (zob. Wykres 5).

Wykres 5. WSS-KF: Krzywa informacyjna testu oraz krzywe informacyjne pozycji dla kondycji fizycznej i witalności



Uwaga. Lewy panel przedstawia krzywą informacyjną całego testu, a prawy panel krzywe informacyjne poszczególnych pozycji. Oś X oznacza poziom cechy ukrytej θ , natomiast oś Y poziom informacji, czyli precyzję pomiaru na danym poziomie cechy. Wyższa wartość informacji oznacza większą dokładność pomiaru w danym zakresie θ . Linia przerywana w lewym panelu wskazuje poziom cechy, przy którym test osiąga maksymalną informację. Oznaczenia pozycji odnoszą się do kolejnych pozycji (itemów) wchodzących w skład analizowanego narzędzia.

Jednocześnie przy łagodniejszym kryterium, czyli rzetelności powyżej 0.60, skala zachowuje użyteczność w szerokim zakresie od około $\theta = -5.42$ do $\theta = 2.13$. W praktyce skala jest więc najbardziej użyteczna w przypadku uczniów deklarujących bardzo słabą kondycję fizyczną i niską witalność, ale dostarcza interpretowalnych wyników także przy przeciętnym

i nieco wyższym poziomie cechy. Największy wkład informacyjny mają pozycje I1 i I5, następnie I4 i I6, natomiast najmniej informacji wnoszą I2 oraz I3.

4.4. Miary rzetelności

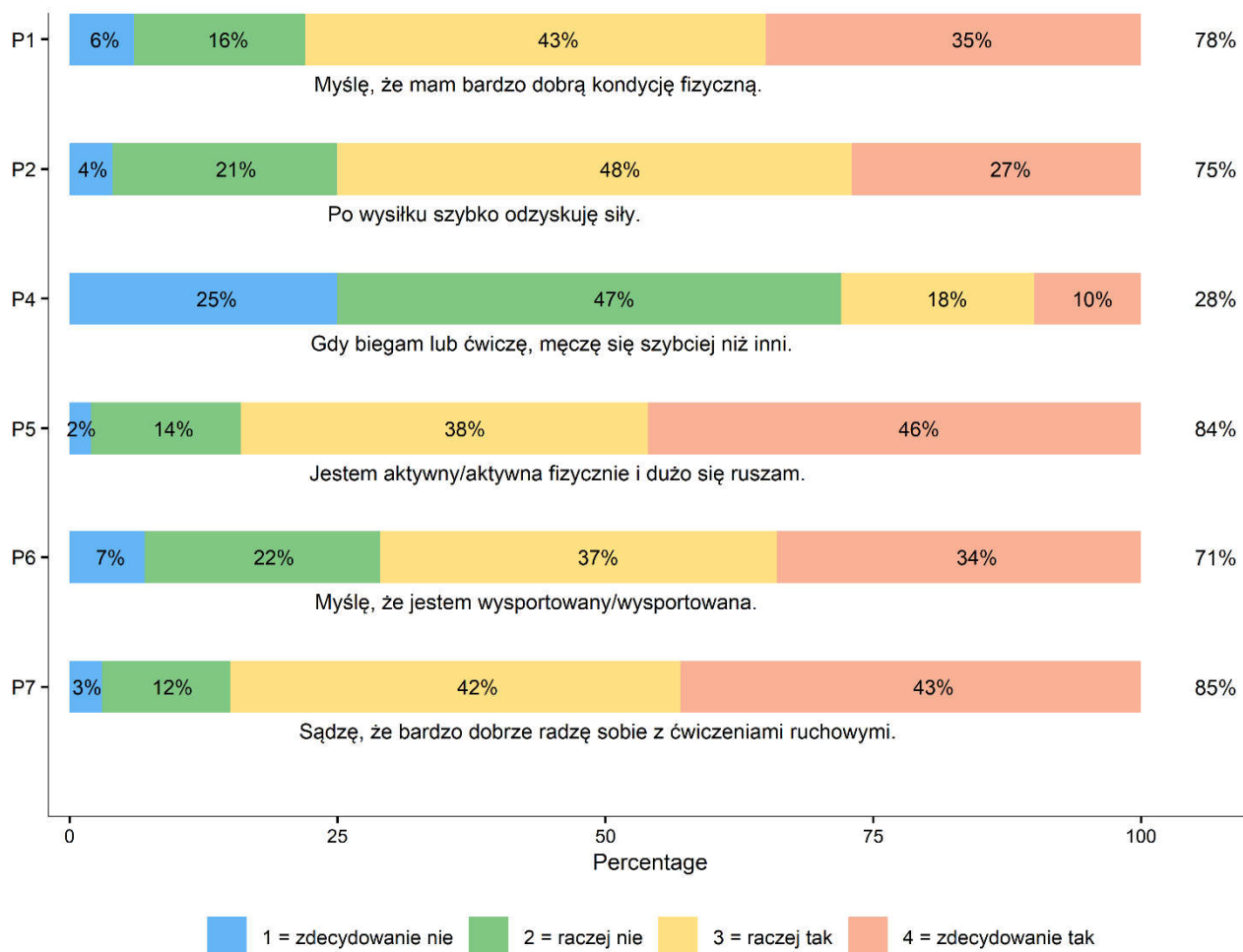
Ostateczna wersja skali samooceny kondycji fizycznej i witalności uzyskała wysokie wskaźniki rzetelności. Alfa porządkowa wyniosła $\alpha_{ord} = 0.908$, a współczynnik omega całkowita McDonalda osiągnął wartość $\omega_t = 0.882$. Oba współczynniki przekraczają wartość 0.80, co wskazuje na bardzo dobrą spójność wewnętrzną skali. Klasyczna alfa Cronbacha również była wysoka i wyniosła $\alpha = 0.871$. Średnia wariancja wyjaśniona (AVE) osiągnęła wartość 0.645.

4.5. Rozkład odpowiedzi na poszczególne twierdzenia skali

Rozkład odpowiedzi wskazuje, że uczniowie przeważnie pozytywnie oceniali własną kondycję fizyczną i witalność. Najwyższy odsetek odpowiedzi pozytywnych uzyskano dla pozycji P7 („Sądzę, że bardzo dobrze radzę sobie z ćwiczeniami ruchowymi”) – łącznie 85% badanych raczej lub zdecydowanie zgodziło się z tym stwierdzeniem. Bardzo wysoki poziom aprobaty dotyczył także pozycji P5 („Jestem aktywny/aktywna fizycznie i dużo się ruszam”) – 84% odpowiedzi pozytywnych. Nieco niższe, ale nadal wysokie odsetki zgody odnotowano dla pozycji P1 („Myślę, że mam bardzo dobrą kondycję fizyczną”) – 78%, P2 („Po wysiłku szybko odzyskuję siły”) – 75%, oraz P6 („Myślę, że jestem wysportowany/wysportowana”) – 71%.

Pozycja P4 („Gdy biegam lub ćwiczę, męczę się szybciej niż inni”) miała charakter negatywny. Z tym stwierdzeniem raczej lub zdecydowanie zgodziło się 28% uczniów, natomiast 72% nie zgodziło się z nim. Wynik ten jest spójny z ogólnie pozytywnym profilem odpowiedzi i wskazuje, że większość badanych nie postrzega siebie jako osób szybciej męczących się podczas aktywności fizycznej niż rówieśnicy. Podsumowując, uczniowie najczęściej deklarowali dobrą sprawność ruchową, aktywność fizyczną oraz pozytywną ocenę własnej kondycji, choć poziom samoidentyfikacji jako osoby wysportowanej był nieco mniej jednoznaczny (zob. Wykres 6).

Wykres 6. WSS-KF: Procentowy rozkład odpowiedzi na poszczególne twierdzenia skali

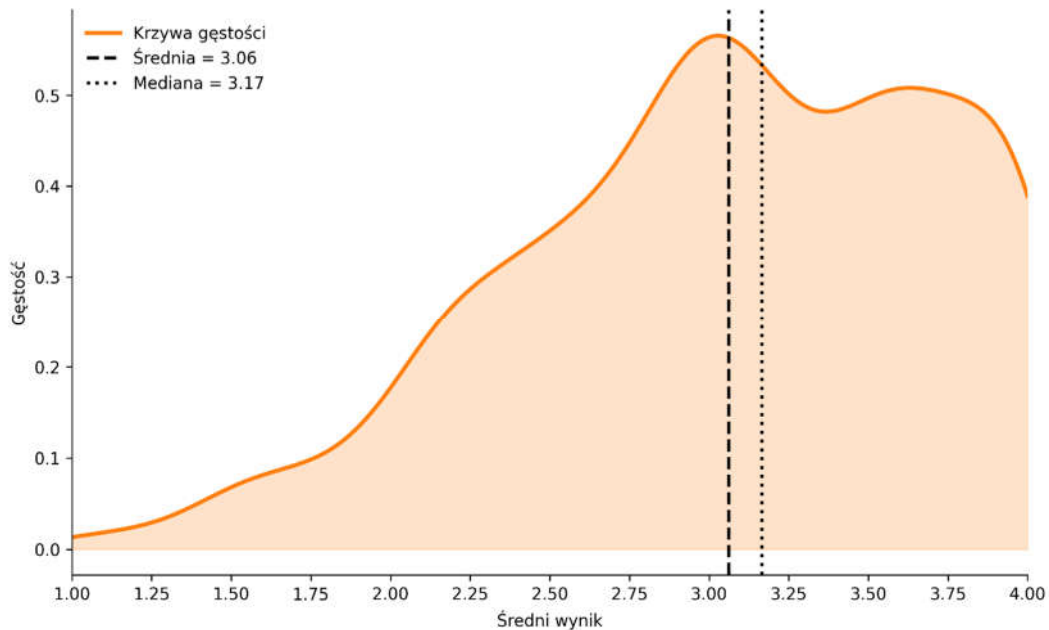


Uwaga. Wartości po prawej stronie wykresu oznaczają łączny odsetek odpowiedzi „raczej tak” i „zdecydowanie tak”. W przypadku pozycji P4, sformułowanej negatywnie, wartość ta odnosi się do zgody ze stwierdzeniem o szybszym męczeniu się podczas biegania lub ćwiczeń.

4.6. Charakterystyka rozkładu wyników (średnia punktów) skali

Ocena kondycji fizycznej i witalności była przeciętnie dość korzystna. Średnia wyniosła $M = 3.06$ przy $SD = 0.66$, a mediana osiągnęła $Me = 3.17$. Wyniki rozciągały się od 1 do 4, przy punkcie środkowym równym 2.50. Ponieważ wyższe wartości oznaczają lepszą ocenę własnej kondycji i witalności, średnia powyżej środka skali wskazuje, że uczniowie generalnie oceniali ten obszar pozytywnie. Rozkład był lekko przesunięty w stronę wyższych wartości, o czym świadczy ujemna skośność wynosząca -0.50. Kurtoza równa -0.34 sugeruje natomiast nieco spłaszczony kształt rozkładu (zob. Wykres 7).

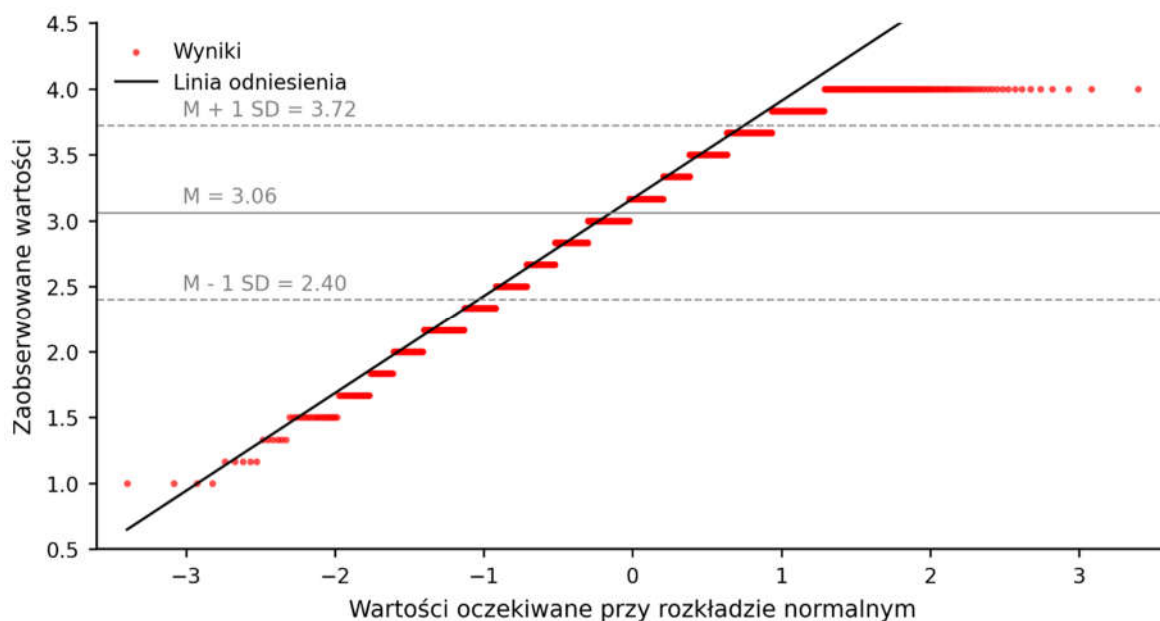
Wykres 7. WSS-KF: Rozkład średnich wyników kondycji fizycznej i witalności



Uwaga. Na wykresie przedstawiono wygładzoną krzywą gęstości średnich wyników. Linia przerywana wyznacza średnią arytmetyczną (M), a linia kropkowana medianę (Me). Wysokość krzywej odzwierciedla względne zagęszczenie obserwacji w poszczególnych zakresach wartości.

Wykres Q–Q doprecyzowuje ten obraz: w centralnej części punkty przebiegają blisko linii odniesienia, co oznacza, że większość wyników układu się względnie regularnie (por. Wykres 8). Odchylenia są bardziej widoczne na krańcach, zwłaszcza przy najwyższych wartościach, gdzie obserwacje tworzą wyraźne wypłaszczenie. Wskazuje to, że część uczniów skupiała się przy bardzo korzystnych ocenach własnej kondycji i witalności, ale bez wyraźnej skrajnej deformacji całego rozkładu.

Wykres 8. WSS-KF: Wykres Q-Q średnich wyników kondycji fizycznej i witalności względem rozkładu normalnego



Uwaga. Wykres przedstawia porównanie kwantylów empirycznych z kwantylami teoretycznymi rozkładu normalnego. Czerwone punkty oznaczają zaobserwowane wartości, a czarna linia pełni funkcję linii odniesienia dla rozkładu normalnego. Odchylenia punktów od tej linii wskazują na odstępstwa rozkładu empirycznego od normalności. Pozioma linia ciągła oznacza średnią arytmetyczną wyniku, oznaczoną jako M . Linie przerywane wyznaczają wartości o jedno odchylenie standardowe poniżej i powyżej średniej, czyli $M - 1 SD$ oraz $M + 1 SD$, gdzie SD oznacza odchylenie standardowe.

4.7. Wnioski

Wyniki przeprowadzonych analiz przemawiają za przyjęciem 6-itemowej wersji WSS-KF jako psychometrycznie uzasadnionej miary kondycji fizycznej i witalności. Model CFA dobrze odtwarzał strukturę danych po uwzględnieniu niewielkiej lokalnej zależności między dwiema treściowo zbliżonymi pozycjami. Skala cechowała się bardzo dobrą rzetelnością, choć największą precyzję osiągała przede wszystkim przy niskim poziomie ocenianej kondycji i witalności. Uzyskane wyniki wspierają wykorzystanie WSS-KF jako spójnego i rzetelnego narzędzia, przy zachowaniu świadomości nierównomiernej precyzji pomiaru w różnych zakresach cechy.

5. Wielowymiarowa Skala Samooceny: Ogólna samoskuteczność (WSS-OS)

5.1. Analiza struktury czynnikowej

1. ~~Zwykle wierzę we własne możliwości.~~
2. Nawet gdy wszystko idzie źle, zwykle znajduję sposób, żeby sobie poradzić.
3. Kiedy coś mnie przerasta, szybko się poddaję.
4. ~~Przeważnie osiągam to, co zaplanowałem/zaplanowałam.~~
5. Potrafię rozwiązać prawie każdy nowy problem.
6. Moje umiejętności pomagają mi radzić sobie w trudnych sytuacjach.
7. Zwykle potrafię znaleźć wyjście nawet z trudnej sytuacji.
8. Zawsze dobrze radzę sobie z różnymi problemami.
9. ~~Czuję, że potrafię osiągnąć to, na czym mi zależy.~~

W pierwszej kolejności zweryfikowano jednoczynnikowy model CFA dla wyjściowej, 9-itemowej wersji skali ogólnej samoskuteczności. Uzyskane wyniki wskazywały, że pełny zestaw pozycji nie zapewniał w pełni satysfakcjonującego dopasowania. Wartości RMSEA w dwóch próbach przekraczały próg 0.08, a w próbie pełnej wartość SRMR była nieznacznie wyższa od przyjętego kryterium 0.08. Należy jednocześnie dodać, że wartości CFI i TLI w każdej z prób przekraczały próg 0.95 (zob. Tabela 13).

Tabela 13. WSS-OS: Wskaźniki dopasowania CFA (model wyjściowy, 9-itemowy)

MODEL	χ^2	df	RMSEA [90% CI]	CFI	TLI	SRMR
Próba treningowa	168.59**	27	0.085 [0.073-0.097]	0.977	0.970	0.031
Próba testowa	128.51**	27	0.072 [0.059-0.084]	0.981	0.974	0.028
Próba pełna	491.05**	36	0.093 [0.086-0.100]	0.960	0.950	0.084
Próba pewna	178.37**	27	0.069 [0.060-0.079]	0.984	0.978	0.027

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; 90% CI = 90% przedział ufności dla RMSEA; CFI = porównawczy indeks dopasowania; TLI = indeks Tuckera–Lewisa; SRMR = standaryzowany pierwiastek średniokwadratowy reszt. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Analiza ładunków czynnikowych wskazała, że wszystkie pozycje były istotnie powiązane z czynnikiem latentnym, ale ich siła była nierówna. Przy redukcji skali kierowano się zasadą usuwania pozycji najslabiej reprezentujących wspólny czynnik, przy jednoczesnym

zachowaniu jednej pozycji sformułowanej negatywnie, zgodnie z założeniami konstrukcyjnymi skali. Z tego względu nie usuwano pozycji 3, mimo że miała ona niższy bezwzględny ładunek niż część pozostałych pozycji. W pierwszym kroku redukcji usunięto pozycję 4, która spośród pozycji sformułowanych pozytywnie miała najniższe ładunki czynnikowe (zob. Tabele 14–15).

Tabela 14. WSS-OS: Ładunki czynnikowe CFA (model wyjściowy, 9-itemowy)

Wskaźnik	Próba treningowa	Próba testowa	Próba pełna	Próba pewna
WSSOS_1	0.731**	0.724**	0.724**	0.726**
WSSOS_2	0.701**	0.728**	0.713**	0.729**
WSSOS_3	-0.533**	-0.511**	-0.521**	-0.615**
WSSOS_4	0.674**	0.659**	0.665**	0.679**
WSSOS_5	0.770**	0.720**	0.747**	0.748**
WSSOS_6	0.791**	0.780**	0.786**	0.773**
WSSOS_7	0.825**	0.826**	0.831**	0.813**
WSSOS_8	0.823**	0.832**	0.828**	0.823**
WSSOS_9	0.723**	0.724**	0.717**	0.729**

Uwaga. Wartości przedstawiają standaryzowane współczynniki STDYX. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Tabela 15. WSS-OS: Wskaźniki dopasowania CFA (model po redukcji, 8-itemowy)

MODEL	χ^2	df	RMSEA [90% CI]	CFI	TLI	SRMR
Próba treningowa	97.04**	20	0.073 [0.059-0.087]	0.987	0.981	0.024
Próba testowa	124.14**	20	0.084 [0.070-0.099]	0.978	0.970	0.028
Próba pełna	181.47**	20	0.074 [0.065-0.084]	0.985	0.978	0.025
Próba pewna	128.92**	20	0.068 [0.057-0.080]	0.987	0.982	0.024

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; 90% CI = 90% przedział ufności dla RMSEA; CFI = porównawczy indeks dopasowania; TLI = indeks Tuckera–Lewisa; SRMR = standaryzowany pierwiastek średniokwadratowy reszt. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Po usunięciu pozycji 4 oszacowano model 8-itemowy. Rozwiązanie to przyniosło częściową poprawę parametrów dopasowania, zwłaszcza w zakresie SRMR, którego wartości we wszystkich próbach znalazły się poniżej progu 0.08. Również wartości CFI i TLI nadal przekraczały próg 0.95. Wartości RMSEA pozostawały jednak w części prób podwyższone,

w tym w próbie testowej przekraczały poziom 0.08. Z tego względu zdecydowano się na dalszą redukcję skali (zob. Tabela 16).

Tabela 16. WSS-OS: Ładunki czynnikowe CFA (model po redukcji, 8-itemowy)

Wskaźnik	Próba treningowa	Próba testowa	Próba pełna	Próba pewna
WSSOS_1	0.729**	0.725**	0.723**	0.725**
WSSOS_2	0.711**	0.728**	0.717**	0.735**
WSSOS_3	-0.536**	-0.519**	-0.526**	-0.620**
WSSOS_5	0.770**	0.721**	0.747**	0.749**
WSSOS_6	0.792**	0.777**	0.785**	0.773**
WSSOS_7	0.834**	0.827**	0.836**	0.818**
WSSOS_8	0.827**	0.834**	0.830**	0.825**
WSSOS_9	0.696**	0.715**	0.699**	0.706**

Uwaga. Wartości przedstawiają standaryzowane współczynniki STDYX. Istotne statystycznie: * $p < .05$. ** $p < .01$.

W kolejnym kroku usunięto pozycję 9. Decyzja ta wynikała z relatywnie niższych ładunków tej pozycji w modelu 8-itemowym. Model 7-itemowy uzyskał lepsze parametry dopasowania niż rozwiązanie 8-itemowe, zwłaszcza w próbie treningowej, pełnej i pewnej. Nadal jednak w próbie testowej RMSEA przekraczało próg 0.08, co wskazywało, że rozwiązanie to nie zapewniało jeszcze w pełni stabilnego dopasowania we wszystkich próbach (zob. Tabele 17–18).

Tabela 17. WSS-OS: Wskaźniki dopasowania CFA (model po redukcji, 7-itemowy)

MODEL	χ^2	df	RMSEA [90% CI]	CFI	TLI	SRMR
Próba treningowa	30.85**	14	0.041 [0.021-0.060]	0.997	0.995	0.016
Próba testowa	87.62**	14	0.085 [0.068-0.102]	0.984	0.975	0.025
Próba pełna	98.63**	14	0.064 [0.053-0.077]	0.991	0.987	0.020
Próba pewna	56.54**	14	0.051 [0.038-0.065]	0.995	0.992	0.017

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; 90% CI = 90% przedział ufności dla RMSEA; CFI = porównawczy indeks dopasowania; TLI = indeks Tuckera–Lewisa; SRMR = standaryzowany pierwiastek średniokwadratowy reszt. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Tabela 18. WSS-OS: Ładunki czynnikowe CFA (model po redukcji, 7-itemowy)

Wskaźnik	Próba treningowa	Próba testowa	Próba pełna	Próba pewna
WSSOS_1	0.690**	0.702**	0.696**	0.698**
WSSOS_2	0.717**	0.728**	0.718**	0.739**
WSSOS_3	-0.531**	-0.519**	-0.523**	-0.619**
WSSOS_5	0.769**	0.728**	0.751**	0.754**
WSSOS_6	0.794**	0.777**	0.785**	0.772**
WSSOS_7	0.848**	0.833**	0.843**	0.823**
WSSOS_8	0.830**	0.838**	0.833**	0.829**

Uwaga. Wartości przedstawiają standaryzowane współczynniki STDYX. Istotne statystycznie: * $p < .05$. ** $p < .01$.

W kolejnym kroku usunięto pozycję 1, która w modelu 7-itemowym miała najniższe ładunki spośród pozycji pozytywnie sformułowanych.

Po tej redukcji model 6-itemowy uzyskał korzystny profil dopasowania. Wartości RMSEA mieściły się w przedziale od 0.023 do 0.054, wskazując na dobre lub akceptowalne dopasowanie we wszystkich analizowanych próbach. Jednocześnie wartości CFI i TLI przekraczały próg 0.95, a wartości SRMR pozostawały wyraźnie poniżej 0.08 (zob. Tabela 19).

Tabela 19. WSS-OS: Wskaźniki dopasowania CFA (model po redukcji, 6-itemowy)

MODEL	χ^2	df	RMSEA [90% CI]	CFI	TLI	SRMR
Próba treningowa	12.38	9	0.023 [0.000-0.051]	0.999	0.999	0.010
Próba testowa	28.39**	9	0.054 [0.032-0.077]	0.995	0.991	0.017
Próba pełna	30.97**	9	0.041 [0.026-0.057]	0.998	0.996	0.013
Próba pewna	32.62**	9	0.047 [0.031-0.065]	0.997	0.994	0.013

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; 90% CI = 90% przedział ufności dla RMSEA; CFI = porównawczy indeks dopasowania; TLI = indeks Tuckera–Lewisa; SRMR = standaryzowany pierwiastek średniokwadratowy reszt. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Analiza ładunków czynnikowych dla końcowego modelu potwierdziła, że pozostawione pozycje dobrze reprezentują wspólny konstrukt. W próbie pełnej bezwzględne wartości ładunków mieściły się w zakresie od 0.520 do 0.851. Najsilniej z czynnikiem związane były pozycje 7 i 8, natomiast najniższy, choć nadal istotny, związek uzyskała pozycja 3. Ujemny

znak ładunku tej pozycji wynikał z jej odwróconego kierunku treściowego. Jej pozostawienie było zgodne z założeniem, że każda skala powinna zawierać jeden wskaźnik o negatywnym kierunku treściowym (zob. Tabela 20).

Tabela 20. WSS-OS: Ładunki czynnikowe CFA (model po redukcji, 6-itemowy)

Wskaźnik	Próba treningowa	Próba testowa	Próba pełna	Próba pewna
WSSOS_2	0.701**	0.697**	0.698**	0.728**
WSSOS_3	-0.530**	-0.514**	-0.520**	-0.614**
WSSOS_5	0.774**	0.739**	0.757**	0.761**
WSSOS_6	0.791**	0.775**	0.783**	0.770**
WSSOS_7	0.853**	0.850**	0.851**	0.828**
WSSOS_8	0.832**	0.838**	0.835**	0.830**

Uwaga. Wartości przedstawiają standaryzowane współczynniki STDYX. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Podsumowując, redukcję skali przeprowadzono etapowo, usuwając pozycje najstabiliej reprezentujące czynnik spośród pozycji pozytywnie sformułowanych, przy zachowaniu jednej pozycji odwróconej. Najlepsze parametry uzyskał model 6-itemowy, obejmujący pozycje 2, 3, 5, 6, 7 i 8. Wyniki CFA wskazują, że jest to rozwiązanie bardziej ekonomiczne i lepiej dopasowane do danych niż wersja wyjściowa, a jednocześnie zgodne z założeniami konstrukcyjnymi skali.

5.2. Analiza niezmienności pomiarowej

Analizę niezmienności pomiarowej skali WSS-OS przeprowadzono ze względu na płeć i klasę, porównując model konfiguralny z modelem skalarnym. W modelu konfiguralnym założono tę samą strukturę czynnikową w grupie chłopców i dziewcząt, natomiast w modelu skalarnym wprowadzono dodatkowe ograniczenia dotyczące ładunków czynnikowych i progów odpowiedzi, konieczne do porównywania średnich latentnych. Wyniki analizy przedstawiono w Tabeli 21 i 22.

Niezmiennność ze względu na płeć

Model konfiguralny charakteryzował się dobrym dopasowaniem do danych ($\chi^2(18) = 50.71$, $p < .01$, RMSEA = 0.050, CFI = 0.996), co wskazuje, że przyjęta struktura

czynnikowa była zasadniczo zgodna z danymi w obu grupach porównywanych ze względu na płeć. Po nałożeniu ograniczeń skalarności uzyskano model o nadal dobrym dopasowaniu: $\chi^2(34) = 100.61$, $p < .01$, RMSEA = 0.052, CFI = 0.992. W porównaniu z modelem konfiguralnym odnotowano jedynie minimalny wzrost RMSEA (Δ RMSEA = 0.002) oraz niewielki spadek CFI (Δ CFI = -0.004); obie zmiany mieściły się w przyjętych granicach interpretacyjnych.

Tabela 21. WSS-OS: Wskaźniki dopasowania modeli niezmienności pomiaru ze względu na płeć

MODEL	Test chi-kwadrat dopasowania modelu ¹					RMSEA	Δ RMSEA	CFI	Δ CFI
	χ^2	df	$\Delta \chi^2$	Δdf	p				
Konfiguralny	50.71**	18	—	—	—	0.050	—	0.996	—
Skalarny	100.61**	34	53.655	16	<.001	0.052	0.002	0.992	-0.004

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; CFI = porównawczy indeks dopasowania. Δ = różnica wartości χ^2 , RMSEA i CFI w odniesieniu do modelu mniej restrykcyjnego. ¹ Wartości $\Delta \chi^2$ i Δdf obliczono z wykorzystaniem procedury DIFFTEST w pakiecie Mplus. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Zgodnie z kryteriami opartymi na zmianach wskaźników dopasowania, uzyskane wartości Δ RMSEA i Δ CFI przemawiają za niezmiennością skalarną pomiaru ze względu na płeć. Jednocześnie wynik procedury DIFFTEST był istotny statystycznie, $\Delta \chi^2(16) = 53.655$, $p < .001$, co wskazuje, że formalnie model skalarny różnił się od modelu konfiguralnego. Ponieważ jednak pogorszenie dopasowania było bardzo niewielkie i nie przekroczyło przyjętych progów dla Δ RMSEA i Δ CFI, a bezwzględne wartości wskaźników dopasowania modelu skalarnego pozostały wysokie, wynik DIFFTEST nie stanowi wystarczającej podstawy do odrzucenia porównywalności pomiaru między grupami.

Na tej podstawie porównano średnie latentne między grupami. Uzyskany wynik wskazuje, że dziewczynki osiągały niższy poziom ogólnej samoskuteczności niż chłopcy, stanowiący grupę referencyjną ($\Delta M = -0.419$, $SE = 0.065$, $p < .001$). Interpretując tę różnicę, należy pamiętać, że wniosek o porównywalności pomiaru opiera się przede wszystkim na kryteriach praktycznych, tj. zmianach RMSEA i CFI, przy jednocześnie istotnym wyniku procedury DIFFTEST.

Niezmienność ze względu na klasę

Model konfiguralny wykazał bardzo dobre dopasowanie do danych: $\chi^2(18) = 37.34$, $p = .005$, RMSEA = 0.038, CFI = 0.998. Po nałożeniu ograniczeń skalarności również uzyskano bardzo dobre dopasowanie modelu: $\chi^2(34) = 77.14$, $p < .001$, RMSEA = 0.042, CFI = 0.996.

Wprowadzenie ograniczeń równości wiązało się jedynie z niewielkimi zmianami praktycznych wskaźników dopasowania. RMSEA wzrosło o 0.004, natomiast CFI obniżyło się o 0.002 (Δ RMSEA = 0.004; Δ CFI = -0.002). Obie zmiany pozostawały wyraźnie w przyjętych granicach, co przemawia za porównywalnością pomiaru ogólnej samoskuteczności między uczniami klas 4–5 i 6–8.

Niemniej jednak formalny test różnicy modeli obliczony w ramach procedury DIFFTEST był istotny statystycznie, $\Delta\chi^2(16) = 42.40$, $p < .001$. Oznacza to, że dowody na rzecz pełnej niezmienności skalarnej nie są całkowicie jednoznaczne. Zmiany RMSEA i CFI wspierają założenie porównywalności pomiaru, jednak istotny wynik DIFFTEST wskazuje na potrzebę ostrożności przy interpretowaniu różnic średnich latentnych.

Z uwzględnieniem tego zastrzeżenia porównano średnie latentne w obu grupach. Uczniowie klas 6–8 osiągnęli istotnie niższy poziom ogólnej samoskuteczności niż uczniowie klas 4–5, przyjęci jako grupa referencyjna ($\Delta M = -0.170$, $SE = 0.057$, $p = .003$). Wynik ten należy interpretować z uwzględnieniem niejednoznacznych przesłanek dotyczących pełnej niezmienności skalarnej.

Tabela 22. WSS-OS: Wskaźniki dopasowania modeli niezmienności pomiaru ze względu na klasę

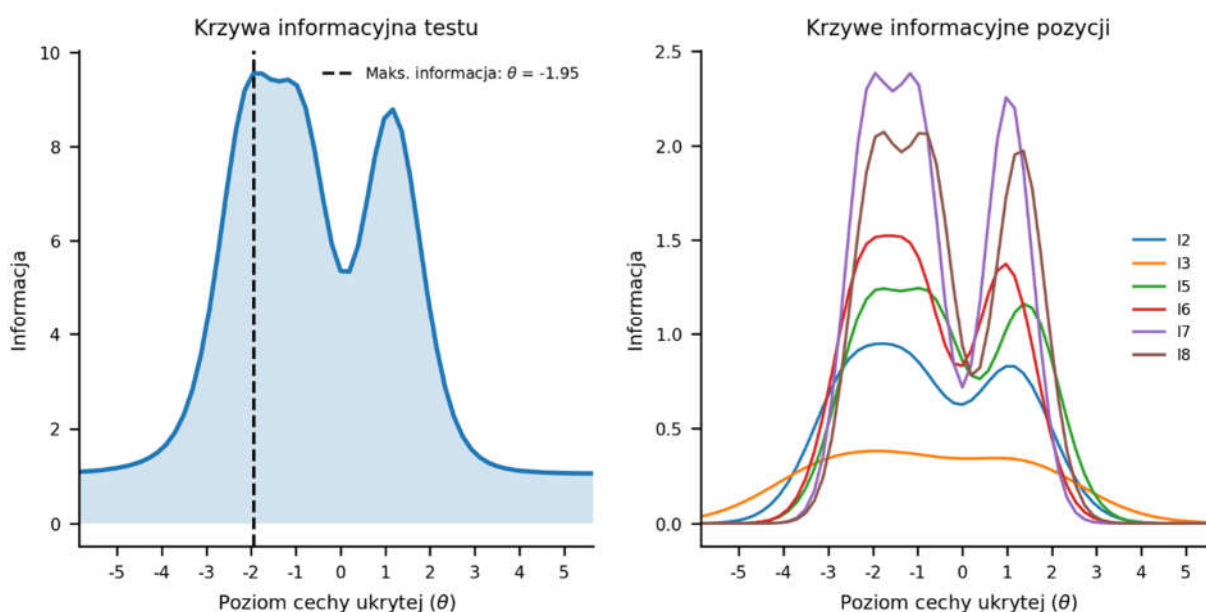
MODEL	Test chi-kwadrat dopasowania modelu ¹					RMSEA	Δ RMSEA	CFI	Δ CFI
	χ^2	df	$\Delta \chi^2$	Δ df	p				
Konfiguralny	37.34**	18	—	—	—	0.038	—	0.998	—
Skalarny	77.14**	34	42.404	16	< .001	0.042	0.004	0.996	-0.002

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; CFI = porównawczy indeks dopasowania. Δ = różnica wartości χ^2 , RMSEA i CFI w odniesieniu do modelu mniej restrykcyjnego. ¹ Wartości $\Delta\chi^2$ i Δdf obliczono z wykorzystaniem procedury DIFFTEST w pakiecie Mplus. Istotne statystycznie: * $p < .05$. ** $p < .01$.

5.3. Krzywa informacyjna testu i pozycji testowych

Skala ogólnej samoskuteczności charakteryzuje się wysoką precyzją pomiaru w szerokim zakresie cechy. Najdokładniej mierzy obniżony poziom samoskuteczności: maksimum krzywej informacyjnej przypada około $\theta = -1.95$, a rzetelność warunkowa osiąga w tym punkcie około 0.91. Wysoka precyzja pomiaru utrzymuje się jednak nie tylko przy niskich wynikach. Rzetelność co najmniej 0.80 obejmuje zakres od $\theta = -3.03$ do $\theta = 2.09$, a po obniżeniu kryterium do 0.70 — od $\theta = -3.49$ do $\theta = 2.51$. Oznacza to, że skala dobrze różnicuje uczniów o niskiej, przeciętnej i umiarkowanie wysokiej samoskuteczności. Największe znaczenie dla precyzji pomiaru mają pozycje I7, I8 i I6, natomiast pozycja I3 odgrywa pod tym względem wyraźnie mniejszą rolę (zob. Wykres 9).

Wykres 9. WSS-OS: Krzywa informacyjna testu oraz krzywe informacyjne pozycji dla ogólnej samoskuteczności



Uwaga. Lewy panel przedstawia krzywą informacyjną całego testu, a prawy panel krzywe informacyjne poszczególnych pozycji. Oś X oznacza poziom cechy ukrytej θ , natomiast oś Y poziom informacji, czyli precyzję pomiaru na danym poziomie cechy. Wyższa wartość informacji oznacza większą dokładność pomiaru w danym zakresie θ . Linia przerywana w lewym panelu wskazuje poziom cechy, przy którym test osiąga maksymalną informację. Oznaczenia pozycji odnoszą się do kolejnych pozycji (itemów) wchodzących w skład analizowanego narzędzia.

5.4. Miary rzetelności

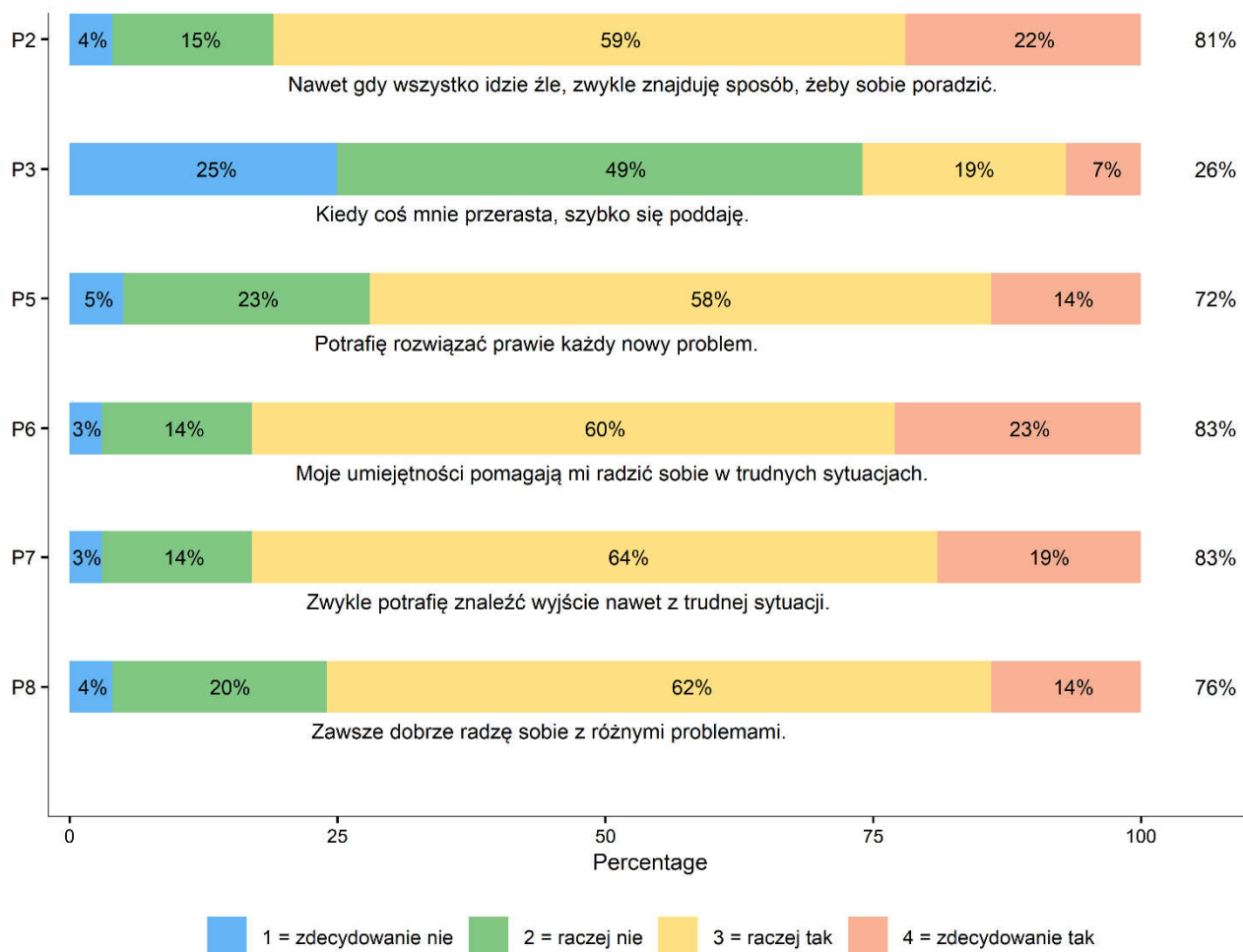
Rzetelność skali ogólnej samoskuteczności można uznać za wysoką. Alfa porządkowa, obliczona z uwzględnieniem porządkowego charakteru odpowiedzi, wyniosła $\alpha_{ord} = 0.878$. Zbliżoną wartość uzyskano dla współczynnika omega całkowita McDonalda, który wyniósł $\omega_t = 0.828$. Klasyczna alfa Cronbacha osiągnęła wartość $\alpha = 0.834$. Łącznie wskaźniki te wskazują na dobrą spójność wewnętrzną skali. Średnia wariancja wyjaśniona (AVA) wyniosła 0.561.

5.5. Rozkład odpowiedzi na poszczególne twierdzenia skali

Rozkład odpowiedzi wskazuje, że uczniowie najczęściej deklarowali umiarkowanie wysokie poczucie ogólnej samoskuteczności. Najwyższy odsetek odpowiedzi pozytywnych uzyskano dla pozycji P6 („Moje umiejętności pomagają mi radzić sobie w trudnych sytuacjach”) oraz P7 („Zwykle potrafię znaleźć wyjście nawet z trudnej sytuacji”) – w obu przypadkach 83% badanych raczej lub zdecydowanie zgodziło się z tymi stwierdzeniami. Wysoki poziom aprobaty dotyczył także pozycji P2 („Nawet gdy wszystko idzie źle, zwykle znajduję sposób, żeby sobie poradzić”) – 81%, P8 („Zawsze dobrze radzę sobie z różnymi problemami”) – 76%, oraz P5 („Potrafię rozwiązać prawie każdy nowy problem”) – 72%.

Odmiennej kierunek interpretacji dotyczy pozycji P3 („Kiedy coś mnie przerasta, szybko się poddam”), która była sformułowana negatywnie. Z tym twierdzeniem raczej lub zdecydowanie zgodziło się 26% uczniów, natomiast 74% nie zgodziło się z nim. Oznacza to, że większość badanych nie deklarowała szybkiego poddawania się w sytuacjach trudnych. Podsumowując, uczniowie najczęściej wskazywali, że potrafią radzić sobie z trudnościami, znajdować wyjście z problemów oraz wykorzystywać własne umiejętności w wymagających sytuacjach, choć przekonanie o rozwiązywaniu „prawie każdego” nowego problemu było nieco mniej jednoznaczne (zob. Wykres 10).

Wykres 10. WSS-OS: Procentowy rozkład odpowiedzi na poszczególne twierdzenia skali

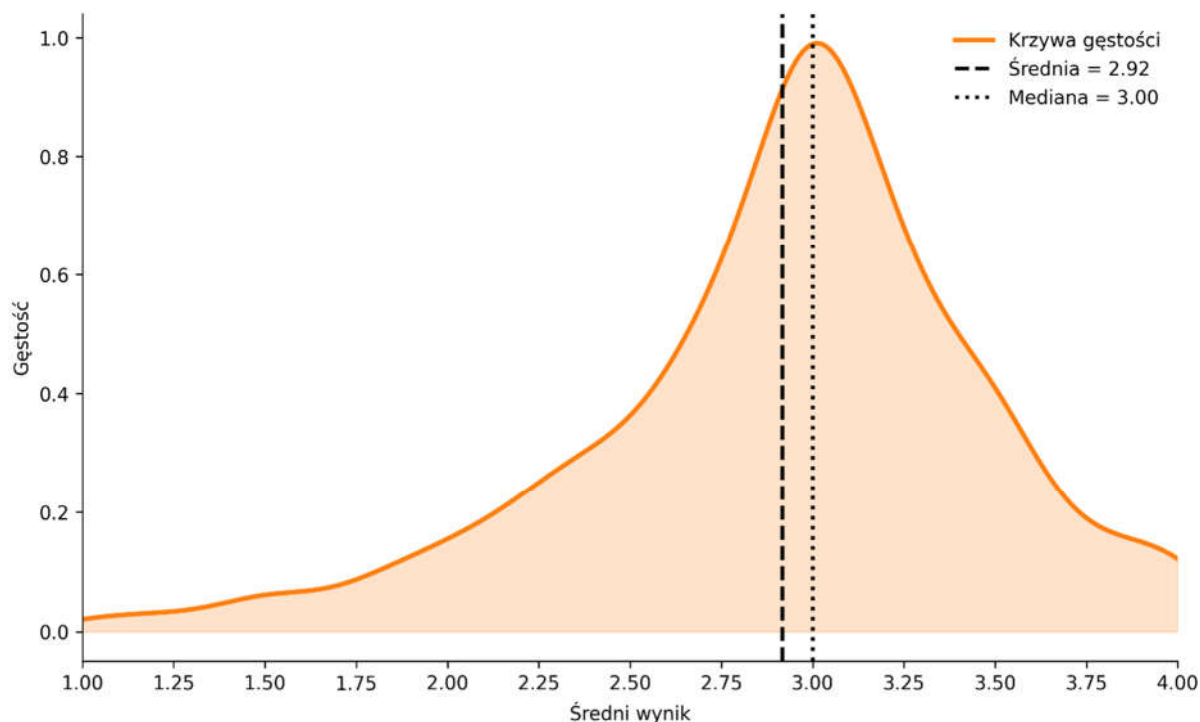


Uwaga. Wartości po prawej stronie wykresu oznaczają łączny odsetek odpowiedzi „raczej tak” i „zdecydowanie tak”. W przypadku pozycji P3, sformułowanej negatywnie, wartość ta odnosi się do zgody ze stwierdzeniem o szybkim poddawaniu się w sytuacjach przekraczających możliwości ucznia.

5.6. Charakterystyka rozkładu wyników (średnia punktów) skali

Przekonanie uczniów o własnej skuteczności kształtowało się na umiarkowanie wysokim poziomie. Średnia wyniosła $M = 2.92$, odchylenie standardowe $SD = 0.55$, a mediana $Me = 3.00$. Zaobserwowany zakres wyników wynosił od 1 do 4; jego środek to 2.50. Wyższe wartości oznaczają silniejsze poczucie skuteczności, dlatego średnia położona powyżej tego punktu wskazuje, że uczniowie dość często postrzegali siebie jako osoby zdolne do radzenia sobie z zadaniami i trudnościami. Ujemna skośność rozkładu, wynosząca -0.70 , pokazuje przesunięcie wyników ku wyższym wartościom, natomiast dodatnia kurtoza, równa 0.91 , sugeruje nieco większą koncentrację obserwacji w centralnej części rozkładu (zob. Wykres 11).

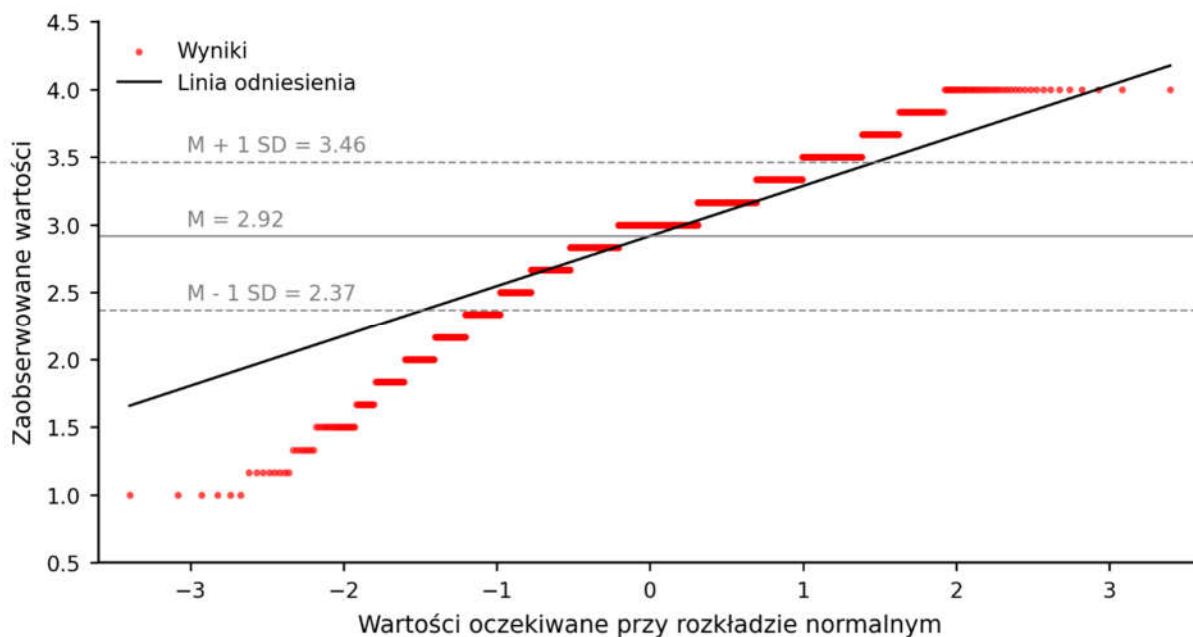
Wykres 11. WSS-OS: Rozkład średnich wyników ogólnej samoskuteczności.



Uwaga. Na wykresie przedstawiono wygładzoną krzywą gęstości średnich wyników. Linia przerywana wyznacza średnią arytmetyczną (M), a linia kropkowana medianę (Me). Wysokość krzywej odzwierciedla względne zagęszczenie obserwacji w poszczególnych zakresach wartości.

Układ punktów na wykresie Q–Q (por. Wykres 12) pokazuje, w jakim stopniu empiryczny rozkład wyników odpowiada rozkładowi normalnemu. W środkowej części obserwacje układają się stosunkowo blisko linii odniesienia, co wskazuje, że zasadnicza część wyników przebiega podobnie do układu oczekiwanego przy normalności. Większe odstępstwa pojawiają się natomiast w obu ogonach rozkładu. W dolnym zakresie punkty znajdują się poniżej linii odniesienia, a w górnym widoczne jest ich wyraźne wypłaszczenie przy najwyższych wartościach. Oznacza to, że rozkład umiarkowanie odbiega od rozkładu normalnego, przede wszystkim ze względu na odchylenia w krańcowych partiach rozkładu. Obraz ten dobrze koresponduje z wcześniejszą informacją o ujemnej skośności oraz podwyższonej kurtozie: większość wyników koncentruje się wokół wartości typowych i relatywnie wysokich, przy ograniczonej liczbie obserwacji w niższym zakresie.

Wykres 12. WSS-OS: Wykres Q-Q średnich wyników ogólnej samoskuteczności względem rozkładu normalnego.



Uwaga. Wykres przedstawia porównanie kwantylów empirycznych z kwantylami teoretycznymi rozkładu normalnego. Czerwone punkty oznaczają zaobserwowane wartości, a czarna linia pełni funkcję linii odniesienia dla rozkładu normalnego. Odchylenia punktów od tej linii wskazują na odstępstwa rozkładu empirycznego od normalności. Pozioma linia ciągła oznacza średnią arytmetyczną wyniku, oznaczoną jako M . Linie przerywane wyznaczają wartości o jedno odchylenie standardowe poniżej i powyżej średniej, czyli $M - 1 SD$ oraz $M + 1 SD$, gdzie SD oznacza odchylenie standardowe.

5.7. Wnioski

Podsumowując, 6-itemowa wersja WSS-OS wykazała dobre właściwości psychometryczne. Analiza CFA potwierdziła jednoczynnikową strukturę skali, a wszystkie pozycje istotnie odzwierciedlały wspólny wymiar ogólnej samoskuteczności. Narzędzie charakteryzowało się wysoką rzetelnością oraz dobrą precyzją pomiaru w szerokim zakresie cechy, ze szczególną wrażliwością na obniżony poziom samoskuteczności. Łącznie wyniki wskazują, że WSS-OS może być traktowana jako spójny, rzetelny i trafny strukturalnie wskaźnik przekonania uczniów o własnej skuteczności.

6. Wielowymiarowa Skala Samooceny: Satysfakcji z wyglądu (WSS-SzW)

6.1. Analiza struktury czynnikowej

1. Podoba mi się, jak wyglądam.
2. Myślę, że mam w sobie coś atrakcyjnego.
3. Jestem zadowolony/zadowolona ze swojego wyglądu.
- ~~4. Czasami wstydzę się swojego wyglądu.~~
5. Myślę, że wyglądam naprawdę dobrze.
6. Uważam, że inni wyglądają lepiej ode mnie.
7. Czuję się dobrze w swoim ciele.

W pierwszym kroku oszacowano jednoczynnikowy model CFA dla oryginalnej, 7-itemowej wersji skali (zob. Tabela 23). We wszystkich analizowanych próbach model ten nie spełnił kryterium akceptowalnego dopasowania w zakresie RMSEA, którego wartości były podwyższone i mieściły się w przedziale od 0.110 do 0.148 (w zależności od rodzaju próby).

Tabela 23. WSS-SzW: Wskaźniki dopasowania dla modelu CFA (model wyjściowy, 7-itemowy)

MODEL	χ^2	df	RMSEA [90% CI]	CFI	TLI	SRMR
Próba treningowa	236.88**	14	0.148 [0.131-0.164]	0.989	0.983	0.036
Próba testowa	142.24**	14	0.112 [0.096-0.129]	0.994	0.991	0.028
Próba pełna	321.42**	14	0.122 [0.111-0.134]	0.992	0.988	0.031
Próba pewna	211.49**	14	0.110 [0.097-0.123]	0.995	0.992	0.023

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; *df* = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; 90% CI = 90% przedział ufności dla RMSEA; CFI = porównawczy indeks dopasowania; TLI = indeks Tuckera-Lewisa; SRMR = standaryzowany pierwiastek średniokwadratowy reszt. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Analiza ładunków czynnikowych wykazała, że wszystkie pozycje były istotnie związane z czynnikiem latentnym, jednak siła tych związków była zróżnicowana. Najstabszym wskaźnikiem okazała się pozycja numer 4, która uzyskała najniższe bezwzględne wartości ładunków we wszystkich próbach. Na tej podstawie pozycję tę usunięto, a następnie ponownie oszacowano jednoczynnikowy model CFA dla wersji 6-itemowej (zob. Tabela 24).

Tabela 24. WSS-SzW: Standaryzowane ładunki czynnikowe dla modelu CFA (model wyjściowy, 7-itemowy)

Wskaźnik	Próba treningowa	Próba testowa	Próba pełna	Próba pewna
WSSSZW_1	0.946**	0.934**	0.939**	0.941**
WSSSZW_2	0.731**	0.789**	0.759**	0.770**
WSSSZW_3	0.960**	0.962**	0.962**	0.970**
WSSSZW_4	-0.637**	-0.662**	-0.648**	-0.691**
WSSSZW_5	0.894**	0.911**	0.902**	0.908**
WSSSZW_6	-0.669**	-0.712**	-0.696**	-0.737**
WSSSZW_7	0.787**	0.842**	0.814**	0.810**

Uwaga. Wartości przedstawiają standaryzowane współczynniki STDYX. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Po usunięciu pozycji numer 4 model 6-itemowy uzyskał wyraźnie lepsze dopasowanie we wszystkich analizowanych próbach. Wartości RMSEA obniżyły się do poziomu akceptowalnego lub dobrego, przy jednocześnie wysokich wartościach CFI i TLI oraz niskich wartościach SRMR (zob. Tabela 25).

Tabela 25. WSS-SzW: Wskaźniki dopasowania dla modelu CFA (model po redukcji, 6-itemowy)

MODEL	χ^2	df	RMSEA [90% CI]	CFI	TLI	SRMR
Próba treningowa	23.57**	9	0.047 [0.024-0.071]	0.999	0.999	0.009
Próba testowa	21.58*	9	0.044 [0.020-0.068]	0.999	0.999	0.008
Próba pełna	37.71**	9	0.047 [0.032-0.063]	0.999	0.999	0.007
Próba pewna	48.49**	9	0.061 [0.045-0.079]	0.999	0.998	0.010

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; 90% CI = 90% przedział ufności dla RMSEA; CFI = porównawczy indeks dopasowania; TLI = indeks Tuckera-Lewisa; SRMR = standaryzowany pierwiastek średniokwadratowy reszt. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Analiza standaryzowanych ładunków czynnikowych w modelu 6-itemowym potwierdziła, że wszystkie pozostawione pozycje silnie reprezentowały wspólny konstrukt. W próbie pełnej bezwzględne wartości ładunków mieściły się w zakresie od 0.658 do 0.963. Dodajmy, że podobny profil wyników uzyskano także w pozostałych próbach (zob. Tabela 26).

Tabela 26. WSS-SzW: Standaryzowane ładunki czynnikowe dla modelu CFA (model po redukcji, 6-itemowy)

Wskaźnik	Próba treningowa	Próba testowa	Próba pełna	Próba pewna
WSSSZW_1	0.948**	0.935**	0.941**	0.942**
WSSSZW_2	0.737**	0.797**	0.768**	0.777**
WSSSZW_3	0.962**	0.962**	0.963**	0.972**
WSSSZW_5	0.897**	0.913**	0.904**	0.909**
WSSSZW_6	-0.623**	-0.691**	-0.658**	-0.718**
WSSSZW_7	0.785**	0.837**	0.811**	0.805**

Uwaga. Wartości przedstawiają standaryzowane współczynniki STDYX. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Podsumowując, redukcja skali poprzez usunięcie najstabilniej ładującej pozycji numer 4 doprowadziła do poprawy dopasowania modelu, przede wszystkim w zakresie RMSEA, przy jednoczesnym zachowaniu bardzo wysokich wartości pozostałych wskaźników dopasowania oraz silnych ładunków czynnikowych. Wyniki te przemawiają za przyjęciem 6-itemowej wersji skali jako rozwiązania lepiej dopasowanego do danych niż pierwotny model 7-itemowy.

6.2. Analiza niezmienności pomiarowej

Analizę niezmienności pomiarowej skali WSS-SzW przeprowadzono ze względu na płeć i klasę, porównując model konfiguralny z modelem skalarnym. W modelu konfiguralnym założono tę samą strukturę czynnikową w grupie chłopców i dziewcząt, natomiast w modelu skalarnym wprowadzono dodatkowe ograniczenia dotyczące ładunków czynnikowych i progów odpowiedzi, konieczne do porównywania średnich latentnych. Wyniki analizy przedstawiono w Tabeli 27 i 28.

Niezmiennność ze względu na płeć

Model konfiguralny wskazywał na akceptowalne dopasowanie przyjętej struktury czynnikowej w grupach wyróżnionych ze względu na płeć: $\chi^2(18) = 77.43$, $p < .01$, RMSEA = 0.068, CFI = 0.998. Następnie oszacowano model skalarny, w którym nałożono ograniczenia równości ładunków czynnikowych i progów odpowiedzi. Tak zdefiniowany model nadal był akceptowalnie dopasowany do danych: $\chi^2(34) = 163.64$, $p < .01$, RMSEA = 0.073,

CFI = 0.996. Zmiana dopasowania względem modelu konfiguralnego była niewielka: $\Delta RMSEA = 0.005$, a $\Delta CFI = -0.002$, co mieściło się w przyjętych granicach interpretacyjnych.

Zgodnie z kryteriami opartymi na zmianach wskaźników dopasowania, uzyskane wartości $\Delta RMSEA$ i ΔCFI przemawiają za niezmiennością skalarną pomiaru ze względu na płeć. Jednocześnie wynik procedury DIFFTEST był istotny statystycznie, $\Delta\chi^2(16) = 91.242$, $p < .001$, co wskazuje, że formalnie model skalarny różnił się od modelu konfiguralnego. Ponieważ jednak pogorszenie dopasowania oceniane na podstawie RMSEA i CFI było niewielkie i nie przekraczało przyjętych progów, a bezwzględne wartości wskaźników dopasowania modelu skalarnego pozostały akceptowalne, wynik DIFFTEST nie stanowi wystarczającej podstawy do odrzucenia porównywalności pomiaru między grupami.

Tabela 27. WSS-SzW: Wskaźniki dopasowania modeli niezmienności pomiaru ze względu na płeć

MODEL	Test chi-kwadrat dopasowania modelu ¹					RMSEA	$\Delta RMSEA$	CFI	ΔCFI
	χ^2	df	$\Delta\chi^2$	Δdf	p				
Konfiguralny	77.43**	18	—	—	—	0.068	—	0.998	—
Skalarny	163.64**	34	91.242	16	< .001	0.073	0.005	0.996	-0.002

Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; df = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; CFI = porównawczy indeks dopasowania. Δ = różnica wartości χ^2 , RMSEA i CFI w odniesieniu do modelu mniej restrykcyjnego. ¹ Wartości $\Delta\chi^2$ i Δdf obliczono z wykorzystaniem procedury DIFFTEST w pakiecie Mplus. Istotne statystycznie: * $p < .05$. ** $p < .01$.

Ponieważ zmiany RMSEA i CFI mieściły się w przyjętych granicach, uznano, że model dostarcza wystarczających podstaw do porównania średnich latentnych. Porównanie to wskazało na niższy poziom satysfakcji z wyglądu u dziewczynek w stosunku do chłopców, przyjętych jako grupa referencyjna ($\Delta M = -1.123$, $SE = 0.221$, $p < .001$). Wniosek ten wymaga jednak pewnej ostrożności interpretacyjnej, ponieważ praktyczne wskaźniki zmiany dopasowania wspierały niezmiennność pomiaru, natomiast wynik DIFFTEST sugerował statystycznie istotne pogorszenie dopasowania modelu skalarnego względem konfiguralnego.

Niezmiennność ze względu na klasę

Model konfiguralny wskazywał na dobre dopasowanie przyjętej struktury czynnikowej do danych: $\chi^2(18) = 71.64$, $p < .001$, RMSEA = 0.064, CFI = 0.999. Po wprowadzeniu ograniczeń skalarności dopasowanie modelu pozostało dobre: $\chi^2(34) = 122.84$, $p < .001$, RMSEA = 0.060, CFI = 0.998.

Przejście do modelu bardziej restrykcyjnego nie wiązało się z pogorszeniem praktycznych wskaźników dopasowania. Wartość RMSEA uległa nawet niewielkiemu obniżeniu ($\Delta\text{RMSEA} = -0.004$), a CFI spadło jedynie o 0.001 ($\Delta\text{CFI} = -0.001$). Zmiana ta mieściła się w przyjętych granicach i wspierała założenie porównywalności pomiaru satysfakcji z wyglądu między uczniami klas 4–5 i 6–8.

Jednocześnie procedura DIFFTEST wykazała istotną statystycznie różnicę między modelami, $\Delta\chi^2(16) = 64.01$, $p < .001$. Tym samym przesłanki dotyczące pełnej niezmienności skalarnej nie były całkowicie spójne: praktyczne kryteria ΔRMSEA i ΔCFI przemawiały za jej przyjęciem, natomiast formalny test różnicy modeli wskazywał na pogorszenie dopasowania po nałożeniu ograniczeń. Porównania średnich latentnych należało zatem interpretować z ostrożnością.

Przy uwzględnieniu tego zastrzeżenia stwierdzono, że uczniowie klas 6–8 charakteryzowali się istotnie niższym poziomem satysfakcji z wyglądu niż uczniowie klas 4–5, przyjęci jako grupa referencyjna ($\Delta M = -1.217$, $SE = 0.176$, $p < .001$).

Tabela 28. WSS-SzW: Wskaźniki dopasowania modeli niezmienności pomiaru ze względu na klasę

MODEL	Test chi-kwadrat dopasowania modelu ¹					RMSEA	Δ RMSEA	CFI	Δ CFI
	χ^2	df	$\Delta\chi^2$	Δ df	p				
Konfiguralny	71.64**	18	—	—	—	0.064	—	0.999	—
Skalarny	122.84**	34	64.006	16	< .001	0.060	-0.004	0.998	-0.001

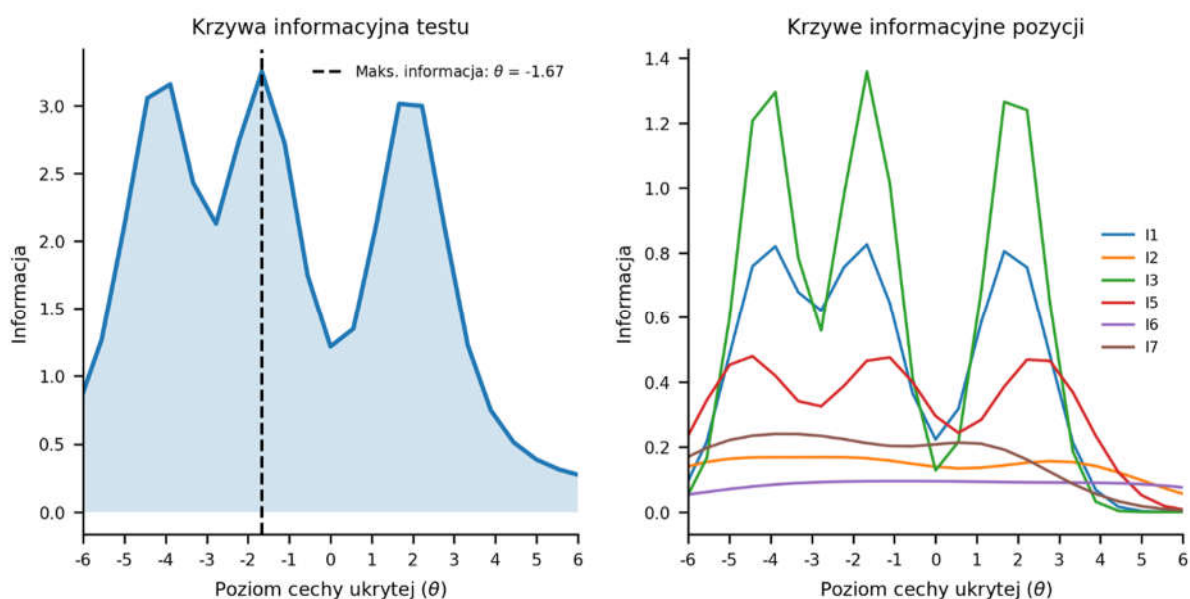
Uwaga. χ^2 = wartość testu chi-kwadrat dopasowania modelu; *df* = stopnie swobody; RMSEA = pierwiastek średniego błędu aproksymacji; CFI = porównawczy indeks dopasowania. Δ = różnica wartości χ^2 , RMSEA i CFI w odniesieniu do modelu mniej restrykcyjnego. ¹ Wartości $\Delta\chi^2$ i Δ df obliczono z wykorzystaniem procedury DIFFTEST w pakiecie Mplus. Istotne statystycznie: * $p < .05$. ** $p < .01$.

6.3. Krzywa informacyjna testu i pozycji testowych

Skala satysfakcji z wyglądu dostarcza najwięcej informacji w dolnej części kontinuum cechy, czyli przede wszystkim przy obniżonym poziomie satysfakcji z własnego wyglądu. Maksimum krzywej informacyjnej przypada około $\theta = -1.67$, a rzetelność warunkowa osiąga w tym punkcie około 0.76. Rzetelność co najmniej 0.70 utrzymuje się w trzech zakresach cechy: od około $\theta = -4.44$ do $\theta = -3.33$, od $\theta = -2.22$ do $\theta = -1.11$ oraz od $\theta = 1.67$ do $\theta = 2.22$. Przy łagodniejszym kryterium, czyli rzetelności powyżej 0.60, zakres użytecznej precyzji obejmuje wartości od $\theta = -5.00$ do $\theta = -0.56$ oraz od $\theta = 1.11$ do $\theta = 2.78$. W praktyce skala najlepiej

różnicuje więc uczniów deklarujących niską satysfakcję z wyglądu, ale zachowuje akceptowalną precyzję także u uczniów uzyskujących wyniki powyżej przeciętnego poziomu cechy. Jednocześnie skala jest wyraźnie mniej precyzyjna w okolicach przeciętnego poziomu cechy. Przy $\theta = 0$ rzetelność warunkowa wynosi około 0.55, co oznacza, że narzędzie słabiej różnicuje uczniów o przeciętnej satysfakcji z wyglądu. Największe znaczenie dla precyzji pomiaru mają pozycje I3, I1 i I4, natomiast najmniej informacyjna jest pozycja I5 (zob. Wykres 13).

Wykres 13. WSS-SzW: Krzywa informacyjna testu oraz krzywe informacyjne pozycji dla satysfakcji z wyglądu



Uwaga. Lewy panel przedstawia krzywą informacyjną całego testu, a prawy panel krzywe informacyjne poszczególnych pozycji. Oś X oznacza poziom cechy ukrytej θ , natomiast oś Y poziom informacji, czyli precyzję pomiaru na danym poziomie cechy. Wyższa wartość informacji oznacza większą dokładność pomiaru w danym zakresie θ . Linia przerywana w lewym panelu wskazuje poziom cechy, przy którym test osiąga maksymalną informację. Oznaczenia pozycji odnoszą się do kolejnych pozycji (itemów) wchodzących w skład analizowanego narzędzia.

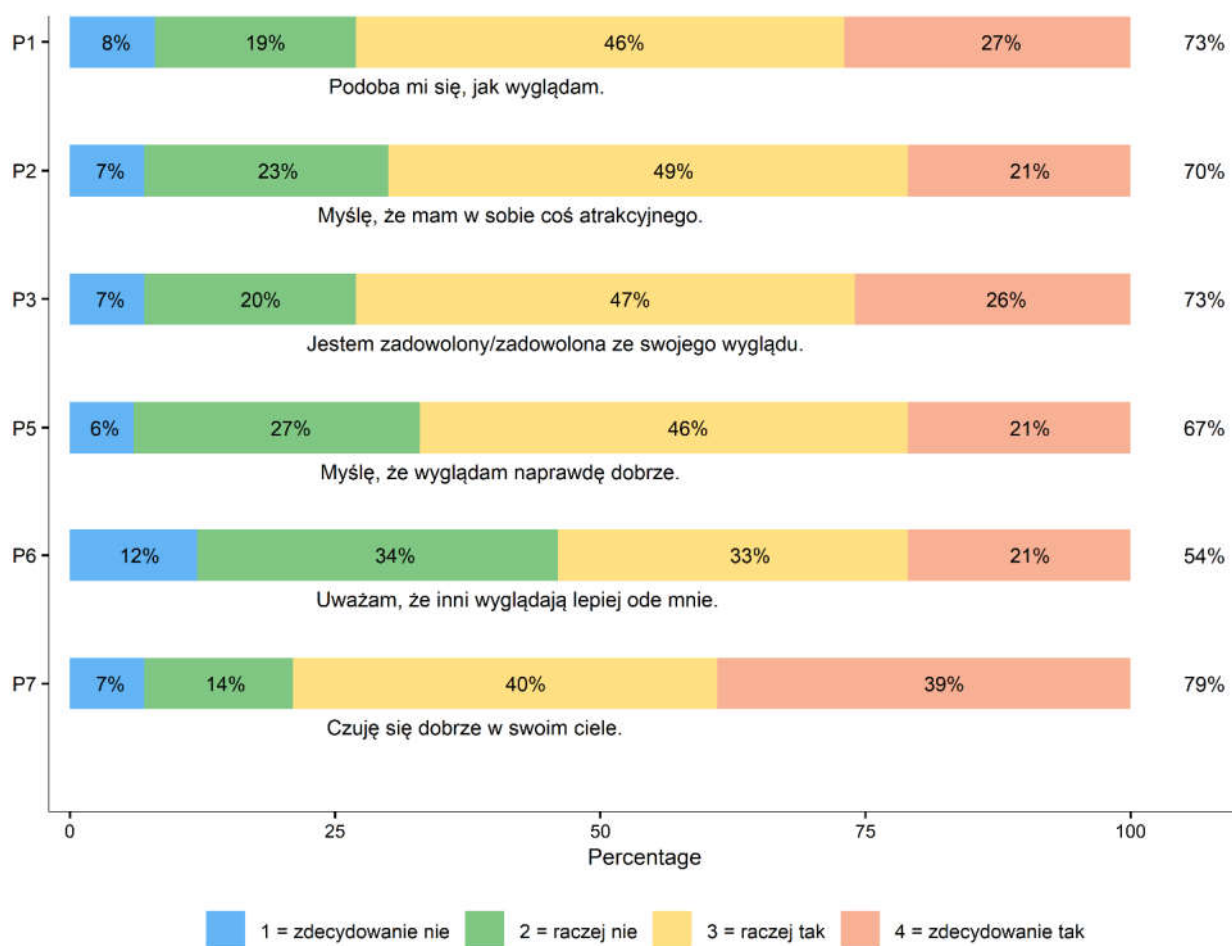
6.4. Miary rzetelności

Skala satysfakcji z wyglądu uzyskała bardzo wysokie wskaźniki rzetelności. Alfa porządkowa wyniosła $\alpha_{ord} = 0.934$, natomiast współczynnik omega całkowita McDonalda osiągnął wartość $\omega_t = 0.908$. Również klasyczna alfa Cronbacha była bardzo wysoka i wyniosła $\alpha = 0.907$. Wszystkie współczynniki wskazują na bardzo wysoką spójność wewnętrzną skali. Dodatkowo wartość wskaźnika AVE wyniosła 0.718.

6.5. Rozkład odpowiedzi na poszczególne twierdzenia skali

Rozkład odpowiedzi wskazuje, że uczniowie przeważnie pozytywnie oceniali własny wygląd, choć odpowiedzi nie były całkowicie jednorodne. Najwyższy odsetek odpowiedzi pozytywnych uzyskano dla pozycji P7 („Czuję się dobrze w swoim ciele”) – 79% badanych raczej lub zdecydowanie zgodziło się z tym stwierdzeniem. Wysoki poziom aprobaty dotyczył także pozycji P1 („Podoba mi się, jak wyglądam”) oraz P3 („Jestem zadowolony/zadowolona ze swojego wyglądu”) – po 73% odpowiedzi pozytywnych. Nieco niższy odsetek zgody odnotowano dla pozycji P2 („Myślę, że mam w sobie coś atrakcyjnego”) – 70%, oraz P5 („Myślę, że wyglądam naprawdę dobrze”) – 67% (zob. Wykres 14).

Wykres 14. WSS-SzW: Procentowy rozkład odpowiedzi na poszczególne twierdzenia skali



Uwaga. Wartości po prawej stronie wykresu oznaczają łączny odsetek odpowiedzi „raczej tak” i „zdecydowanie tak”. W przypadku pozycji P6, sformułowanej negatywnie, wartość ta odnosi się do zgody ze stwierdzeniem, że inni wyglądają lepiej od osoby badanej.

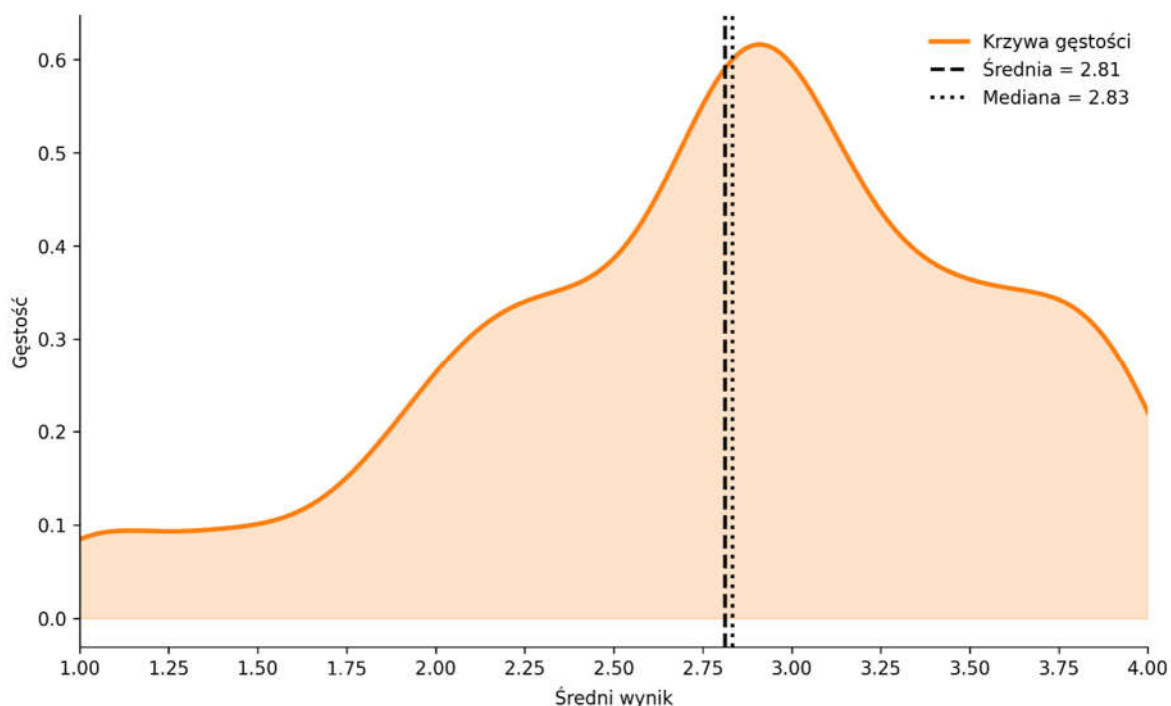
Szczegółnej uwagi wymaga pozycja P6 („Uważam, że inni wyglądają lepiej ode mnie”), która była sformułowana negatywnie. Z tym stwierdzeniem raczej lub zdecydowanie zgodziło

się 54% uczniów, natomiast 46% wybrało odpowiedzi wskazujące na brak zgody. Oznacza to, że mimo ogólnie pozytywnej oceny własnego wyglądu znaczna część badanych porównuje się z innymi w sposób niekorzystny dla siebie. Podsumowując, uczniowie najczęściej deklarowali zadowolenie z wyglądu i dobre samopoczucie we własnym ciele, jednak odpowiedzi na pozycję odwróconą wskazują na wyraźną obecność porównań społecznych dotyczących atrakcyjności fizycznej.

6.6. Charakterystyka rozkładu wyników (średnia punktów) skali

Wyniki dotyczące satysfakcji z wyglądu wskazują na umiarkowanie pozytywną ocenę tego obszaru. Średnia wyniosła $M = 2.81$ przy $SD = 0.73$, a mediana $Me = 2.83$. Wyniki mieściły się między 1 a 4, o punkcie środkowym równym 2.50. Ponieważ wyższe wartości odpowiadają większej satysfakcji z wyglądu, średnia powyżej tego punktu pokazuje, że przeciętna ocena była raczej korzystna. Lekko ujemna skośność (-0.48) wskazuje, że wyniki były nieco częściej przesunięte ku wyższym wartościom. Kurtoza bliska zeru, choć nieznacznie ujemna (-0.23), sugeruje natomiast rozkład zbliżony pod względem „wysmukłości” do normalnego, jedynie lekko spłaszczony (zob. Wykres 15).

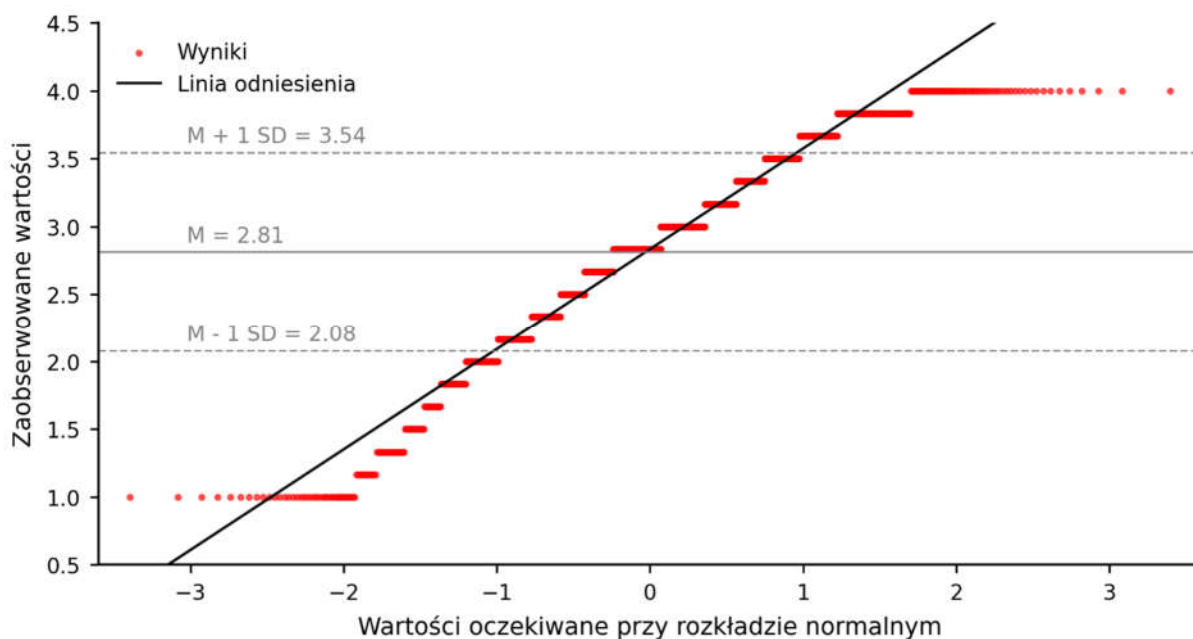
Wykres 15. WSS-SzW: Rozkład średnich wyników satysfakcji z wyglądu



Uwaga. Na wykresie przedstawiono wygładzoną krzywą gęstości średnich wyników. Linia przerywana wyznacza średnią arytmetyczną (M), a linia kropkowana medianę (Me). Wysokość krzywej odzwierciedla względne zagęszczenie obserwacji w poszczególnych zakresach wartości.

Na wykresie Q-Q (por. Wykres 16) widać, że środkowa część rozkładu satysfakcji z wyglądu dość dobrze podąża za linią odniesienia. Oznacza to, że wyniki zbliżone do przeciętnych układają się w sposób względnie zgodny z rozkładem normalnym. Wyraźniejsze odstępstwa pojawiają się natomiast na krańcach rozkładu. W dolnym zakresie widoczne jest skupienie obserwacji przy najniższych wartościach, a w górnym — wypłaszczenie przy wynikach najwyższych. Sugeruje to, że rozkład umiarkowanie odbiega od rozkładu normalnego, przede wszystkim ze względu na zachowanie ogonów. Taki układ jest spójny z wcześniejszym opisem rozkładu: przeciętna satysfakcja z wyglądu była raczej korzystna, ale wyniki były dość zróżnicowane, a część uczniów lokowała się zarówno przy niskich, jak i bardzo wysokich ocenach tego obszaru.

Wykres 16. WSS-SzW: Wykres Q-Q średnich wyników satysfakcji z wyglądu względem rozkładu normalnego



Uwaga. Wykres przedstawia porównanie kwantyli empirycznych z kwantylami teoretycznymi rozkładu normalnego. Czerwone punkty oznaczają zaobserwowane wartości, a czarna linia pełni funkcję linii odniesienia dla rozkładu normalnego. Odchylenia punktów od tej linii wskazują na odstępstwa rozkładu empirycznego od normalności. Pozioma linia ciągła oznacza średnią arytmetyczną wyniku, oznaczoną jako M . Linie przerywane wyznaczają wartości o jedno odchylenie standardowe poniżej i powyżej średniej, czyli $M - 1 SD$ oraz $M + 1 SD$, gdzie SD oznacza odchylenie standardowe.

6.7. Wnioski

W świetle przeprowadzonych analiz najbardziej uzasadnione jest przyjęcie 6-itemowej wersji WSS-SzW. Redukcja pozycji poprawiła dopasowanie modelu CFA, przy jednoczesnym zachowaniu silnego związku pozostałych twierdzeń ze wspólnym czynnikiem satysfakcji