

Audyt jakości ankiety

Aktywność fizyczna i samopoczucie studentów

2026-10-11 16:14

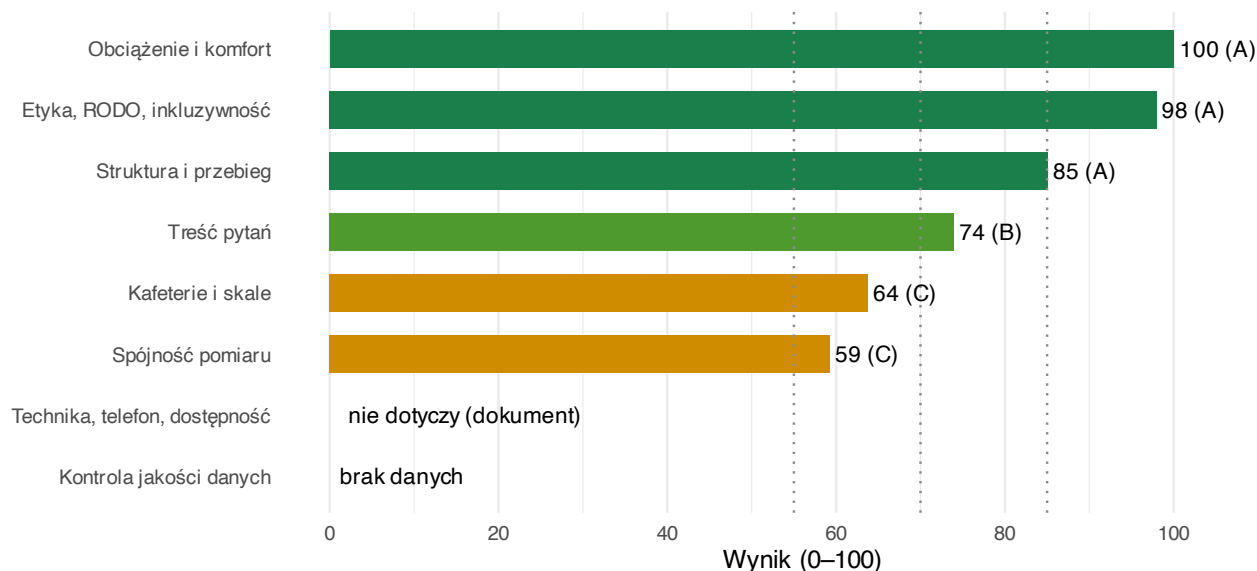
Spis treści

1 Podsumowanie	1
1.1 Najważniejsze poprawki	3
2 Jak powstał ten audyt	4
3 Mapa ankiety	5
4 Pytania: problemy i propozycje poprawek	6
5 Kafeterie i skale odpowiedzi	11
6 Kontrola jakości danych	11
7 Obciążenie i komfort badanego	12
8 Spójność i rzetelność pomiaru	13
9 Struktura i przebieg	14
10 Technika, telefon i dostępność	14
11 Etyka, RODO i inkluzyność	14
12 Gotowość do analizy danych	15
13 Aneksy	16
13.1 Pełna lista problemów	16
13.2 Kandydaci odrzuceni przez panel AI	17
13.3 Słownik kodów problemów	17
13.4 Parametry audytu	20

1 Podsumowanie

Indeks jakości ankiety: 80/100 — ocena B

dobra — wymaga drobnych poprawek



Parametr	Wartość
Ankieta	Aktywność fizyczna i samopoczucie studentów
Dokument	Aktywność fizyczna i samopoczucie studentów.docx (DOCX)
Data audytu	2026-10-11 16:14
Tryb audytu	szybki
Źródła danych	Kwestionariusz odczytano z dokumentu „Aktywność fizyczna i samopoczucie studentów.docx” (DOCX): model wyodrębnił 10 pytań, a każde pytanie, odpowiedź i wiersz macierzy porównano z tekstem dokumentu.
Części z pytaniami	3
Pytania	10
Pozycje (z wierszami macierzy)	14
Szacowany czas wypełniania	≈ 2,1 min
Pytania obowiązkowe	nie oznaczono w dokumencie
Wykryte problemy	29 (krytyczne: 0, ważne: 13, drobne: 16)
Model AI	qwen3.6:35b-mlx

Kluczowym problemem jest Q06, które łączy dwa różne stany (zmęczenie i brak energii), co utrudnia precyzyjną interpretację odpowiedzi i zwiększa ryzyko błędu pomiaru. Q02 wymaga od respondentów precyzyjnej pamięci i liczenia dni w poprzednim tygodniu, co generuje wysoki błąd recall i zachęca do zadowalania się byle jaką odpowiedzią. W Q01 występuje presja społecznej pożądaności oraz niejasna definicja progu aktywności, co może zafałszować dane o frekwencji ćwiczeń. Q04 stanowi heterogeniczną macierz, w której miesza się konstrukty afektywne, postawowe i barierowe, uniemożliwiając sensowne sumowanie wyników. Q09 brakuje instrukcji co do formatu wpisywanego roku, co stwarza ryzyko błędnych danych (np. wpisania wieku zamiast roku). Q10 zawiera niepełną listę opcji, wykluczającą studentów studiów niestacjonarnych i tych w innych formach kształcenia. Q01 zawiera również niepotrzebny filtr oparty na specyficznej definicji WHO, który może niepasować do celów badawczych. Te błędy w doborze słów i strukturze pytań obniżają wiarygodność mierzonych zmiennych.

Mocne strony

- Q05 i Q06 stosują jasne okresy referencyjne (ostatnie 30 dni, ostatnie 2 tygodni), co poprawia trafność recallu.
- Q07 (ranking) jest dobrze zdefiniowanym zadaniem, które pozwala na precyzyjne ustalenie priorytetów respondentów.
- Pytania demograficzne (Q08, Q09) są standardowe i łatwe do przetworzenia, co zmniejsza obciążenie ankiety.
- Krótka estymowana czas trwania (2 min) i przejrzysta struktura sprzyjają utrzymaniu uwagi respondentów.
- Zastosowanie skali 11-punktowej w Q05 może zwiększyć wariancję odpowiedzi w porównaniu do standardowych skal 5- lub 7-punktowych.

1.1 Najważniejsze poprawki

Nr	Waga	Pytanie	Problem	Co zrobić
1	ważny	Q06	Dwa pytania w jednym	Pytanie łączy dwa różne stany: zmęczenie (fizyczne/psychiczne) i brak energii. Respondent może czuć się zmęczony, ale mieć energię, lub odwrotnie. Odpowiedź 'Często' jest niejednoznaczna. Poprawka: Rozdziel na dwa pytania: 'Jak często czułeś(-aś) się zmęczony(-a)?' oraz 'Jak często miałeś(-aś) mało energii?' lub zmień formułkę na 'czułeś(-aś) się wyczerpany(-a) lub pozbawiony(-a) energii' (jeden stan).
2	ważny	Q09	Brak jednostki lub zakresu	Brak jednostki lub zakresu. Respondent może wpisać rok (np. 1995) lub wiek (np. 25), co zepsuje dane. Pole liczbowe bez instrukcji jest ryzykowne. Poprawka: Dodaj instrukcję: 'Podaj rok urodzenia (np. 1995)' lub zmień na 'Wiek w latach' z polem tekstowym i walidacją.
3	ważny	Q06	Nieostre określenia	Słowo 'mało' jest względne i nieprecyzyjne. Różni respondenci mają różne progi postrzegania 'niewielkiej ilości' energii. Poprawka: Użyj standardowych, zweryfikowanych sformułowań, np. 'czułeś(-aś) się wyczerpany(-a)' lub 'czułeś(-aś) brak sił', które są lepiej zdefiniowane w literaturze psychometrycznej.
4	ważny	Q02	Trudne przypominanie lub szacowanie	Wymaga precyzyjnej pamięci i liczenia dni z poprzedniego tygodnia, co jest trudne i obciążone błędem recall (Tourangeau). Respondenci często zgadywają lub stosują heurystykę. Poprawka: Rozważ użycie kalendarza lub listy dni tygodnia do zaznaczenia, aby ułatwić recall, lub zmień pytanie na oceny przedziałów (np. 0, 1-2, 3-4, 5-7 dni).
5	ważny	Q04	Przeczenie lub podwójne przeczenie	Stwierdzenie jest negatywne i merytorycznie odwrócone względem pozostałych (pozytywnych). W skali zgodności respondent musi wykonać dodatkową operację kognitywną (negację), co zwiększa ryzyko błędu i straightliningu. Poprawka: Zmień twierdzenie na pozytywne: 'Aktywność fizyczna jest dla mnie wartościowa/cenna' lub odwróć skalę dla tego wiersza (jeśli platforma pozwala), ale lepiej unikać negacji w twierdzeniach.
6	ważny	Q10	Luki, niepełna lista odpowiedzi	Brakuje opcji dla studentów studiów niestacjonarnych, zaocznych lub tych, którzy ukończyli licencjat/inżyniera, a kontynuują naukę w innej formie, co uniemożliwia im udzielenie poprawnej odpowiedzi. Poprawka: Dodać opcję 'Inne (proszę określić)' lub rozszerzyć listę o 'studia niestacjonarne/zaoczne' oraz 'po ukończeniu licencjatu/inżyniera'.
7	ważny	Q01	Nieostre określenia	Sformułowanie 'łącznie co najmniej 30 minut dziennie' jest mylące: sugeruje, że 30 minut to próg konieczny do uznania aktywności za 'tak', podczas gdy w badaniach epidemiologicznych często liczy się sumaryczny czas tygodniowy lub dni z aktywnością >30 min. Respondent może nie wiedzieć, czy ma oceniać każdy dzień z osobna, czy cały tydzień. Poprawka: Zmień pytanie na: 'Czy w ciągu ostatnich 7 dni podejmowałeś(-aś) aktywność fizyczną, która trwała co najmniej 30 minut w przynajmniej jednym dniu?' lub precyzuj, czy chodzi o sumaryczny czas, czy dni z aktywnością.
8	ważny	Q04	Niejednorodna bateria	Wiersze mierzą różne konstrukty: przyjemność (afekt), samopoczucie po ćwiczeniach (efekt fizjologiczny/psychologiczny), postrzeganie wartości czasu (postawa), dostępność czasu (bariera), motywacja (cel). Sumowanie ich do jednej skali byłoby nieuzasadnione. Poprawka: Traktuj te pytania jako osobne wskaźniki lub podskale, a nie jako jedną skalę 'postawy wobec aktywności'. W analizie nie twórz sumarycznego wyniku bez weryfikacji jednowymiarowości.
9	ważny	Q01	Brak filtra	Pytanie filtrujące zakłada, że 'aktywność fizyczna' oznacza 30 min dziennie, co jest specyficzną definicją WHO, nie powszechną. Poprawka: Dodać opcję 'Nie wiem' lub wyjaśnić definicję aktywności, aby uniknąć błędnych odpowiedzi 'Nie'.
10	ważny	Q01	Presja aprobaty społecznej	Pytanie o aktywność fizyczną jest silnie obciążone presją społeczną pożądaności; odpowiedź 'Tak' jest pożądana społecznie, co może prowadzić do zaniżania bezruchu. Poprawka: Zastosuj neutralizację presji społecznej, np. dodając informację, że 'wiele osób ma różne nawyki' lub używając techniki 'randomized response' jeśli to możliwe, a w ankiecie internetowej upewniając o anonimowości i braku oceny.

Następne kroki

- Rozdzielić Q06 na dwa osobne pytania: jedno o częstotliwość odczuwania zmęczenia, drugie o brak energii, aby uniknąć efektu podwójnego pytania.
- Przepisać Q02, dodając instrukcję pomocniczą lub zmieniając okres referencyjny na 'ostatni tydzień' z pytaniem o liczbę dni, ale z uproszczoną formułą (np. 'czy ćwiczyłeś w poniedziałek?'), aby zmniejszyć obciążenie pamięci.
- Zmienić Q04 na osobne pozycje lub grupy pytań, aby oddzielić konstrukty afektywne od postawowych, lub usunąć pozycje mierzące inne konstrukty niż zamierzony.
- Dodać do Q09 instrukcję 'wpisz rok, np. 1995' oraz opcję 'nie chcę podać' dla zgodności z RODO i etyką.
- Rozszerzyć listę w Q10 o opcje 'studia niestacjonarne', 'zaoczne' oraz 'ukończyłem(-a) licencjat/inżyniera, kontynuuję naukę w innej formie'.

- Przeprowadzić pilotaż z udziałem 5-10 studentów, aby zweryfikować zrozumienie sformułowań i czas potrzebny na odpowiedź.

Wymiar	Waga	Wynik	Ocena	Kryt.	Ważne	Drobne	Podstawa
Treść pytań	20	74	B	0	6	0	10 pytań
Kafeteria i skale	15	64	C	0	4	10	10 pytań
Kontrola jakości danych	10	—	—	0	0	0	za mało kryteriów
Obciążenie i komfort	15	100	A	0	0	2	55/55 pkt
Spójność pomiaru	10	59	C	0	2	1	40/60 pkt
Struktura i przebieg	10	85	A	0	1	2	92/100 pkt
Technika, telefon, dostępność	10	—	—	0	0	0	nie dotyczy (dokument)
Etyka, RODO, inkluzywność	10	98	A	0	0	1	100/100 pkt

2 Jak powstał ten audyt

Audyt łączy cztery niezależne źródła informacji, a każde ustalenie ma podaną wagę (krytyczny, ważny, drobny), pewność (0–1) i źródło.

1. **Odczyt kwestionariusza z dokumentu.** Kwestionariusz odczytano z dokumentu „Aktywność fizyczna i samopoczucie studentów.docx” (DOCX): model wyodrębnił 10 pytań, a każde pytanie, odpowiedź i wiersz macierzy porównano z tekstem dokumentu. Tekst (akapity, automatyczna numeracja, tabele, pola odpowiedzi) odczytano bezpośrednio z pliku; model rozpoznał w nim pytania, ich typy, odpowiedzi, wiersze i kolumny macierzy, instrukcje, filtry i oznaczenia pytań obowiązkowych. Pytania, odpowiedzi i zdania, których nie ma w dokumencie, odrzucono, a wprowadzenie pochodzi wprost z dokumentu.
2. **Reguły automatyczne.** Wzorce językowe (przeczenia, nieostre określenia, brak okresu odniesienia, sformułowania sugerujące, skróty, pytania hipotetyczne), analiza kafeterii (nakładanie się i luki w przedziałach, zrównoważenie skal, opcje „nie wiem/wolę nie odpowiadać”, „inne”), metryki obciążenia (szacowany czas, macierze, pytania otwarte, czytelność FOG-PL), wykrywanie pytań kontrolnych, elementów zgody i RODO.
3. **Panel ekspertów AI** (model qwen3.6:35b-mlx): metodolog i psychometra, rzecznik badanego (wywiady poznawcze, prosty język). Każda rola niezależnie przegląda wszystkie pytania; ustalenia tej samej wady z kilku źródeł są łączone, a pewność rośnie ze zgodnością ($1 - \prod(1 - p_i)$). Kandydaci z reguł, których panel nie potwierdził, trafiają do aneksu. Panel ocenia też całość ankiety (wprowadzenie, przebieg, kontrolę danych, baterie) i proponuje nowe brzmienie każdego wadliwego pytania.
4. **Testy techniczne.** Audyt dokumentu nie obejmuje dostępności (WCAG), wersji na telefon, komunikatów walidacji, podziału na strony ani paska postępu — zależą od platformy, na której ankieta zostanie zaprogramowana. Po jej zaprogramowaniu warto zaudytować link do ankiety.

Punktacja. Wymiary dotyczące pojedynczych pytań (treść, kafeteria): każde pytanie zaczyna od 100% i traci za najpoważniejszy problem 100% (krytyczny), 60% (ważny) lub 20% (drobny) oraz 10% za każdy kolejny (maks. 30%), z wagą pewności (pewność $\geq 0,7$: $\times 1$; $0,45-0,7$: $\times 0,7$; $< 0,45$: $\times 0,4$); wynik wymiaru to średnia po pytaniach minus problemy dotyczące całej ankiety (15/6/2 pkt). Wymiary oparte na kryteriach (etyka, przebieg, kontrola danych, spójność pomiaru, obciążenie) to odsetek punktów za spełnione kryteria (częściowo = połowa). Dostępność: 100 minus kary za każdą regułę axe-core (12/8/4/2 zależnie od wpływu) i problemy na telefonie. Jeden pewny problem krytyczny ogranicza wynik do 84 (brak oceny A), trzy lub więcej — do 69. Indeks ogólny to średnia ważona dostępnych wymiarów.

Oceny: A ≥ 85 (bardzo dobra), B 70–84 (dobra, drobne poprawki), C 55–69 (przeciętna, wymaga poprawek), D < 55 (słaba, wymaga przebudowy).

Odczyt dokumentu. Model mógł pominąć pytanie albo źle rozpoznać jego typ (np. jedną odpowiedź zamiast kilku); sprawdź mapę ankiety poniżej. Pytania obowiązkowe nie są oznaczone w dokumencie, więc audyt nie ocenia wymuszania odpowiedzi.

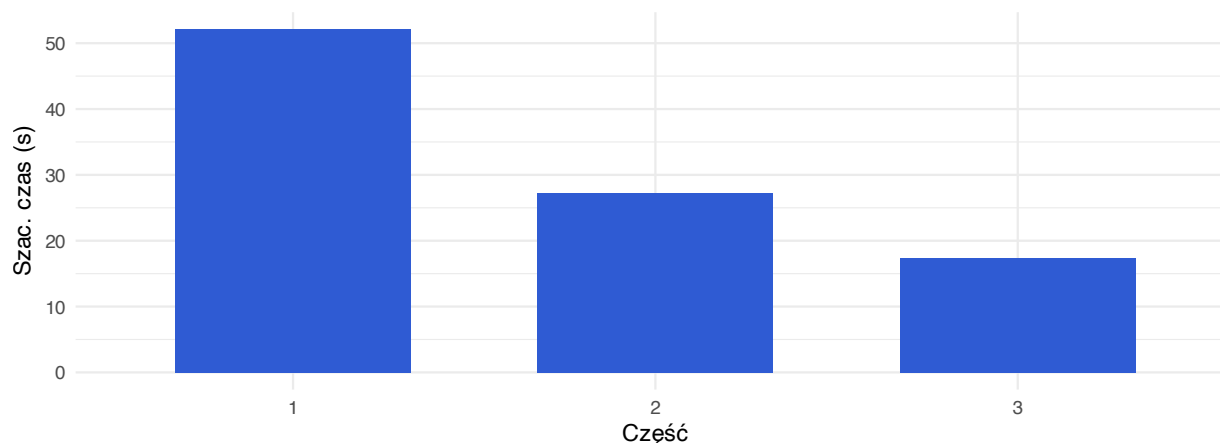
Ograniczenia. Ocena AI i reguł to ekspertyza wstępna, a nie zastępstwo pilotażu z ludźmi: model może przeoczyć problem albo zgłosić fałszywy alarm (dlatego podawana jest pewność i źródła). Syntetyczni respondenci odpowiadają spójniej niż ludzie. Test axe-core wykrywa ok. 30–40% problemów dostępności; pełna ocena wymaga testu z czytnikiem ekranu. Szacowany czas pochodzi z modelu tempa czytania i odpowiadania (ok. 240 słów/min).

Wymiar	Waga	Komentarz
Treść pytań	20	Czy pytania są jasne, neutralne, jednoznaczne i dotyczą jednej rzeczy naraz.
Kafeteria i skale	15	Czy odpowiedzi są rozłączne, wyczerpujące, zrównoważone i pasują do pytań.
Kontrola jakości danych	10	Pytania sprawdzające uwagę, pozycje odwrócone, pary kontrolne, pytanie o rzetelność.
Obciążenie i komfort	15	Długość, macierze, pytania otwarte, wymuszanie odpowiedzi, czytelność, postęp.
Spójność pomiaru	10	Wielopozycyjne, jednorodne skale i przewidywana spójność odpowiedzi.
Struktura i przebieg	10	Wprowadzenie, kolejność pytań, filtry, terminologia, zakończenie.
Technika, telefon, dostępność	10	Naruszenia WCAG (axe-core), wygoda na telefonie, błędy techniczne.
Etyka, RODO, inkluzywność	10	Świadoma zgoda, informacja o danych, dane wrażliwe, język włączający.

Waga	Kara (pytanie)	Kara (ankieta)	Komentarz
krytyczny	100%	15	unieważnia dane z pytania, uniemożliwia odpowiedź albo narusza etykę
ważny	60%	6	prawdopodobnie zniekształca odpowiedzi lub myli wielu badanych
drobny	20%	2	styl, drobna niejasność, wygoda

3 Mapa ankiety

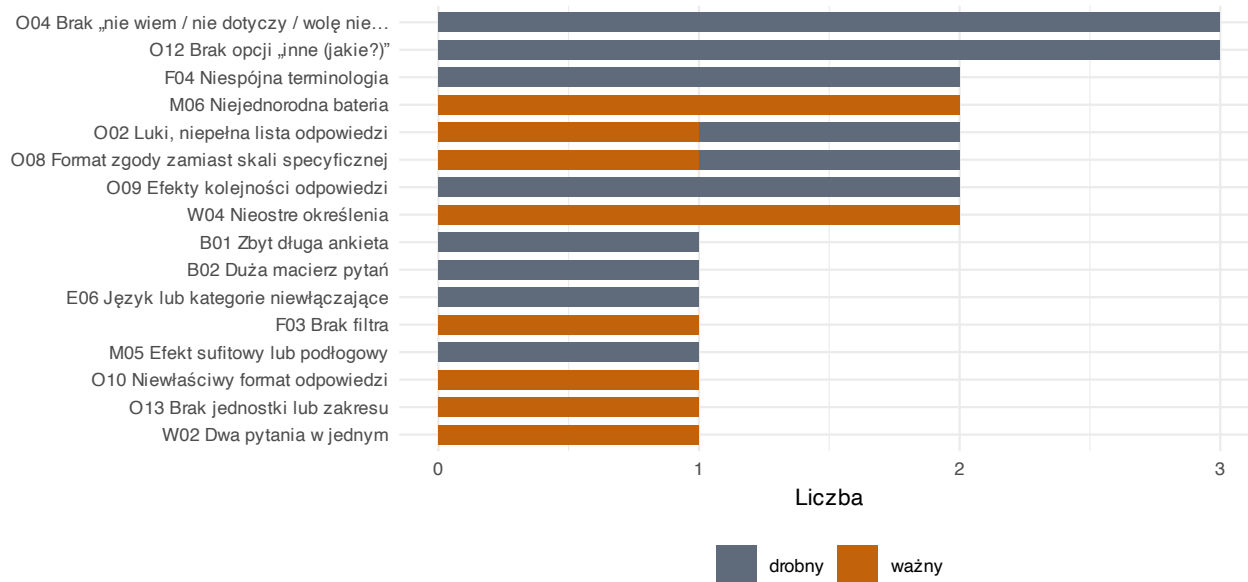
Część	Tytuł w dokumencie	Pytań	Pozycji	Szac. czas (s)
wprowadzenie		0	0	3
Część 1	Część A. Aktywność fizyczna	4	8	52
Część 2	Część B. Samopoczucie	3	3	27
Część 3	Metryczka	3	3	17
zakończenie		0	0	3



ID (nr)	cz.	Typ	Pytanie	Odp.	Obow.	jakość
Q01 (A1)	1	jednokrotny wybór	Czy w ciągu ostatnich 7 dni podejmowałeś(-aś) aktywność fizyczną trwającą łącznie co najmniej ...	2		20
Q02 (A2)	1	liczba	W ile dniach w ciągu ostatnich 7 dni podejmowałeś(-aś) taką aktywność?	0		40
Q03 (A3)	1	wielokrotny wybór	Jakie formy aktywności podejmowałeś(-aś) w ciągu ostatnich 7 dni?	6		80
Q04 (A4)	1	macierz	Na ile zgadzasz się z poniższymi stwierdzeniami?	5x5		10
Q05 (B1)	2	jednokrotny wybór	Jak oceniasz swoje ogólne samopoczucie w ciągu ostatnich 30 dni?	11		60
Q06 (B2)	2	jednokrotny wybór	Jak często w ciągu ostatnich 2 tygodni czułeś(-aś) się zmęczony(-a) lub miałeś(-aś) mało energ...	5		10
Q07 (B3)	2	ranking	Uzereguj poniższe powody podejmowania aktywności od najważniejszego (1) do najmniej ważnego (...)	4		20
Q08 (M1)	3	jednokrotny wybór	Płeć	4		80
Q09 (M2)	3	liczba	Rok urodzenia:	0		30
Q10 (M3)	3	jednokrotny wybór	Rok studiów	6		20

W nawiasie: numer pytania w dokumencie (jeśli różni się od identyfikatora). Kolumna „Obow.” pokazuje pytania oznaczone w dokumencie jako obowiązkowe.

4 Pytania: problemy i propozycje poprawek



Q01 (nr w dokumencie A1)

Czy w ciągu ostatnich 7 dni podejmowałeś(-aś) aktywność fizyczną trwającą łącznie co najmniej 30 minut dziennie?

cz. 1 · jednokrotny wybór · jakość 20/100

Filtr / przejście: Nie -> przejdź do pytania A4

Odpowiedzi: Tak / Nie

- **[ważny] Nieostre określenia** (W04; pewność 0,94; AI: metodolog + AI: rzecznik badanego). Sformułowanie ‘łącznie co najmniej 30 minut dziennie’ jest mylące: sugeruje, że 30 minut to próg konieczny do uznania aktywności za ‘tak’, podczas gdy w badaniach epidemiologicznych często liczy się sumaryczny czas tygodniowy lub dni z aktywnością >30 min. Respondent może nie wiedzieć, czy ma oceniać każdy dzień z osobna, czy cały tydzień. *Dowód:* trwającą łącznie co najmniej 30 minut dziennie *Poprawka:* Zmień pytanie na: ‘Czy w ciągu ostatnich 7 dni podejmowałeś(-aś) aktywność fizyczną, która trwała co najmniej 30 minut w przynajmniej jednym dniu?’ lub precyzuj, czy chodzi o sumaryczny czas, czy dni z aktywnością.
- **[ważny] Brak filtra** (F03; pewność 0,75; AI: przegląd całości). Pytanie filtrujące zakłada, że ‘aktywność fizyczna’ oznacza 30 min dziennie, co jest specyficzną definicją WHO, nie powszechną. *Dowód:* Q01 ... trwającą łącznie co najmniej 30 minut dziennie? *Poprawka:* Dodać opcję ‘Nie wiem’ lub wyjaśnić definicję aktywności, aby uniknąć błędnych odpowiedzi ‘Nie’.
- **[ważny] Presja aprobaty społecznej** (W09; pewność 0,75; AI: metodolog). Pytanie o aktywność fizyczną jest silnie obciążone presją społeczną pożądaności; odpowiedź ‘Tak’ jest pożądana społecznie, co może prowadzić do zaniżania bezruchu. *Dowód:* Czy w ciągu ostatnich 7 dni podejmowałeś(-aś) aktywność fizyczną? *Poprawka:* Zastosuj neutralizację presji społecznej, np. dodając informację, że ‘wiele osób ma różne nawyki’ lub używając techniki ‘randomized response’ jeśli to możliwe, a w ankiecie internetowej upewniając o anonimowości i braku oceny.

Propozycja: Czy w ciągu ostatnich 7 dni wykonywałeś(-aś) aktywność fizyczną, która trwała łącznie co najmniej 30 minut w jednym dniu?

Uzasadnienie: Usunięto mylące sformułowanie ‘dziennie’ odnoszące się do całego tygodnia, zastępując je precyzyjnym odniesieniem do ‘jednego dnia’, co ułatwia ocenę; pytanie pozostaje zamknięte, by minimalizować presję społeczną poprzez jasne kryterium ‘tak/nie’ oparte na konkretnym zdarzeniu.

Q02 (nr w dokumencie A2)

W ile dniach w ciągu ostatnich 7 dni podejmowałeś(-aś) taką aktywność?

cz. 1 · liczba · jakość 40/100

- **[ważny] Trudne przypominanie lub szacowanie** (W07; pewność 0,97; AI: metodolog + AI: rzecznik badanego). Wymaga precyzyjnej pamięci i liczenia dni z poprzedniego tygodnia, co jest trudne i obciążone błędem recall (Tourangeau). Respondenci często zgadywają lub stosują heurystykę. *Dowód:* W ile dniach w ciągu ostatnich 7 dni podejmowałeś(-aś) taką aktywność? *Poprawka:* Rozważ użycie kalendarza lub listy dni tygodnia do zaznaczenia, aby ułatwić recall, lub zmień pytanie na oceny przedziałów (np. 0, 1-2, 3-4, 5-7 dni).

Propozycja: W ilu dniach w ciągu ostatnich 7 dni wykonywałeś(-aś) aktywność fizyczną trwającą co najmniej 30 minut?

Uzasadnienie: Usunięto słowo 'taką' odnoszące się do niejasnego 'dniennie' z poprzedniego pytania, zastępując je powtórzeniem jasnego kryterium 'trwającą co najmniej 30 minut', co zmniejsza obciążenie pamięci i ryzyko błędu recall poprzez odniesienie do konkretnego, łatwiejszego do zapamiętania zdarzenia.

Q03 (nr w dokumencie A3)

Jakie formy aktywności podejmowałeś(-aś) w ciągu ostatnich 7 dni?

cz. 1 · wielokrotny wybór · jakość 80/100

Odpowiedzi: Bieganie lub nordic walking / Ćwiczenia na siłowni / Pływanie / Jazda na rowerze / Zajęcia fitness lub taniec / Inne (jakie?)

- **[drobny] Brak opcji „inne (jakie?)”** (O12; pewność 0,96; AI: metodolog + AI: rzecznik badanego). Opcja 'Inne' istnieje, ale brak pola tekstowego do wpisania odpowiedzi uniemożliwia zebranie danych o nieprzewidzianych formach aktywności, co ogranicza kompletność danych. *Dowód:* Inne (jakie?) *Poprawka:* Dodaj pole tekstowe obok opcji 'Inne', aby respondent mógł wpisać konkretną formę aktywności.

Propozycja: Jakie formy aktywności fizycznej podejmowałeś(-aś) w ciągu ostatnich 7 dni? (Można zaznaczyć wiele odpowiedzi)

Nowe odpowiedzi: Bieganie lub nordic walking / Ćwiczenia na siłowni / Pływanie / Jazda na rowerze / Zajęcia fitness lub taniec / Inne (proszę wpisać)

Uzasadnienie: Dodano pole tekstowe do opcji 'Inne', aby umożliwić respondentom podanie nieprzewidzianych form aktywności, co zwiększa kompletność i wiarygodność danych.

Q04 (nr w dokumencie A4)

Na ile zgadzasz się z poniższymi stwierdzeniami?

cz. 1 · macierz · jakość 10/100

Odpowiedzi: Zdecydowanie się nie zgadzam / Raczej się nie zgadzam / Ani się zgadzam, ani nie zgadzam / Raczej się zgadzam / Zdecydowanie się zgadzam

Wiersze: (1) Ruch sprawia mi przyjemność.; (2) Po ćwiczeniach czuję się lepiej.; (3) Aktywność fizyczna to dla mnie strata czasu.; (4) Łatwo znajduję czas na ruch w ciągu tygodnia.; (5) Ćwiczę, bo chcę dbać o zdrowie.

- **[ważny] Przeczenie lub podwójne przeczenie** (W03, wiersz 3; pewność 0,97; AI: metodolog + AI: rzecznik badanego). Stwierdzenie jest negatywne i merytorycznie odwrócone względem pozostałych (pozytywnych). W skali zgodności respondent musi wykonać dodatkową operację kognitywną (negację), co zwiększa ryzyko błędu i straightliningu. *Dowód:* Aktywność fizyczna to dla mnie strata czasu. *Poprawka:* Zmień twierdzenie na pozytywne: 'Aktywność fizyczna jest dla mnie wartościowa/cenna' lub odwróć skalę dla tego wiersza (jeśli platforma pozwala), ale lepiej unikać negacji w twierdzeniach.
- **[ważny] Format zgody zamiast skali specyficznej** (O08, wiersz 3; pewność 0,75; AI: metodolog). Użycie skali zgodności dla twierdzenia negatywnego ('strata czasu') jest mniej precyzyjne niż skala 'jak bardzo to prawda' lub 'jak bardzo to opisuje'. Skala zgodności jest podatna na akwiescencję. *Dowód:* Zdecydowanie się nie zgadzam / Raczej się nie zgadzam... *Poprawka:* Zmień skalę na 'W jakim stopniu to zdanie opisuje Twoje odczucia?' lub 'Jak bardzo to prawda?' dla wszystkich wierszy, aby ujednolicić format i zmniejszyć błąd metodologiczny.

- **[ważny] Niejednorodna bateria** (M06; pewność 0,73; AI: przegląd całości + AI: metodolog). Wiersze mierzą różne konstrukty: przyjemność (afekt), samopoczucie po ćwiczeniach (efekt fizjologiczny/psychologiczny), przestrzeganie wartości czasu (postawa), dostępność czasu (bariera), motywacja (cel). Sumowanie ich do jednej skali byłoby nieuzasadnione. *Dowód:* Grid z 5 wierszami o różnych aspektach aktywności *Poprawka:* Traktuj te pytania jako osobne wskaźniki lub podskale, a nie jako jedną skalę 'postawy wobec aktywności'. W analizie nie twórz sumarycznego wyniku bez weryfikacji jednowymiarowości.
- **[drobny] Duża macierz pytań** (B02; pewność 0,56; AI: przegląd całości). Krótka bateria 5 wierszy w formie siatki jest akceptowalna, ale na telefonie może być trudna w obsłudze. *Dowód:* Q04 [grid (one answer per row)] ... 5 wierszy *Poprawka:* Na urządzeniach mobilnych wyświetlać pytania jako pojedyncze sekwencje zamiast siatki.

Propozycja: Na ile poniższe stwierdzenia opisują Twoją sytuację?

Nowe odpowiedzi: Wcale nie opisują / Raczej nie opisują / Trudno powiedzieć / Ani tak, ani nie / Raczej opisują / W pełni opisują

Nowe wiersze: (1) Ruch sprawia mi przyjemność.; (2) Po ćwiczeniach czuję się lepiej.; (3) Aktywność fizyczna to dla mnie strata czasu.; (4) Łatwo znajduję czas na ruch w ciągu tygodnia.; (5) Ćwiczę, bo chcę dbać o zdrowie.

Uzasadnienie: Zmieniono skalę z 'zgadzam się/nie zgadzam' na 'opisują sytuację', co redukuje ryzyko akwiescencji i ułatwia przetwarzanie kognitywne, szczególnie dla twierdzeń negatywnych, które nie wymagają już dodatkowej operacji negacji w umyśle respondenta.

Q05 (nr w dokumencie B1)

Jak oceniasz swoje ogólne samopoczucie w ciągu ostatnich 30 dni?

cz. 2 · jednokrotny wybór · jakość 60/100

Odpowiedzi: 0 / 1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10

- **[drobny] Brak „nie wiem / nie dotyczy / wolę nie odpowiadać”** (O04; pewność 0,97; AI: metodolog + AI: rzecznik badanego + reguła). Brak opcji 'Nie dotyczy' lub 'Wolę nie odpowiadać' może zmusić respondentów, którzy nie chcą lub nie mogą ocenić samopoczucia, do losowego zaznaczenia liczby, co wprowadza szum do danych. *Dowód:* Zaznacz jedną liczbę na skali od 0 do 10. *Poprawka:* Dodaj opcję 'Wolę nie odpowiadać' lub 'Nie dotyczy' na końcu listy.
- **[drobny] Efekty kolejności odpowiedzi** (O09; pewność 0,93; AI: metodolog + AI: rzecznik badanego + reguła). Długa lista 11 opcji może powodować efekt primacy (wybieranie początkowych wartości) lub fatigue. W ankietach mobilnych może to być trudne do klikania. *Dowód:* 0 / 1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 *Poprawka:* Rozważ użycie suwaka (slider) lub uproszczenie skali do 5-7 punktów, jeśli nie jest wymagana precyzja 0-10, lub zapewnij łatwy interfejs do klikania.
- **[drobny] Efekt sufitowy lub podłogowy** (M05; pewność 0,60; AI: przegląd całości). Skala 0-10 może prowadzić do efektu sufitowego u studentów, jeśli większość jest zadowolona. *Dowód:* Q05 Jak oceniasz swoje ogólne samopoczucie ... 0 – bardzo złe 10 – bardzo dobre *Poprawka:* Upewnić się, że skala jest dobrze zinterpretowana; ewentualnie rozważyć 5-stopniową dla lepszej dystrybucji.

Propozycja: Jak oceniasz swoje ogólne samopoczucie w ciągu ostatnich 30 dni?

Nowe odpowiedzi: 0 - Bardzo złe / 1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 - Bardzo dobre / Nie dotyczy / Wolę nie odpowiadać

Uzasadnienie: Dodano opcję 'Nie dotyczy / Wolę nie odpowiadać', aby uniknąć zmuszania respondentów do losowego zaznaczania wartości, co wprowadzałoby szum do danych, oraz upewniono się, że etykiety skali są widoczne dla wszystkich opcji.

Q06 (nr w dokumencie B2)

Jak często w ciągu ostatnich 2 tygodni czuleś(-aś) się zmęczony(-a) lub miałeś(-aś) mało energii?

cz. 2 · jednokrotny wybór · jakość 10/100

Odpowiedzi: Wcale / Kilka dni / Więcej niż połowę dni / Prawie codziennie / Wolę nie odpowiadać

- **[ważny]** Dwa pytania w jednym (W02; pewność 0,97; AI: metodolog + AI: rzecznik badanego + reguła). Pytanie łączy dwa różne stany: zmęczenie (fizyczne/psychiczne) i brak energii. Respondent może czuć się zmęczony, ale mieć energię, lub odwrotnie. Odpowiedź ‘Często’ jest niejednoznaczna. *Dowód:* czułeś(-aś) się zmęczony(-a) lub miałeś(-aś) mało energii? *Poprawka:* Rozdziel na dwa pytania: ‘Jak często czułeś(-aś) się zmęczony(-a)?’ oraz ‘Jak często miałeś(-aś) mało energii?’ lub zmień formułkę na ‘czułeś(-aś) się wyczerpany(-a) lub pozbawiony(-a) energii’ (jeden stan).
- **[ważny]** Nieostre określenia (W04; pewność 0,97; AI: metodolog + AI: rzecznik badanego + reguła). Słowo ‘mało’ jest względne i nieprecyzyjne. Różni respondenci mają różne progi postrzegania ‘niewielkiej ilości’ energii. *Dowód:* mało energii *Poprawka:* Użyj standardowych, zweryfikowanych sformułowań, np. ‘czułeś(-aś) się wyczerpany(-a)’ lub ‘czułeś(-aś) brak sił’, które są lepiej zdefiniowane w literaturze psychometrycznej.
- **[drobny]** Brak „nie wiem / nie dotyczy / wolę nie odpowiadać” (O04; pewność 0,60; AI: rzecznik badanego). Obecność tej opcji jest dobra, ale warto upewnić się, że jest widoczna i zrozumiała. *Dowód:* Wolę nie odpowiadać *Poprawka:* Upewnij się, że opcja ta jest wyraźnie oznaczona i dostępna dla wszystkich respondentów.
- **[drobny]** Brak opcji „inne (jakie?)” (O12; pewność 0,50; reguła). Lista odpowiedzi nie ma opcji „Inne (jakie?)” — może nie obejmować wszystkich sytuacji. *Poprawka:* Dodaj „Inne (jakie?)” z polem do wpisania.

Propozycja: Jak często w ciągu ostatnich 2 tygodni odczuwałeś(-aś) zmęczenie?

Nowe odpowiedzi: Wcale / W kilka dni / W więcej niż połowie dni / Prawie codziennie / Wolę nie odpowiadać

Uzasadnienie: Rozdzielono pytanie dwuelementowe na ‘zmęczenie’, usuwając nieprecyzyjne ‘mało energii’, co zwiększa precyzję pomiaru i redukuje dwuznaczność odpowiedzi, zachowując przy tym opcję ‘Wolę nie odpowiadać’ dla etykiety.

Q07 (nr w dokumencie B3)

Uzereguj poniższe powody podejmowania aktywności od najważniejszego (1) do najmniej ważnego (4).

cz. 2 · ranking · jakość 20/100

Odpowiedzi: zdrowie / wygląd / kontakty towarzyskie / odstresowanie się

- **[ważny]** Niewłaściwy format odpowiedzi (O10; pewność 0,75; AI: metodolog). Ranking wymusza arbitralny wybór, nawet gdy powody są równoważne. To obciąża kognitywnie respondenta i może nie odzwierciedlać rzeczywistych motywacji (np. dwa główne powody są równie ważne). *Dowód:* Uzereguj poniższe powody... od najważniejszego (1) do najmniej ważnego (4). *Poprawka:* Zmień na ‘wybierz do 2 najważniejszych powodów’ lub poproś o ocenę każdego z osobna (np. skala 1-5), co jest mniej obciążające i bardziej miarodajne.
- **[drobny]** Efekty kolejności odpowiedzi (O09; pewność 0,75; AI: metodolog). Brak rotacji kolejności powodów może prowadzić do efektu pozycji (pierwsze opcje są częściej wybierane jako ‘najważniejsze’ w rankingach). *Dowód:* zdrowie / wygląd / kontakty towarzyskie / odstresowanie się *Poprawka:* Zaimplementuj losową kolejność (randomizację) opcji w rankingu dla różnych respondentów.
- **[drobny]** Luki, niepełna lista odpowiedzi (O02; pewność 0,60; AI: rzecznik badanego). Lista może nie obejmować wszystkich ważnych powodów, np. ‘relaks’ lub ‘zdrowie psychiczne’. *Dowód:* zdrowie / wygląd / kontakty towarzyskie / odstresowanie się *Poprawka:* Rozważ dodanie opcji ‘Inne’ lub upewnij się, że lista jest wyczerpująca.

Propozycja: Proszę ocenić, jak bardzo każdy z poniższych powodów motywuje Cię do podejmowania aktywności fizycznej, wybierając jedną opcję dla każdego wiersza.

Nowe odpowiedzi: Bardzo mało / Mało / Średnio / Dużo / Bardzo dużo

Nowe wiersze: (1) Zdrowie; (2) Wygląd; (3) Kontakty towarzyskie; (4) Odstresowanie się

Uzasadnienie: Zmieniono format z rankingowego na ocenę intensywności każdego powodu z osobna (Likert), co eliminuje wymuszony arbitralny wybór i obciążenie kognitywne związane z porównywaniem powodów, pozwalając na równoważną ocenę motywacji.

Q08 (nr w dokumencie M1)

Płeć

cz. 3 · jednokrotny wybór · jakość 80/100

Odpowiedzi: Kobieta / Mężczyzna / Inna / Wolę nie odpowiadać

- **[drobny] Język lub kategorie niewłączające** (E06; pewność 0,96; AI: przegląd całości + AI: metodolog + AI: rzecznik badanego). Opcje 'Inna' i 'Wolę nie odpowiadać' poprawiają inkluzję, co jest pozytywne, ale 'Inna' jest nieprecyzyjna. *Dowód:* Q08 Płeć ... Inna / Wolę nie odpowiadać *Poprawka:* Rozważyć dodanie pola 'Inne (wymień)' dla opcji 'Inna', jeśli liczba odpowiedzi będzie istotna.

Propozycja: Płeć

Nowe odpowiedzi: Kobieta / Mężczyzna / Inna / Nie chcę podać

Uzasadnienie: Zmieniono 'Wolę nie odpowiadać' na 'Nie chcę podać', co jest standardową, neutralną formą w polskich badaniach demograficznych, oraz zachowano 'Inna' jako opcję nieprecyzyjną, ale konieczną dla inkluzji, bez dodawania dodatkowych podkategorii, które mogłyby być mylące.

Q09 (nr w dokumencie M2)

Rok urodzenia:

cz. 3 · liczba · jakość 30/100

- **[ważny] Brak jednostki lub zakresu** (O13; pewność 0,97; AI: metodolog + AI: rzecznik badanego + reguła). Brak jednostki lub zakresu. Respondent może wpisać rok (np. 1995) lub wiek (np. 25), co zepsuje dane. Pole liczbowe bez instrukcji jest ryzykowne. *Dowód:* Rok urodzenia: *Poprawka:* Dodaj instrukcję: 'Podaj rok urodzenia (np. 1995)' lub zmień na 'Wiek w latach' z polem tekstowym i walidacją.
- **[drobny] Niespójna terminologia** (F04; pewność 0,60; AI: przegląd całości). Pytanie o rok urodzenia jest precyzyjne, ale w kontekście studentów może być mniej użyteczne niż 'Rok studiów'. *Dowód:* Q09 Rok urodzenia: *Poprawka:* Upewnić się, że zmienna 'Rok studiów' (Q10) pokrywa wiek respondentów adekwatnie.

Propozycja: W którym roku urodziłeś(-aś)?

Uzasadnienie: Dodano słowo 'roku' do pytania, aby jednoznacznie wskazać, że respondent ma wpisać rok (np. 1990), a nie wiek, co zapobiega błędom w kodowaniu danych.

Q10 (nr w dokumencie M3)

Rok studiów

cz. 3 · jednokrotny wybór · jakość 20/100

Odpowiedzi: I / II / III / IV / V / studia doktoranckie

- **[ważny] Luki, niepełna lista odpowiedzi** (O02; pewność 0,97; AI: metodolog + AI: rzecznik badanego). Brakuje opcji dla studentów studiów niestacjonarnych, zaocznych lub tych, którzy ukończyli licencjat/inżyniera, a kontynuują naukę w innej formie, co uniemożliwia im udzielenie poprawnej odpowiedzi. *Dowód:* I / II / III / IV / V / studia doktoranckie *Poprawka:* Dodać opcję 'Inne (proszę określić)' lub rozszerzyć listę o 'studia niestacjonarne/zaoczne' oraz 'po ukończeniu licencjatu/inżyniera'.
- **[drobny] Brak „nie wiem / nie dotyczy / wolę nie odpowiadać”** (O04; pewność 0,96; AI: metodolog + AI: rzecznik badanego). Brak opcji 'Nie dotyczy' lub 'Preferuję nie odpowiadać', co zmusza respondentów do wyboru niepasującej kategorii lub pominięcia pytania, co zaburza dane. *Dowód:* I / II / III / IV / V / studia doktoranckie *Poprawka:* Dodać opcję 'Nie dotyczy' (np. dla osób niebędących studentami) lub 'Preferuję nie odpowiadać'.
- **[drobny] Brak opcji „inne (jakie?)”** (O12; pewność 0,50; reguła). Lista odpowiedzi nie ma opcji „Inne (jakie?)” — może nie obejmować wszystkich sytuacji. *Poprawka:* Dodaj „Inne (jakie?)” z polem do wpisania.

Propozycja: Jaki jest Twój aktualny rok studiów?

Nowe odpowiedzi: I rok / II rok / III rok / IV rok / V rok / Studia doktoranckie / Inne (proszę wskazać) / Nie dotyczy / Preferuję nie odpowiadać

Uzasadnienie: Rozszerzono listę opcji o brakujące kategorie (np. studia doktoranckie, inne formy studiów) oraz dodano opcje 'Inne', 'Nie dotyczy' i 'Preferuję nie odpowiadać', aby zapewnić kompletność i uniknąć zmuszania respondentów do wyboru niepasującej odpowiedzi.

Uwagi dotyczące wszystkich pytań

- **[drobny] Format zgody zamiast skali specyficznej (O08).** Format „zgadzam się / nie zgadzam się” w 1 macierzy(ach) (Q04). Skale specyficzne dla treści (np. „jak często”, „jak bardzo”) są zwykle rzetelniejsze i mniej podatne na potakiwanie — chyba że to wystandaryzowany kwestionariusz.

5 Kafeterie i skale odpowiedzi

ID	Typ	Pkt	+/-	Zrównoważ.	Poza skalą	Kierunek	Odpowiedzi
Q01	nominalna	2					Tak / Nie
Q03	nominalna	6					Bieganie lub nordic walking / Ćwiczenia na siłowni / Pływanie / Jazda...
Q04	zgoda	5	2/1/2	tak		od negat.	Zdecydowanie się nie zgadzam / Raczej się nie zgadzam / Ani się zgadz...
Q05	liczbowa	11	0/0/0				0 / 1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10
Q06	częstość	4			Wolę nie odpowiadać		Wcale / Kilka dni / Więcej niż połowę dni / Prawie codziennie / Wolę ...
Q07	nominalna	4					zdrowie / wygląd / kontakty towarzyskie / odstresowanie się
Q08	nominalna	3			Wolę nie odpowiadać		Kobieta / Mężczyzna / Inna / Wolę nie odpowiadać
Q10	nominalna	6					I / II / III / IV / V / studia doktoranckie

ID	Waga	Problem	Komentarz
Q03	drobny	Brak opcji „inne (jakie?)”	Opcja 'Inne' istnieje, ale brak pola tekstowego do wpisania odpowiedzi uniemożliwia zebranie danych o nieprzewidywanych formach aktywności, co ogranicza kompletność danych.
Q04	ważny	Format zgody zamiast skali specyficznej	Użycie skali zgodności dla twierdzenia negatywnego ('strata czasu') jest mniej precyzyjne niż skala 'jak bardzo to prawda' lub 'jak bardzo to opisuje'. Skala zgodności jest podatna na akwiescencję.
Q05	drobny	Brak „nie wiem / nie dotyczy / wolę nie odpowiadać”	Brak opcji 'Nie dotyczy' lub 'Wolę nie odpowiadać' może zmusić respondentów, którzy nie chcą lub nie mogą ocenić samopoczucia, do losowego zaznaczenia liczby, co wprowadza szum do danych.
Q05	drobny	Efekty kolejności odpowiedzi	Długa lista 11 opcji może powodować efekt primacy (wybieranie początkowych wartości) lub fatigue. W ankietach mobilnych może to być trudne do klikania.
Q06	drobny	Brak „nie wiem / nie dotyczy / wolę nie odpowiadać”	Obecność tej opcji jest dobra, ale warto upewnić się, że jest widoczna i zrozumiała.
Q06	drobny	Brak opcji „inne (jakie?)”	Lista odpowiedzi nie ma opcji „Inne (jakie?)” — może nie obejmować wszystkich sytuacji.
Q07	ważny	Niewłaściwy format odpowiedzi	Ranking wymusza arbitralny wybór, nawet gdy powody są równoważne. To obciąża kognitywnie respondenta i może nie odzwierciedlać rzeczywistych motywacji (np. dwa główne powody są równie ważne).
Q07	drobny	Efekty kolejności odpowiedzi	Brak rotacji kolejności powodów może prowadzić do efektu pozycji (pierwsze opcje są częściej wybierane jako 'najważniejsze' w rankingach).
Q07	drobny	Luki, niepełna lista odpowiedzi	Lista może nie obejmować wszystkich ważnych powodów, np. 'relaks' lub 'zdrowie psychiczne'.
Q09	ważny	Brak jednostki lub zakresu	Brak jednostki lub zakresu. Respondent może wpisać rok (np. 1995) lub wiek (np. 25), co zepsuje dane. Pole liczbowe bez instrukcji jest ryzykowne.
Q10	ważny	Luki, niepełna lista odpowiedzi	Brakuje opcji dla studentów studiów niestacjonarnych, zaocznych lub tych, którzy ukończyli licencjat/inżyniera, a kontynuują naukę w innej formie, co uniemożliwia im udzielenie poprawnej odpowiedzi.
Q10	drobny	Brak „nie wiem / nie dotyczy / wolę nie odpowiadać”	Brak opcji 'Nie dotyczy' lub 'Preferuję nie odpowiadać', co zmusza respondentów do wyboru niepasującej kategorii lub pominięcia pytania, co zaburza dane.
Q10	drobny	Brak opcji „inne (jakie?)”	Lista odpowiedzi nie ma opcji „Inne (jakie?)” — może nie obejmować wszystkich sytuacji.
—	drobny	Format zgody zamiast skali specyficznej	Format „zgadzam się / nie zgadzam się” w 1 macierzy(ach) (Q04). Skale specyficzne dla treści (np. „jak często”, „jak bardzo”) są zwykle rzetelniejsze i mniej podatne na potakiwanie — chyba że to wystandaryzowany kwestio...

6 Kontrola jakości danych

Kryterium	Status	Pkt	Komentarz
Pytanie sprawdzające uwagę (w dłuższej ankiecie)	nie dotyczy	—	Krótki czas trwania nie wymaga kontroli uwagi. [ALL]
Pozycje odwrócone w bateriach	tak	25/25	W baterii Q04 występuje pozycjonowanie odwrócone bez oznaczenia. [Q04: Q04_3]
Pary sprawdzające spójność	nie dotyczy	—	Krótki czas trwania nie wymaga par spójnościowych. [ALL]
Pytanie o rzetelność / szczerłość odpowiedzi	nie dotyczy	—	Krótki czas trwania nie wymaga kontroli poważności. [ALL]

Wykryte elementy kontrolne

Brak.

Propozycje pytań kontrolnych do dodania

Typ	Po pytaniu	Treść	Oczekiwana odpowiedź	Uzasadnienie
pytanie z instrukcją	Q03	Proszę zaznaczyć tylko odpowiedź: Pływanie.	Pływanie	Element ten służy do identyfikacji respondentów, którzy nie czytają instrukcji ani nie uważnie zaznaczają odpowiedzi w pytaniach wielokrotnego wyboru, co jest częstym źródłem szumu w danych.
para spójności	Q02	Czy w ciągu ostatnich 7 dni nie podejmowałeś żadnej aktywności fizycznej trwającej co najmniej 30 minut?	Tak	Pozwala to zweryfikować spójność między odpowiedzią 'Nie' w Q01 a liczbą dni w Q02. Jeśli w Q01 jest 'Nie', a w Q02 wpisano liczbę >0, występuje sprzeczność logiczna.
pytanie o rzetelność	Q07	Czy uważasz, że regularna aktywność fizyczna jest korzystna dla zdrowia?	Tak	Proste pytanie o zgodność z normą społeczną, które wymaga minimalnego wysiłku poznawczego, służy jako filtr uwagi (attention check) po trudnym zadaniu rankingowym.

Do wykrywania nieuważnych odpowiedzi w danych z badania właściwego warto policzyć wskaźniki: najdłuższy ciąg identycznych odpowiedzi (*long-string*), zmienność odpowiedzi w obrębie osoby (*IRV*), odległość Mahalanobisa, spójność parzysta-nieparzysta oraz czas wypełniania (*speeding*, np. < 2 s na pozycję). Poniższy kod liczy je dla danych w formacie Respondent AI lub eksportu z platformy (kolumny = pozycje skal):

```
d <- read.csv("audit_data/items_values.csv", fileEncoding = "UTF-8") # or your platform export
items <- setdiff(names(d), c("persona_id", "style")) # scale items only
x <- as.matrix(d[items])
longstring <- apply(x, 1, function(r) max(rle(as.vector(r[!is.na(r)]))$lengths))
irv <- apply(x, 1, sd, na.rm = TRUE)
mahal <- if (nrow(x) > ncol(x) + 5) mahalanobis(x, colMeans(x, na.rm = TRUE), cov(x, use = "pairwise"))
flag <- longstring >= 0.8 * ncol(x) | irv < 0.3
data.frame(id = d[[1]], longstring, irv = round(irv, 2), mahal, flag)
```

7 Obciążenie i komfort badanego

Kryterium	Status	Pkt	Komentarz
Szacowany czas wypełniania	2,1 min	25/25	
Największa macierz (wiersze)	5	15/15	
Obowiązkowe pytania otwarte	brak danych	—	Dokument nie oznacza pytań obowiązkowych.
Odsetek pytań obowiązkowych	brak danych	—	Dokument nie oznacza pytań obowiązkowych.
Czytelność (FOG-PL)	7,3	15/15	
Wskaźnik postępu	nie dotyczy	—	Nie dotyczy dokumentu — zależy od platformy ankiety.

(cd.)

Kryterium	Status	Pkt	Komentarz
Najbardziej obciążona strona (pytania / pozycje)	nie dotyczy	—	Nie dotyczy dokumentu — podział na strony ustala się na platformie ankiety.

Czytelność pytań: FOG-PL = **7,3** (łatwy), średnio 7,5 słów w zdaniu, 11% wyrazów trudnych (4+ sylaby). Wartości do ok. 9 są zrozumiałe dla większości dorosłych; powyżej 12 tekst wymaga wykształcenia wyższego.

Najtrudniejsze językowo pytania

ID	FOG	Pytanie
Q09	20,8	Rok urodzenia:
Q07	20,0	Uszereguj poniższe powody podejmowania aktywności od najważniejszego (1) do najmniej ważnego (4).
Q03	12,5	Jakie formy aktywności podejmowałeś(-aś) w ciągu ostatnich 7 dni?
Q01	8,7	Czy w ciągu ostatnich 7 dni podejmowałeś(-aś) aktywność fizyczną trwającą łącznie co najmniej 30 minut dziennie?
Q02	8,0	W ile dniach w ciągu ostatnich 7 dni podejmowałeś(-aś) taką aktywność?
Q05	8,0	Jak oceniasz swoje ogólne samopoczucie w ciągu ostatnich 30 dni?

ID	Waga	Problem	Komentarz
Q04	drobny	Duża macierz pytań	Krótką baterię 5 wierszy w formie siatki jest akceptowalna, ale na telefonie może być trudna w obsłudze.
—	drobny	Zbyt długa ankieta	Różnica między deklarowanym czasem (8 min) a rzeczywistym (2 min) może zniechęcić lub wprowadzić w błąd.

8 Spójność i rzetelność pomiaru

Kryterium	Status	Pkt	Komentarz
Konstrukty mierzone kilkoma pozycjami	tak	25/25	Kluczowe konstrukty (postawy, samopoczucie) mierzone są wieloelementowo. [Q04, Q06, Q07]
Jednorodnie baterie	nie	0/20	Bateria Q04 łączy przyjemność z dostępem czasowym, co tworzy heterogeniczny konstruk. [Q04 Rows: (1) Ruch sprawia mi przyjemność ... (4) Łatwo znajduję czas]
Spójny format i kierunek skal	tak	15/15	
Brak problemów z kluczem (dane)	brak danych	—	
Pozycje różnicują badanych (dane)	brak danych	—	
Spójność wewnętrzna baterii (dane)	brak danych	—	

Przewidywana spójność odpowiedzi ludzi. Największym zagrożeniem dla wiarygodności danych jest bateria Q04, która łączy różne konstrukty (afektywne, poznawcze, behawioralne), co może obniżyć spójność wewnętrzną (alpha Cronbacha) i utrudniać interpretację jednego wskaźnika. Pytania Q05 i Q06 są pojedynczymi miernikami złożonych stanów (samopoczucie, zmęczenie), co czyni je podatnymi na błędy pomiaru i subiektywną interpretację, choć są to standardowe instrumenty w epidemiologii. Odpowiedzi w Q01-Q03 mogą być obciążone błędem pamięci (recall bias) ze względu na krótki okres odniesienia (7 dni) i brak konkretnych definicji 'aktywności fizycznej' (czy liczy się praca domowa, spacer?). Ryzyko znużenia (satisficing) jest umiarkowane, szczególnie w Q07 (ranking), które jest kognitywnie wymagające. Brak elementów kontrolnych zwiększa ryzyko wyłapania respondentów nieprzestrzegających instrukcji.

Przewidywane α dla ludzi z formuły Spearmana-Browna: $\alpha = k \cdot \bar{r} / (1 + (k-1) \cdot \bar{r})$, gdzie k to liczba pozycji, a $\bar{r}$ to typowa korelacja między pozycjami o danej jednorodności (wysoka $\approx 0,45$; średnia $\approx 0,30$; niska $\approx 0,15$).

Bateria	Konstrukt	k	Jednorodność	α ludzie (przew.)	α synt.	Odwrócone	Komentarz
Q04	Postawy i motywacja do aktywności fizycznej (postawy, odczucia po wysiłku, postrzegane bariery czasowe, instrumentalna motywacja zdrowotna)	5	średnia	0,68 (0,56–0,77)	—	Q04_3	Skala miesza konstrukty o różnej naturze psychometrycznej: emocjonalną (przyjemność, samopoczucie), poznawczą (strata czasu) i behawioralną/barierową (znajdowanie czasu). Elementy r1 i r2 są silnie s...

9 Struktura i przebieg

Kryterium	Status	Pkt	Komentarz
Wprowadzenie i instrukcja	tak	20/20	Wprowadzenie zawiera powitanie, cel, instrukcje i zgodę. [108 słów przed pierwszym pytaniem]
Zakończenie i podziękowanie	tak	10/10	Obecna jest strona zamykająca z podziękowaniami. [Dziękujemy za udział w badaniu! Jeśli chcesz poznać jego wyniki, napisz do nas na adres podany na początku ankiety.]
Logiczna kolejność, łagodne otwarcie, bloki tematyczne	tak	20/20	Pytania są pogrupowane tematycznie, demografia na końcu. [Część A. Aktywność fizyczna ... Część 3 ... Część 4]
Pytania drażliwe i metryczka bliżej końca	tak	20/20	Pytania demograficzne i wrażliwe umieszczone są na końcu. [Q08 Płeć, Q09 Rok urodzenia, Q10 Rok studiów]
Spójna terminologia	częściowo	8/15	Terminologia jest spójna, choć Q04 używa słowa 'Ruch' zamiast 'aktywność fizyczna'. [Q01: aktywność fizyczną ... Q03: formy aktywności ... Q04: Ruch]
Brak zbędnych powtórzeń	tak	15/15	Nie występują powtórzenia pytań. [ALL]

ID	Waga	Problem	Komentarz
Q01	ważny	Brak filtra	Pytanie filtrujące zakłada, że 'aktywność fizyczna' oznacza 30 min dziennie, co jest specyficzną definicją WHO, nie powszechną. Poprawka: Dodać opcję 'Nie wiem' lub wyjaśnić definicję aktywności, aby uniknąć błędnych odpowiedzi 'Nie'.
Q09	drobny	Niespójna terminologia	Pytanie o rok urodzenia jest precyzyjne, ale w kontekście studentów może być mniej użyteczne niż 'Rok studiów'. Poprawka: Upewnić się, że zmienna 'Rok studiów' (Q10) pokrywa wiek respondentów adekwatnie.
—	drobny	Niespójna terminologia	Kryterium spełnione częściowo: Spójna terminologia. Terminologia jest spójna, choć Q04 używa słowa 'Ruch' zamiast 'aktywność fizyczna'.

10 Technika, telefon i dostępność

Audyt dotyczy dokumentu, więc dostępności (WCAG), działania na telefonie i komunikatów walidacji nie sprawdzano – zależą od platformy, na której ankieta zostanie zaprogramowana. Ten wymiar nie wlicza się do indeksu jakości. Po zaprogramowaniu ankiety zaaudituj jej link: przejście na komputerze i telefonie z testem axe-core pokaże problemy techniczne.

11 Etyka, RODO i inkluzywność

Kryterium	Status	Pkt	Komentarz
Cel badania	tak	12/12	Cel badania jest jasno określony w tekście wprowadzającym. [...gii Zdrowia Uniwersytetu Przykładowego. Celem badania jest poznanie związku między aktywnością fizyczną a samo...]
Kto prowadzi badanie (instytucja, zespół)	tak	8/8	Wprowadzenie zawiera nazwę zespołu badawczego oraz instytucji. [...aństwo, zapraszamy do udziału w badaniu prowadzonym przez Zespół Psychologii Zdrowia Uniwersytetu Przykładowego...]

(cd.)

Kryterium	Status	Pkt	Komentarz
Kontakt do badaczy	tak	10/10	Adres e-mail do kontaktu z zespołem badawczym jest podany. [...nie w celach naukowych, zgodnie z RODO. Kontakt z zespołem badawczym: badanie.aktywnosc@example.edu. Klikni...]
Dobrowolność udziału	tak	10/10	Zaznaczono, że udział jest dobrowolny. [...ie około 8 minut. Udział w badaniu jest dobrowolny i anonimowy — nie pytamy o imię, nazwisko ani adres e-mail...]
Anonimowość lub poufność odpowiedzi	tak	12/12	Gwarancja anonimowości i brak pytań o dane identyfikujące. [...nut. Udział w badaniu jest dobrowolny i anonimowy — nie pytamy o imię, nazwisko ani adres e-mail. W każdej c...]
Informacja o przetwarzaniu danych (RODO)	tak	14/14	Wzmianka o zgodności z RODO oraz podanie administratora danych. [...nianie ankiety bez podawania przyczyny. Administratorem danych jest Uniwersytet Przykładowy; dane będą przetwa...]
Możliwość przerwania w każdej chwili	tak	8/8	Uczestnik ma prawo do przerwania badania w dowolnym momencie. [...tamy o imię, nazwisko ani adres e-mail. W każdej chwili możesz przerwać wypełnianie ankiety bez podawania przy...]
Wyraźne wyrażenie zgody	tak	12/12	Obecna jest klauzula wyrażenia zgody poprzez kliknięcie przycisku. [...(albo wypełnienie tej ankiety) oznacza zgodę na udział w badaniu. Badanie jest przeznaczone dla osób peł...]
Przewidywany czas wypełniania	tak	6/6	Podano szacowany czas, ale jest on rozbieżny z modelem (8 vs 2 min). [...entów. Wypełnienie ankiety zajmie około 8 minut. Udział w badaniu jest dobrowolny i anonimowy — nie pytam...]
Warunek wieku / grupa docelowa	tak	8/8	Określono grupę docelową jako osoby pełnoletnie. [...niu. Badanie jest przeznaczone dla osób pełnoletnich. Część A. Aktywność fizyczna]

ID	Waga	Problem	Komentarz
Q08	drobny	Język lub kategorie niewłaściwe	Opcje 'Inna' i 'Wolę nie odpowiadać' poprawiają inkluzję, co jest pozytywne, ale 'Inna' jest nieprecyzyjna. Poprawka: Rozważyć dodanie pola 'Inne (wymień)' dla opcji 'Inna', jeśli liczba odpowiedzi będzie istotna.

12 Gotowość do analizy danych

Proponowana książka kodów: nazwy zmiennych zgodne z identyfikatorami pytań (wiersze macierzy jako Qnn_r), wartości kodowane od 1 w kolejności odpowiedzi. Pytania otwarte wymagają kodowania treści; pozycje odwrócone trzeba przekodować przed liczeniem wyników.

Zmienna	Etykieta	Typ	Wartości
Q01	Czy w ciągu ostatnich 7 dni podejmowałeś(-aś) aktywność fizyczną trwającą łącznie co najm...	nominalna	1 = Tak; 2 = Nie
Q02	W ile dniach w ciągu ostatnich 7 dni podejmowałeś(-aś) taką aktywność?	liczba	
Q03_1	Jakie formy aktywności podejmowałeś(-aś) w ciągu ...: Bieganie lub nordic walking	wielokrotny (0/1)	0 = nie, 1 = tak
Q03_2	Jakie formy aktywności podejmowałeś(-aś) w ciągu ...: Ćwiczenia na siłowni	wielokrotny (0/1)	0 = nie, 1 = tak
Q03_3	Jakie formy aktywności podejmowałeś(-aś) w ciągu ...: Pływanie	wielokrotny (0/1)	0 = nie, 1 = tak
Q03_4	Jakie formy aktywności podejmowałeś(-aś) w ciągu ...: Jazda na rowerze	wielokrotny (0/1)	0 = nie, 1 = tak
Q03_5	Jakie formy aktywności podejmowałeś(-aś) w ciągu ...: Zajęcia fitness lub taniec	wielokrotny (0/1)	0 = nie, 1 = tak
Q03_6	Jakie formy aktywności podejmowałeś(-aś) w ciągu ...: Inne (jakie?)	wielokrotny (0/1)	0 = nie, 1 = tak
Q04_1	Ruch sprawia mi przyjemność.	porządkowa	1 = Zdecydowanie się nie zgadzam; 2 = Raczej się nie zgadzam; 3 = Ani się zgadzam, ani nie zgadzam; 4 = Raczej się zgad...
Q04_2	Po ćwiczeniach czuję się lepiej.	porządkowa	1 = Zdecydowanie się nie zgadzam; 2 = Raczej się nie zgadzam; 3 = Ani się zgadzam, ani nie zgadzam; 4 = Raczej się zgad...
Q04_3	Aktywność fizyczna to dla mnie strata czasu.	porządkowa	1 = Zdecydowanie się nie zgadzam; 2 = Raczej się nie zgadzam; 3 = Ani się zgadzam, ani nie zgadzam; 4 = Raczej się zgad...

(cd.)

Zmienna	Etykieta	Typ	Wartości
Q04_4	Łatwo znajduję czas na ruch w ciągu tygodnia.	porządkowa	1 = Zdecydowanie się nie zgadzam; 2 = Raczej się nie zgadzam; 3 = Ani się zgadzam, ani nie zgadzam; 4 = Raczej się zgad...
Q04_5	Ćwiczę, bo chcę dbać o zdrowie.	porządkowa	1 = Zdecydowanie się nie zgadzam; 2 = Raczej się nie zgadzam; 3 = Ani się zgadzam, ani nie zgadzam; 4 = Raczej się zgad...
Q05	Jak oceniasz swoje ogólne samopoczucie w ciągu ostatnich 30 dni?	porządkowa	1 = 0; 2 = 1; 3 = 2; 4 = 3; 5 = 4; 6 = 5; 7 = 6; 8 = 7; 9 = 8; 10 = 9; 11 = 10
Q06	Jak często w ciągu ostatnich 2 tygodni czułeś(-aś) się zmęczony(-a) lub miałeś(-aś) mało ...	nominalna	1 = Wcale; 2 = Kilka dni; 3 = Więcej niż połowę dni; 4 = Prawie codziennie; 5 = Wolę nie odpowiadać
Q07	Uzereguj poniższe powody podejmowania aktywności od najważniejszego (1) do najmniej ważn...	nominalna	1 = zdrowie; 2 = wygląd; 3 = kontakty towarzyskie; 4 = odstresowanie się
Q08	Płeć	nominalna	1 = Kobieta; 2 = Mężczyzna; 3 = Inna; 4 = Wolę nie odpowiadać
Q09	Rok urodzenia:	liczba	
Q10	Rok studiów	nominalna	1 = I; 2 = II; 3 = III; 4 = IV; 5 = V; 6 = studia doktoranckie

Szkic kodu R do liczenia wyników baterii i ich rzetelności na danych z badania (psych::alpha z kluczem pozycji odwróconych):

```
library(psych)
d <- read.csv("dane.csv", fileEncoding = "UTF-8") # eksport z platformy / platform export
# Q04
b1_items <- c("Q04_1", "Q04_2", "Q04_3", "Q04_4", "Q04_5")
b1_keys <- ifelse(b1_items %in% c("Q04_3"), -1, 1)
print(alpha(d[b1_items], keys = b1_keys))
d$b1 <- rowMeans(reverse.code(b1_keys, d[b1_items]), na.rm = TRUE)
```

13 Aneksy

13.1 Pełna lista problemów

ID	Kod	Problem	Waga	Pewn.	Źródło	Komentarz
Q01	W04	Nieostre określenia	ważny	0,94	AI: metodolog + AI: rzecznik badanego	Sformułowanie 'łącznie co najmniej 30 minut dziennie' jest mylące: sugeruje, że 30 minut to próg konieczny do uznania aktywności za 'tak', podczas gdy w badaniach epidemiologicznych często liczy się ...
Q01	F03	Brak filtra	ważny	0,75	AI: przegląd całości	Pytanie filtrujące zakłada, że 'aktywność fizyczna' oznacza 30 min dziennie, co jest specyficzną definicją WHO, nie powszechną.
Q01	W09	Presja aprobaty społecznej	ważny	0,75	AI: metodolog	Pytanie o aktywność fizyczną jest silnie obciążone presją społecznej pożądanoci; odpowiedź 'Tak' jest pożądana społecznie, co może prowadzić do zaniżania bezruchu.
Q02	W07	Trudne przypominanie lub szacowanie	ważny	0,97	AI: metodolog + AI: rzecznik badanego	Wymaga precyzyjnej pamięci i liczenia dni z poprzedniego tygodnia, co jest trudne i obciążone błędem recall (Tourangeau). Respondenci często zgadują lub stosują heurystykę.
Q03	O12	Brak opcji „inne (jakie?)”	drobny	0,96	AI: metodolog + AI: rzecznik badanego	Opcja 'Inne' istnieje, ale brak pola tekstowego do wpisania odpowiedzi uniemożliwia zebranie danych o nieprzewidzianych formach aktywności, co ogranicza kompletność danych.
Q04/3	W03	Przeczenie lub podwójne przeczenie	ważny	0,97	AI: metodolog + AI: rzecznik badanego	Stwierdzenie jest negatywne i merytorycznie odwrócone względem pozostałych (pozytywnych). W skali zgodności respondent musi wykonać dodatkową operację kognitywną (negację), co zwiększa ryzyko błędu i...
Q04/3	O08	Format zgody zamiast skali specyficznej	ważny	0,75	AI: metodolog	Użycie skali zgodności dla twierdzenia negatywnego ('strata czasu') jest mniej precyzyjne niż skala 'jak bardzo to prawda' lub 'jak bardzo to opisuje'. Skala zgodności jest podatna na akwiescencję.
Q04	M06	Niejednorodna bateria	ważny	0,73	AI: przegląd całości + AI: metodolog	Wiersze mierzą różne konstrukty: przyjemność (afekt), samopoczucie po ćwiczeniach (efekt fizjologiczny/psychologiczny), postrzeganie wartości czasu (postawa), dostępność czasu (bariera), motywacja (c...
Q04	B02	Duża macierz pytań	drobny	0,56	AI: przegląd całości	Krótką baterię 5 wierszy w formie siatki jest akceptowalna, ale na telefonie może być trudna w obsłudze.
Q05	O04	Brak „nie wiem / nie dotyczy / wolę nie odpowiadać”	drobny	0,97	AI: metodolog + AI: rzecznik badanego + reguła	Brak opcji 'Nie dotyczy' lub 'Wolę nie odpowiadać' może zmusić respondentów, którzy nie chcą lub nie mogą ocenić samopoczucia, do losowego zaznaczenia liczby, co wprowadza szum do danych.

(cd.)

ID	Kod	Problem	Waga	Pewn.	Źródło	Komentarz
Q05	O09	Efekty kolejności odpowiedzi	drobny	0,93	AI: metodolog + AI: rzecznik badanego + reguła	Długa lista 11 opcji może powodować efekt primacy (wybieranie początkowych wartości) lub fatigue. W ankietach mobilnych może to być trudne do klikania.
Q05	M05	Efekt sufitowy lub podłogowy	drobny	0,60	AI: przegląd całości	Skala 0-10 może prowadzić do efektu sufitowego u studentów, jeśli większość jest zadowolona.
Q06	W02	Dwa pytania w jednym	ważny	0,97	AI: metodolog + AI: rzecznik badanego + reguła	Pytanie łączy dwa różne stany: zmęczenie (fizyczne/psychiczne) i brak energii. Respondent może czuć się zmęczony, ale mieć energię, lub odwrotnie. Odpowiedź 'Często' jest niejednoznaczna.
Q06	W04	Nieostre określenia	ważny	0,97	AI: metodolog + AI: rzecznik badanego + reguła	Słowo 'mało' jest względne i nieprecyzyjne. Różni respondenci mają różne progi postrzegania 'niewielkiej ilości' energii.
Q06	O04	Brak „nie wiem / nie dotyczy / wolę nie odpowiadać”	drobny	0,60	AI: rzecznik badanego	Obecność tej opcji jest dobra, ale warto upewnić się, że jest widoczna i zrozumiała.
Q06	O12	Brak opcji „inne (jakie?)”	drobny	0,50	reguła	Lista odpowiedzi nie ma opcji „Inne (jakie?)” — może nie obejmować wszystkich sytuacji.
Q07	O10	Niewłaściwy format odpowiedzi	ważny	0,75	AI: metodolog	Ranking wymusza arbitralny wybór, nawet gdy powody są równoważne. To obciąża kognitywnie respondenta i może nie odzwierciedlać rzeczywistych motywacji (np. dwa główne powody są równie ważne).
Q07	O09	Efekty kolejności odpowiedzi	drobny	0,75	AI: metodolog	Brak rotacji kolejności powodów może prowadzić do efektu pozycji (pierwsze opcje są częściej wybierane jako 'najważniejsze' w rankingach).
Q07	O02	Luki, niepełna lista odpowiedzi	drobny	0,60	AI: rzecznik badanego	Lista może nie obejmować wszystkich ważnych powodów, np. 'relaks' lub 'zdrowie psychiczne'.
Q08	E06	Język lub kategorie niewłaściwe	drobny	0,96	AI: metodolog + AI: rzecznik badanego	Opcje 'Inna' i 'Wolę nie odpowiadać' poprawiają inkluzję, co jest pozytywne, ale 'Inna' jest nieprecyzyjna.
Q09	O13	Brak jednostki lub zakresu	ważny	0,97	AI: metodolog + AI: rzecznik badanego + reguła	Brak jednostki lub zakresu. Respondent może wpisać rok (np. 1995) lub wiek (np. 25), co zepsuje dane. Pole liczbowe bez instrukcji jest ryzykowne.
Q09	F04	Niespójna terminologia	drobny	0,60	AI: przegląd całości	Pytanie o rok urodzenia jest precyzyjne, ale w kontekście studentów może być mniej użyteczne niż 'Rok studiów'.
Q10	O02	Luki, niepełna lista odpowiedzi	ważny	0,97	AI: metodolog + AI: rzecznik badanego	Brakuje opcji dla studentów studiów niestacjonarnych, zaocznych lub tych, którzy ukończyli licencjat/inżyniera, a kontynuują naukę w innej formie, co uniemożliwia im udzielenie poprawnej odpowiedzi.
Q10	O04	Brak „nie wiem / nie dotyczy / wolę nie odpowiadać”	drobny	0,96	AI: metodolog + AI: rzecznik badanego	Brak opcji 'Nie dotyczy' lub 'Preferuję nie odpowiadać', co zmusza respondentów do wyboru niepasującej kategorii lub pominięcia pytania, co zaburza dane.
Q10	O12	Brak opcji „inne (jakie?)”	drobny	0,50	reguła	Lista odpowiedzi nie ma opcji „Inne (jakie?)” — może nie obejmować wszystkich sytuacji.
—	M06	Niejednorodna bateria	ważny	0,75	AI: przegląd całości	Kryterium niespełnione: Jednorodne baterie. Bateria Q04 łączy przyjemność z dostępem czasowym, co tworzy heterogeniczny konstruk.
—	B01	Zbyt długa ankieta	drobny	0,90	AI: przegląd całości	Rozbieżność między deklarowanym czasem (8 min) a rzeczywistym (2 min) może zniechęcić lub wprowadzić w błąd.
—	F04	Niespójna terminologia	drobny	0,75	AI: przegląd całości	Kryterium spełnione częściowo: Spójna terminologia. Terminologia jest spójna, choć Q04 używa słowa 'Ruch' zamiast 'aktywność fizyczna'.
—	O08	Format zgody zamiast skali specyficznej	drobny	0,40	reguła	Format „zgadzam się / nie zgadzam się” w 1 macierzy(ach) (Q04). Skale specyficzne dla treści (np. „jak często”, „jak bardzo”) są zwykle rzetelniejsze i mniej podatne na potakiwanie — chyba że to wyst...

13.2 Kandydaci odrzuceni przez panel AI

Problemy wskazane przez reguły z niską pewnością, których żaden ekspert AI nie potwierdził. Nie wliczają się do wyników; warto je przejrzeć.

Brak.

13.3 Słownik kodów problemów

Kod	Wymiar	Problem	Komentarz
W01	Treść pytań	Pytanie sugerujące	Sformułowanie podsuwa „właściwą” odpowiedź (np. „Czy zgadzasz się, że...”, wartościujące przymiotniki, powoływanie się na autorytety lub większość).

(cd.)

Kod	Wymiar	Problem	Komentarz
W02	Treść pytań	Dwa pytania w jednym	Pytanie lub twierdzenie dotyczy dwóch rzeczy naraz; nie wiadomo, do której odnosi się odpowiedź.
W03	Treść pytań	Przeczenie lub podwójne przeczenie	Przeczenie w twierdzeniu ocenianym na skali zgody („nie zgadzam się, że nie...”) utrudnia zrozumienie.
W04	Treść pytań	Nieostre określenia	Słowa typu „często”, „regularnie”, „ostatnio”, „dużo” każdy rozumie inaczej.
W05	Treść pytań	Brak okresu odniesienia	Pytanie o zachowanie lub częstotliwość bez wskazania okresu (np. „w ciągu ostatnich 7 dni”).
W06	Treść pytań	Żargon, skróty, trudne słowa	Terminy specjalistyczne lub skróty, których część badanych może nie znać.
W07	Treść pytań	Trudne przypominanie lub szacowanie	Wymaga przypomnienia lub policzenia czegoś, czego badany zwykle nie pamięta (długi okres, dokładne liczby).
W08	Treść pytań	Ukryte założenie	Pytanie zakłada coś, co nie musi dotyczyć badanego (np. „Jak bardzo lubisz...”, „Dlaczego przestałeś...”).
W09	Treść pytań	Presja aprobaty społecznej	Jedną z odpowiedzi jest wyraźnie „lepiej widziana”, co zachęca do upiększania.
W10	Treść pytań	Pytanie drażliwe bez złagodzenia	Temat intymny lub obciążający zadany wprost, bez wprowadzenia, normalizacji i możliwości pominięcia.
W11	Treść pytań	Zbyt długie lub złożone	Długie zdanie, wiele warunków lub zdań podrzędnych; badany gubi sens przed odpowiedzią.
W12	Treść pytań	Niejednoznaczne	Pytanie można zrozumieć na kilka sposobów (niejasne słowo, niejasny podmiot lub zakres).
W13	Treść pytań	Pytanie hipotetyczne	Pytanie o zachowanie w wyobrażonej sytuacji; deklaracje słabo przewidują rzeczywiste zachowania.
W14	Treść pytań	Brak lub niejasna instrukcja	Nie wiadomo, ile odpowiedzi zaznaczyć, w jakim formacie wpisać odpowiedź lub czego dotyczy skala.
W15	Treść pytań	Błędy językowe	Literówki, błędy gramatyczne lub interpunkcyjne obniżające wiarygodność i zrozumiałość.
W16	Treść pytań	Zakłada wiedzę badanego	Pytanie wymaga wiedzy lub informacji, których badany może nie mieć.
O01	Kafeteria i skale	Nakładające się kategorie	Te same wartości mieszczą się w dwóch odpowiedziach (np. „18–25” i „25–35”).
O02	Kafeteria i skale	Luki, niepełna lista odpowiedzi	Części badanych nie pasuje żadna odpowiedź (luka w przedziałach, brakujące kategorie).
O03	Kafeteria i skale	Nierównoważona skala	Więcej odpowiedzi pozytywnych niż negatywnych (lub odwrotnie) przesuwają wyniki.
O04	Kafeteria i skale	Brak „nie wiem / nie dotyczy / wolę nie odpowiadać”	Badany bez zdania, którego pytanie nie dotyczy, lub który nie chce odpowiadać, musi zgadywać.
O05	Kafeteria i skale	Odpowiedzi nie pasują do pytania	Kategorie odpowiedzi odpowiadają na inne pytanie niż zadane (np. pytanie „jak często”, skala zgody).
O06	Kafeteria i skale	Nieodpowiednia liczba punktów skali	Zbyt mało punktów (słabe różnicowanie) albo zbyt wiele (pozorna precyzja).
O07	Kafeteria i skale	Niejasne etykiety punktów skali	Etykiety nieostre, nierównomierne lub tylko na końcach skali tam, gdzie potrzeba pełnych opisów.
O08	Kafeteria i skale	Format zgody zamiast skali specyficznej	Skale „zgadzam się / nie zgadzam się” sprzyjają potakiwaniu; skale specyficzne dla pytań są zwykle rzetelniejsze (Saris i in., 2010), o ile nie jest to wystandaryzowane narzędzie.
O09	Kafeteria i skale	Efekty kolejności odpowiedzi	Długa lista bez rotacji — pierwsze pozycje są wybierane częściej (efekt pierwszeństwa).
O10	Kafeteria i skale	Niewłaściwy format odpowiedzi	Np. jednokrotny wybór tam, gdzie pasuje kilka odpowiedzi, albo pole tekstowe zamiast listy lub liczby.
O11	Kafeteria i skale	Niespójny kierunek skal	Skale w ankiecie raz zaczynają się od wartości negatywnych, raz od pozytywnych — rośnie ryzyko pomyłek.
O12	Kafeteria i skale	Brak opcji „inne (jakie?)”	Lista nominalna nie obejmuje wszystkich możliwości i nie pozwala dopisać własnej.
O13	Kafeteria i skale	Brak jednostki lub zakresu	Pole liczbowe bez jednostki (godziny? minuty? złote?) lub bez sensownego zakresu wartości.
C01	Kontrola jakości danych	Brak pytań sprawdzających uwagę	W dłuższej ankiecie brakuje pytania z instrukcją (np. „zaznacz »raczej się nie zgadzam«”).
C02	Kontrola jakości danych	Wadliwe pytanie kontrolne	Pytanie kontrolne jest zbyt oczywiste, jednoznaczne, obraźliwe albo źle umieszczone.
C03	Kontrola jakości danych	Brak pozycji odwróconych	Wszystkie twierdzenia baterii są sformułowane w tym samym kierunku — nie da się wykryć potakiwania.
C04	Kontrola jakości danych	Brak par sprawdzających spójność	Brak pytań o to samo zadanych dwukrotnie innymi słowami, które pozwoliłyby wykryć losowe odpowiedzi.
C05	Kontrola jakości danych	Brak pytania o rzetelność odpowiedzi	Brak pytania na końcu, czy badany odpowiadał uważnie i szczerze (np. „Czy możemy użyć Twoich danych?”).
B01	Obciążenie i komfort	Zbyt długa ankieta	Przewidywany czas wypełniania jest długi — rośnie odsetek porzuceń i odpowiedzi „byłoby jakich”.
B02	Obciążenie i komfort	Duża macierz pytań	Wiele wierszy na jednej skali sprzyja jednakowym odpowiedziom (straightlining) i męczą, zwłaszcza na telefonie.
B03	Obciążenie i komfort	Obowiązkowe pytania otwarte	Wymagane pytania otwarte wydłużają czas i zwiększają porzucenia oraz odpowiedzi-wypełniacze.
B04	Obciążenie i komfort	Wymuszanie odpowiedzi	Wszystkie (lub prawie wszystkie) pytania są obowiązkowe, także te drażliwe, bez możliwości odmowy.
B05	Obciążenie i komfort	Trudny język (czytelność)	Wysoki indeks mglistości (FOG-PL) — długie zdania i trudne wyrazy.
B06	Obciążenie i komfort	Brak wskaźnika postępu	Badany nie wie, ile jeszcze zostało do końca.

(cd.)

Kod	Wymiar	Problem	Komentarz
B07	Obciążenie i komfort	Przeładowana strona	Bardzo wiele pytań na jednej stronie.
B08	Obciążenie i komfort	Powtarzalność, znużenie	Seria bardzo podobnych pytań lub skal męczy i obniża jakość odpowiedzi pod koniec.
M01	Spójność pomiaru	Konstrukt mierzony jednym pytaniem	Złożone zjawisko (postawa, cecha, stan) mierzone jednym pytaniem — niska rzetelność.
M02	Spójność pomiaru	Pytanie nie różnicuje	Prawie wszyscy odpowiadają tak samo — pytanie nie wnosi informacji.
M03	Spójność pomiaru	Problem z kluczem (pozycja odwrócona?)	Pozycja koreluje ujemnie z resztą skali — trzeba ją odwrócić w kluczu albo przeformułować.
M04	Spójność pomiaru	Słaba zgodność z resztą skali	Niska korelacja pozycji z sumą pozostałych — mierzy coś innego albo jest niejasna.
M05	Spójność pomiaru	Efekt sufitowy lub podłogowy	Większość odpowiedzi w skrajnej kategorii — skala nie różnicuje w tej części.
M06	Spójność pomiaru	Niejednorodna bateria	Pozycje jednej skali dotyczą różnych zjawisk — wynik sumaryczny będzie trudny do interpretacji.
M07	Spójność pomiaru	Niestabilne odpowiedzi	Te same profile respondentów odpowiadają za każdym razem inaczej.
M08	Spójność pomiaru	Różne rozumienie pytania	W wywiadach poznawczych badani różnie rozumieją pytanie lub jego kluczowe pojęcia.
F01	Struktura i przebieg	Drażliwe lub metryczkowe pytania na początku	Pytania osobiste przed zbudowaniem zaangażowania zwiększają porzucenia i odmowy.
F02	Struktura i przebieg	Wpływ wcześniejszych pytań	Wcześniejsze pytanie może ukierunkować odpowiedź na późniejsze.
F03	Struktura i przebieg	Brak filtra	Pytanie nie dotyczy wszystkich, a nie ma pytania filtrującego ani opcji „nie dotyczy”.
F04	Struktura i przebieg	Niespójna terminologia	To samo pojęcie nazywane różnie w różnych miejscach (lub różne pojęcia tak samo).
F05	Struktura i przebieg	Brak lub słabe wprowadzenie	Brak powitania, instrukcji, informacji o czasie trwania i sposobie odpowiadania.
F06	Struktura i przebieg	Brak zakończenia	Brak podziękowania i informacji, że odpowiedzi zostały zapisane.
F07	Struktura i przebieg	Nielogiczna kolejność	Skoki tematyczne, pytania szczegółowe przed ogólnymi, brak bloków tematycznych.
F08	Struktura i przebieg	Powtórzone pytanie	To samo pytanie (prawie dosłownie) występuje dwukrotnie — chyba że to celowa para kontrolna.
F09	Struktura i przebieg	Nielosowy przydział do warunków	Badany sam wybiera wariant lub warunek badania (np. jednakowe odpowiedzi „xxx”, „strzelaj”) — przydział nie jest losowy, grupy mogą się systematycznie różnić; użyj randomizacji wbudowanej w platformę.
A01	Technika, telefon, dostępność	Naruszenie WCAG (axe-core)	Automatyczny test dostępności axe-core wykrył naruszenie wytycznych WCAG 2.x.
A02	Technika, telefon, dostępność	Przewijanie w poziomie na telefonie	Strona jest szersza niż ekran telefonu.
A03	Technika, telefon, dostępność	Za małe pola dotyku	Pola wyboru mniejsze niż 24x24 px (WCAG 2.2, kryterium 2.5.8) — łatwo trafić w złą odpowiedź.
A04	Technika, telefon, dostępność	Za mała czcionka	Tekst mniejszy niż 12 px jest trudny do czytania.
A05	Technika, telefon, dostępność	Macierz nieczytelna na telefonie	Tabela z wieloma kolumnami nie mieści się na ekranie telefonu albo ma bardzo wąskie kolumny.
A06	Technika, telefon, dostępność	Pole bez etykiety	Pole formularza nie ma etykiety dostępnej dla czytników ekranu.
A07	Technika, telefon, dostępność	Błąd techniczny lub utknięcie	Syntetyczni respondenci nie mogli przejść ankiety (błąd, zawieszenie, nieprzechodząca walidacja).
A08	Technika, telefon, dostępność	Niejasna walidacja	Formularz odrzuca odpowiedź, a wymóg (format, zakres) nie jest opisany przy pytaniu.
A09	Technika, telefon, dostępność	Strona niedostosowana do telefonu	Brak znacznika viewport — telefon pokazuje pomniejszoną wersję komputerową.
E01	Etyka, RODO, inkluzywność	Brak informacji o celu badania	Badany nie wie, po co jest badanie i do czego posłużą odpowiedzi.
E02	Etyka, RODO, inkluzywność	Brak świadomej zgody	Brak informacji o dobrowolności i wyraźnego wyrażenia zgody na udział.
E03	Etyka, RODO, inkluzywność	Brak informacji RODO	Brak administratora danych, celu i podstawy przetwarzania, okresu przechowywania i praw badanego.
E04	Etyka, RODO, inkluzywność	Dane szczególnej kategorii	Pytania o zdrowie, seksualność, religię, poglądy polityczne lub pochodzenie (art. 9 RODO) wymagają uzasadnienia, wyraźnej zgody i zabezpieczeń.
E05	Etyka, RODO, inkluzywność	Dane identyfikujące	Imię, nazwisko, e-mail, telefon lub adres odbierają anonimowość — zbieraj tylko, gdy to konieczne.
E06	Etyka, RODO, inkluzywność	Język lub kategorie niewłączające	Np. tylko dwie kategorie płci bez „wolę nie odpowiadać”, stereotypy, forma tylko męska.
E07	Etyka, RODO, inkluzywność	Brak kontaktu do badacza	Badany nie ma kogo zapytać o badanie ani do kogo zgłosić wątpliwości.
E08	Etyka, RODO, inkluzywność	Brak informacji o możliwości rezygnacji	Brak informacji, że można przerwać udział w każdej chwili bez konsekwencji.
E09	Etyka, RODO, inkluzywność	Ryzyko dyskomfortu bez wsparcia	Pytania o trudne doświadczenia bez ostrzeżenia, możliwości pominięcia i informacji, gdzie szukać pomocy.
E10	Etyka, RODO, inkluzywność	Brak informacji o anonimowości	Badany nie wie, czy odpowiedzi są anonimowe lub poufne.
E11	Etyka, RODO, inkluzywność	Nie wiadomo, kto prowadzi badanie	Brak nazwy instytucji, zespołu lub osoby odpowiedzialnej za badanie.

13.4 Parametry audytu

Parametr	Wartość
Wersja Respondent AI / version	0.1.0
Model	qwen3.6:35b-mlx
Backend	ollama
Tryb / mode	quick
Języki / languages	pl, en
Role AI / AI roles	metodolog i psychometra, rzecznik badanego (wywiady poznawcze, prosty język)
Przebiegi na żywo / live	both
Urządzenie mobilne / device	Pixel 7
axe-core	4.12.1
Wywołania modelu / LLM calls	14
Nieudane wywołania / failed calls	0
Odrzuceni kandydaci / rejected candidates	0
Czas audytu (min) / duration	3.6
Źródła / sources	Aktywność fizyczna i samopoczucie studentów.docx