

Audyt jakości ankiety

Cześć! Witaj w ankiecie poświęconej memom!

2026-10-11 17:10

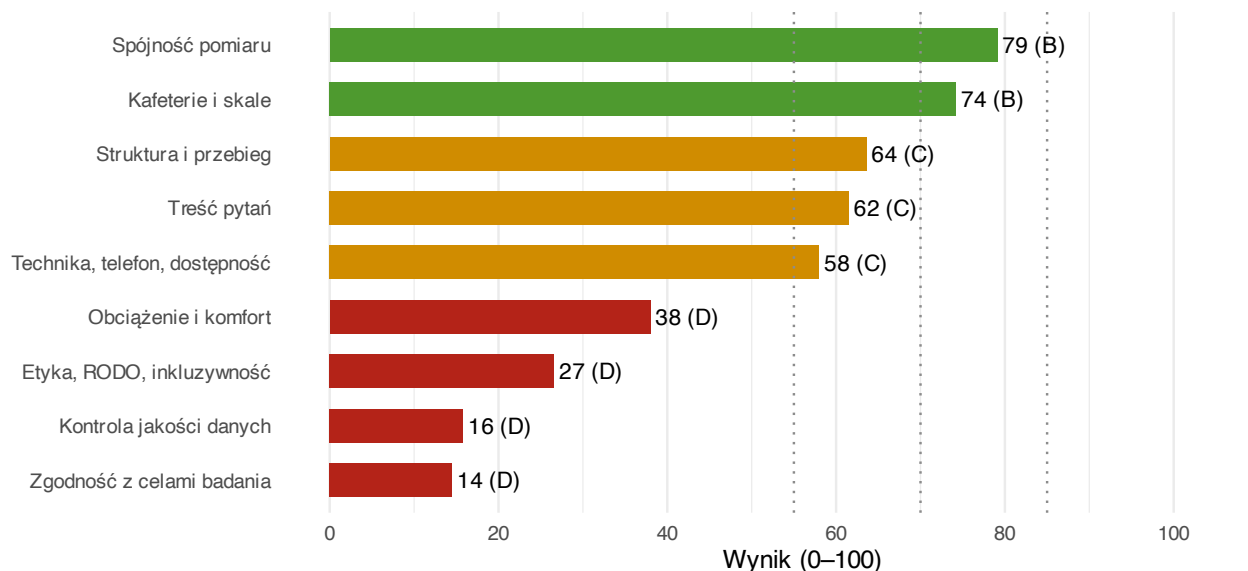
Spis treści

1 Podsumowanie	1
1.1 Najważniejsze poprawki	3
2 Jak powstał ten audyt	4
3 Mapa ankiety	5
4 Zgodność z celami badania	6
5 Pytania: problemy i propozycje poprawek	10
6 Kafeterie i skale odpowiedzi	15
7 Kontrola jakości danych	15
8 Obciążenie i komfort badanego	17
9 Spójność i rzetelność pomiaru	17
10 Struktura i przebieg	18
11 Technika, telefon i dostępność	19
12 Etyka, RODO i inkluzywność	19
13 Gotowość do analizy danych	20
14 Aneksy	22
14.1 Pełna lista problemów	22
14.2 Kandydaci odrzuceni przez panel AI	24
14.3 Słownik kodów problemów	24
14.4 Parametry audytu	27

1 Podsumowanie

Indeks jakości ankiety: 48/100 — ocena D

słaba — wymaga przebudowy



Parametr	Wartość
Ankieta	Cześć! Witaj w ankiecie poświęconej memom!
Link	https://www.metodolog.pl/limesurvey/index.php?r=survey/index&sid=862578&lang=pl
Data audytu	2026-10-11 17:10
Tryb audytu	szybki
Źródła danych	Kwestionariusz odtworzono z 2 przejść ankiety (przebiegi audytowe: komputer, telefon).
Strony z pytaniami	3
Pytania	20
Pozycje (z wierszami macierzy)	210
Szacowany czas wypełniania	≈ 16,8 min
Pytania obowiązkowe	18 (90%)
Wykryte problemy	53 (krytyczne: 4, ważne: 31, drobne: 18)
Cele i hipotezy	7 (pokrycie pełne: 0, częściowe: 6, brak: 1)
Model AI	qwen3.6:35b-mlx

Kluczowym problemem ankiety jest Q01, w którym badany sam wybiera warunek eksperymentalny, co uniemożliwia losowy przydział i wprowadza błąd autoselekcji, czyniąc niemożliwym testowanie wniosków przyczynowych. Drugim krytycznym błędem jest Q02 (powtarzana macierz 20x10), które generuje ekstremalne obciążenie poznawcze, ryzyko odpowiadania w jednej linii (straightlining) oraz szybkie porzucenie ankiety przez respondentów. Wskaźniki jakościowe (Q16, Q17) są mierzone pojedynczymi pozycjami o niskiej wiarygodności, co osłabia kontrolę jakości danych. Hipotezy H1 i H2 są mierzone pojedynczymi itemami, co zwiększa błędy pomiaru i utrudnia weryfikację relacji między zmiennymi. Hipoteza H4 nie może być zweryfikowana, ponieważ brak jest pytania o aktywność twórczą memów (V13). Demografia (Q13, Q18) jest niekompletna i wykluczająca, co narusza zasady inkluzji i uniemożliwia analizę subpopulacji. Należy natychmiast przeredagować Q01, aby warunek był przypisywany losowo przez system, oraz skrócić lub zmienić format Q02, aby zmniejszyć obciążenie respondenta.

Mocne strony

- Ankieta zawiera elementy kontroli jakości (Q16, Q17), które pozwalają na wstępną weryfikację staranności odpowiedzi, choć ich format wymaga poprawy.
- Struktura ankiety jest logiczna, z wyraźnym podziałem na sekcję badawczą i demograficzną, co ułatwia nawigację po wprowadzeniu zmian.
- Zastosowanie skali oceny dla wielu pozycji w Q02 pozwala na zbieranie danych o natężeniu opinii, pod warunkiem redukcji liczby wierszy.
- Obecność pytań otwartych (Q19, Q20) daje możliwość uzyskania kontekstowych informacji o przebiegu badania i ewentualnych zakłóceniach.

1.1 Najważniejsze poprawki

Nr	Waga	Pytanie	Problem	Co zrobić
1	krytyczny	Q02	Duża macierz pytań	Macierz ma 20 wierszy — rośnie ryzyko odpowiadania „w jednej linii” (straightlining) i zmęczenia, zwłaszcza na telefonie. AI: 20 wierszy w siatce, powtórzonych 10 razy, spowoduje silne zmęczenie, prostoliniowość i odrzucenie ankiety, szczególnie na urządzeniach mobilnych. Poprawka: Podziel na mniejsze bloki (do ok. 6–8 wierszy) na osobnych stronach.
2	krytyczny	Q01	Nielosowy przydział do warunków	Odpowiedzi nie różnią się treścią (xxx / xxx / xxx / xxx / xxx) — to wygląda na przydział do warunków badania przez samego badanego. Taki przydział nie jest losowy (ludzie wybierają np. pierwszą lub środkową opcję), więc grupy mogą się systematycznie różnić, a badany widzi „dziwne” pytanie. AI: Instrukcja zachęca do zgadywania, co wprowadza losowy szum do danych i dyskwalifikuje tę pozycję jako miarodajnego filtru lub warunku. Poprawka: Usuń pytanie i użyj randomizacji platformy (np. w LimeSurvey: losowa grupa pytań / równanie z rand(), w Qualtrics: Randomizer z „Evenly present”).
3	krytyczny	Q02	Powtarzalność, znużenie	Ten sam blok (20 pozycji) badany wypełnia 10 razy (np. po każdym materiale; kopie: Q03, Q04, Q05, Q06, Q07, Q08, Q09, Q10...) — razem 200 ocen; pod koniec rośnie zmęczenie, odpowiadanie „w jednej linii” i porzucenia. AI: Powtórzenie tej samej długiej siatki 10 razy (200 ocen) jest metodologicznie nieakceptowalne i niszczy wiarygodność danych przez efekt zmęczenia. Poprawka: Skróć blok do kilku kluczowych twierdzeń, rozłóż materiały na osobne strony lub losuj podzbiór materiałów dla każdego badanego.
4	krytyczny	Q01	Plan badania nie pozwala sprawdzić hipotezy	Plan badania zakłada losowy przydział do warunków, a w ankiecie badany sam wybiera wariant (Q01) — porównania warunków będą obciążone autoselekcją. Poprawka: Przydzielaj warunki losowo na platformie (np. w LimeSurvey: losowa grupa pytań lub równanie z rand()) i usuń pytanie wyboru.
5	ważny	Q13	Język lub kategorie niewłączające	Brak opcji 'Inna' lub 'Nie chcę podawać' wyklucza respondentów niebinarnych lub tych, którzy nie chcą ujawniać płci, naruszając zasady inkluzji. Poprawka: Dodaj opcje 'Inna' lub 'Nie chcę podawać' oraz rozważ rozdzielenie płci biologicznej od tożsamości płciowej.
6	ważny	Q02	Dwa pytania w jednym	Twierdzenie łączy dwie różne cechy (prostotę i zapamiętywalność); respondent może zgadzać się z jedną, a nie z drugą, co zniekształca wynik. Poprawka: Rozdziel na dwa osobne pytania: jedno o prostotę, drugie o łatwość zapamiętania.
7	ważny	Q02	Dwa pytania w jednym	Połączenie 'lub' tworzy podwójne przesłanie; respondent może nie utożsamiać się z żadnym z tych elementów, mimo że jeden z nich występuje. Poprawka: Rozdziel na trzy osobne twierdzenia lub zmień na 'element humoru (np. ironia, sarkazm)' jako jedną kategorię.
8	ważny	Q12	Niewłaściwy format odpowiedzi	Pole tekstowe bez walidacji liczbowej przyjmie odpowiedzi tekstowe ('około 2', 'wiele'), co uniemożliwi analizę statystyczną. Poprawka: Użyj pola numerycznego z wymaganym formatem liczbowym lub suwaka z przedziałem.
9	ważny	Q14	Niewłaściwy format odpowiedzi	Pole tekstowe bez walidacji liczbowej przyjmie odpowiedzi nieliczbowe, co utrudni kodowanie danych demograficznych. Poprawka: Użyj pola numerycznego z zakresem wiekowym lub listy rozwijanej z przedziałami wiekowymi.
10	ważny	Q18	Luki, niepełna lista odpowiedzi	Kategorie są niepełne (brak 'rozwiedziony', 'wdowa/wdowiec', 'w separacji') i używają potocznego określenia 'Singiel', które nie jest standardem akademickim. Poprawka: Rozszerz listę o wszystkie stany cywilne i użyj standardowych terminów (np. 'w związku małżeńskim', 'w konkubinacie').

Następne kroki

- Przerebadaj Q01: zastąp wybór przez respondenta losowym przydziałem do grup eksperymentalnych (np. z ironią vs bez ironii) na poziomie systemu ankiety.
- Zredukuj liczbę wierszy w Q02 lub podziel macierz na mniejsze bloki, aby uniknąć zmęczenia i straightlingu; rozważ zmianę formatu na pytania pojedyncze lub krótsze macierze.
- Rozbuduj Q13 i Q18 o opcje 'Inna', 'Nie chcę podać', 'Rozwiedziony', 'Wdowa/Wdowiec', aby zapewnić inkluzję i kompletność danych demograficznych.
- Dodaj walidację liczbową do pól tekstowych Q12 i Q14, aby uniemożliwić wpisanie odpowiedzi nieliczbowych i ułatwić kodowanie danych.
- Dodaj pozycje mierzące aktywność twórczą memów (V13), aby umożliwić testowanie hipotezy H4 i podział respondentów na grupy 'twórcy' i 'nie-twórcy'.
- Przeprowadź pilotaż (pretest) z realnymi respondentami, aby zweryfikować zrozumienie instrukcji, czas trwania i reakcję na zmodyfikowane elementy ankiety.

Wymiar	Waga	Wynik	Ocena	Kryt.	Ważne	Drobne	Podstawa
Zgodność z celami badania	15	14	D	1	10	1	cele i hipotezy: 7

(cd.)

Wymiar	Waga	Wynik	Ocena	Kryt.	Ważne	Drobne	Podstawa
Treść pytań	20	62	C	0	3	1	20 pytań
Kafeteria i skale	15	74	B	0	5	2	20 pytań
Kontrola jakości danych	10	16	D	0	2	2	20/100 pkt
Obciążenie i komfort	15	38	D	2	1	4	53/100 pkt
Spójność pomiaru	10	79	B	0	1	1	50/60 pkt
Struktura i przebieg	10	64	C	1	0	3	72/90 pkt
Technika, telefon, dostępność	10	58	C	0	4	2	axe-core + przebieg mobilny
Etyka, RODO, inkluzywność	10	27	D	0	5	2	30/92 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. **Rekonstrukcja kwestionariusza.** Kwestionariusz odtworzono z 2 przejść ankiety (przebiegi audytowe: komputer, telefon). Pytania, odpowiedzi, wiersze macierzy, komunikaty walidacji i teksty stron zostały scalone z przebiegów wszystkich respondentów (z uwzględnieniem ścieżek warunkowych).
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. **Dane i testy techniczne.** Brak danych z przebiegów person. Przebiegi na żywo (bez wysyłania) na komputerze i telefonie (Pixel 7) z testem dostępności axe-core 4.12.1 (WCAG 2.0–2.2 A/AA) i pomiarem układu strony.
5. **Zgodność z celami badania.** Model wyodrębnił z tekstu badacza (memy_plan_badania.docx) cele, problemy i hipotezy oraz zmienne potrzebne do każdego z nich, przypisał zmienne do pytań ankiety (z poziomem pomiaru) i ocenił, czy dane pozwolą na zaplanowane analizy. Elementy, których nie ma w tekście badacza, odrzucono; zmienne uznane za niemierzone sprawdzono jeszcze w brzmieniu pytań. Każdy cel i każda hipoteza to kryterium wymiaru „Zgodność z celami badania” (pokrycie pełne 100%, częściowe 50%, brak 0%).

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).

Zakres. Ankieta deklaruje 60 pytań, a w przejściach widoczne było 20 — pozostałe należą do ścieżek lub warunków badania, których przejścia nie odwiedziły; audyt ich nie obejmuje. Aby je uwzględnić, zaudytuj przebieg kilku person (różne ścieżki).

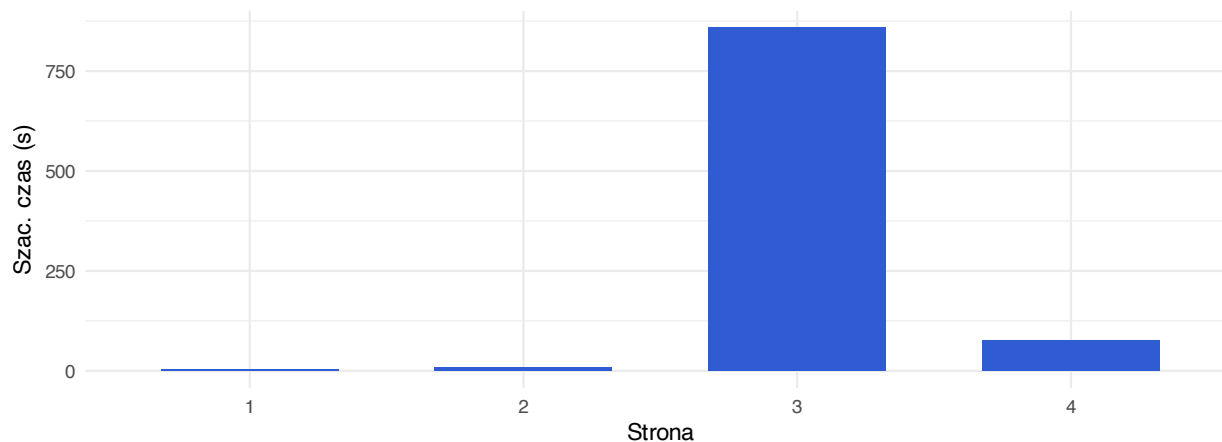
Ograniczenia. Ocena AI i reguł to ekspertyza wstępna, a nie zastępstwo pilotażu z ludźmi: model może przeoczyć problem albo zgłosić fałszywy alarm (dlatego podawana jest pewność i źródła). Syntetyczni respondenci odpowiadają spójniej niż ludzie. Test axe-core wykrywa ok. 30–40% problemów dostępności;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

Wymiar	Waga	Komentarz
Zgodność z celami badania	15	Czy ankieta mierzy zmienne potrzebne do celów, problemów i hipotez badania (podanych przez badacza), na poziomie pomiaru pozwalającym na zaplanowane analizy, i czy nie ma pytań bez celu.
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, macierz, 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	uniważnia 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

Strona	Typ strony	Pytań	Pozycji	Szac. czas (s)	Postęp
1	wprowadzenie	0	0	3	Wypełniono 0% tej ankiety
2	pytania	1	1	10	Wypełniono 0% tej ankiety
3	pytania	10	200	858	Wypełniono 14% tej ankiety
4	pytania	9	9	77	Wypełniono 86% tej ankiety



ID	str.	Typ	Pytanie	Odp.	Obow.	jakość
Q01	2	jednokrotny wybór	Wybierz jedną z poniższych odpowiedzi. Jeśli nie wiesz jaką wybrać, to "strzelaj".	5	•	0
Q02	3	macierz	Nie ma tutaj odpowiedzi złych/błędnych lub dobrych/prawidłowych. Liczy się Twoja wyjątkowa opi...	20×4	•	0
Q03	3	macierz	= Q02 (powtórzenie bloku)	20×4	•	0
Q04	3	macierz	= Q02 (powtórzenie bloku)	20×4	•	0
Q05	3	macierz	= Q02 (powtórzenie bloku)	20×4	•	0
Q06	3	macierz	= Q02 (powtórzenie bloku)	20×4	•	0
Q07	3	macierz	= Q02 (powtórzenie bloku)	20×4	•	0
Q08	3	macierz	= Q02 (powtórzenie bloku)	20×4	•	0
Q09	3	macierz	= Q02 (powtórzenie bloku)	20×4	•	0
Q10	3	macierz	= Q02 (powtórzenie bloku)	20×4	•	0
Q11	3	macierz	= Q02 (powtórzenie bloku)	20×4	•	0
Q12	4	tekst krótki	Ile czasu dziennie spędzasz w mediach społecznościowych (podaj średni czas spędzany dziennie w...	0	•	30

(cd.)

ID	str.	Typ	Pytanie	Odp.	Obow.	jakość
Q13	4	jednokrotny wybór	Jaka jest Twoja płeć biologiczna?	2	•	40
Q14	4	tekst krótki	Ile masz lat?	0	•	30
Q15	4	jednokrotny wybór	Jaki jest Twój poziom wykształcenia?	3	•	100
Q16	4	jednokrotny wybór	Jak bardzo byłeś/byłaś zmotywowany/a do ukończenia badania?	3	•	40
Q17	4	jednokrotny wybór	Jak szczerze były Twoje odpowiedzi podczas wypełniania badania?	3	•	40
Q18	4	jednokrotny wybór	Jaki jest Twój status związku?	2	•	40
Q19	4	tekst krótki	Jaki był Twoim zdaniem cel tego badania?	0		100
Q20	4	tekst krótki	Czy w trakcie badania w Twoim otoczeniu wydarzyło się coś, co mogło mieć wpływ na jego przebieg...	0		30

4 Zgodność z celami badania

Cele, problemy i hipotezy podał badacz (memy_plan_badania.docx). Dla każdego elementu sprawdzono, czy ankieta mierzy potrzebne zmienne, na jakim poziomie pomiaru i czy dane pozwolą na zaplanowaną analizę. **Pokrycie:** pełne — 0, częściowe — 6, brak — 1 (z 7). **Plan badania:** quasi-eksperymentalny. **Populacja:** Osoby w wieku 18–35 lat, które korzystają z mediów społecznościowych co najmniej raz dziennie..

ID	Rodzaj	Cel / hipoteza	Pokrycie	Co pozwalają dane, problemy
P1	pytanie badawcze	Które cechy materiału (humor, zrozumiałość, krótkość, emocje, efekt „wow”) najsilniej przewidują deklarowaną gotowość do udostępnienia?	częściowe	Regresja wielokrotna lub modele wielopoziomowe z użyciem zmiennych V1-V6. Wszystkie zmienne są mierzone pojedynczymi itemami lub krótkimi bateriami w skali porządkowej. Analiza będzie opierała się na korelacjach rangowych (Spearman) lub regresji z użyciem metod robustnych, ze względu na naruszenie założeń liniowości i normalności...
P2	pytanie badawcze	Czy materiały zawierające ironię lub parodię są oceniane jako bardziej wartościowe niż materiały bez tych elementów?	częściowe	Test t dla prób niezależnych lub ANOVA porównująca V7 (wartość) w grupach utworzonych na podstawie V8 (obecność ironii). Grupy będą oparte na postrzeganej obecności ironii (V8), a nie na manipulacji eksperymentalnej. Problemy: Brak randomizacji warunków (Q01 ma błąd F09). Warunek V8 (ironia) nie jest manipulacją, lecz cechą pos...
H1	hipoteza	Im bardziej materiał jest postrzegany jako humorystyczny, tym wyższa deklarowana gotowość do jego udostępnienia.	częściowe	Regresja lub korelacja między V2 (humor) a V1 (gotowość do udostępnienia). Oczekiwana pozytywna zależność. Problemy: Oba zmienne są mierzone pojedynczymi itemami w skali 4-punktowej. Relacja będzie silnie obciążona błędem pomiaru. Pojedynczy item nie jest miarodajnym wskaźnikiem złożonego konstru 'humor'. Ryzyko fałszywych wnio...
H2	hipoteza	Osoby spędzające więcej czasu w mediach społecznościowych lepiej rozumieją odniesienia zawarte w memach.	częściowe	Korelacja między V9 (czas w mediach) a V10 (zrozumienie odniesień). V10 jest mierzony itemami o prostocie i szybkim zrozumieniu. Problemy: V9 (czas) jest mierzony jako 'średni czas dzienny', co jest obciążone błędem pamięci. V10 (zrozumienie odniesień) jest mylone z ogólną zrozumiałością/prostotą. Konstrukty nie są adekwatnie m...
H3	hipoteza	Kobiety częściej niż mężczyźni deklarują, że memy wzbudzają w nich emocje.	częściowe	Test t dla prób niezależnych porównujący V12 (emocje) między grupami V11 (płeć). Problemy: V11 (płeć) pomija grupy niebinarne, co może zniekształcić wyniki w populacji młodych (18-35 lat). V12 (emocje) jest mierzony jednym itemem, który odnosi się do 'materiału wideo', a nie specyficznie do 'memów' (jak w hipotezie). Może to pr...
H4	hipoteza	Osoby, które same tworzą memy, wyżej oceniają ich potencjał wiralowy niż osoby, które ich nie tworzą.	brak	Brak danych. Nie można przeprowadzić analizy. Problemy: Zmienna V13 (tworzenie memów) nie jest mierzona w ankiecie. Brak pytania o aktywność twórczą respondenta uniemożliwia podział na grupy 'twórcy' i 'nie-twórcy'. Hipoteza nie może być zweryfikowana.
C1	cel pomiarowy	Utworzenie wskaźnika „potencjału wiralowego” materiału jako skali sumarycznej z 20 twierdzeń (alfa Cronbacha co najmniej 0,80).	częściowe	Analiza wiarygodności (alfa Cronbacha) dla skali sumarycznej zbudowanej z itemów Q02/r1, r10, r14, r16, r19 (potencjał wiralowy). Problemy: Bateria jest heterogeniczna (mieszanie intencji udostępnienia z zaangażowaniem i ponownym obejrzeniem). Uwagi wskazują na niską jednorodność baterii. Alfa Cronbacha może być niska (<0.80), ...

Zmienne i pytania, które je mierzą

Zmienna	Rola	Dla	Pytania	Pokrycie	Poziom pomiaru	Komentarz
V1 gotowość do udostępnienia	zmienna zależna	P1, H1	Q02_10	częściowe	porządkowy	Zmienna gotowość do udostępnienia mierzona jest pojedynczym twierdzeniem w skali 4-punktowej. Brak jest odwróconych pozycji lub kontekstu, co zwiększa ryzyko błędu metody pomiaru (method bias). Pojedynczy item nie jest miarodajny...
V2 humor	predyktor	P1, H1	Q02_8	częściowe	porządkowy	Humor mierzony jednym itemem (Q02/r8). Konstrukty percepcyjne takie jak humor są złożone i wymagają wieloelementowej skali dla wiarygodności. Pojedynczy item jest narażony na subiektywną interpretację respondenta.
V3 zrozumiałość	predyktor	P1	Q02_5, Q02_13, Q02_20	częściowe	porządkowy	Zrozumiałość mierzona trzema itemami (prostota, zrozumiałość dla grup, szybkie zrozumienie przesłania). Itemy te są do siebie bardzo podobne (redundancja), co może sztucznie zawyżać współczynnik alfy Cronbacha, ale nie mierzą sze...
V4 krótkość	predyktor	P1	Q02_18	częściowe	porządkowy	Krótkość mierzona jednym itemem. Jest to cecha obiektywna (czas trwania), a nie tylko percepcyjna, co może prowadzić do rozbieżności między faktyczną długością a percepcją respondenta.
V5 emocje	predyktor	P1	Q02_12	częściowe	porządkowy	Emocje mierzone jednym itemem. Konstrukty emocjonalne są wielowymiarowe (np. emocje pozytywne vs negatywne, intensywność). Pojedynczy item 'wzbudził emocje' jest zbyt ogólny i nieprecyzyjny.
V6 efekt wow	predyktor	P1	Q02_17	częściowe	porządkowy	Efekt 'wow' mierzony jednym itemem. Jest to konstruktywny, subiektywny i trudny do operacjonalizacji jednym stwierdzeniem. Ryzyko błędu poznawczego jest wysokie.
V7 wartość materiału	zmienna zależna	P2	Q02_11	częściowe	porządkowy	Wartość materiału mierzona jednym itemem. Jest to ocena ogólna, która może być zafałszowana przez inne cechy (np. humor), co utrudnia izolację samej 'wartościowości' w analizie.
V8 obecność ironii lub parodii	warunek / manipulacja	P2	Q02_15	częściowe	porządkowy	Obecność ironii/parodii mierzona jednym itemem. Respondenci mogą mylić ironię z sarkazmem lub parodią. Brak jest manipulacji warunkowej (video jest to samo dla wszystkich), więc jest to postrzegana obecność, a nie faktyczna cecha...
V9 czas spędzany w mediach społecznościowych	predyktor	H2	Q12	częściowe	ilorazowy	Czas w mediach mierzony jako średni czas dzienny w godzinach wpisywany swobodnie. Format tekstowy wymaga czyszczenia danych. Odpowiedź 'średni czas' jest obciążona błędem pamięci i subiektywną oceną respondenta, a nie danymi z pl...
V10 zrozumienie odniesień	zmienna zależna	H2	Q02_4, Q02_20	częściowe	porządkowy	Zrozumienie odniesień mierzone itemami o prostocie i szybkim zrozumieniu przesłania. Nie mierzy to bezpośrednio 'zrozumienia odniesień' (referencji), lecz ogólną łatwość odbioru. Konstrukty są mylone.
V11 płeć	zmienna grupująca	H3	Q13	częściowe	nominalny	Płeć mierzona jako biologiczna (kobieta/mężczyzna). Brak opcji 'inna' lub 'wolę nie mówić' narusza zasady etyczne i może wykluczać respondenci niebinarnych, co wprowadza błąd doboru próby w kontekście współczesnych badań społecznych...
V12 emocje wzbudzone przez memy	zmienna zależna	H3	Q02_12	częściowe	porządkowy	Emocje wzbudzone przez memy mierzone tym samym itemem co V5 (Q02/r12). Pytanie odnosi się do 'materiału wideo', a nie specyficznie do 'memów'. Może to wprowadzać błąd kontekstowy, jeśli badany materiał nie jest memem w ścisłym se...
V13 tworzenie memów	zmienna grupująca	H4	—	brak	—	Tworzenie memów nie jest mierzone w dostarczonym fragmencie ankiety. Brak pytania o aktywność twórczą respondenta uniemożliwia testowanie H4.
V14 potencjał wiralowy	zmienna zależna	H4, C1	Q02_1, Q02_10, Q02_14, Q02_16, Q02_19	częściowe	porządkowy	Potencjał wiralowy mierzony baterią 5 itemów (udostępnienie, uożsamienie, komentarz, ponowne obejrzenie, przyciągnięcie uwagi). Itemy te są heterogeniczne (mieszają intencję udostępnienia z zaangażowaniem). Skala może mieć niską...
V15 wiek	cecha populacji	populacja	Q14	częściowe	ilorazowy	Wiek mierzony jako liczba lat wpisywana swobodnie. Format tekstowy wymaga walidacji. Brak ograniczeń wiekowych w pytaniu (choć target to 18-35) może prowadzić do błędnych wpisów.

(cd.)

Zmienna	Rola	Dla	Pytania	Pokrycie	Poziom pomiaru	Komentarz
V16 częstotliwość korzystania z mediów społecznościowych	cecha populacji	populacja	—	brak	—	Częstotliwość korzystania z mediów (co najmniej raz dziennie) nie jest mierzona w ankiecie. Brak weryfikacji kryterium włączenia do badania. Wiek i czas (V9) nie zastępują częstotliwości.

Proponowane pytania do dodania (dla zmiennych niemierzonych lub mierzonych częściowo)

Zmienna (dla)	Proponowane pytanie	Odpowiedzi	Po pytaniu	Uzasadnienie
tworzenie memów (H4)	Czy regularnie tworzysz własne memy internetowe (np. dodajesz napisy, edytujesz obrazy, tworzysz nowe formaty)?		Q12	Zmienna V13 nie jest mierzona w obecnej wersji ankiety, co uniemożliwia weryfikację hipotezy H4 dotyczącej twórców memów.
częstotliwość korzystania z mediów społecznościowych (NA)	Jak często korzystasz z mediów społecznościowych? (wybierz jedną odpowiedź)	Raz dziennie lub częściej / Kilka razy w tygodniu / Raz w tygodniu / Rzadziej niż raz w tygodniu	Q12	Zmienna V16 jest kluczowa dla definicji populacji (codzienne korzystanie), a jej brak uniemożliwia weryfikację warunku wstępnej selekcji próby.
gotowość do udostępnienia (P1, H1)	Materiał jest tak angażujący, że chciałbym/chciałabym go udostępnić znajomym.	W ogóle się nie zgadzam / Trochę się nie zgadzam / Trochę się zgadzam / Całkowicie się zgadzam	Q02	Obecny item V1 jest narażony na błąd pomiaru; dodanie drugiego, precyzyjnie sformułowanego twierdzenia zwiększa wiarygodność pomiaru intencji udostępnienia.
humor (P1, H1)	Materiał jest zabawny i wywołuje uśmiech lub śmiech.	W ogóle się nie zgadzam / Trochę się nie zgadzam / Trochę się zgadzam / Całkowicie się zgadzam	Q02	Humor mierzony pojedynczym itemem jest słabo operacjonalizowany; dodanie drugiego itemu poprawia wiarygodność pomiaru tego złożonego konstru.
zrozumiałość (P1)	Treść materiału jest prosta i łatwa do zrozumienia bez dodatkowego wysiłku.	W ogóle się nie zgadzam / Trochę się nie zgadzam / Trochę się zgadzam / Całkowicie się zgadzam	Q02	Zrozumiałość mierzona fragmentarycznie wymaga uzupełnienia o precyzyjny item, aby poprawić wiarygodność pomiaru predyktora.
krótkość (P1)	Materiał jest krótki i nie wymaga długiego czasu na przyswojenie.	W ogóle się nie zgadzam / Trochę się nie zgadzam / Trochę się zgadzam / Całkowicie się zgadzam	Q02	Krótkość mierzona pojedynczym itemem jest narażona na błąd pomiaru; drugi item zwiększa wiarygodność pomiaru tej cechy.
emocje (P1)	Materiał wywołuje u mnie silne odczucia emocjonalne (np. zdziwienie, rozbawienie, wzruszenie).	W ogóle się nie zgadzam / Trochę się nie zgadzam / Trochę się zgadzam / Całkowicie się zgadzam	Q02	Emocje mierzone pojedynczym itemem są słabo operacjonalizowane; dodanie drugiego itemu poprawia wiarygodność pomiaru predyktora.
efekt wow (P1)	Materiał zaskakuje mnie i wyróżnia się na tle innych treści, które widzę w mediach społecznościowych.	W ogóle się nie zgadzam / Trochę się nie zgadzam / Trochę się zgadzam / Całkowicie się zgadzam	Q02	Efekt wow mierzony pojedynczym itemem jest narażony na błąd pomiaru; drugi item zwiększa wiarygodność pomiaru tej cechy.

Pytania, których nie wykorzystuje żaden cel ani hipoteza

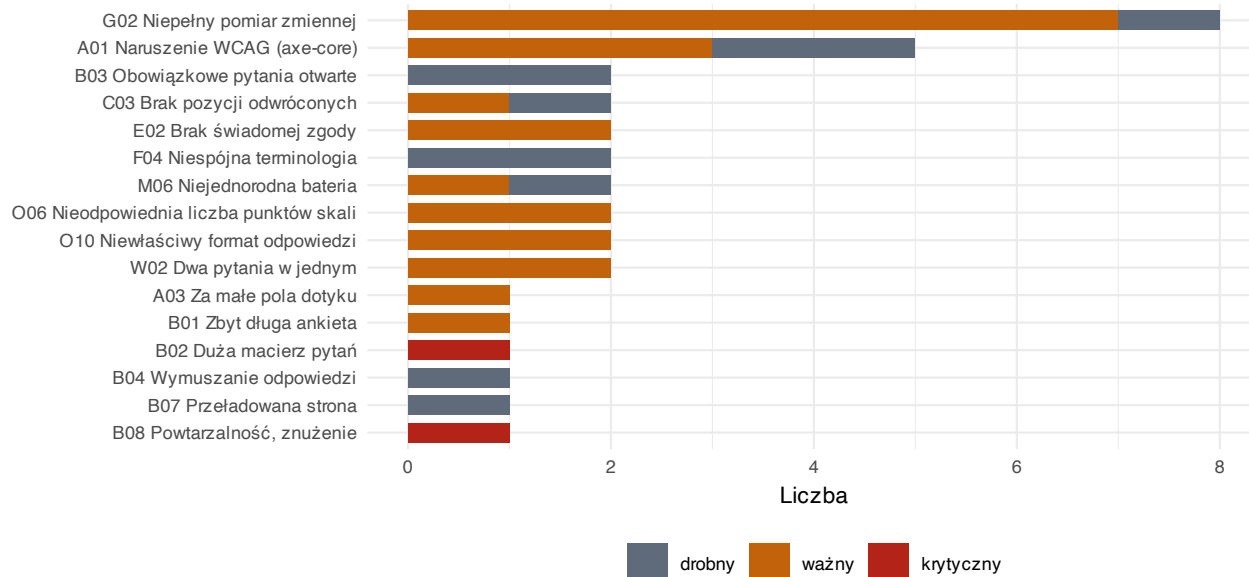
ID	Pytanie	Funkcja	Komentarz
Q15	Jaki jest Twój poziom wykształcenia?	metryczka	Pytanie o wykształcenie jest typowym elementem demograficznym opisującym próbę, ale nie służy do testowania hipotez ani kontroli metodologicznej w kontekście badanych zmiennych.
Q18	Jaki jest Twój status związku?	metryczka	Status związku jest zmienną demograficzną, która nie ma bezpośredniego związku z badanymi konstrukcjami medialnymi ani nie służy jako kontrola metodologiczna w tym modelu.

Metryczka, pytania selekcyjne i kontrolne mają swoją funkcję; jako problem liczą się tylko pytania „bez celu”.

Problemy zgodności z celami

ID	Waga	Problem	Komentarz
Q01	krytyczny	Plan badania nie pozwala sprawdzić hipotezy	Plan badania zakłada losowy przydział do warunków, a w ankiecie badany sam wybiera wariant (Q01) — porównania warunków będą obciążone autoselekcją. Poprawka: Przydzielaj warunki losowo na platformie (np. w LimeSurvey: losowa grupa pytań lub równanie z rand()) i usuń pytanie wyboru.
Q02	ważny	Niepełny pomiar zmiennej	gotowość do udostępnienia (zmienna zależna; dla P1, H1) mierzona tylko częściowo (Q02_10). Zmienna gotowość do udostępnienia mierzona jest pojedynczym twierdzeniem w skali 4-punktowej. Brak jest odwróconych pozycji lub kontekstu, co zwiększa ryzyko błędu metody pomiaru (method bias). Pojedynczy item nie jest miarodajny dla skom...
Q02	ważny	Niepełny pomiar zmiennej	wartość materiału (zmienna zależna; dla P2) mierzona tylko częściowo (Q02_11). Wartość materiału mierzona jednym itemem. Jest to ocena ogólna, która może być zafałszowana przez inne cechy (np. humor), co utrudnia izolację samej 'wartościowości' w analizie. Poprawka: Dodaj pytanie (lub sprawdzoną skalę) mierzące tę zmienną.
Q02	ważny	Niepełny pomiar zmiennej	zrozumienie odniesień (zmienna zależna; dla H2) mierzona tylko częściowo (Q02_4, Q02_20). Zrozumienie odniesień mierzane itemami o prostocie i szybkim zrozumieniu przesłania. Nie mierzy to bezpośrednio 'zrozumienia odniesień' (referencji), lecz ogólną łatwość odbioru. Konstrukty są mylone. Poprawka: Dodaj pytanie (lub sprawdzon...
Q02	ważny	Niepełny pomiar zmiennej	emocje wzbudzone przez memy (zmienna zależna; dla H3) mierzona tylko częściowo (Q02_12). Emocje wzbudzone przez memy mierzane tym samym itemem co V5 (Q02/r12). Pytanie odnosi się do 'materiału wideo', a nie specyficznie do 'memów'. Może to wprowadzać błąd kontekstowy, jeśli badany materiał nie jest memem w ścisłym sensie. Popra...
Q02	ważny	Niepełny pomiar zmiennej	potencjał wiralowy (zmienna zależna; dla H4, C1) mierzona tylko częściowo (Q02_1, Q02_10, Q02_14, Q02_16, Q02_19). Potencjał wiralowy mierzony baterią 5 itemów (udostępnienie, utożsamienie, komentarz, ponowne obejrzenie, przyciągnięcie uwagi). Itemy te są heterogeniczne (miesząc intencję udostępnienia z zaangażowaniem). Skala ...
Q12	ważny	Poziom pomiaru nie pozwala na analizę	czas spędzany w mediach społecznościowych (H2) jest wpisywana jako tekst — przed analizą trzeba ją oczyścić i zamieniać na liczby, część odpowiedzi przepadnie. Poprawka: Użyj pola liczbowego z jednostką i zakresem.
Q12	ważny	Niepełny pomiar zmiennej	czas spędzany w mediach społecznościowych (predyktor; dla H2) mierzona tylko częściowo (Q12). Czas w mediach mierzony jako średni czas dzienny w godzinach wpisywany swobodnie. Format tekstowy wymaga czyszczenia danych. Odpowiedź 'średni czas' jest obciążona błędem pamięci i subiektywną oceną respondenta, a nie danymi z platform...
Q13	ważny	Niepełny pomiar zmiennej	pleć (zmienna grupująca; dla H3) mierzona tylko częściowo (Q13). Płeć mierzona jako biologiczna (kobieta/mężczyzna). Brak opcji 'inna' lub 'wolę nie mówić' narusza zasady etyczne i może wykluczać respondenci niebinarnych, co wprowadza błąd doboru próby w kontekście współczesnych badań społecznych. Poprawka: Dodaj pytanie (lub s...
—	ważny	Brak pomiaru zmiennej z hipotezy	tworzenie memów (zmienna grupująca; potrzebna do H4) — żadne pytanie ankiety jej nie mierzy. Tworzenie memów nie jest mierzone w dostarczonym fragmencie ankiety. Brak pytania o aktywność twórczą respondenta uniemożliwia testowanie H4. Część pytań ankiety nie była widoczna w przejściach — zmienna może być w pominiętych ścieżkach...
—	ważny	Niezgodność z badaną populacją	częstotliwość korzystania z mediów społecznościowych (cecha populacji; potrzebna do populacji badania) — żadne pytanie ankiety jej nie mierzy. Częstotliwość korzystania z mediów (co najmniej raz dziennie) nie jest mierzona w ankiecie. Brak weryfikacji kryterium włączenia do badania. Wiek i czas (V9) nie zastępują częstotliwości...
Q14	drobny	Niepełny pomiar zmiennej	wiek (cecha populacji; dla populacji badania) mierzona tylko częściowo (Q14). Wiek mierzony jako liczba lat wpisywana swobodnie. Format tekstowy wymaga walidacji. Brak ograniczeń wiekowych w pytaniu (choć target to 18-35) może prowadzić do błędnych wpisów. Poprawka: Dodaj pytanie (lub sprawdzoną skalę) mierzące tę zmienną.

5 Pytania: problemy i propozycje poprawek



Q01

Wybierz jedną z poniższych odpowiedzi. Jeśli nie wiesz jaką wybrać, to “strzelaj”.

str. 2 · jednokrotny wybór · obowiązkowe · jakość 0/100

Odpowiedzi: xxx / xxx / xxx / xxx / xxx

- **[krytyczny] Nielosowy przydział do warunków** (F09; pewność 0,97; AI: metodolog + AI: rzecznik badanego + reguła). Odpowiedzi nie różnią się treścią (xxx / xxx / xxx / xxx / xxx) — to wygląda na przydział do warunków badania przez samego badanego. Taki przydział nie jest losowy (ludzie wybierają np. pierwszą lub środkową opcję), więc grupy mogą się systematycznie różnić, a badany widzi „dziwne” pytanie. AI: Instrukcja zachęca do zgadywania, co wprowadza losowy szum do danych i dyskwalifikuje tę pozycję jako miarodajnego filtru lub warunku. *Dowód:* xxx / xxx / xxx / xxx / xxx *Poprawka:* Usuń pytanie i użyj randomizacji platformy (np. w LimeSurvey: losowa grupa pytań / równanie z rand(), w Qualtrics: Randomizer z „Evenly present”).
- **[krytyczny] Plan badania nie pozwala sprawdzić hipotezy** (G07; pewność 0,85; reguła). Plan badania zakłada losowy przydział do warunków, a w ankiecie badany sam wybiera wariant (Q01) — porównania warunków będą obciążone autoselekcją. *Dowód:* Q01 *Poprawka:* Przydzielaj warunki losowo na platformie (np. w LimeSurvey: losowa grupa pytań lub równanie z rand()) i usuń pytanie wyboru.

Propozycja: (usuń to pytanie — przydział do warunków przez losowanie platformy)

Uzasadnienie: Badany nie powinien sam wybierać warunku; w LimeSurvey użyj losowej grupy pytań albo równania z rand() i warunków wyświetlania.

Q02

Nie ma tutaj odpowiedzi złych/błędnych lub dobrych/prawidłowych. Liczy się Twoja wyjątkowa opinia na temat materiału.

To samo pytanie powtarza się jako: Q03, Q04, Q05, Q06, Q07, Q08, Q09, Q10, Q11

str. 3 · macierz · obowiązkowe · jakość 0/100

Odpowiedzi: W ogóle się nie zgadzam / Trochę się nie zgadzam / Trochę się zgadzam / Całkowicie się zgadzam

Wiersze: (1) Zachęciłbym/abym rodzinę/przyjaciół do obejrzenia tego materiału.; (2) Czuję, że inni mogliby się utożsamiać z tym co jest przedstawione na filmie.; (3) Materiał zachęca do wyrażenia opinii (np. w formie oceny, komentarza).; (4) Materiał wideo zawiera odniesienia/znaczenia które rozumiem.; (5) Treść filmu jest prosta i łatwa do zapamiętania.; (6) Forma jest dopasowana do realiów platform społecznościowych.; (7) Obejrzałbym/abym ponownie ten materiał.;

(8) Wideo zawiera element humoru.; (9) Wideo posługuje się stylem który pasuje do internetu.; (10) Mógłbym/abym udostępnić taki materiał w swoich social mediach.; (11) Uważam, że wideo jest wartościowe.; (12) Materiał wzbudził we mnie emocje.; (13) W mojej opinii film jest zrozumiały dla wielu grup społecznych w moim kraju.; (14) Chciałbym/abym obejrzeć więcej filmów stworzonych w podobnym stylu.; (15) W filmie jest element ironii, sarkazmu lub parodii.; (16) Wideo jest "bezpieczne" do udostępnienia.; (17) Materiał wywołuje efekt „wow”.; (18) Film jest na tyle krótki, że łatwo go obejrzeć w dowolnej chwili.; (19) Materiał przykuł moją uwagę.; (20) Szybko zrozumiałem/am główne przesłanie filmu.

- **[krytyczny] Duża macierz pytań** (B02; pewność 0,97; AI: przegląd całości + AI: metodolog + AI: rzecznik badanego + reguła). Macierz ma 20 wierszy — rośnie ryzyko odpowiadania „w jednej linii” (straightlining) i zmęczenia, zwłaszcza na telefonie. AI: 20 wierszy w siatce, powtórzonych 10 razy, spowoduje silne zmęczenie, prostolinijność i odrzucenie ankiety, szczególnie na urządzeniach mobilnych. *Dowód: 20 wierszy Poprawka:* Podziel na mniejsze bloki (do ok. 6–8 wierszy) na osobnych stronach.
- **[krytyczny] Powtarzalność, znużenie** (B08; pewność 0,97; AI: metodolog + AI: rzecznik badanego + reguła). Ten sam blok (20 pozycji) badany wypełnia 10 razy (np. po każdym materiale; kopie: Q03, Q04, Q05, Q06, Q07, Q08, Q09, Q10...) — razem 200 ocen; pod koniec rośnie zmęczenie, odpowiadanie „w jednej linii” i porzucenia. AI: Powtórzenie tej samej długiej siatki 10 razy (200 ocen) jest metodologicznie nieakceptowalne i niszczy wiarygodność danych przez efekt zmęczenia. *Dowód: 200 ocen Poprawka:* Skróć blok do kilku kluczowych twierdzeń, rozłóż materiały na osobne strony lub losuj podzbiór materiałów dla każdego badanego.
- **[ważny] Dwa pytania w jednym** (W02, wiersz 5; pewność 0,97; AI: metodolog + AI: rzecznik badanego + reguła). Twierdzenie łączy dwie różne cechy (prostotę i zapamiętywalność); respondent może zgadzać się z jedną, a nie z drugą, co zniekształca wynik. *Dowód: Treść filmu jest prosta i łatwa do zapamiętania. Poprawka:* Rozdziel na dwa osobne pytania: jedno o prostotę, drugie o łatwość zapamiętania.
- **[ważny] Dwa pytania w jednym** (W02, wiersz 15; pewność 0,97; AI: metodolog + AI: rzecznik badanego + reguła). Połączenie 'lub' tworzy podwójne przesłanie; respondent może nie utożsamiać się z żadnym z tych elementów, mimo że jeden z nich występuje. *Dowód: W filmie jest element ironii, sarkazmu lub parodii. Poprawka:* Rozdziel na trzy osobne twierdzenia lub zmień na 'element humoru (np. ironia, sarkazm)' jako jedną kategorię.
- **[ważny] Brak pozycji odwróconych** (C03; pewność 0,68; AI: przegląd całości). Brak odwróconych pozycji w siatkach, co uniemożliwia wykrycie efektu przychylności (acquiescence bias). *Dowód: Wszystkie twierdzenia w siatce są skierowane pozytywnie. Poprawka:* Dodaj 2-3 odwrócone pozycje w siatkach, aby wykryć respondentów stosujących ślepe zaznaczanie.
- **[ważny] Niepełny pomiar zmiennej** (G02; pewność 0,65; ai:alignment). gotowość do udostępnienia (zmienna zależna; dla P1, H1) mierzona tylko częściowo (Q02_10). Zmienna gotowość do udostępnienia mierzona jest pojedynczym twierdzeniem w skali 4-punktowej. Brak jest odwróconych pozycji lub kontekstu, co zwiększa ryzyko błędu metody pomiaru (method bias). Pojedynczy item nie jest miarodajny dla skomplikowanego konstru behavioral intention. *Dowód: Q02_10 Poprawka:* Dodaj pytanie, np.: „Materiał jest tak angażujący, że chciałbym/chciałabym go udostępnić znajomym.” (W ogóle się nie zgadzam / Trochę się nie zgadzam / Trochę się zgadzam / Całkowicie się zgadzam).
- **[ważny] Niepełny pomiar zmiennej** (G02; pewność 0,65; ai:alignment). wartość materiału (zmienna zależna; dla P2) mierzona tylko częściowo (Q02_11). Wartość materiału mierzona jednym itemem. Jest to ocena ogólna, która może być zafałszowana przez inne cechy (np. humor), co utrudnia izolację samej 'wartościowości' w analizie. *Dowód: Q02_11 Poprawka:* Dodaj pytanie (lub sprawdzoną skalę) mierzące tę zmienną.
- **[ważny] Niepełny pomiar zmiennej** (G02; pewność 0,65; ai:alignment). zrozumienie odniesień (zmienna zależna; dla H2) mierzona tylko częściowo (Q02_4, Q02_20). Zrozumienie odniesień mierzone itemami o prostocie i szybkim zrozumieniu przesłania. Nie mierzy to bezpośrednio 'zrozumienia odniesień' (referencji), lecz ogólną łatwość odbioru. Konstrukty są mylone. *Dowód: Q02_4, Q02_20 Poprawka:* Dodaj pytanie (lub sprawdzoną skalę) mierzące tę zmienną.
- **[ważny] Niepełny pomiar zmiennej** (G02; pewność 0,65; ai:alignment). emocje wzbudzone przez memy (zmienna zależna; dla H3) mierzona tylko częściowo (Q02_12). Emocje wzbudzone przez memy mierzone tym samym itemem co V5 (Q02/r12). Pytanie odnosi się do 'materiału wideo', a nie specyficznie do 'memów'. Może to wprowadzać błąd kontekstowy, jeśli badany materiał nie jest memem w ścisłym sensie. *Dowód: Q02_12 Poprawka:* Dodaj pytanie (lub sprawdzoną skalę) mierzące tę zmienną.
- **[ważny] Niepełny pomiar zmiennej** (G02; pewność 0,65; ai:alignment). potencjał wiralowy (zmienna zależna; dla H4, C1) mierzona tylko częściowo (Q02_1, Q02_10, Q02_14, Q02_16, Q02_19). Potencjał wiralowy mierzony baterią 5 itemów (udostępnienie, utożsamienie, komentarz, ponowne obejrzanie, przyciągnięcie uwagi). Itemy te są heterogeniczne (mieszą intencję udostępnienia z zaangażowaniem). Skala może mieć niską wiarygodność (alfa Cronbacha), co jest problemem wymienionym w uwagach. *Dowód: Q02_1, Q02_10, Q02_14,*

Q02_16, Q02_19 *Poprawka*: Dodaj pytanie (lub sprawdzoną skalę) mierzące tę zmienną.

- **[ważny] Niejednorodna bateria** (M06; pewność 0,56; AI: przegląd całości). Siatka łączy różne aspekty (humor, zrozumiałość, udostępnianie), co może czynić ją heterogeniczną i trudną do interpretacji. *Dowód*: Siatka łączy humor, zrozumiałość i udostępnianie. *Poprawka*: Rozdziel siatkę na podskale mierzące konkretne konstrukty (np. humor, zrozumiałość, wartość) lub упростить treść.
- **[drobny] Brak „nie wiem / nie dotyczy / wolę nie odpowiadać”** (O04; pewność 0,90; AI: metodolog). Brak opcji ‘Nie dotyczy’ lub ‘Nie mam zdania’ zmusza respondentów do zgadywania, jeśli materiał jest im nieznany lub nieistotny. *Dowód*: W ogóle się nie zgadzam / ... / Całkowicie się zgadzam *Poprawka*: Dodaj opcję ‘Nie dotyczy’ lub ‘Nie mam zdania’ na końcu skali.
- **[drobny] Niespójna terminologia** (F04; pewność 0,75; AI: przegląd całości). Terminologia jest spójna, choć występuje drobna zmiana od ‘materiał’ do ‘film’ w siatce, co może być mylące. *Dowód*: Zmiana terminu od ‘materiał’ do ‘film’ w Q02. *Poprawka*: Ujednolicić terminologię, używając jednego słowa (np. ‘materiał’ lub ‘film’) dla wszystkich pozycji.

Propozycja: Proszę ocenić poniższe twierdzenia dotyczące załączonego materiału. Odtwórz materiał, a następnie zaznacz jedną odpowiedź dla każdego wiersza.

Nowe wiersze: (1) Zachęciłbym/abym rodzinę/przyjaciół do obejrzenia tego materiału.; (2) Czuję, że inni mogliby się utożsamić z tym, co jest przedstawione na filmie.; (3) Materiał zachęca do wyrażenia opinii (np. w formie oceny, komentarza).; (4) Materiał wideo zawiera odniesienia/znaczenia, które rozumiem.; (5) Treść filmu jest prosta.; (6) Forma jest dopasowana do realiów platform społecznościowych.; (7) Obejrzałbym/abym ponownie ten materiał.; (8) Wideo zawiera element humoru.; (9) Wideo posługuje się stylem, który pasuje do internetu.; (10) Mogłbym/abym udostępnić taki materiał w swoich social mediach.; (11) Uważam, że wideo jest wartościowe.; (12) Materiał wzbudził we mnie emocje.; (13) W mojej opinii film jest zrozumiały dla wielu grup społecznych w moim kraju.; (14) Chciałbym/abym obejrzeć więcej filmów stworzonych w podobnym stylu.; (15) W filmie jest element ironii, sarkazmu lub parodii.; (16) Wideo jest „bezpieczne” do udostępnienia.; (17) Materiał przykuł moją uwagę.; (18) Film jest na tyle krótki, że łatwo go obejrzeć w dowolnej chwili.; (19) Materiał przykuł moją uwagę.; (20) Szybko zrozumiałem/am główne przesłanie filmu.

Uzasadnienie: Rozdzielono podwójne twierdzenie w wierszu 5 na dwie osobne kwestie, usunięto niejasne sformułowania w wierszu 15, dodano opcję ‘Nie mam zdania’ do skali odpowiedzi oraz poprawiono interpunkcję i styl w wierszach 18 i 19 dla lepszej czytelności.

Q12

Ile czasu dziennie spędzasz w mediach społecznościowych (podaj średni czas spędzany dziennie w godzinach)?

str. 4 · tekst krótki · obowiązkowe · jakość 30/100

- **[ważny] Niewłaściwy format odpowiedzi** (O10; pewność 0,97; AI: metodolog + AI: rzecznik badanego + reguła). Pole tekstowe bez walidacji liczbowej przyjmie odpowiedzi tekstowe (‘około 2’, ‘wiele’), co uniemożliwi analizę statystyczną. *Dowód*: podaj średni czas spędzany dziennie w godzinach *Poprawka*: Użyj pola numerycznego z wymaganym formatem liczbowym lub suwaka z przedziałem.
- **[ważny] Poziom pomiaru nie pozwala na analizę** (G03; pewność 0,80; reguła). czas spędzany w mediach społecznościowych (H2) jest wpisywana jako tekst — przed analizą trzeba ją czyścić i zamieniać na liczby, część odpowiedzi przypadnie. *Dowód*: Q12: a number typed as free text (needs cleaning) *Poprawka*: Użyj pola liczbowego z jednostką i zakresem.
- **[ważny] Niepełny pomiar zmiennej** (G02; pewność 0,65; ai:alignment). czas spędzany w mediach społecznościowych (predyktor; dla H2) mierzona tylko częściowo (Q12). Czas w mediach mierzony jako średni czas dzienny w godzinach wpisywany swobodnie. Format tekstowy wymaga czyszczenia danych. Odpowiedź ‘średni czas’ jest obciążona błędem pamięci i subiektywną oceną respondenta, a nie danymi z platform. *Dowód*: Q12 *Poprawka*: Dodaj pytanie (lub sprawdzoną skalę) mierzące tę zmienną.
- **[drobny] Obowiązkowe pytania otwarte** (B03; pewność 0,75; AI: przegląd całości). Wymagane otwarte pytanie o czas może prowadzić do nieprecyzyjnych odpowiedzi i zwiększyć czas wypełniania. *Dowód*: Ile czasu dziennie spędzasz w mediach społecznościowych (podaj średni czas...) *Poprawka*: Rozważyć zamianę na pytanie zamknięte z przedziałami czasowymi, aby ułatwić analizę i zmniejszyć obciążenie.

Propozycja: Podaj, ile godzin średnio dziennie spędzasz w mediach społecznościowych.

Uzasadnienie: Zmieniono instrukcję na jasną prośbę o podanie liczby godzin, co wymusza odpowiedź numeryczną i ułatwia dalszą analizę statystyczną bez konieczności ręcznego czyszczenia danych.

Q13

Jaka jest Twoja płeć biologiczna?

str. 4 · jednokrotny wybór · obowiązkowe · jakość 40/100

Odpowiedzi: Kobieta / Mężczyzna

- **[ważny] Język lub kategorie niewłączające** (E06; pewność 0,97; AI: przegląd całości + AI: metodolog + AI: rzecznik badanego + reguła). Brak opcji 'Inna' lub 'Nie chcę podawać' wyklucza respondentów niebinarnych lub tych, którzy nie chcą ujawniać płci, naruszając zasady inkluzji. *Dowód:* Kobieta / Mężczyzna *Poprawka:* Dodaj opcje 'Inna' lub 'Nie chcę podawać' oraz rozważ rozdzielenie płci biologicznej od tożsamości płciowej.
- **[ważny] Niepełny pomiar zmiennej** (G02; pewność 0,65; ai:alignment). płeć (zmienna grupująca; dla H3) mierzona tylko częściowo (Q13). Płeć mierzona jako biologiczna (kobieta/mężczyzna). Brak opcji 'inna' lub 'wolę nie mówić' narusza zasady etyczne i może wykluczać respondenci niebinarnych, co wprowadza błąd doboru próby w kontekście współczesnych badań społecznych. *Dowód:* Q13 *Poprawka:* Dodaj pytanie (lub sprawdzoną skalę) mierzące tę zmienną.

Propozycja: Jaka jest Twoja płeć?

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

Uzasadnienie: Usunięto określenie 'biologiczna', które jest kontrowersyjne i niejasne, oraz dodano opcje 'Inna' i 'Nie chcę podać', aby zapewnić inkluzję i poszanowanie prywatności respondentów.

Q14

Ile masz lat?

str. 4 · tekst krótki · obowiązkowe · jakość 30/100

- **[ważny] Niewłaściwy format odpowiedzi** (O10; pewność 0,97; AI: metodolog + AI: rzecznik badanego + reguła). Pole tekstowe bez walidacji liczbowej przyjmie odpowiedzi nieliczbowe, co utrudni kodowanie danych demograficznych. *Dowód:* Ile masz lat? *Poprawka:* Użyj pola numerycznego z zakresem wiekowym lub listy rozwijanej z przedziałami wiekowymi.
- **[drobny] Obowiązkowe pytania otwarte** (B03; pewność 0,75; AI: przegląd całości). Wymagane otwarte pytanie o wiek może prowadzić do błędów wpisania i zwiększyć czas wypełniania. *Dowód:* Ile masz lat? (wymagane) *Poprawka:* Rozważyć zamianę na pytanie zamknięte z przedziałami wiekowymi, aby ułatwić analizę i zmniejszyć obciążenie.
- **[drobny] Niepełny pomiar zmiennej** (G02; pewność 0,65; ai:alignment). wiek (cecha populacji; dla populacji badania) mierzona tylko częściowo (Q14). Wiek mierzony jako liczba lat wpisywana swobodnie. Format tekstowy wymaga walidacji. Brak ograniczeń wiekowych w pytaniu (choć target to 18-35) może prowadzić do błędnych wpisów. *Dowód:* Q14 *Poprawka:* Dodaj pytanie (lub sprawdzoną skalę) mierzące tę zmienną.

Propozycja: Podaj swój wiek w pełnych latach.

Uzasadnienie: Sformułowano pytanie tak, aby wymuszać podanie konkretnej liczby całkowitej, co eliminuje ryzyko wpisania odpowiedzi nieliczbowych i ułatwia kodowanie danych demograficznych.

Q16

Jak bardzo byłeś/byłaś zmotywowany/a do ukończenia badania?

str. 4 · jednokrotny wybór · obowiązkowe · jakość 40/100

Odpowiedzi: Brak motywacji / Umiarkowana motywacja / Wysoka motywacja

- **[ważny] Nieodpowiednia liczba punktów skali** (O06; pewność 0,94; AI: metodolog + AI: rzecznik badanego + reguła). Skala 3-punktowa ma niską dyskryminację i zwiększa ryzyko satysfakcjonowania (Krosnick); trudniej

wykryć subtelne różnice w motywacji. *Dowód:* Brak motywacji / Umiarkowana motywacja / Wysoka motywacja
Poprawka: Rozszerz skalę do 5 lub 7 punktów, aby zwiększyć precyzję pomiaru.

Propozycja: Jak bardzo byłeś/byłaś zmotywowany/a do ukończenia tego badania?

Nowe odpowiedzi: Wcale nie byłem/byłam zmotywowany/a / Byłem/byłam słabo zmotywowany/a / Byłem/byłam umiarkowanie zmotywowany/a / Byłem/byłam mocno zmotywowany/a / Byłem/byłam bardzo mocno zmotywowany/a

Uzasadnienie: Rozszerzono skalę do 5 punktów, aby zwiększyć dyskryminację odpowiedzi i zmniejszyć ryzyko satysfakcjonowania, zastępując nieprecyzyjne etykiety bardziej precyzyjnymi stopniami motywacji.

Q17

Jak szczerze były Twoje odpowiedzi podczas wypełniania badania?

str. 4 · jednokrotny wybór · obowiązkowe · jakość 40/100

Odpowiedzi: Niski / Umiarkowany / Wysoki

- **[ważny] Nieodpowiednia liczba punktów skali** (O06; pewność 0,94; AI: metodolog + AI: rzecznik badanego + reguła). Skala 3-punktowa do pomiaru szczerości jest zbyt gruba i podatna na błędy pomiaru; 'umiarkowany' jest pojęciem niejasnym. *Dowód:* Niski / Umiarkowany / Wysoki *Poprawka:* Rozszerz skalę do 5 lub 7 punktów i używaj etykiet opisowych na każdym poziomie.

Propozycja: Jak szczerze były Twoje odpowiedzi podczas wypełniania tego badania?

Nowe odpowiedzi: Wcale nie szczerze / Słabo szczerze / Umiarkowanie szczerze / Bardzo szczerze / Wyjątkowo szczerze

Uzasadnienie: Rozszerzono skalę do 5 punktów, aby zapewnić lepszą wiarygodność pomiaru i uniknąć niejasności związanych z pojęciem 'umiarkowany' w skali 3-punktowej.

Q18

Jaki jest Twój status związku?

str. 4 · jednokrotny wybór · obowiązkowe · jakość 40/100

Odpowiedzi: W związku / Singiel

- **[ważny] Luki, niepełna lista odpowiedzi** (O02; pewność 0,97; AI: przegląd całości + AI: metodolog + AI: rzecznik badanego). Kategorie są niepełne (brak 'rozwiedziony', 'wdowa/wdowiec', 'w separacji') i używają potocznego określenia 'Singiel', które nie jest standardem akademickim. *Dowód:* W związku / Singiel *Poprawka:* Rozszerz listę o wszystkie stany cywilne i użyj standardowych terminów (np. 'w związku małżeńskim', 'w konkubinacie').

Propozycja: Jaki jest Twój aktualny status związku?

Nowe odpowiedzi: W związku małżeńskim / W związku cywilnym / W związku partnerskim (nieformalnym) / Rozwiedziony/a / Wdowa/Wdowiec / Rozdzielony/a / Pozostaję singlem / Inna odpowiedź

Uzasadnienie: Rozszerzono listę opcji o brakujące kategorie (rozwiedzeni, wdowy, rozdzieloni) i zastąpiono potoczne 'Singiel' standardowym określeniem 'Pozostaję singlem', aby lista była wyczerpująca i akademicka.

Q20

Czy w trakcie badania w Twoim otoczeniu wydarzyło się coś, co mogło mieć wpływ na jego przebieg?

str. 4 · tekst krótki · jakość 30/100

- **[ważny] Ukryte założenie** (W08; pewność 0,90; AI: metodolog). Pytanie zawiera ukryte założenie, że coś się wydarzyło, co może zmusić respondentów do wymyślania zdarzeń lub podawania nieistotnych informacji, jeśli nic się nie stało. *Dowód:* Czy w trakcie badania w Twoim otoczeniu wydarzyło się coś... *Poprawka:* Zmienić na neutralne pytanie: 'Czy w trakcie wypełniania ankiety wydarzyło się w Twoim otoczeniu coś, co miało wpływ na jej przebieg?' z opcją 'Nic'.

- **[drobny] Zbyt długie lub złożone** (W11; pewność 0,60; AI: rzecznik badanego). Pytanie jest nieco zawile i abstrakcyjne; 'przebieg badania' może być mylony z przebiegiem samego procesu oglądania memów, a nie zewnętrznymi zakłóceniami. *Dowód:* Czy w trakcie badania w Twoim otoczeniu wydarzyło się coś, co mogło mieć wpływ na jego przebieg? *Poprawka:* Uprościć sformułowanie, np.: 'Czy coś z zewnątrz (hałas, rozmowa) odwracało Twoją uwagę podczas wypełniania ankiety?'.

Propozycja: Czy w trakcie wypełniania tego ankiety wydarzyły się jakiekolwiek zdarzenia lub zakłócenia (np. rozmowy, dzwonek telefonu, praca innych osób), które mogły wpłynąć na sposób, w jaki odpowiadałeś/aś na pytania?

Uzasadnienie: Usunięto ukryte założenie, że coś się wydarzyło, poprzez sformułowanie pytania w sposób otwarty i neutralny, a także dodano konkretne przykłady zakłóceń, aby zwiększyć zrozumiałość i odnieść pytanie do procesu wypełniania ankiety, a nie abstrakcyjnego 'przebiegu badania'.

Bez uwag: Q15, Q19

Uwagi dotyczące wszystkich pytań

- **[drobny] Format zgody zamiast skali specyficznej** (O08). Format „zgadzam się / nie zgadzam się” w 10 macierzy(ach) (Q02, Q03, Q04, Q05, Q06, Q07). Skale specyficzne dla treści (np. „jak często”, „jak bardzo”) są zwykle rzetelniejsze i mniej podatne na potakiwanie — chyba że to wystandaryzowany kwestionariusz.

6 Kafeterie i skale odpowiedzi

ID	Typ	Pkt	+/-	Zrównoważ.	Poza skalą	Kierunek	Odpowiedzi
Q01	nominalna	5					xxx / xxx / xxx / xxx / xxx
Q02	zgoda	4	2/0/2	tak		od negat.	W ogóle się nie zgadzam / Trochę się nie zgadzam / Trochę się zgadzam...
Q13	nominalna	2					Kobieta / Mężczyzna
Q15	nominalna	3					Podstawowe / Średnie / Wyższe
Q16	natężenie	3	1/1/1	tak		od negat.	Brak motywacji / Umiarkowana motywacja / Wysoka motywacja
Q17	natężenie	3	1/1/1			od negat.	Niski / Umiarkowany / Wysoki
Q18	nominalna	2					W związku / Singiel

ID	Waga	Problem	Komentarz
Q02	drobny	Brak „nie wiem / nie dotyczy / wolę nie odpowiadać”	Brak opcji 'Nie dotyczy' lub 'Nie mam zdania' zmusza respondentów do zgadywania, jeśli materiał jest im nieznany lub nieistotny.
Q12	ważny	Niewłaściwy format odpowiedzi	Pole tekstowe bez walidacji liczbowej przyjmie odpowiedzi tekstowe ('około 2', 'wiele'), co uniemożliwi analizę statystyczną.
Q14	ważny	Niewłaściwy format odpowiedzi	Pole tekstowe bez walidacji liczbowej przyjmie odpowiedzi nieliczbowe, co utrudni kodowanie danych demograficznych.
Q16	ważny	Nieodpowiednia liczba punktów skali	Skala 3-punktowa ma niską dyskryminację i zwiększa ryzyko satysfakcjonowania (Krosnick); trudniej wykryć subtelne różnice w motywacji.
Q17	ważny	Nieodpowiednia liczba punktów skali	Skala 3-punktowa do pomiaru szczerości jest zbyt gruba i podatna na błędy pomiaru; 'umiarkowany' jest pojęciem niejasnym.
Q18	ważny	Luki, niepełna lista odpowiedzi	Kategorie są niepełne (brak 'rozwódziona', 'wdowa/wdowiec', 'w separacji') i używają potocznego określenia 'Singiel', które nie jest standardem akademickim.
—	drobny	Format zgody zamiast skali specyficznej	Format „zgadzam się / nie zgadzam się” w 10 macierzy(ach) (Q02, Q03, Q04, Q05, Q06, Q07). Skale specyficzne dla treści (np. „jak często”, „jak bardzo”) są zwykle rzetelniejsze i mniej podatne na potakiwanie — chyba że t...

7 Kontrola jakości danych

Kryterium	Status	Pkt	Komentarz
Pytanie sprawdzające uwagę (w dłuższej ankiecie)	nie	0/35	Dłuższe badanie nie zawiera instrukcji odpowiedzi uwagi, co zwiększa ryzyko nieostrożnego wypełniania. [Brak instrukcji odpowiedzi uwagi (np. wybierz 'trochę się nie zgadzam').]

(cd.)

Kryterium	Status	Pkt	Komentarz
Pozycje odwrócone w bateriach	nie	0/25	Brak odwróconych pozycji w siatkach, co uniemożliwia wykrycie efektu przychylności. [Q02: —]
Pary sprawdzające spójność	nie	0/20	Brak par spójnościowych, które pozwoliłyby wykryć losowe odpowiadanie. [Brak par pytań o tę samą treść w różnych miejscach.]
Pytanie o rzetelność / szczerłość odpowiedzi	tak	20/20	Obecne pytania o motywację i szczerłość służą jako kontrola poważności traktowania zadania. [Q17; Q19; Q20]

Wykryte elementy kontrolne

ID	Typ	Treść	Oczekiwana odpowiedź	Źródło
Q17	pytanie o rzetelność	Jak szczerze były Twoje odpowiedzi podczas wypełniania badania?		rule
Q19	pytanie o cel badania	Jaki był Twoim zdaniem cel tego badania?		rule
Q20	pytanie o zakłócenia	Czy w trakcie badania w Twoim otoczeniu wydarzyło się coś, co mogło mieć wpływ na jego przebieg?		rule
Q16	pytanie o rzetelność	Pytanie o motywację (Q16) i szczerłość (Q17) jest podatne na efekt społecznej pożądaności. Respondenci będą chcieli wygl...	Wysoka motywacja	rule

Propozycje pytań kontrolnych do dodania

Typ	Po pytaniu	Treść	Oczekiwana odpowiedź	Uzasadnienie
pytanie z instrukcją	Q02	Aby upewnić się, że czytasz instrukcje, wybierz opcję 'Trochę się nie zgadzam' dla tej pozycji.	Wybór 'Trochę się nie zgadzam'	Element ten służy do wyłapania respondentów, którzy nie czytają instrukcji przed wypełnieniem długiej siatki (Q02-Q11). Zgodnie z zaleceniami Tourangeau i Krosnicka, takie elementy są kluczowe w długich bateriach, aby zmniejszyć ryzyko wysycenia satysfakcjonującego (satisficing) i losowego zaznaczania odpowiedzi.
pytanie z instrukcją	Q16	Wybierz opcję 'Umiarkowana motywacja', aby potwierdzić, że przeczytałeś/aś to pytanie.	Wybór 'Umiarkowana motywacja'	Pytania o motywację i szczerłość (Q16, Q17) są silnie obciążone błędem systematycznym wynikającym z efektu społecznej pożądaności. Respondenci będą chcieli wyglądać na zaangażowanych i szczerych. Element 'instructed response' pomaga oddzielić staranne odpowiedzi od tych, które są po prostu 'pozytywne' ze względu na presję społeczną, zwiększając wiarygodność danych o jakości odpowiedzi.
para spójności	Q12	Czy korzystasz z mediów społecznościowych? (Tak / Nie)	Tak	Dodanie pytania o ogólną aktywność w mediach społecznościowych przed Q12 (Ile czasu...) pozwala na weryfikację spójności odpowiedzi. Jeśli respondent odpowie 'Nie', a następnie poda czas spędzony w mediach, dane te należy uznać za błędne. To standardowa procedura kontrolna w badaniach dotyczących używania mediów.

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 = "pairwis
flag <- longstring >= 0.8 * ncol(x) | irv < 0.3
data.frame(id = d[[1]], longstring, irv = round(irv, 2), mahal, flag)
```

8 Obciążenie i komfort badanego

Kryterium	Status	Pkt	Komentarz
Szacowany czas wypełniania	16,8 min	12/25	
Największa macierz (wiersze)	20	0/15	
Obowiązkowe pytania otwarte	0	10/10	
Odsetek pytań obowiązkowych	90%	6/10	
Czytelność (FOG-PL)	8,6	15/15	
Wskaźnik postępu	tak	10/10	
Najbardziej obciążona strona (pytania / pozycje)	10 / 200	0/15	

Czytelność pytań: FOG-PL = **8,6** (łatwy), średnio 7,3 słów w zdaniu, 14% 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
Q17	13,2	Jak szczerze były Twoje odpowiedzi podczas wypełniania badania?
Q16	12,5	Jak bardzo byłeś/byłaś zmotywowany/a do ukończenia badania?
Q20	11,5	Czy w trakcie badania w Twoim otoczeniu wydarzyło się coś, co mogło mieć wpływ na jego przebieg?
Q13	10,0	Jaka jest Twoja płeć biologiczna?
Q15	10,0	Jaki jest Twój poziom wykształcenia?
Q02	8,7	Nie ma tutaj odpowiedzi złych/błędnych lub dobrych/prawidłowych. Liczy się Twoja wyjątkowa opinia na temat materiału.

ID	Waga	Problem	Komentarz
Q02	krytyczny	Duża macierz pytań	Macierz ma 20 wierszy — rośnie ryzyko odpowiadania „w jednej linii” (straightlining) i zmęczenia, zwłaszcza na telefonie. AI: 20 wierszy w siatce, powtórzonych 10 razy, spowoduje silne zmęczenie, prostolinijność i odrzu...
Q02	krytyczny	Powtarzalność, znużenie	Ten sam blok (20 pozycji) badany wypełnia 10 razy (np. po każdym materiale; kopie: Q03, Q04, Q05, Q06, Q07, Q08, Q09, Q10...) — razem 200 ocen; pod koniec rośnie zmęczenie, odpowiadanie „w jednej linii” i porzucenia. AI: ...
Q12	drobny	Obowiązkowe pytania otwarte	Wymagane otwarte pytanie o czas może prowadzić do nieprecyzyjnych odpowiedzi i zwiększyć czas wypełniania.
Q14	drobny	Obowiązkowe pytania otwarte	Wymagane otwarte pytanie o wiek może prowadzić do błędów wpisania i zwiększyć czas wypełniania.
—	ważny	Zbyt długa ankieta	Szacowany czas wypełniania ok. 17 min (210 pozycji) — wysokie ryzyko porzuceń.
—	drobny	Wymuszanie odpowiedzi	90% pytań jest obowiązkowych.
—	drobny	Przeładowana strona	Strona 3: 10 pytań, 200 pozycji.

9 Spójność i rzetelność pomiaru

Kryterium	Status	Pkt	Komentarz
Konstrukty mierzone kilkoma pozycjami	tak	25/25	Konstrukty są mierzone wieloma pozycjami w siatkach, co poprawia niezawodność. [Siatki Q02-Q11 zawierają wiele pozycji oceniających materiał.]
Jednorodne baterie	częściowo	10/20	Siatka łączy różne aspekty (humor, zrozumiałość, udostępnianie), co może czynić ją heterogeniczną. [Siatka Q02 łączy humor, zrozumiałość i udostępnianie.]
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. Odpowiedzi w baterii Q02-Q11 są prawdopodobnie mało spójne wewnętrznie (niska wiarygodność alfa Cronbacha) ze względu na heterogeniczność konstruktywną (mieszanie konstruktywnych, emocjonalnych i behawioralnych) oraz brak skal odwróconych, co utrudnia wykrycie losowych odpowiedzi. Respondenci mogą wykazywać silny efekt akwiescencji (zgadzania się z większością twierdzeń) ze względu na długą listę i brak 'neutralnych' opcji (brak 'Nie mam zdania' lub 'Rozumiem, ale nie zgadzam się'). Pytania Q16 i Q17 są silnie obciążone błędem systematycznym wynikającym z efektu społecznej pożądaności, co może prowadzić do przesadnej oceny motywacji i szczerości. Długie siatki (Q02-Q11) są podatne na wysycenie satysfakcjonującego (satisficing), co może objawiać się monotonnym zaznaczaniem tej samej opcji (np. 'Całkowicie się zgadzam') bez głębszego przetwarzania treści. Brak opcji 'Preferuję nie odpowiadać' lub 'Nie wiem' w Q01-Q11 może zmuszać respondentów do podawania fałszywych odpowiedzi, co obniża wiarygodność danych.

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
Q02	Ocena materiału wideo (zainteresowanie, zrozumienie, wartość, udostępnialność)	20	niska	0,78 (0,51–0,87)	—		Bateria jest heterogeniczna i mierzy wiele różnych konstruków (emocje, zrozumienie, humor, udostępnialność, wartość). Elementy takie jak 'humor' (r8) czy 'ironia' (r15) są specyficzne i mogą nie kore...

10 Struktura i przebieg

Kryterium	Status	Pkt	Komentarz
Wprowadzenie i instrukcja	tak	20/20	Wstęp jest przyjazny i zawiera podstawowe informacje. [60 słów przed pierwszym pytaniem]
Zakończenie i podziękowanie	brak danych	—	
Logiczna kolejność, łagodne otwarcie, bloki tematyczne	częściowo	10/20	Kolejność jest logiczna (tematyczne -> demografia), ale długa sekcja siatkowa może męczyć przed pytaniami demograficznymi. [Pytania demograficzne i wrażliwe na końcu, ale siatki przed nimi.]
Pytania drażliwe i metryczka bliżej końca	tak	20/20	Pytania demograficzne i wrażliwe zostały umieszczone na końcu, co jest zalecaną praktyką. [Pytania o płeć, wiek i wykształcenie są na końcu (Q13-Q15).]
Spójna terminologia	częściowo	8/15	Terminologia jest spójna, choć występuje drobna zmiana od 'materiał' do 'film' w siatce, co może być mylące. [Zmiana terminu od 'materiał' do 'film' w Q02.]
Brak zbędnych powtórzeń	tak	15/15	Nie stwierdzono niepotrzebnych powtórzeń pytań w widocznej ścieżce. [Brak powtórzeń pytań w widocznej ścieżce.]

ID	Waga	Problem	Komentarz
Q01	krytyczny	Nielosowy przydział do warunków	Odpowiedzi nie różnią się treścią (xxx / xxx / xxx / xxx / xxx) — to wygląda na przydział do warunków badania przez samego badanego. Taki przydział nie jest losowy (ludzie wybierają np. pierwszą lub środkową opcję), więc grupy mogą się systematycznie różnić, a badany widzi „dziwne” pytanie. AI: Ins...
Q02	drobny	Niespójna terminologia	Terminologia jest spójna, choć występuje drobna zmiana od 'materiał' do 'film' w siatce, co może być mylące. Poprawka: Ujednolicić terminologię, używając jednego słowa (np. 'materiał' lub 'film') dla wszystkich pozycji.
—	drobny	Niespójna terminologia	Kryterium spełnione częściowo: Spójna terminologia. Terminologia jest spójna, choć występuje drobna zmiana od 'materiał' do 'film' w siatce, co może być mylące.
—	drobny	Nielogiczna kolejność	Kryterium spełnione częściowo: Logiczna kolejność, łagodne otwarcie, bloki tematyczne. Kolejność jest logiczna (tematyczne -> demografia), ale długa sekcja siatkowa może męczyć przed pytaniami demograficznymi.

11 Technika, telefon i dostępność

Strony przetestowano na komputerze (1280×900) i na telefonie (Pixel 7). Test axe-core 4.12.1 sprawdza reguły WCAG 2.0, 2.1 i 2.2 (poziomy A i AA) oraz dobre praktyki.

Reguła	Problem	Wpływ	Gdzie	Elem.	Przykład
button-name	Przycisk bez nazwy	critical	komputer: str. 1, 2, 3, 4; telefon: str. 1, 2, 3, 4	8	#navbar-toggler
color-contrast	Za słaby kontrast tekstu	serious	komputer: str. 1, 2, 3, 4; telefon: str. 1, 2, 3, 4	608	#ls-button-submit
listitem	Element listy poza listą	serious	komputer: str. 2, 4; telefon: str. 2, 4	32	#jvatbd862578X73X35392
region	Treść poza obszarami nawigacyjnymi	moderate	komputer: str. 1, 2, 3, 4; telefon: str. 1, 2, 3, 4	8	.top-container
page-has-heading-one	Brak nagłówka pierwszego stopnia	moderate	komputer: str. 2, 3, 4; telefon: str. 2, 3, 4	6	html

Pomiar układu stron na telefonie

Strona	Szer. układu (px)	Przewijanie w poziomie (px)	Małe pola dotyku	Tekst < 12 px	Mediana czcionki	Za szerokie elementy
1	412	0	0	0	21	
2	412	0	5	0	18	
3	412	0	0	0	18	
4	412	0	11	0	18	

Waga	Problem	Komentarz
ważny	Za małe pola dotyku	16 pól odpowiedzi ma obszar dotyku mniejszy niż 24×24 px. Poprawka: Powiększ klikalny obszar (cała etykieta lub komórka klikalna, min. 24–44 px).

12 Etyka, RODO i inkluzywność

Kryterium	Status	Pkt	Komentarz
Cel badania	tak	12/12	Cel badania jest jasno określony w tekście wprowadzającym. [...] ankiety IMF_3 Cześć! Witaj w ankiecie poświęconej memom! Chcielibyśmy poznać Twoje zdanie na temat ich ogląd...]
Kto prowadzi badanie (instytucja, zespół)	nie	0/8	Nie podano nazwy badacza, zespołu ani instytucji, co jest wymagane przez standardy etyczne i RODO. [Brak informacji o autorze lub instytucji prowadzącej badanie.]
Kontakt do badaczy	nie	0/10	Respondenci nie mają możliwości skontaktowania się z badaczem w przypadku pytań lub wątpliwości. [Brak adresu e-mail lub telefonu kontaktowego w tekście.]

(cd.)

Kryterium	Status	Pkt	Komentarz
Dobrowolność udziału	nie	0/10	Nie zaznaczono wyraźnie, że udział jest dobrowolny, co jest podstawowym wymogiem etycznym. [Brak informacji o dobrowolnym charakterze udziału.]
Anonimowość lub poufność odpowiedzi	tak	12/12	Anonimowość danych została wyraźnie zapewniona w tekście wprowadzającym. [...mie około 20 minut. Twoje odpowiedzi są anonimowe, dlatego możesz być szczerzy Z góry dziękujemy za Twój udział...]
Informacja o przetwarzaniu danych (RODO)	nie	0/14	Brak klauzuli informacyjnej zgodnej z RODO dotyczącej przetwarzania danych osobowych. [Brak informacji o administratorze danych i podstawie prawnej.]
Możliwość przerwania w każdej chwili	nie	0/8	Nie poinformowano respondentów o prawie do wycofania zgody lub przerwania badania w dowolnym momencie. [Brak informacji o prawie do wycofania się z badania.]
Wyraźne wyrażenie zgody	nie	0/12	Nie ma wyraźnego mechanizmu wyrażenia zgody (np. checkbox), a samo wypełnienie ankiety nie zawsze jest wystarczającym potwierdzeniem świadomej zgody w kontekście RODO. [Brak wyraźnego oświadczenia zgody na udział w badaniu.]
Przewidywany czas wypełniania	tak	6/6	Szacowany czas wypełnienia został podany w tekście wprowadzającym. [...życiu. Wypełnienie ankiety zajmie około 20 minut. Twoje odpowiedzi są anonimowe, dlatego możesz być szczerzy...]
Warunek wieku / grupa docelowa	nie dotyczy	—	Nie dotyczy, brak informacji o ograniczeniach wiekowych, ale pytanie o wiek jest wymagane.

ID	Waga	Problem	Komentarz
Q13	ważny	Język lub kategorie niewłaściwe	Brak opcji 'Inna' lub 'Nie chcę podawać' wyklucza respondentów niebinarnych lub tych, którzy nie chcą ujawniać płci, naruszając zasady inkluzji. Poprawka: Dodaj opcje 'Inna' lub 'Nie chcę podawać' oraz rozważ rozdzielenie płci biologicznej od tożsamości płciowej.
—	ważny	Brak świadomej zgody	Brak wyraźnego mechanizmu wyrażenia zgody (np. checkbox) jest naruszeniem standardów etycznych i RODO. Poprawka: Dodaj wyraźne oświadczenie zgody na udział w badaniu oraz checkbox potwierdzający zapoznanie się z regulaminem.
—	ważny	Brak informacji RODO	Brak klauzuli informacyjnej zgodnej z RODO dotyczącej przetwarzania danych osobowych. Poprawka: Dodaj klauzulę informacyjną RODO z danymi administratora, celem przetwarzania i prawami użytkownika.
—	ważny	Brak kontaktu do badacza	Respondenci nie mają możliwości skontaktowania się z badaczem w przypadku pytań lub wątpliwości. Poprawka: Dodaj adres e-mail do kontaktu z zespołem badawczym w sekcji wprowadzającej lub zamykającej.
—	ważny	Brak świadomej zgody	Kryterium niespełnione: Wyraźne wyrażenie zgody. Nie ma wyraźnego mechanizmu wyrażenia zgody (np. checkbox), a samo wypełnienie ankiety nie zawsze jest wystarczającym potwierdzeniem świadomej zgody w kontekście RODO.
—	drobny	Brak informacji o możliwości rezygnacji	Nie poinformowano respondentów o prawie do wycofania zgody lub przerwania badania w dowolnym momencie. Poprawka: Dodaj informację o prawie do wycofania zgody i przerwania badania w dowolnym momencie bez konsekwencji.
—	drobny	Nie wiadomo, kto prowadzi badanie	Nie podano nazwy badacza, zespołu ani instytucji, co jest wymagane przez standardy etyczne i RODO. Poprawka: Podaj nazwę instytucji lub zespołu badawczego prowadzącego badanie w tekście wprowadzającym.

13 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	Wybierz jedną z poniższych odpowiedzi. Jeśli nie wiesz jaką wybrać, to "strzelaj".	nominalna	1 = xxx; 2 = xxx; 3 = xxx; 4 = xxx; 5 = xxx
Q02_1	Zachęciłbym/abym rodzinę/przyjaciół do obejrzenia tego materiału.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam

(cd.)

Zmienna	Etykieta	Typ	Wartości
Q02_2	Czuję, że inni mogliby się utożsamić z tym co jest przedstawione na filmie.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_3	Materiał zachęca do wyrażenia opinii (np. w formie oceny, komentarza).	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_4	Materiał wideo zawiera odniesienia/znaczenia które rozumiem.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_5	Treść filmu jest prosta i łatwa do zapamiętania.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_6	Forma jest dopasowana do realiów platform społecznościowych.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_7	Obejrzałbym/abym ponownie ten materiał.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_8	Wideo zawiera element humoru.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_9	Wideo posługuje się stylem który pasuje do internetu.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_10	Mógłbym/abym udostępnić taki materiał w swoich social mediach.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_11	Uważam, że wideo jest wartościowe.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_12	Materiał wzbudził we mnie emocje.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_13	W mojej opinii film jest zrozumiały dla wielu grup społecznych w moim kraju.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_14	Chciałbym/abym obejrzeć więcej filmów stworzonych w podobnym stylu.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_15	W filmie jest element ironii, sarkazmu lub parodii.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_16	Wideo jest "bezpieczne" do udostępnienia.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_17	Materiał wywołuje efekt „wow”.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_18	Film jest na tyle krótki, że łatwo go obejrzeć w dowolnej chwili.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_19	Materiał przykuł moją uwagę.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q02_20	Szybko zrozumiałem/am główne przesłanie filmu.	porządkowa	1 = W ogóle się nie zgadzam; 2 = Trochę się nie zgadzam; 3 = Trochę się zgadzam; 4 = Całkowicie się zgadzam
Q03_1– Q03_20	= Q02 (kolejny materiał / next material)	jak Q02	jak Q02 / as Q02
Q04_1– Q04_20	= Q02 (kolejny materiał / next material)	jak Q02	jak Q02 / as Q02
Q05_1– Q05_20	= Q02 (kolejny materiał / next material)	jak Q02	jak Q02 / as Q02
Q06_1– Q06_20	= Q02 (kolejny materiał / next material)	jak Q02	jak Q02 / as Q02
Q07_1– Q07_20	= Q02 (kolejny materiał / next material)	jak Q02	jak Q02 / as Q02
Q08_1– Q08_20	= Q02 (kolejny materiał / next material)	jak Q02	jak Q02 / as Q02
Q09_1– Q09_20	= Q02 (kolejny materiał / next material)	jak Q02	jak Q02 / as Q02

(cd.)

Zmienna	Etykieta	Typ	Wartości
Q10_1– Q10_20	= Q02 (kolejny materiał / next material)	jak Q02	jak Q02 / as Q02
Q11_1– Q11_20	= Q02 (kolejny materiał / next material)	jak Q02	jak Q02 / as Q02
Q12	Ile czasu dziennie spędzasz w mediach społecznościowych (podaj średni czas spędzany dzień...	tekst	
Q13	Jaka jest Twoja płeć biologiczna?	nominalna	1 = Kobieta; 2 = Mężczyzna
Q14	Ile masz lat?	tekst	
Q15	Jaki jest Twój poziom wykształcenia?	nominalna	1 = Podstawowe; 2 = Średnie; 3 = Wyższe
Q16	Jak bardzo byłeś/byłaś zmotywowany/a do ukończenia badania?	porządkowa	1 = Brak motywacji; 2 = Umiarkowana motywacja; 3 = Wysoka motywacja
Q17	Jak szczerze były Twoje odpowiedzi podczas wypełniania badania?	porządkowa	1 = Niski; 2 = Umiarkowany; 3 = Wysoki
Q18	Jaki jest Twój status związku?	nominalna	1 = W związku; 2 = Singiel
Q19	Jaki był Twoim zdaniem cel tego badania?	tekst	
Q20	Czy w trakcie badania w Twoim otoczeniu wydarzyło się coś, co mogło mieć wpływ na jego pr...	tekst	

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
# Q02
b1_items <- c("Q02_1", "Q02_2", "Q02_3", "Q02_4", "Q02_5", "Q02_6", "Q02_7", "Q02_8", "Q02_9", "Q02_10", "Q02_11", "Q02_12", "Q02_13", "Q02_14", "Q02_15", "Q02_16", "Q02_17", "Q02_18", "Q02_19", "Q02_20")
print(alpha(d[b1_items]))
d$b1 <- rowMeans(d[b1_items], na.rm = TRUE)
```

14 Aneksy

14.1 Pełna lista problemów

ID	Kod	Problem	Waga	Pewn.	Źródło	Komentarz
Q01	F09	Nielosowy przydział do warunków	krytyczny	0,97	AI: metodolog + AI: rzecznik badanego + reguła	Odpowiedzi nie różnią się treścią (xxx / xxx / xxx / xxx / xxx) — to wygląda na przydział do warunków badania przez samego badanego. Taki przydział nie jest losowy (ludzie wybierają np. pierwszą lub ...
Q01	G07	Plan badania nie pozwala sprawdzić hipotezy	krytyczny	0,85	reguła	Plan badania zakłada losowy przydział do warunków, a w ankiecie badany sam wybiera wariant (Q01) — porównania warunków będą obciążone autoselekcją.
Q02	B02	Duża macierz pytań	krytyczny	0,97	AI: przegląd całości + AI: metodolog + AI: rzecznik badanego + reguła	Macierz ma 20 wierszy — rośnie ryzyko odpowiadania „w jednej linii” (straightlining) i zmęczenia, zwłaszcza na telefonie. AI: 20 wierszy w siatce, powtórzonych 10 razy, spowoduje silne zmęczenie, pro...
Q02	B08	Powtarzalność, znużenie	krytyczny	0,97	AI: metodolog + AI: rzecznik badanego + reguła	Ten sam blok (20 pozycji) badany wypełnia 10 razy (np. po każdym materiale; kopie: Q03, Q04, Q05, Q06, Q07, Q08, Q09, Q10...) — razem 200 ocen; pod koniec rośnie zmęczenie, odpowiadanie „w jednej linii”...
Q02/5	W02	Dwa pytania w jednym	ważny	0,97	AI: metodolog + AI: rzecznik badanego + reguła	Twierdzenie łączy dwie różne cechy (prostotę i zapamiętywalność); respondent może zgadzać się z jedną, a nie z drugą, co zniekształca wynik.
Q02/15	W02	Dwa pytania w jednym	ważny	0,97	AI: metodolog + AI: rzecznik badanego + reguła	Połączenie 'lub' tworzy podwójne przesłanie; respondent może nie utożsamiać się z żadnym z tych elementów, mimo że jeden z nich występuje.
Q02	C03	Brak pozycji odwróconych	ważny	0,68	AI: przegląd całości	Brak odwróconych pozycji w siatkach, co uniemożliwia wykrycie efektu przychylności (acquiescence bias).
Q02	G02	Niepełny pomiar zmiennej	ważny	0,65	ai:alignment	gotowość do udostępnienia (zmienna zależna; dla P1, H1) mierzona tylko częściowo (Q02_10). Zmienna gotowość do udostępnienia mierzona jest pojedynczym twierdzeniem w skali 4-punktowej. Brak jest odwr...
Q02	G02	Niepełny pomiar zmiennej	ważny	0,65	ai:alignment	wartość materiału (zmienna zależna; dla P2) mierzona tylko częściowo (Q02_11). Wartość materiału mierzona jednym itemem. Jest to ocena ogólna, która może być zafałszowana przez inne cechy (np. humor)...

(cd.)

ID	Kod	Problem	Waga	Pewn.	Źródło	Komentarz
Q02	G02	Niepełny pomiar zmiennej	ważny	0,65	ai:alignment	zrozumienie odniesień (zmienna zależna; dla H2) mierzona tylko częściowo (Q02_4, Q02_20). Zrozumienie odniesień mierzone itemami o prostocie i szybkim zrozumieniu przesłania. Nie mierzy to bezpośrednio...
Q02	G02	Niepełny pomiar zmiennej	ważny	0,65	ai:alignment	emocje wzbudzone przez memy (zmienna zależna; dla H3) mierzona tylko częściowo (Q02_12). Emocje wzbudzone przez memy mierzone tym samym itemem co V5 (Q02/r12). Pytanie odnosi się do 'materiału wideo'...
Q02	G02	Niepełny pomiar zmiennej	ważny	0,65	ai:alignment	potencjał wiralowy (zmienna zależna; dla H4, C1) mierzona tylko częściowo (Q02_1, Q02_10, Q02_14, Q02_16, Q02_19). Potencjał wiralowy mierzony baterią 5 itemów (udostępnienie, utożsamienie, komentarz...
Q02	M06	Niejednorodna bateria	ważny	0,56	AI: przegląd całości	Siatka łączy różne aspekty (humor, zrozumiałość, udostępnianie), co może czynić ją heterogeniczną i trudną do interpretacji.
Q02	O04	Brak „nie wiem / nie dotyczy / wolę nie odpowiadać”	drobny	0,90	AI: metodolog	Brak opcji 'Nie dotyczy' lub 'Nie mam zdania' zmusza respondentów do zgadywania, jeśli materiał jest im niezny lub nieistotny.
Q02	F04	Niespójna terminologia	drobny	0,75	AI: przegląd całości	Terminologia jest spójna, choć występuje drobna zmiana od 'materiał' do 'film' w siatce, co może być mylące.
Q12	O10	Niewłaściwy format odpowiedzi	ważny	0,97	AI: metodolog + AI: rzecznik badanego + reguła	Pole tekstowe bez walidacji liczbowej przyjmie odpowiedzi tekstowe ('około 2', 'wiele'), co uniemożliwi analizę statystyczną.
Q12	G03	Poziom pomiaru nie pozwala na analizę	ważny	0,80		czas spędzany w mediach społecznościowych (H2) jest wpisywana jako tekst — przed analizą trzeba ją czyścić i zamieniać na liczby, część odpowiedzi przypadnie.
Q12	G02	Niepełny pomiar zmiennej	ważny	0,65	ai:alignment	czas spędzany w mediach społecznościowych (predyktor; dla H2) mierzona tylko częściowo (Q12). Czas w mediach mierzony jako średni czas dzienny w godzinach wpisywany swobodnie. Format tekstowy wymaga ...
Q12	B03	Obowiązkowe pytania otwarte	drobny	0,75	AI: przegląd całości	Wymagane otwarte pytanie o czas może prowadzić do nieprecyzyjnych odpowiedzi i zwiększyć czas wypełniania.
Q13	E06	Język lub kategorie niewłaściwe	ważny	0,97	AI: przegląd całości + AI: rzecznik badanego + reguła	Brak opcji 'Inna' lub 'Nie chcę podawać' wyklucza respondentów niebinarnych lub tych, którzy nie chcą ujawniać płci, naruszając zasady inkluzji.
Q13	G02	Niepełny pomiar zmiennej	ważny	0,65	ai:alignment	płeć (zmienna grupująca; dla H3) mierzona tylko częściowo (Q13). Płeć mierzona jako biologiczna (kobieta/mężczyzna). Brak opcji 'inna' lub 'wolę nie mówić' narusza zasady etyczne i może wykluczać res...
Q14	O10	Niewłaściwy format odpowiedzi	ważny	0,97	AI: metodolog + AI: rzecznik badanego + reguła	Pole tekstowe bez walidacji liczbowej przyjmie odpowiedzi nieliczbowe, co utrudni kodowanie danych demograficznych.
Q14	B03	Obowiązkowe pytania otwarte	drobny	0,75	AI: przegląd całości	Wymagane otwarte pytanie o wiek może prowadzić do błędów wpisania i zwiększyć czas wypełniania.
Q14	G02	Niepełny pomiar zmiennej	drobny	0,65	ai:alignment	wiek (cecha populacji; dla populacji badania) mierzona tylko częściowo (Q14). Wiek mierzony jako liczba lat wpisywana swobodnie. Format tekstowy wymaga walidacji. Brak ograniczeń wiekowych w pytaniu ...
Q16	O06	Nieodpowiednia liczba punktów skali	ważny	0,94	AI: metodolog + AI: rzecznik badanego + reguła	Skala 3-punktowa ma niską dyskryminację i zwiększa ryzyko satysfakcjonowania (Krosnick); trudniej wykryć subtelne różnice w motywacji.
Q17	O06	Nieodpowiednia liczba punktów skali	ważny	0,94	AI: metodolog + AI: rzecznik badanego + reguła	Skala 3-punktowa do pomiaru szczerości jest zbyt gruba i podatna na błędy pomiaru; 'umiarkowany' jest pojęciem niejasnym.
Q18	O02	Luki, niepełna lista odpowiedzi	ważny	0,97	AI: przegląd całości + AI: metodolog + AI: rzecznik badanego	Kategorie są niepełne (brak 'rozwidziony', 'wdowa/wdowiec', 'w separacji') i używają potocznego określenia 'Singiel', które nie jest standardem akademickim.
Q20	W08	Ukryte założenie	ważny	0,90	AI: metodolog	Pytanie zawiera ukryte założenie, że coś się wydarzyło, co może zmusić respondentów do wymyślenia zdarzeń lub podawania nieistotnych informacji, jeśli nic się stało.
Q20	W11	Zbyt długie lub złożone	drobny	0,60	AI: rzecznik badanego	Pytanie jest nieco zawile i abstrakcyjne; 'przebieg badania' może być mylony z przebiegiem samego procesu oglądania memów, a nie zewnętrznymi zakłóceniami.
—	A01	Naruszenie WCAG (axe-core)	ważny	0,90	axe-core	Przycisk bez nazwy (button-name, wpływ: critical); desktop: str. 1, 2, 3, 4; mobile: str. 1, 2, 3, 4.
—	A01	Naruszenie WCAG (axe-core)	ważny	0,90	axe-core	Za słaby kontrast tekstu (color-contrast, wpływ: serious); desktop: str. 1, 2, 3, 4; mobile: str. 1, 2, 3, 4.
—	A01	Naruszenie WCAG (axe-core)	ważny	0,90	axe-core	Element listy poza listą (listitem, wpływ: serious); desktop: str. 2, 4; mobile: str. 2, 4.
—	E02	Brak świadomej zgody	ważny	0,90	AI: przegląd całości	Brak wyraźnego mechanizmu wyrażenia zgody (np. checkbox) jest naruszeniem standardów etycznych i RODO.
—	E03	Brak informacji RODO	ważny	0,90	AI: przegląd całości	Brak klauzuli informacyjnej zgodnej z RODO dotyczącej przetwarzania danych osobowych.

(cd.)

ID	Kod	Problem	Waga	Pewn.	Źródło	Komentarz
—	E07	Brak kontaktu do badaczy	ważny	0,90	AI: przegląd całości	Respondenci nie mają możliwości skontaktowania się z badaczem w przypadku pytań lub wątpliwości.
—	A03	Za małe pola dotyku	ważny	0,75	test telefonu	16 pól odpowiedzi ma obszar dotyku mniejszy niż 24×24 px.
—	C01	Brak pytań sprawdzających uwagę	ważny	0,75	AI: przegląd całości	Kryterium niespełnione: Pytanie sprawdzające uwagę (w dłuższej ankiecie). Dłuższe badanie nie zawiera instrukcji odpowiedzi uwagi, co zwiększa ryzyko nieostrożnego wypełniania.
—	E02	Brak świadomej zgody	ważny	0,75	AI: przegląd całości	Kryterium niespełnione: Wyraźne wyrażenie zgody. Nie ma wyraźnego mechanizmu wyrażenia zgody (np. checkbox), a samo wypełnienie ankiety nie zawsze jest wystarczającym potwierdzeniem świadomej zgody w...
—	B01	Zbyt długa ankiet	ważny	0,70	reguła	Szacowany czas wypełniania ok. 17 min (210 pozycji) — wysokie ryzyko porzucenia.
—	G01	Brak pomiaru zmiennej z hipotezy	ważny	0,70	ai:alignment	tworzenie memów (zmienna grupująca; potrzebna do H4) — żadne pytanie ankiety jej nie mierzy. Tworzenie memów nie jest mierzone w dostarczonym fragmencie ankiety. Brak pytania o aktywność twórczą resp...
—	G06	Niezgodność z badaną populacją	ważny	0,70	ai:alignment	częstotliwość korzystania z mediów społecznościowych (cecha populacji; potrzebna do populacji badania) — żadne pytanie ankiety jej nie mierzy. Częstotliwość korzystania z mediów (co najmniej raz dzie...
—	A01	Naruszenie WCAG (axe-core)	drobny	0,90	axe-core	Treść poza obszarami nawigacyjnymi (landmarks) (region, wpływ: moderate); desktop: str. 1, 2, 3, 4; mobile: str. 1, 2, 3, 4.
—	A01	Naruszenie WCAG (axe-core)	drobny	0,90	axe-core	Brak nagłówka pierwszego stopnia (page-has-heading-one, wpływ: moderate); desktop: str. 2, 3, 4; mobile: str. 2, 3, 4.
—	E08	Brak informacji o możliwości rezygnacji	drobny	0,90	AI: przegląd całości	Nie poinformowano respondentów o prawie do wycofania zgody lub przerwania badania w dowolnym momencie.
—	E11	Nie wiadomo, kto prowadzi badanie	drobny	0,90	AI: przegląd całości	Nie podano nazwy badacza, zespołu ani instytucji, co jest wymagane przez standardy etyczne i RODO.
—	C03	Brak pozycji odwróconych	drobny	0,75	AI: przegląd całości	Kryterium niespełnione: Pozycje odwrócone w bateriach. Brak odwróconych pozycji w siatkach, co uniemożliwia wykrycie efektu przychylności.
—	C04	Brak par sprawdzających spójność	drobny	0,75	AI: przegląd całości	Kryterium niespełnione: Pary sprawdzające spójność. Brak par spójnościowych, które pozwoliłyby wykryć losowe odpowiadanie.
—	F04	Niespójna terminologia	drobny	0,75	AI: przegląd całości	Kryterium spełnione częściowo: Spójna terminologia. Terminologia jest spójna, choć występuje drobna zmiana od 'material' do 'film' w siatce, co może być mylące.
—	F07	Nielogiczna kolejność	drobny	0,75	AI: przegląd całości	Kryterium spełnione częściowo: Logiczna kolejność, łagodne otwarcie, bloki tematyczne. Kolejność jest logiczna (tematyczne -> demografia), ale długa sekcja siatkowa może męczyć przed pytaniami demogr...
—	M06	Niejednorodna bateria	drobny	0,75	AI: przegląd całości	Kryterium spełnione częściowo: Jednorodne baterie. Siatka łączy różne aspekty (humor, zrozumiałość, udostępnianie), co może czynić ją heterogeniczną.
—	B04	Wymuszanie odpowiedzi	drobny	0,60	reguła	90% pytań jest obowiązkowych.
—	B07	Przeładowana strona	drobny	0,55	reguła	Strona 3: 10 pytań, 200 pozycji.
—	O08	Format zgody zamiast skali specyficznej	drobny	0,40	reguła	Format „zgadzam się / nie zgadzam się” w 10 macierzy(ach) (Q02, Q03, Q04, Q05, Q06, Q07). Skale specyficzne dla treści (np. „jak często”, „jak bardzo”) są zwykle rzetelniejsze i mniej podatne na pota...

14.2 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.

14.3 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).
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.

(cd.)

Kod	Wymiar	Problem	Komentarz
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	Jedna 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	Niezrównoważona 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 pytania 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, niejednoznaczne, 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 ankiet	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 męczy, 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.
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.

(cd.)

Kod	Wymiar	Problem	Komentarz
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 24×24 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.
G01	Zgodność z celami badania	Brak pomiaru zmiennej z hipotezy	Zmienna potrzebna do celu, pytania badawczego lub hipotezy nie jest mierzona żadnym pytaniem — tej części badania nie da się zrealizować.
G02	Zgodność z celami badania	Niepełny pomiar zmiennej	Zmienna jest mierzona tylko częściowo: jednym pytaniem zamiast skali, bez ważnego aspektu, w innym okresie lub u innej grupy niż zakłada hipoteza.
G03	Zgodność z celami badania	Poziom pomiaru nie pozwala na analizę	Format odpowiedzi (np. przedziały, dwie kategorie, pole tekstowe) nie pozwala na zaplanowaną analizę (korelację, regresję, porównanie grup).
G04	Zgodność z celami badania	Pytanie bez związku z celami	Pytanie nie służy żadnemu celowi, problemowi ani hipotezie (i nie jest metryczką ani pytaniem kontrolnym) — wydłuża ankietę bez pożytku dla analiz.
G05	Zgodność z celami badania	Brak zmiennych kontrolnych	Brakuje zmiennych zakłócających lub kontrolnych potrzebnych do rzetelnego testu hipotezy (albo sprawdzenia manipulacji w eksperymencie).
G06	Zgodność z celami badania	Niezgodność z badaną populacją	Ankieta nie pozwala sprawdzić, czy badany należy do populacji opisanej w celach (brak pytań selekcyjnych), albo pyta o inną grupę.
G07	Zgodność z celami badania	Plan badania nie pozwala sprawdzić hipotezy	Np. hipoteza przyczynowa przy jednorazowym badaniu korelacyjnym, eksperyment bez losowego przydziału do warunków albo bez grupy porównawczej.
G08	Zgodność z celami badania	Cel lub hipoteza nietestowalna	Sformułowanie celu lub hipotezy jest zbyt ogólne (brak kierunku, nieokreślone zmienne), by sprawdzić je danymi z ankiety.
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 płci bez „wolę nie odpowiadać”, stereotypy, forma tylko męska.	Np. tylko dwie kategorie płci bez „wolę nie odpowiadać”, stereotypy, forma tylko męska.
E07	Etyka, RODO, inkluzywność	Brak kontaktu do badaczy	Badany nie ma kogo zapytać o badanie ani do kogo zgłosić wątpliwości.

(cd.)

Kod	Wymiar	Problem	Komentarz
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.

14.4 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	20
Nieudane wywołania / failed calls	0
Odrzuceni kandydaci / rejected candidates	0
Czas audytu (min) / duration	5.6
Źródła / sources	przebieg_desktop, przebieg_mobile