BCN PEG Survey QC BDPs

List of QC Flags | Open text questions

⬤ Yellow Flag ⬤ Red Flag

Flag	Criticality	Description	Time investment
Unrealistic numbers	⬤ (red)	Unrealistic values in open-text number write-ins; e.g. 20 strollers purchased for a child	◕ (quick win)
Open-text straightlining	⬤ (yellow)	Same text answer provided by the same respondent across different open-text questions	◕ (quick win)
Fake brand in unaided	⬤ (red)	Respondent mentioning fake brand (from aided list) in the preceding unaided awareness question (bot/ cheater)	◑ (mid level)
Language check	⬤ (yellow)	Respondent open-text response in different language vs. survey language (bot/ cheater)	◑ (mid level)
Same open-text across diff. resp.	⬤ (red)	Same text answer provided by different respondents. (bot/ cheater) _Threshold_: Text contains at last 20 letters _Caveat_: Text may be repeated in different questions (different columns in raw dataset); bots may use spelling mistakes	◕ (high/manual)
Open text – Manual scan (not exhaustive)	⬤ (red)	1. Copy-pasted answer from the company website or other survey questions (e.g. KPC) 2. Repeat respondents (same unaided brands with different order; similar writing style in reasoning-related questions) 3. Low quality/ gibberish / not relevant answer	◕ (high/manual)

◕ Quick win | ◑ Mid level / requires refreshing | ◕ High / Manual

Note: Number / type of QC flags and robustness is subject to survey purpose/ target audience. Reach out to BCNPEGStaffing@Bain.com to discuss specific cases

List of QC Flags | Fixed response questions

Yellow Flag ⬤	Red Flag ⬤

Flag	Criticality	Description	Time investment
Speeder	🔴	Respondent completing survey quicker than 40-50% of median completion time	◔
Fake vendor	🔴	Respondent citing "considered/ used" any of the fake brands	◔
Brand awareness	🟡	Respondent aware of all/ 90% aided brands. _Threshold_: Typically used for surveys with 15+ brands/ competitors	◔
Usage inconsistency	🟡	Respondent cites "currently using" for a particular brand, but unaware of any specific products of the same brand	◔
Straightliners	🟡	Respondent providing same answers in matrix type questions (e.g. same rating for a brand across all KPCs; same level of likelihood across several following questions; same NPS for all brands)	◑
Switching mismatch	🟡	Respondent selecting "switched in the past" but has no brand selected under "Used in the past, but not anymore" in the awareness funnel	◔
High share of "IDK"	🟡	Respondent with high share of "I don't know" responses across multiple question. _Threshold_: 4-5 questions	◔
Age check	🟡	1) Respondent age and birth year not matching. _Caveat:_ Keep the two questions apart - presumably at the beginning and end 2) Respondent indicating unusually low/high age OR mismatch in age and experience	◔
HH income check	🟡	1) Mismatch of spending level and income (ratio) 2) Respondent indicating unusually low/high income	◔
Other (case specific)	🟡	Respondent mentioning contradictory answers across diff. questions	◔

◔ Quick win	◑ Mid level / requires refreshing	◕ High / Manual

> **Note**: Number / type of QC flags and robustness is subject to survey purpose/ target audience. Reach out to BCNPEGStaffing@Bain.com to discuss specific cases

This information is confidential and was prepared by Bain & Company solely for the use of our client; it is not to be relied on by any 3rd party without Bain's prior written consent

BAIN & COMPANY 3

Key ways to identify bots / cheat respondents | Open-text QC

/ NOT EXHAUSTIVE

	Description	**Example**
Repeated answers across respondents	• Different respondents with the exact same verbatim – Answer can be repeated in a different open-text question/ brand – Set threshold to ~20 characters or above	Respondent ID Why did you buy this product? 1 There was a 10% discount on the xyz website 2 Because the price was right 3 There was a 10% discount on the xyz website
Public information copied	• Respondent copying text from the company / other website	Respondent ID Why this NPS score? 1 Babyzen YOYO2 Folds & Unfolds in a Moment While You Hold Your Baby & Can be Worn on the Shoulder. For Parents Who Need to Fold Their Stroller at Home, Hop on a Bus, Take a Train or Plane 2 Very durable stroller. Easy to clean, handles really well. Comfortable for a growing child
Different language used	• Respondent answering in a different language vs. survey language – Past example: ~20% responses in a German survey answered free text responses in Dutch	Respondent ID Language Text input Question 1 English The product is very reliable 2 English Ich liebe dieses Produkt 3 English I like the design
Imitated responses	• Fake brand from aided awareness question mentioned by respondent in unaided awareness (free-text) question – Indicates that respondent had access to the survey Q	Respondent ID Brand 1 Brand 2 Brand 3 1 Mylan Cybex Egg r) Mylan [TAG: US;GB;DE;fake] 2 Easywalker Mamas Cybex s) PureTech [TAG: US;GB;DE;fake] 3 Cybex Stokke iCandy
Similar writing style	• Different respondents have the same set of brands in unaided awareness with minor alterations in order, coupled with similar responses for drivers of NPS	Respondent ID Brand 1 Brand 2 Brand 3 Why did you buy this product? 1 Brand A Brand B Brand C IT HAS A GOOD QUALITY 2 Brand B Brand C Brand A IT HAS A AWSOME 3 Brand A Brand C Brand B IT IS A REALLY AWESOME 4 Brand A Brand B Brand C IT IS A COOL 5 Brand B Brand A Brand C IT HAS A GOOD 6 Brand A Brand C Brand B IT HAS A COOL

Note: Cheaters / bots typically make slight alterations in the text due to which QC based formulas might not be exhaustive.

Key ways to identify low quality respondents | Open-text QC

/ NOT EXHAUSTIVE

	Description	**Example**
Non-relevant brands	• Irrelevant brands mentioned e.g., answer related to software in the survey related to consumer goods	Respondent ID, Brand 1, Brand 2, Brand 3 1 Lego, Amazon, Troop 2 Baby jogger, Mamas and papas moon, Silver cross 3 kidskraft, hartan [Irrelevant brand]
Same answer across questions	• Same answers across multiple open text question (within a single respondent)	Respondent ID, Why do you like product A?, Why do you like product B? 1 Because it's the best, Because it's the best 2 OK, OK 3 I like the size, I like the size
Non-meaningful answer	• Gibberish answers/ Non-relevant answer • Monosyllabic answers which don't (clearly) answer the question – Example: Good in drivers of NPS	Respondent ID Open text 1 asdzxckashdgy 2 gasdjagsdjyg 3 The product is very reliable Respondent ID Why did you buy this product? 1 The design convinced me 2 The price was right 3 The weather in Paris is nice Respondent ID What software do you use? 1 Windows 2 Office chair 3 Microsoft Office Respondent ID Wy this NPS score? 1 because it has a great features 2 good 3 My Baby fell comfort with here, and this brand is best for my

Note: As a standard approach we don't blacklist low quality open text responses if they're not being flagged under any other QC criteria