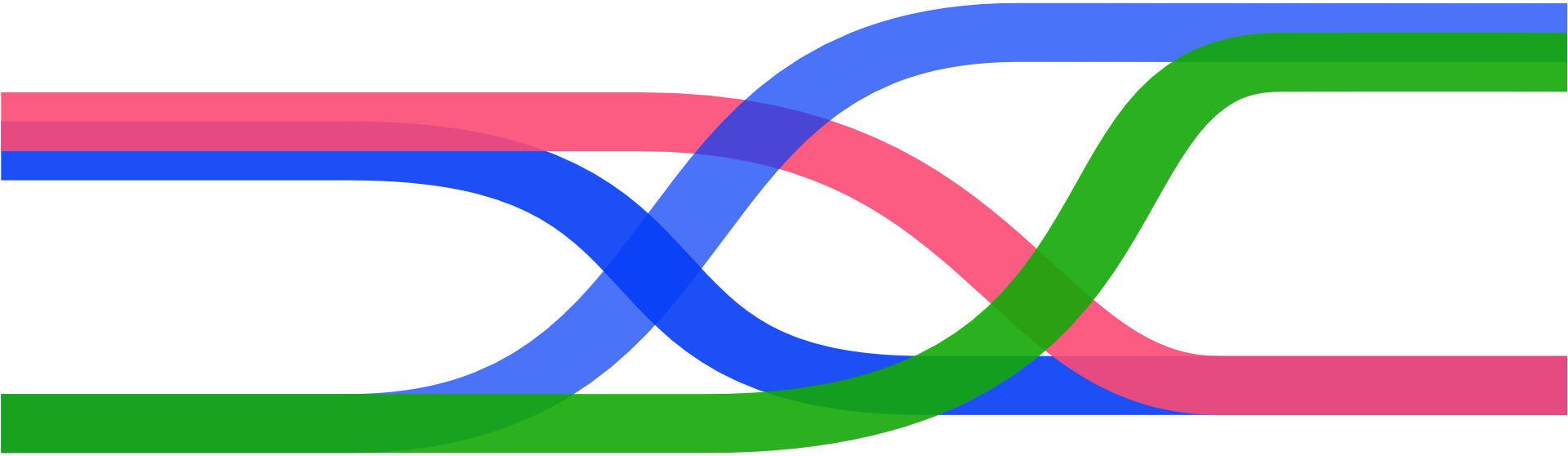


# Survey Scripting Practices in Market Research

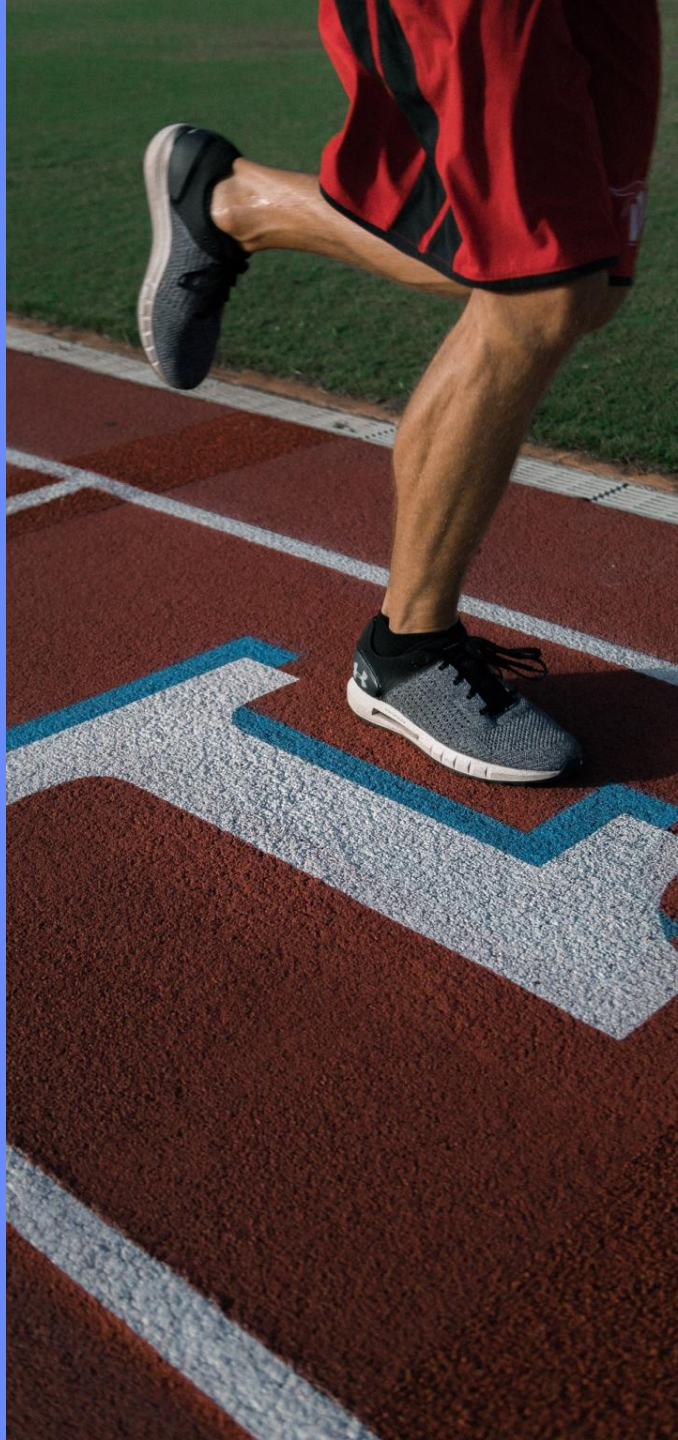
22 July 2025: A guide to optimise survey question and flow design



conjointly

conjointly

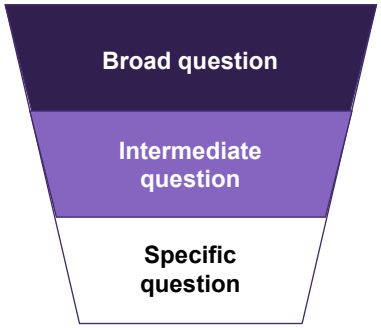
# Overview



# Establishing logical order, grouping related topics, using the funnel approach

OVERVIEW | GENERAL PRACTICES

- Start with engaging, simple questions to build rapport and ease into the study
  - Our screening questions can also act as warm-in questions for better data quality further into the survey
- Place demographic questions strategically
  - Beginning → demographics used for screening
  - End → anything not used for screening to avoid contributing to survey fatigue early on
- Avoid sensitive questions at the beginning to prevent early dropout
  - Place income questions at the end if not used for qualification purposes
- Group related topics, use the funnel approach within topics



Funnel approach

- Test questions on yourself first for **readability**, **ease of comprehension**, and **ability to recall the information**
  - Or better yet, share your survey link with colleagues along with a heads-up on what to look for
  - If you struggle to understand a sentence, don't know exactly what a word means, or the terms used seem too technical; **respondents will struggle too**
    - > ❌ "To what extent do you find the product's tactile interface conducive to your usage patterns?"
    - > ✅ "How do you feel about the product's touch controls?"
- Use the soft launch as **a pilot test** for difficult questions

[26] Krosnick & Presser (2010). [52] Taherdoost (2022).

# Provide memory aids for ease of recall

OVERVIEW | COMPREHENSION & RECALL

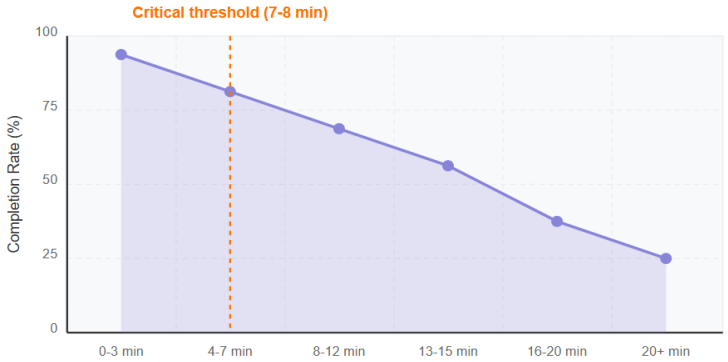
- Anchor timelines! Time is better perceived based on **temporal landmarks** (cultural, seasonal, etc.)
  - Bad: “How often did you use the product in Q1?”
  - Meh: “How often did you use the product between January and March 2025?”
  - Good: “How often did you use the product this past winter (Jan-Mar 2025)?”
- Prepare respondents for recalling information
  - By dividing into sections: “The next section will ask about your recent purchases”
  - By providing context: “Think(ing) about your purchases in the past six months…”
  - By providing memory aids: “Below is a list of cleaning products commonly used in restaurants. Do you purchase any of these for your establishment?”
- Use specific time frames instead of vague time quantifiers. The definition of *rarely* will depend on the individual perception of time, lifestyle, etc.
  - “How often do you exercise?”
    - > Bad:
      - Rarely
      - Sometimes
      - Often
      - Very often
    - > Good:
      - Never
      - 1-2 times per month
      - 1-2 times per week
      - 3 or more times per week
      - Every day

[29] Larsen & Thompson (1995). [47] Schwarz & Oyserman (2001). [48] Shum (1998). [52] Taherdoost (2022).

# Keep surveys short to avoid fatigue and increase data quality

OVERVIEW | SURVEY LENGTH

- The ideal maximum survey length has a median of **10 minutes**
- The maximum recommended length is **20 minutes**
  - Some precautions are necessary to maintain data quality for 20-minute surveys
  - Abandonment rates increase substantially after 7 mins




## Example survey planning by length

**Short survey** (5-7 min)
- ✓ Single topic focus
- ✓ 15-20 questions max
- ✓ Can include complex question types (e.g., recall, several grids)

**Standard survey** (8-12 min)
- ✓ 1-2 related topics
- ✓ 20-30 questions
- ✓ Mix of simple and complex questions

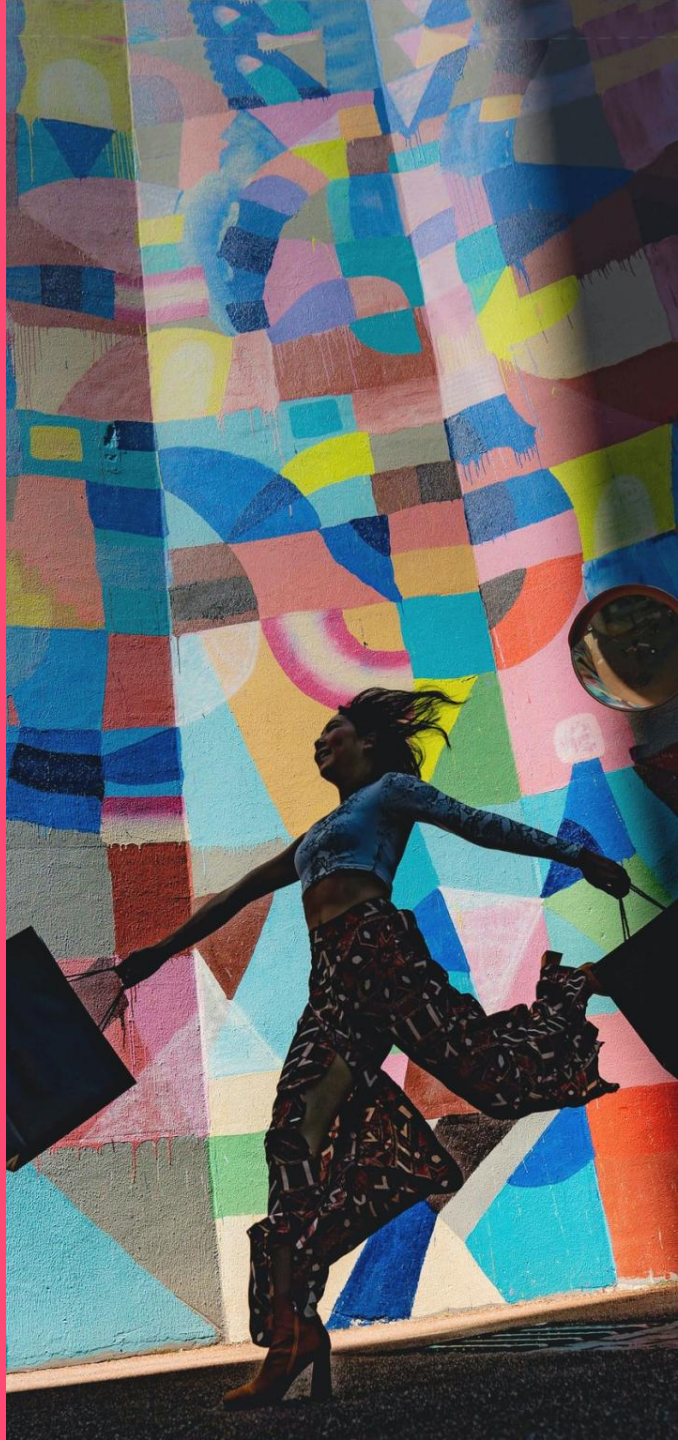
**Long survey** (13-20 min)
- ✓ Consider splitting into multiple surveys (sections if not possible)
- ✓ Requires stronger incentives
- ✓ Needs engaging design

- Place important questions before the 7-min mark
  - First 1-3 min: Screening and warm-up
  - 3-7: Critical research questions
  - 7+: Secondary questions, demographics, nice-to-have insights

[34] Liu & Wronski (2018). [45] Revilla & Ochoa (2017).

conjointly

# Question types



# Good practices for asking questions

QUESTION TYPES | GOOD PRACTICES

- Remember that we are typically limited to holding and processing ~7 (+/- 2) items in working memory at once
  - For **Likert scales**, **Matrix grids**, **Multiple choice questions**:
    - > Optimal number of alternatives is 7
    - > Efficiency decreases significantly at 11 alternatives
  - We might think that adding more categories or broadening the scale will contribute to the granularity of the data — but this might end up overloading the working memory, leading to low data quality
- Avoid negation: processing can be difficult and highly dependent on crosslinguistic factors
- Clearly communicate the scale, the grounds for evaluation, and the product/concept/SKU of interest
  - Do not rely on recall or indexicals (“us”, “our product”)
- Make sure that the respondent is informed of the stimulus in some capacity
  - They should have either (i) used the product, (ii) are familiar with the product, (iii) you have given enough information in the experiment that enables the respondent to judge the product
  - If experiencing the product is not necessary to answer the questions (e.g., concept tests, rating the design), clearly state that they should answer based on the information provided in the survey
- For pricing studies, include a price knowledge question for current users
  - Helps understand price sensitivity and useful for data quality

[26] Krosnick & Presser (2010). [36] Miller (1956).

# Net Promoter Score (NPS)

QUESTION TYPES | NET PROMOTER SCORE

## Standard practices

- ✅ On the scale of 0-10
- ✅ Clearly stated question
  - "On a scale of 0 to 10, how likely are you to recommend [product / service / concept]?"
  - State the product clearly in the question text, avoid relying on recall or using indexical such as "us" or "our product"
- ✅ Clearly stated scale with labelled extremes
  - "Extremely likely" to "not at all likely"
- ✅ Three groups of respondents (this grouping might depend on the researcher or the study):
  - Detractors: 0-6
  - Passives: 7-8
  - Promoters: 9-10

## Recommended practices

- Include an **open-ended** question to probe into the respondent's NPS rating
  - "Why did you rate this [product / service / concept] a 6 out of 10?"
- Keep the survey short to prevent survey fatigue, don't include more than three questions for each brand.

### Fallouts

- NPS is argued to be a poor measure of loyalty
  - Low predictability: fails to correlate with other metrics
  - It is argued that repurchase is a better indicator
- The lack of a no-answer option
- The standard grouping does not form statistically homogeneous groups and can also vastly differ across cultural settings

[3] American Customer Satisfaction Index (n.d.). [5] Baehre, O'Dwyer, O'Malley, & Lee (2022). [25] Kristensen & Eskildsen (2014).

# Likert scale questions

QUESTION TYPES | LIKERT SCALE

## Standard practices

- ✅ 5-7 points
  - The optimal length was found to be 7, which maximises validity and reliability
- ✅ Scale points should be clearly labelled with appropriate anchors
  - Ideally all points should be anchored
- ✅ The scale length can depend on what the research goals are
  - Longer scales will provide better discrimination power
  - Shorter scales will decrease cognitive fatigue
  - Even-numbered scales will force respondents to lean a certain way by removing the centre option
  - Odd-numbered scales allow for neutrality but increase the bias for central tendency

## Recommended practices

- 5-, 7-, and 10-point scales are easiest for respondents
- For maximum discriminating power, use 6+ response categories
- For non-Western respondents, supplement scales with:
  - Anchoring (providing standardised examples as reference points; e.g., “1 means I regret purchasing this product and would not buy it again”)
  - Quantifying responses directly (e.g., “3 means I feel this way a few days a year")
  - Collecting additional qualitative responses
- For Chinese respondents, consider 4-point scales as this would increase construct validity

**Fallouts**

- Respondents take significantly longer to select the middle category, suggesting higher cognitive load
- Less clear items lead to increased middle category endorsement, potentially skewing results
- Many non-Western respondents find the ordinal scaling structure unfamiliar
- Scale length decisions affect multiple aspects of research quality simultaneously (validity, reliability, respondent burden, etc.), difficult to optimise for all factors

[21] Flaskerud (2012). [28] Kulas & Stachowski (2009). [31] Lee, Jones, Mineyama, & Zhang (2002). [32] Lewis (1993). [36] Miller (1956). [41] Preston & Colman (2000). [51] Symonds (1924).

# Matrix grids

QUESTION TYPES | MATRIX GRIDS

## Standard practices

- ✅ 5-7 items per row and column each
- ✅ Clear row and column labels
- ✅ Randomised item order (if not ordered list)
- ✅ Use to ask about similar items on the same scale
- ✅ Useful for comparisons, offering a visual ground for contrasting different concepts

## Recommended practices

- Scale direction should remain consistent
- Use straightforward, predictable wording; avoid using negative items with positive statements in the same grid
  - Example of **reverse-wording**:
    1. The customer service team was helpful.
    2. The product met my expectations.
    3. **The website was difficult to navigate.**
    4. I would recommend this company to others.
    5. **The pricing was unclear.**
- Apply when survey length is a significant concern
- Ensure mobile compatibility, test on multiple devices
- Avoid in critical questions requiring maximum attention

### Fallouts

- Increased behaviours such as straightlining or patterned responses, with respondents being more likely to choose the same option for multiple stimuli
- Grids perform poorly on smartphones
- Might increase cognitive burden despite their visual efficiency, as respondents must continuously reference both row and column information
- Can be confusing for respondents with lower digital literacy

[8] Beatty, Collins, Kaye, Padilla, Willis, & Wilmot (2019). [15] Couper, Tourangeau, Conrad, & Zhang (2013). [16] Debell, Wilson, Jackman, & Figueroa (2019). [24] Grady, Greenspan, & Liu (2019). [33] Liu & Cernat (2018). [39] Neuert, Joss Roßmann, & Silber (2023). [44] Revilla & Couper (2018). [49] Smyth & Olson (2019). [53] Toepoel, Das, & Van Soest (2009). [54] Tourangeau, Couper, & Conrad (2004). [57] Vehovar, Couper, & Cehovin (2023).

# Multiple-choice questions

QUESTION TYPES | MULTIPLE-CHOICE

## Standard practices

- ✅ Each question should have a specific objective, clearly stated in the question text
- ✅ All necessary information needed to answer the question should be included, anything not directly needed should be avoided
- ✅ Relevant, to-the-point, and precise, mutually-exclusive alternatives. Avoid ambiguity
- ✅ Alternatives of the same type (e.g., brands, products, reasons)
- ✅ Randomise alternatives, unless ordered

## Recommended practices

- Use distractor alternatives
  - Fake brands, implausible reasons, non-existent products etc. will be helpful to weed out careless responses.
- An optimal MCQ should ideally contain 5-7 alternatives, efficiency decreases after 11 alternatives
- Use accessible grammar and language. Avoid using “big words”
- Almost always include “No answer” and “other” options
  - People will choose randomly if no option can accommodate their lack of position, usage, etc. which will lead to low quality data
- Avoid using negative items in the question text
  - “Which of these brands would you be least likely to buy?”

**Fallouts**

- Satisficing behaviour is commonly used to conserve cognitive effort
- The depth of insight is limited for this question type
- Forced choice that mismatches real-life scenarios
- Random selection can limit the possibility of a researcher determining whether the respondent gives meaningful answers

[11] Burton, Sudweeks, Merrill, & Wood (1990). [17] Dolnicar (2013). [55] Tourangeau, Rips, & Rasinski (2000).

# When MCQs fail: Satisficing behaviour

QUESTION TYPES | SATISFICING BEHAVIOUR

## What is satisficing behaviour?

Respondents might select things in MCQs not because they considered all options and thought the selected ones apply to them, but because **the options met an acceptable threshold**.

Respondents might:

- skip items entirely if allowed
- choose the first reasonable, “close enough” option rather than considering all options listed
- select middle response options to avoid extreme positions
- agree with statements regardless of content (known as the acquiescence bias, or “yea-saying”, leading to a disproportionate number of positive responses)
- respond randomly

## Recommended practices

- Reduce cognitive burden for respondents
  - keep question texts and response options clear, concise, and straightforward
- Vary question formats to maintain engagement with the experiment
- Include dummy options in multiple-choice questions (e.g., fake brand names, non-existing products, nonsensical reasons)
- Keep surveys short and as easy and low-effort as possible

[6] Barge & Gehlbach (2012).

conjointly

# Sensitive questions

# Sensitive questions entails three different dimensions: Intrusiveness, threat of disclosure, and social desirability

| SENSITIVE QUESTIONS | SENSITIVITY DIMENSION | |
|---|---|---|
| **Intrusiveness** | **Threat of disclosure** | **Social desirability** |
| Questions that touch on topics typically considered inappropriate for everyday conversations and are viewed as **invasions of privacy, regardless of the respondent's actual answer.** | Concerns about the **possible consequences** of giving a truthful answer should the information become known to a third party. | Respondents tend to **over-report "good behaviours"** (e.g., recycling) and **underreport "bad" ones** (e.g., smoking) to align with societal expectations. |
| Sensitivity stems from the **question** | Sensitivity stems from the **question** | Sensitivity stems from the **answer** |
| *Income*<br>*Religion* | *Teenagers' use of marijuana*<br>*Tax fraud* | *Frequency of exercise*<br>*Sustainability behaviours* |
| ***Response error*** | | |

[56] Tourangeau & Yan (2007).

# Gender and sex are two different concepts and thus shouldn’t be conflated with each other

SENSITIVE QUESTIONS | GENDER

## Sex

is usually assigned and registered **at birth** based on primary sex characteristics (genitalia) and **reproductive functions**.

## Gender

often expressed in terms of masculinity and femininity, is largely **culturally determined** and reflects a person's **internal sense** of identity.

## Notes

- Poorly designed gender questions can lead to potential dropouts and degrade data quality
  - Misclassification of gender identities may cause confusion
  - Binary-only options can alienate non-binary/gender non-confirming respondents, leading to distrust or potential participant dropouts
- Allowing multiple gender identity selections can:
  - improve inclusivity
  - but introduce significant challenges for data check, coding and analysis

[7] Bauer, Braimoh, Scheim & Dharma (2017). [12] Cartwright & Nancarrow (2022).

# Gender questions

SENSITIVE QUESTIONS | GENDER

## Standard practices

- ✅ Use clear, respectful language
  - **Good:** “How do you identify your gender?”
  - **Good:** “What is your gender (identity)?”
- ❌ Do not conflate sex and gender or use them both in the same question
  - **Bad:** “What is your gender/sex?”
  - **Bad:** “What is your gender or sex?”
- ✅ Include options beyond binary man/woman
- ✅ Order answer options alphabetically/randomly to avoid bias
- ✅ Placement in survey flow
  - At the beginning: for screening needs
  - At the end: prevent biases that might influence responses to the main questions

## Recommended practices

- When it’s necessary to get the full picture, include a write-in option
  - “Other, please specify”
- Include a "Prefer not to say" option to respect privacy
- Adapt terminology for regional/cultural context
  - "Hijra" in South Asian surveys

**Example question**

What is your gender?

- Woman
- Man
- Other
- Prefer not to say

*(Answer options ordered alphabetically/randomly)*

[12] Cartwright & Nancarrow (2022). [22] Garvey, Hart, Metcalfe & Fellabaum-Toston (2019). [42] Puckett, Brown, Dunn, Mustanski, & Newcomb (2020).

# Income questions are highly sensitive and prone to response error due to privacy concerns and social desirability bias

SENSITIVE QUESTIONS | INCOME

## Standard practices

- ✅ Use **income ranges or brackets** to reduce survey sensitivity and non-response
  - Non-overlapping and exhaustive (cover all potential income levels)
  - Start from zero to avoid presumptions about minimum income
- ❌ Don't ask for exact figures of income (can be intrusive, cause lower response rates or inaccurate answers)
- ✅ Provides **context** for accurate data collection by explicitly specifying:
  - Individual vs. household income
  - Pre-tax vs. post-tax income
  - Annual vs. monthly income reporting
  - Provides critical context for accurate data interpretation
- ✅ Localise currency (e.g., USD, EUR)
- ✅ Placement in survey flow
  - Position income questions later in the survey (preferably toward the end) to minimise discomfort and dropout

## Recommended practices

- Update income ranges regularly to account for inflation, especially for countries with high inflation rates (e.g., Argentina, Turkey)
- Income reporting might differ by country, check the standard for each country in your survey: **monthly** (e.g., Poland) vs. **annually** (e.g., United States)
- Define income ranges based on each country's income distribution
  - median income, quartiles/deciles from government statistics or World Bank data
- In most contexts, **household income** is a better economic indicator compared to **individual income**
- Include a "Prefer not to answer" option to give respondents the autonomy to opt out of the question rather than drop out the whole survey

[35] Micklewright & Schnepf (2010). [38] Moore & Loomis (2001). [59] Yan, Curtin, & Jans (2010).

# Concepts of race and ethnicity evolve as complex social constructs rather than simple biological realities

SENSITIVE QUESTIONS | RACE & ETHNICITY

**Race** used to be associated with physical/biological characteristics (skin colour, facial features, etc.) and presumed to have genetic basis

**Ethnicity** used to be more commonly associated with cultural identity, heritage, language, customs, and social groupings

**Both race and ethnicity** are now understood primarily as **social constructs** rather than biological realities

*The American Anthropological Association (1998) explicitly rejects racial categories as biologically meaningful and emphasises cultural factors.*

As racial and ethnic diversity grows, it is important to prioritise respondents' **autonomy**.

- Poorly designed questions can lead to potential dropouts and inaccurate data
  - Outdated or limited race/ethnicity categorisation schemes can alienate respondents
  - Forcing respondents to select a single category when they identify with multiple groups creates frustration and data inaccuracy
  - Using outdated or offensive terms can decrease completion rates and damage research credibility
  - Attention to cross-country design: Categories appropriate in one cultural context may be meaningless or inappropriate in others

[2] American Anthropological Association (1998). [14] Cornell & Hartmann (2006). [46] Robinson & Leonard (2024). [58] Winker (2004).

# Race / ethnicity questions

SENSITIVE QUESTIONS | RACE & ETHNICITY

## Standard practices

- ✅ Carefully consider **whether collecting this data is essential to your research goals**
- ✅ Stay relevant to and inclusive of the racial identities of the market of interest.
- ✅ Questions should allow for **flexible self-identification** through:
  - multiple selection options
  - “other” option with a text field for self-identification
  - open-ended responses
- ✅ Categories must reflect the cultural and demographic context of your target respondents.
- ✅ Placement in survey flow
  - Placed in the middle or latter part of the demographic section

## Recommended practices

- Use locally relevant categories
  - Racial/Ethnic categories vary significantly by country
- Open-ended questions can capture more nuanced identities but creates categorisation challenges
- Use current, culturally appropriate terminology
  - Hispanic or Latino
  - Native American or First Nations
- Consider local laws about ethnic data collection
  - French laws prohibit the collection of data revealing racial or ethnic origin

**Example question** (U.S. Census Bureau, 2025)

What is your race and/or ethnicity?

*Select all that apply.*

- American Indian or Alaska Native
- Asian
- Black or African American
- Hispanic or Latino
- Middle Eastern or North African
- Native Hawaiian or Pacific Islander
- White
- Other: _____ (text input)

*(Answer options ordered alphabetically or by population size)*

[10] Bureau, U. C. (2024). [43] République française (2019).

# Respondents tend to show social desirability biases in religion questions by overreporting their attendance or beliefs

SENSITIVE QUESTIONS | RELIGION

## Standard practices

- ✅ Be clear about why this information is relevant and being collected, as well as how it will be used and that it will stay anonymous
- ❌ Avoid asking “what is your religion” as this would presume that the respondent has one
- ✅ Use neutral language.
  - ❌ “Do you identify with Christianity?”
  - ✅ “What religion, if any, do you identify with?”
- ✅ Make sure to not conflate the cultural identity and religious practices
- ✅ For answer options:
  - Use broad categories but also make sure to include options for smaller groups.
  - Always include a “Prefer not to say” option.
  - ❌ Avoid excessively detailed religious categories
- ❌ Do not use the phrase “religious preference” as the origin of religious practice can be subject to debate
- ✅ Placement in survey flow
  - Placed toward the end of the demographic section

## Recommended practices

- If not necessary, avoid asking religion in detail.
  - If all you need to know is whether they identify with a/any religion rather than what religion they identify with, keep the options broad
- Sample alternative set:
  - Major religions (Christianity, Islam, Judaism, Hinduism, Buddhism, etc.)
  - Other religion (write-in)
  - Spiritual but not religious
  - Agnostic
  - Atheist
  - No religion
  - Prefer not to answer
- For broad categories, consider adding details about what denominations would count into the broad category
- Consider local laws about ethnic data collection
  - French laws prohibit the collection of data revealing racial or ethnic origin

[40] Pew Research Center (2018). [43] République française (2019).

# Medical questions

SENSITIVE QUESTIONS | MEDICAL QUESTIONS

## Standard practices

- ✅ Use simple, clear language comprehensible to the respondents, avoid using overly complex, medical vernacular
- ✅ Consider using visual aids for body parts or pain scales if applicable
- ✅ Clearly explain data use and storage, comply with local government regulations, obtain explicit consent
- ✅ Placement in survey flow
  - In the middle of the survey rather than beginning or end

## Recommended practices

- **Ethics** and informed **consent**:
  - Make sure that the respondent is fully informed of the purpose and confidentiality of the study
  - Give explicit details about the use of the data provided by the respondent before they agree to participate
- **Reduce harm**, **increase benefit**:
  - Make sure that the benefit of the participation in the survey outweighs the harm. Asking medical questions (life threatening or difficult topics) can bring forth memories or emotions that are hard to handle
- **Rights** of confidentiality:
  - Consider the rights of the respondent in their country of residence. Countries may have specific guidelines on how to handle medical information beyond data privacy
- **Different cultures** may have different boundaries on what medical information you can gather
  - Be aware of community stigmas around specific medical conditions
  - Consider alternative approach (indirect questioning)
  - Avoid direct cross-cultural health comparisons without accounting for cultural patterns

[1] Aday & Cornelius (2011). [13] Clark (1983).

# Social desirability: standards and recommended practices

SENSITIVE QUESTIONS | SOCIAL DESIRABILITY

Respondents often **over-report "good behaviours"** (e.g., recycling) and **underreport "bad" ones** (e.g., smoking) to align with societal expectations. This is referred to as the **social desirability bias (SDB).**

## Standard practices

❌ Avoid direct self-assessment
- ❌ "Do you act sustainably?"
- ✅ "Which of these actions have you taken in the past month?"
  - I recycled electronics
  - I used reusable bags

✅ Use neutral, non-judgmental wording to neutralise/normalise all potential responses
- For example, use phrasing like: "Many people..."
- Use phrasing like: "Some people smoke occasionally, others don't. How often do you smoke?" to reduce perceived judgment

✅ Avoid loaded terms that suggest "right" answers

## Recommended practices

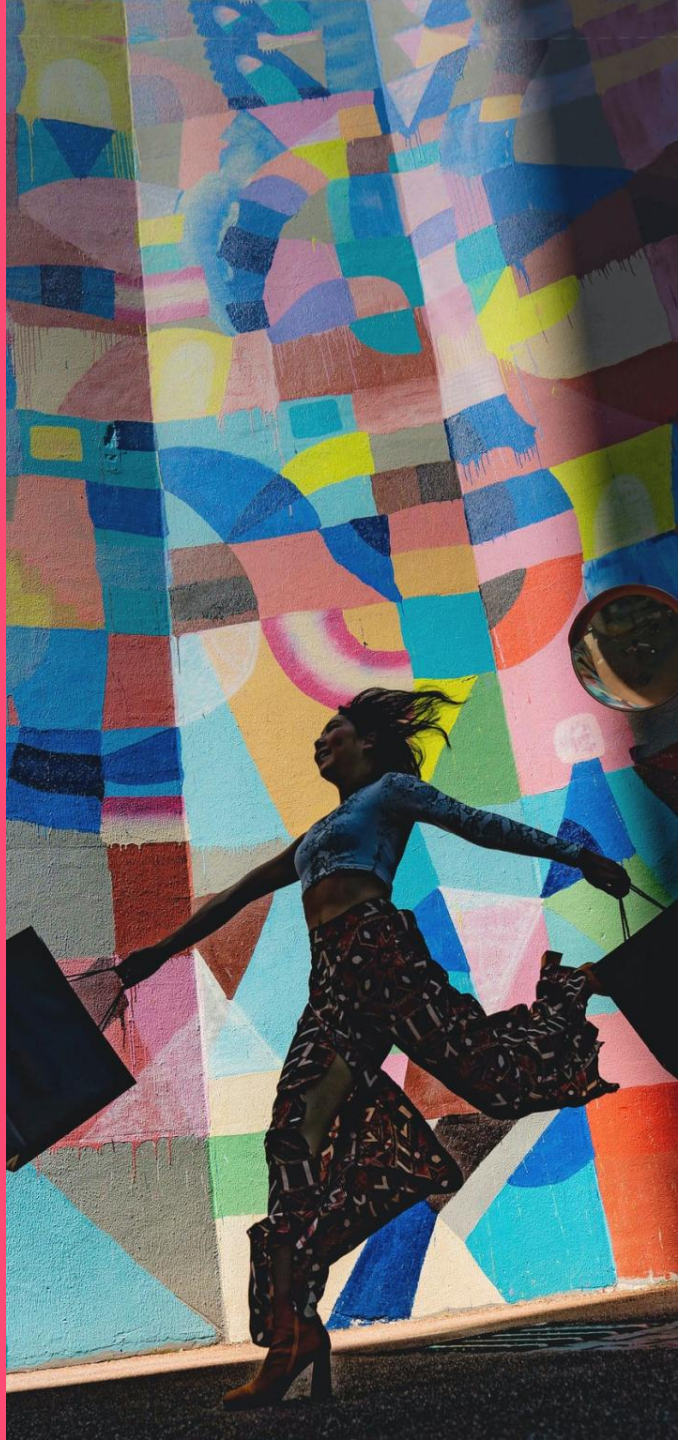
- Carefully embedding the sensitive question in a **constructed context** to lower the respondent's feelings of jeopardy
  - Embedding the question in a series of questions starting with unoffending, general questions connected to the topic of interest, and then gradually narrowing the focus to more specific behaviours
- Consider **indirect questioning techniques** for sensitive topics.
  - Asks respondents to make predictions about how a similar other would think or act in a particular situation. But in some situations, indirect questions may introduce other forms of error
  - "How common do you think it is for people in your community to buy bio products?"

SDB can significantly **skew insights for market research,** for example:
- **Sustainability**: Respondents may claim to buy eco-friendly products more often than they do, leading businesses to overestimate demand for green initiatives and potentially misallocate resources
- **Smoking/Drinking**: Underreporting of frequency can mislead companies about the true market size for cessation programs, low-alcohol beverages, or health-focused products, resulting in flawed market entry strategies

[9] Brace (2018). [18] Durmaz, Dursun & Kabadayi (2020). [19] Fırat (2019). [20] Fisher & Tellis (1998). [27] Krumpal (2013). [30] Larson (2019). [37] Miller, Sudman, Bradburn & Belson (1985). [50] Snijders & Matzat (2007).

conjointly

# User experience



# User experience

USER EXPERIENCE | ACCESSIBILITY & MOBILE COMPATIBILITY

## Accessibility

- Establish clear information hierarchy
  - e.g. **Title** vs. Content
- Sans-serif fonts are easier to read and more accessible
- Avoid all caps as it reduces reading speed / comprehension
- Don't rely on **colour only** to differentiate text
  - It can help comprehension, but it should not be the only mode of differentiation
  - If you are using colour, test your design by using a colour blindness simulator
- Support resizing, try to have image zoom enabled for all images included in the survey
- Be descriptive on error messages. Clearly state what needs to be fixed, and where
- Make sure all text is accessible via screen readers
- Try to use what has been tested and tried
  - If you are using a new question type, make sure to clearly communicate how it works and offer examples.
  - Consider cultural differences when choosing question types and scales for the markets of interest
- Include progress bars

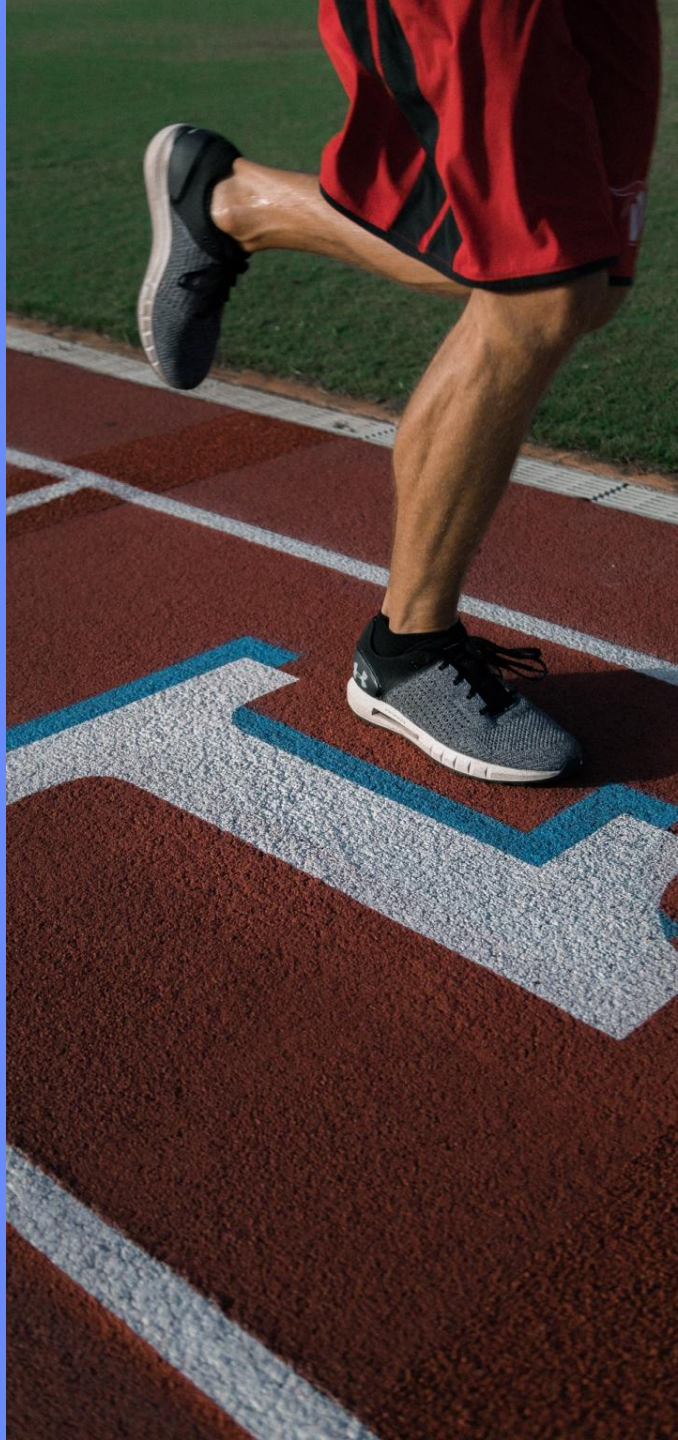
## Mobile compatibility

- Conserve the cognitive burden for the respondents.
  - keep question texts and response options clear, concise, and straightforward
- Vary question formats to maintain engagement with the experiment
- Include dummy options in multiple-choice questions (e.g., fake brand names, non-existing products, nonsensical reasons)
- Keep surveys short and as easy and low-effort as possible

[4] Apple (n.d.). [23] Google (n.d.).

conjointly

# References



# Further reading

REFERENCES | FURTHER READING

1. Brace, I., & Bolton, K. (2022). *Questionnaire Design: How to plan, structure and write survey material for effective market research* (Fifth edition). KoganPage.

   - It provides practical guidelines for planning, structuring, and composing questionnaires in the field of market research.
   - It covers multiple aspects of survey design, from crafting question text and answer options to exploring different questionnaire types, addressing ethical issues, and designing questionnaires for multi-country surveys.
   - It contains case studies that help readers apply theoretical knowledge to real-world scenarios.
   - For an easy start, it systematically guides you through the entire survey design process (from planning to piloting) while also introducing basic statistical concepts necessary for effective research.

2. Bradburn, N. M., Sudman, S., & Wansink, B. (2004). *Asking Questions: The definitive guide to questionnaire design--for market research, political polls, and social and health questionnaires*. John Wiley & Sons.

3. Breakwell, G. M., Barnett, J., & Wright, D. B. (2020). *Research Methods in Psychology.*

   - Focuses on research methods (both qualitative and quantitative) rather than solely on survey design.
   - Covers a wide range of methodologies, from designing experiments to conducting various types of interviews.
   - Includes multiple aspects of the research process, such as study design, data collection techniques, and data analysis methods (e.g., content analysis and statistical approaches).

4. Robinson, S. B., & Leonard, K. F. (2024). *Designing Quality Survey Questions*. Sage.

   - Topics are up-to-date and quite relevant for market research.
   - Covers the entire survey design process and addresses multiple aspects (such as avoiding survey fatigue, managing social desirability bias, etc.) with a clear structure and easy-to-understand language.
   - Includes up-to-date topics and addresses modern challenges, such as appropriate language preferences for sensitive questions (gender, income) and survey design considerations for respondents with low vision or blindness.
   - Uses real-world examples/stories of poorly constructed questions and inadequate response options/scales from actual surveys, demonstrating common pitfalls to avoid.
   - Includes a Checklist for Quality Question Design as a practical tool.

# References


[1] Aday, L. A., & Cornelius, L. J. (2011). Designing and conducting health surveys: a comprehensive guide. John Wiley & Sons.
[2] American Anthropological Association. (1998, May 17). AAA Statement on Race. The American Anthropological Association. https://americananthro.org/about/policies/statement-on-race/
[3] American Customer Satisfaction Index. (n.d.). Retrieved March 6, 2025, from https://theacsi.org/
[4] Apple. (n.d.). Human Interface Guidelines. https://developer.apple.com/design/human-interface-guidelines
[5] Baehre, S., O'Dwyer, M., O'Malley, L., & Lee, N. (2022). The use of Net Promoter Score (NPS) to predict sales growth: insights from an empirical investigation. Journal of the Academy of Marketing Science, 50(1), 67-84.
[6] Barge, S., & Gehlbach, H. (2012). Using the theory of satisficing to evaluate the quality of survey data. Research in Higher Education, 53(2), 182-200.
[7] Bauer, G. R., Braimoh, J., Scheim, A. I., & Dharma, C. (2017). Transgender-inclusive measures of sex/gender for population surveys: Mixed-methods evaluation and recommendations. PLOS ONE, 12(5).
[8] Beatty, P. C., Collins, D., Kaye, L., Padilla J.-L., Willis, G. B., & Wilmot, A. (2019). Advances in questionnaire design, development, evaluation and testing. John Wiley & Sons, Inc.
[9] Brace, I. (2018). Questionnaire design: How to plan, structure and write survey material for effective market research. Kogan Page.
[10] Bureau, U. C. (2024, April 8). What Updates to OMB's Race/Ethnicity Standards Mean for the Census Bureau. Census.gov.
[11] Burton, S. J., Sudweeks, R. R., Merrill, P. F., & Wood, B. (1990). How to prepare better multiple-choice test items: Guidelines for university faculty (Doctoral dissertation, Brigham Young University. Department of Instructional Science).
[12] Cartwright, T., & Nancarrow, C. (2022). A Question of Gender: Gender classification in international research. International Journal of Market Research, 64(5), 575–593.
[13] Clark, M. M. (1983). Cultural context of medical practice. Western Journal of Medicine, 139(6), 806.
[14] Cornell, S., & Hartmann, D. (2006). Ethnicity and race: Making Identities in a Changing World (2nd ed.). Pine Forge Press.
[15] Couper, M. P., Tourangeau, R., Conrad, F. G., & Zhang, C. (2013). The design of grids in web surveys. Social Science Computer Review, 31(3), 322–345.
[16] Debell, M., Wilson, C., Jackman, S., & Figueroa, L. (2021). Optimal response formats for online surveys: Branch, grid, or single item? Journal of Survey Statistics and Methodology, 9(1), 1–24.
[17] Dolnicar, S. (2013). Asking good survey questions. Journal of Travel Research, 52(5), 551-574.
[18] Durmaz, A., Dursun, İ., & Kabadayi, E. T. (2020). Mitigating the effects of social desirability bias in self-report surveys: Classical and new techniques. In Applied social science approaches to mixed methods research (pp. 146-185).
[19] Fırat, D. Y., Osman; Smilkova, Detelina. (2019). Business & Management Practices. IJOPEC.
[20] Fisher, R. J., & Tellis, G. J. (1998). Removing social desirability bias with indirect questioning: Is the cure worse than the disease? Advances in consumer research, 25(1).
[21] Flaskerud, J. H. (2012). Cultural bias and Likert-type scales revisited. Issues in mental health nursing, 33(2), 130-132.
[22] Garvey, J. C., Hart, J., Metcalfe, A. S., & Fellabaum-Toston, J. (2019). Methodological Troubles with Gender and Sex in Higher Education Survey Research. The Review of Higher Education, 43(1), 1–24.
[23] Google. (n.d.). Interactive Canvas Design Guidelines. https://developers.google.com/assistant/interactivecanvas/design
[24] Grady, R. H., Greenspan, R. L., & Liu, M. (2019). What Is the best size for matrix-style questions in online surveys? Social Science Computer Review, 37(3), 435–445.
[25] Kristensen, K., & Eskildsen, J. (2014). Is the NPS a trustworthy performance measure? The TQM Journal, 26(2), 202-214.
[26] Krosnick, J. A., & Presser, A. (2010). Question and questionnaire design. In P. V. Marsden & J. D. Wright (Eds.), Handbook of Survey Research (pp. 264–313).
[27] Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. Quality & Quantity, 47(4), 2025–2047.
[28] Kulas, J. T., & Stachowski, A. A. (2009). Middle category endorsement in odd-numbered Likert response scales: Associated item characteristics, cognitive demands, and preferred meanings. Journal of Research in Personality, 43(3), 489-493.
[29] Larsen, S. F., & Thompson, C. P. (1995). Reconstructive memory in the dating of personal and public news events. Memory & Cognition, 23(6), 780-790.
[30] Larson, R. B. (2019). Controlling social desirability bias. International Journal of Market Research, 61(5), 534–547.

# References, cont.


[31] Lee, J. W., Jones, P. S., Mineyama, Y., & Zhang, X. E. (2002). Cultural differences in responses to a Likert scale. Research in nursing & health, 25(4), 295-306.
[32] Lewis, J. R. (1993). Multipoint scales: Mean and median differences and observed significance levels. International Journal of Human-Computer Interaction, 5(4), 383-392.
[33] Liu, M., & Cernat, A. (2018). Item-by-item Versus Matrix Questions. Social Science Computer Review, 36(6), 690–706.
[34] Liu, M., & Wronski, L. (2018). Examining completion rates in web surveys via over 25,000 real-world surveys. Social Science Computer Review, 36(1), 116-124.
[35] Micklewright, J., & Schnepf, S. V. (2010). How reliable are income data collected with a single question?. Journal of the Royal Statistical Society Series A: Statistics in Society, 173(2), 409-429.
[36] Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychological Review, 63, 81-97.
[37] Miller, P. V., Sudman, S., Bradburn, N. M., & Belson, W. A. (1985). Asking Questions: A Practical Guide to Questionnaire Design. Journal of the American Statistical Association, 80(389), 235.
[38] Moore, J. C., & Loomis, L. S. (2001). Using alternative question strategies to reduce income nonresponse. Survey Methodology, 3.
[39] Neuert, C., Joss Roßmann, & Silber, H. (2023). Using Eye-Tracking Methodology to Study Grid Question Designs in Web Surveys. Journal of Official Statistics, 39(1), 79–101.
[40] Pew Research Center. (2018, July 5). How does Pew Research Center measure the religious composition of the U.S.? Answers to frequently asked questions. Pew Research Center. https://www.pewresearch.org/religion/2018/07/05/how-does-pew-research-center-measure-the-religious-composition-of-the-u-s-answers-to-frequently-asked-questions/
[41] Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. Acta Psychologica, 104(2000), 1-15.
[42] Puckett, J. A., Brown, N. C., Dunn, T., Mustanski, B., & Newcomb, M. E. (2020). Perspectives from transgender and gender diverse people on how to ask about gender. LGBT health, 7(6), 305-311.
[43] République française. (2019, March 18). Article 8 [Interdiction de traitement de certaines catégories de données à caractère personnel]. *Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés*. Légifrance. https://www.legifrance.gouv.fr/loda/article_lc/LEGIARTI000037090124/2019-03-18
[44] Revilla, M., & Couper, M. P. (2018). Comparing grids with vertical and horizontal item-by-item formats for PCs and smartphones. Social Science Computer Review, 36(3), 349-368.
[45] Revilla, M., & Ochoa, C. (2017). Ideal and maximum length for a web survey. International Journal of Market Research, 59(5).
[46] Robinson, S. B., & Leonard, K. F. (2024). Demographic and other sensitive topics. In Designing Quality Survey Questions. SAGE Publications.
[47] Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. The American Journal of Evaluation, 22(2), 127-160.
[48] Shum, M. S. (1998). The role of temporal landmarks in autobiographical memory processes. Psychological bulletin, 124(3), 423.
[49] Smyth, J. D., & Olson, K. (2019). A comparison of fully labeled and top-labeled grid question formats. In P. Beatty, D. Collins, L. Kaye, J. L. Padilla, G. Willis, & A. Wilmo (Eds.), Advances in Questionnaire Design, Development, Evaluation and Testing. John Wiley & Sons.
[50] Snijders, C. C. P., & Matzat, U. (2007). Reducing social desirability bias through indirect questioning in scenarios: When does it work in online surveys?
[51] Symonds, P. M. (1924). On the Loss of Reliability in Ratings Due to Coarseness of the Scale. Journal of Experimental Psychology, 7(6), 456.
[52] Taherdoost, H. (2022). Designing a Questionnaire for a Research Paper: a Comprehensive Guide to Design and Develop an Effective Questionnaire. Asian Journal of Managerial Science, 11(1), 8–16.
[53] Toepoel, V., Das, M., & Van Soest, A. (2009). Design of web questionnaires: The effects of the number of items per screen. Field Methods, 21(2), 200–213.
[54] Tourangeau, R., Couper, M. P., & Conrad, F. (2004). Spacing, position, and order: Interpretive heuristics for visual features of survey questions. Public Opinion Quarterly, 68(3), 368-393.
[55] Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). The psychology of survey response.
[56] Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. Psychological bulletin, 133(5), 859.
[57] Vehovar, V., Couper, M. P., & Cehovin, G. (2023). Alternative Layouts for Grid Questions in PC and Mobile Web Surveys: An Experimental Evaluation Using Response Quality Indicators and Survey Estimates. Social Science Computer Review, 41(6).
[58] Winker, M. A. (2004). Measuring Race and Ethnicity: Why and How? JAMA, 292(13), 1612.
[59] Yan, T., Curtin, R., & Jans, M. (2010). Trends in income nonresponse over two decades. Journal of Official Statistics, 26(1), 145-164.