SURVEYS · QUESTIONNAIRE DESIGN

Question Order Effects: How Earlier Items Shape Later Answers

The same question can produce different answers depending on what was asked before it. This article explains the main kinds of question order effect, the psychology behind them, the classic experiments and the practical steps that reduce unwanted effects.

ModQu research article · 2,603 words · 23 references (APA 7th edition)

Survey questions are never answered in isolation. Each one is read after the questions that came before it, and respondents carry ideas, feelings and comparisons from one question to the next. As a result, the same question can produce different answers depending on where it appears in a questionnaire. These **question order effects** (also called context effects) are among the best-documented findings in survey research. Some are large enough to change the conclusions of a study.

This article explains the main types of order effect, the psychological processes behind them and the classic experiments that revealed them. It then sets out what questionnaire designers can do to prevent unwanted effects and to detect them when they occur.

A classic demonstration

The best-known order effect was found in the United States shortly after the Second World War. Respondents were asked two questions: whether the United States should let communist newspaper reporters from other countries come in and send back the news as they saw it, and whether a communist country such as Russia should let American newspaper reporters in to report the news as they saw it. When the question about communist reporters came first, a minority supported admitting them. When it came after the question about American reporters, support roughly doubled (Hyman & Sheatsley, 1950). Having agreed that Russia should admit American reporters, many respondents felt bound by fairness to agree that the United States should admit communist ones.

The effect also worked in the other direction: support for letting American reporters into Russia was lower when that question came second. Schuman and Presser repeated the experiment decades later and found the same pattern, though support for both had changed over time (Schuman & Presser, 1981). Neither order gives the "true" level of support. Each tells us how people answer in a particular context.

Types of order effect

Order effects can be described by their direction and by the relationship between the questions involved.

Direction: assimilation and contrast

In an **assimilation** effect, the answer to a later question moves towards the answer to, or the content of, an earlier one. In a **contrast** effect, the later answer moves away from it. The communist reporters example is an assimilation effect driven by consistency. Contrast effects occur when an earlier item serves as a comparison standard: a respondent who has just rated an excellent product may rate an ordinary one lower than they would otherwise have done.

Relationship: general and specific questions

Many of the most important order effects involve a general question, such as overall satisfaction, and one or more specific questions about the same topic, such as satisfaction with particular features. When the specific questions come first, they can change the answer to the general question in two opposite ways.

They can bring the specific topic to mind so that it weighs more heavily in the general judgement. Strack, Martin and Schwarz asked students about their dating frequency and their general life satisfaction. When the dating question came first, the correlation between the two answers rose sharply, because students were thinking about their dating lives when they assessed their lives as a whole (Strack et al., 1988).

Alternatively, specific questions can lead respondents to leave the specific topic out of the general answer, because they have already reported on it. In a related study of marriage and life satisfaction, introducing the two questions as being about separate topics reduced their correlation, because respondents excluded their marriage from their assessment of the rest of their lives (Schwarz et al., 1991). This is sometimes called a **subtraction** or redundancy effect.

A classic subtraction effect

An experiment on abortion attitudes illustrates the subtraction effect clearly. Respondents were asked whether a pregnant woman should be able to obtain a legal abortion if she was married and did not want any more children, a general question, and whether this should be possible if there was a strong chance of a serious defect in the baby, a specific one. Support for the general item was noticeably lower when it followed the specific item (Schuman et al., 1981). One interpretation is that respondents who had just expressed support for abortion in the more sympathetic specific case read the general question as excluding that case, and so considered only the remaining, less sympathetic reasons.

Additive and subtractive effects between parallel questions

Order effects also occur between parallel questions about different objects. David Moore analysed several Gallup experiments in which two similar questions were asked in different orders (Moore, 2002). In one, respondents were asked whether two baseball players, Pete Rose and Shoeless Joe Jackson, should be eligible for the Hall of Fame. Support for Jackson, whose case was less familiar, was higher when the question came after the question about Rose, as respondents applied a common standard to both. Moore described this as an additive effect, in which ratings of the second object rise, and distinguished it from subtractive effects, in which they fall. In some pairs, both ratings moved closer together.

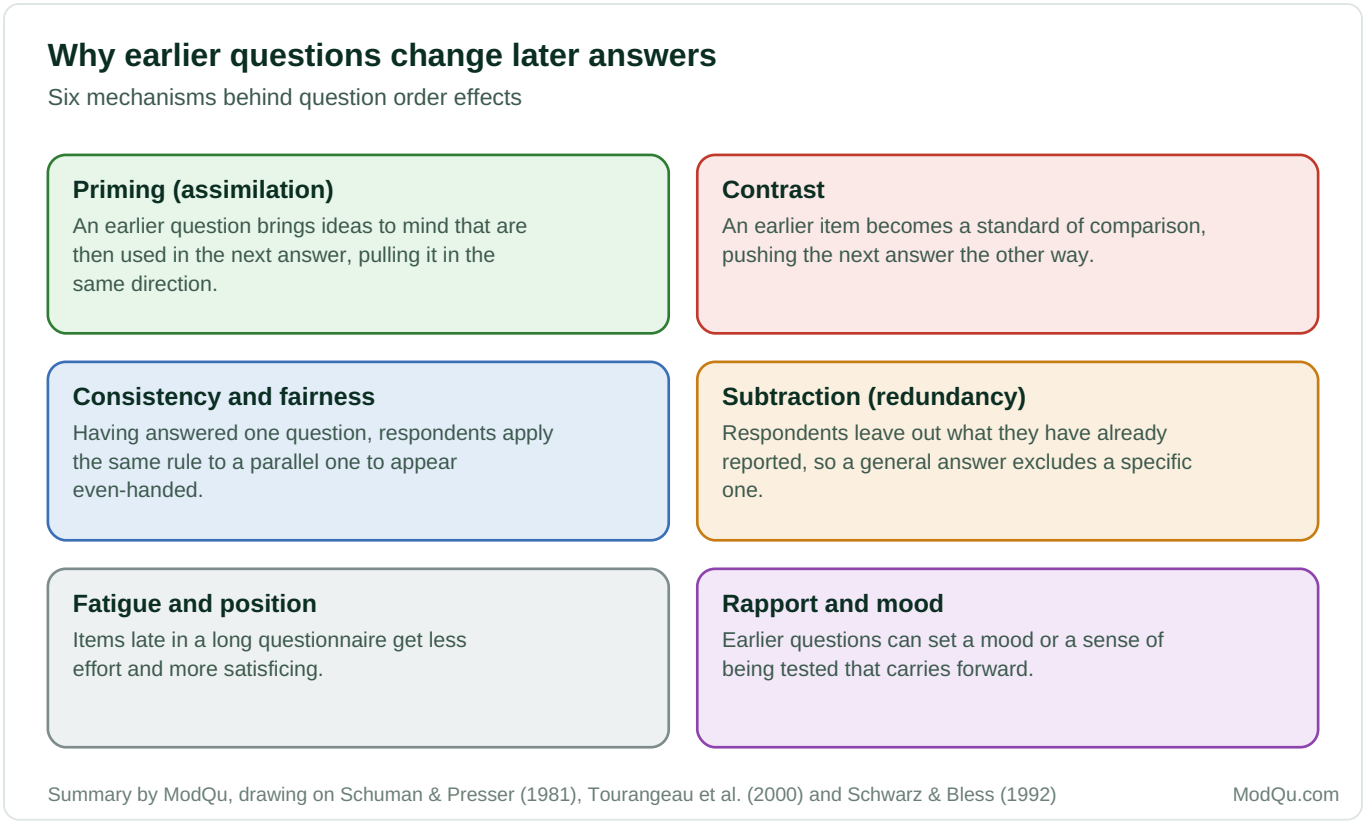


Figure 1. Six mechanisms that produce question order effects. Summary by ModQu, drawing on Schuman and Presser (1981), Tourangeau et al. (2000) and Schwarz and Bless (1992).

Why order effects happen

Order effects follow from the way people answer survey questions. Answering involves understanding the question, retrieving relevant information, forming a judgement and reporting it (Tourangeau et al., 2000). Earlier questions can affect each step.

Comprehension

Earlier questions help respondents interpret later ones. Respondents assume that each question asks for new information, so they may read a general question that follows a specific one as asking about "everything else" (Schwarz et al., 1991). Earlier questions also define the topic: a question about "the economy" that follows questions about unemployment may be read as a question about jobs.

Retrieval

Most opinions are not stored as ready-made answers. Respondents build them from whatever considerations come to mind, and recently activated ideas are more likely to come to mind (Zaller & Feldman, 1992). An earlier question makes certain considerations more accessible, which then shape the later answer. Tourangeau and colleagues found that such carryover effects were strongest for respondents who held related but conflicting beliefs about a topic, because the earlier question determined which side of their ambivalence came to the fore (Tourangeau et al., 1989).

Judgement

Whether accessible information produces assimilation or contrast depends on how it is used. According to the inclusion/exclusion model, information that respondents include in their mental picture of the target pulls the judgement towards it, while information that is excluded and used as a comparison standard pushes the judgement away (Schwarz & Bless, 1992). Strack and Martin described the same distinction in terms of how respondents decide what information is relevant and appropriate to use (Strack & Martin, 1987). Small features of the questionnaire, such as an introduction or a transition sentence, can tip respondents one way or the other.

Reporting

At the reporting stage, respondents may want to appear consistent, fair or knowledgeable. The communist reporters example reflects a norm of even-handedness. In another study, respondents who were first asked difficult knowledge questions about their member of Congress later reported less interest in politics, apparently inferring from their inability to answer that they could not be very interested (Bishop et al., 1984).

How large and how common are order effects?

Order effects are not universal. Many pairs of questions show no detectable effect, and early studies sometimes found little evidence of order effects at all (Bradburn & Mason, 1964). Effects are most likely when the questions are closely related in content, when the later question is general or ambiguous, when respondents do not have firm opinions and when the earlier question makes a particular consideration or norm prominent (Tourangeau & Rasinski, 1988; Sudman et al., 1996). Effects can be large, sometimes changing results by 10 or 20 percentage points, but they are hard to predict in advance.

Order effects are also not limited to opinion questions. Experiments with self-rated health, one of the most widely used survey measures, show that answers depend partly on whether the general health question comes before or after questions about specific conditions, and on the order of the response options (Garbarski et al., 2015). The position of the general health question can even change comparisons between groups. One study found that the gap in self-rated health between Hispanic and non-Hispanic White respondents in the United States depended partly on whether the general question was asked before or after questions about chronic conditions (Lee & Schwarz, 2014). This kind of effect matters for any study that compares groups.

Questions about local services are affected as well. A local government survey found that general ratings of government changed depending on whether they came before or after a series of questions about specific services (Benton & Daly, 1991). Order effects between general and specific evaluations of other topics, including driving standards, were reported in early British experiments (Kalton et al., 1978), and specific questions asked first have been found to change general ratings of the same topic (McFarland, 1981).

Response order effects

A related but distinct phenomenon concerns the order of response options within a single question. In self-administered surveys, where respondents can see all the options, options near the top of a list tend to be chosen more often (a primacy effect). In telephone interviews, where respondents hear the options, those heard last tend to be chosen more often (a recency effect). Krosnick and Alwin explained these patterns in terms of satisficing: respondents choose the first acceptable option they consider carefully, and which options receive careful attention depends on the mode (Krosnick & Alwin, 1987). Randomising or rotating the order of unordered options spreads these effects across options instead of letting them favour one.

Designing to manage order effects

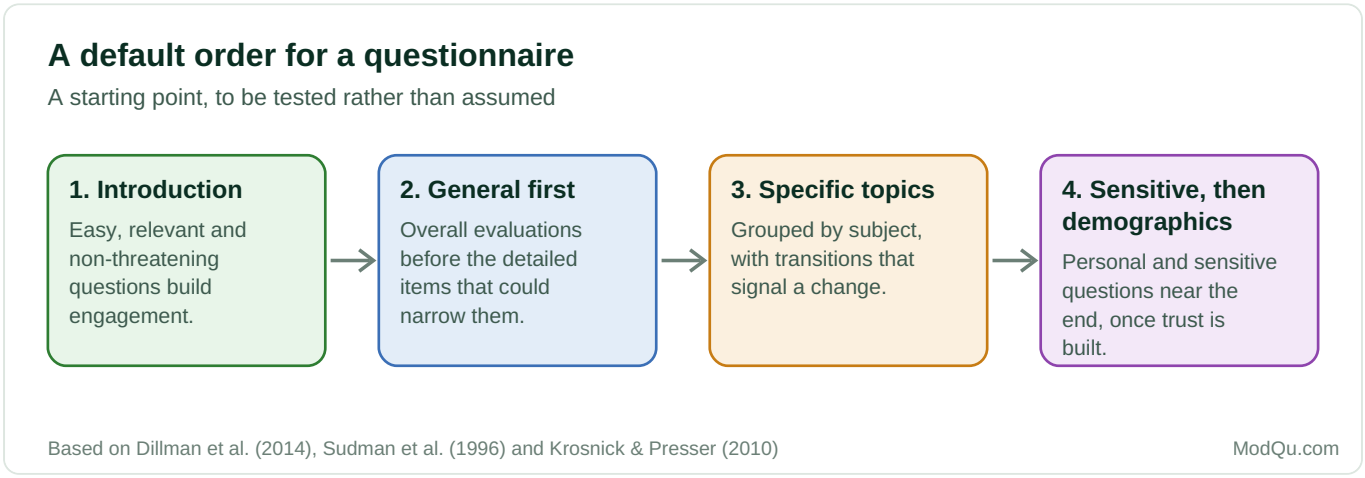


Figure 2. A default order for a questionnaire. Based on Dillman et al. (2014), Sudman et al. (1996) and Krosnick and Presser (2010).

Order effects cannot be eliminated, because every question has to come somewhere. The aim is to choose an order that produces answers suited to the research purpose and to know where effects are likely.

Put general questions before specific ones

When an overall evaluation is the main measure, such as overall satisfaction with a product or overall rating of a service, it should usually come before detailed questions about features. Asked first, the general question reflects the respondent's own overall view. Asked after a series of specific questions, it may be distorted by whichever features were just mentioned, or may exclude them (Krosnick & Presser, 2010).

Keep order the same over time

When a question is repeated to track change, it should appear in the same position and after the same preceding questions in every wave. A new block of questions inserted before a tracking question can create an apparent change that is really an order effect. This is one of the most common causes of unexplained shifts in long-running business trackers.

Separate related questions or signal a change of topic

If two related questions must both be asked but should not influence each other, they can be separated by unrelated material. Transition statements, such as "Now we would like to ask about a different topic," signal that the next question should be answered on its own terms (Schwarz et al., 1991).

Group questions by topic and build trust before sensitive items

Questionnaires are easier to answer when questions on the same topic are grouped together, in a logical order. Sensitive questions are best placed after rapport has been built, and demographic questions are usually asked at the end, except where they are needed earlier for routing (Dillman et al., 2014).

Randomise when there is no natural order

For sets of parallel items, such as ratings of several brands, randomising the order across respondents prevents any one brand from consistently benefiting or suffering from its position. Randomisation does not remove order effects; it spreads them evenly so that they do not bias comparisons. It also allows researchers to estimate the size of the effect by comparing respondents who received different orders.

Test important orders experimentally

When a key result might depend on question order, the most reliable approach is a split-ballot experiment in which random halves of the sample receive different orders (Schuman & Presser, 1981). The difference between the halves measures the order effect directly. Online survey platforms make such experiments easy and inexpensive.

Consider layout as well as sequence

In web surveys, order is not only a matter of which question comes first. Questions shown together on one screen are read as a set, so respondents are more likely to compare them and to answer them consistently than when each appears on its own page. Grids of items with a common scale encourage respondents to answer them as a block, which can

increase assimilation between items and make straightlining easier (Dillman et al., 2014). Where independence between items matters, placing them on separate pages, or interleaving them with unrelated items, reduces these effects at some cost in length and time.

Analyse and report order effects

When order has been randomised, the analysis should check for order effects rather than simply pooling the results. A simple comparison of each key result between order groups shows whether the effect is large enough to matter. If it is, the report should say so and explain which order is the better basis for the main conclusions, usually the one in which the key question was asked first. Hiding an order effect by averaging across versions gives a number that no respondent actually experienced.

A short illustration shows why this matters. Suppose a utility company asks customers to rate its overall service and, separately, the clarity of its bills. Half the sample sees the overall question first and half sees it after the billing question. If the overall score averages 7.4 out of 10 when asked first and 6.8 when asked second (illustrative numbers), the billing question is pulling the overall rating down. The company learns two things: that billing is a sore point, and that its long-running tracker should keep the overall question in first position so that results remain comparable over time.

Implications for business surveys

Order effects are especially relevant for business surveys that combine overall ratings with detailed diagnostic questions. A customer satisfaction survey that asks about delivery problems, billing errors and support waiting times before asking about overall satisfaction may produce a lower overall score than one that asks the overall question first. Neither is wrong, but the two are not comparable, and changing the order between survey waves will create a false trend.

The same applies to brand and advertising research. Asking about a competitor before asking about a client's brand can change how the client's brand is rated, through either comparison or consistency. Asking whether respondents have seen an advertisement before asking about brand attitudes can prime the advertisement's message. In each case, the safest design puts the measures that matter most for decisions first, keeps their context constant over time and uses randomisation and experiments where order is uncertain.

Conclusion

Respondents read each question in the light of the ones before it. Earlier questions shape how later questions are understood, what comes to mind, which standards of comparison are used and what answers seem consistent or fair. The result can be assimilation, contrast, consistency or subtraction, and effects can be large enough to change conclusions. They are most likely between closely related questions, especially general and specific ones. Questionnaire designers manage them by asking general questions first, keeping the context of tracking questions fixed, separating or signalling unrelated topics, randomising parallel items and testing important orders experimentally.

References

Benton, J. E., & Daly, J. L. (1991). A question order effect in a local government survey. *Public Opinion Quarterly, 55*(4), 640–642. https://doi.org/10.1086/269285

Bishop, G. F., Oldendick, R. W., & Tuchfarber, A. J. (1984). What must my interest in politics be if I just told you "I don't know"? *Public Opinion Quarterly, 48*(2), 510–519.

Bradburn, N. M., & Mason, W. M. (1964). The effect of question order on responses. *Journal of Marketing Research, 1*(4), 57–61. https://doi.org/10.1177/002224376400100410

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method* (4th ed.). Wiley.

Garbarski, D., Schaeffer, N. C., & Dykema, J. (2015). The effects of response option order and question order on self-rated health. *Quality of Life Research, 24*(6), 1443–1453. https://doi.org/10.1007/s11136-014-0861-y

Hyman, H. H., & Sheatsley, P. B. (1950). The current status of American public opinion. In J. C. Payne (Ed.), *The teaching of contemporary affairs: Twenty-first yearbook of the National Council for the Social Studies* (pp. 11–34). National Council for the Social Studies.

Kalton, G., Collins, M., & Brook, L. (1978). Experiments in wording opinion questions. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 27*(2), 149–161. https://doi.org/10.2307/2346942

Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly, 51*(2), 201–219. https://doi.org/10.1086/269029

Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (2nd ed., pp. 263–313). Emerald.

Lee, S., & Schwarz, N. (2014). Question context and priming meaning of health: Effect on differences in self-rated health between Hispanics and non-Hispanic Whites. *American Journal of Public Health, 104*(1), 179–185. https://doi.org/10.2105/AJPH.2012.301055

McFarland, S. G. (1981). Effects of question order on survey responses. *Public Opinion Quarterly, 45*(2), 208–215. https://doi.org/10.1086/268651

Moore, D. W. (2002). Measuring new types of question-order effects: Additive and subtractive. *Public Opinion Quarterly, 66*(1), 80–91. https://doi.org/10.1086/338631

Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. Academic Press.

Schuman, H., Presser, S., & Ludwig, J. (1981). Context effects on survey responses to questions about abortion. *Public Opinion Quarterly, 45*(2), 216–223. https://doi.org/10.1086/268652

Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Erlbaum.

Schwarz, N., Strack, F., & Mai, H.-P. (1991). Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *Public Opinion Quarterly, 55*(1), 3–23. https://doi.org/10.1086/269239

Strack, F., & Martin, L. L. (1987). Thinking, judging, and communicating: A process account of context effects in attitude surveys. In H.-J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 123–148). Springer.

Strack, F., Martin, L. L., & Schwarz, N. (1988). Priming and communication: Social determinants of information use in judgments of life satisfaction. *European Journal of Social Psychology, 18*(5), 429–442. https://doi.org/10.1002/ejsp.2420180505

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. Jossey-Bass.

Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin, 103*(3), 299–314. https://doi.org/10.1037/0033-2909.103.3.299

Tourangeau, R., Rasinski, K. A., Bradburn, N., & D'Andrade, R. (1989). Carryover effects in attitude surveys. *Public Opinion Quarterly, 53*(4), 495–524. https://doi.org/10.1086/269169

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press. https://doi.org/10.1017/CBO9780511819322

Zaller, J., & Feldman, S. (1992). A simple theory of the survey response: Answering questions versus revealing preferences. *American Journal of Political Science, 36*(3), 579–616. https://doi.org/10.2307/2111583