Satisficing behavior in online panelists

Theo Downes-Le Guin

MRA Annual Conference & Symposium

2 June 2005

Background on Doxus

Primary market research consultancy catering to technology (hardware, services, software) and financial services companies.

Founded in 2000. Offices in Portland and Atlanta, and an established network of partners in clients' mature and emerging overseas markets.

Mix of work:

- About two-thirds quantitative (web, phone, hybrid), one-third qualitative (groups, IDIs, online threaded discussions, ethnography-based methods).
- For many clients we are mainly enterprise or business-focused, but we have a substantial flow of consumer and microbusiness research as well.

Half of our research takes place outside of the US.

Focus markets and categories in high tech:

- Productivity, collaboration and enterprise software; management applications; middleware.
- PCs and servers.
- Core and edge fixed-wire and wireless networking.
- Imaging and printing, from low-end devices to production and commercial printing/copying.
- Web stores and e-services.
- Developer and channel programs and services.
- Technical support.

Background

Web surveys + web panels offer immediate research gratification based on cost and time-to-data.

“Doing a web survey” is simultaneously a choice of sample source and mode.

As a result, potential sources of survey error are minimized or overlooked.

Numerous studies have focused on optimal web survey design, but few have examined how the response pattern of panelists differs.

Web panels are no different from any other form of panel research: require special care in the panel management process, survey design and data cleaning to balance cost advantages against potential error.

Agenda

What is satisficing?

Opportunities to reduce satisficing in:

- Panel building/management
- Survey design
- Data cleaning

Cost implications

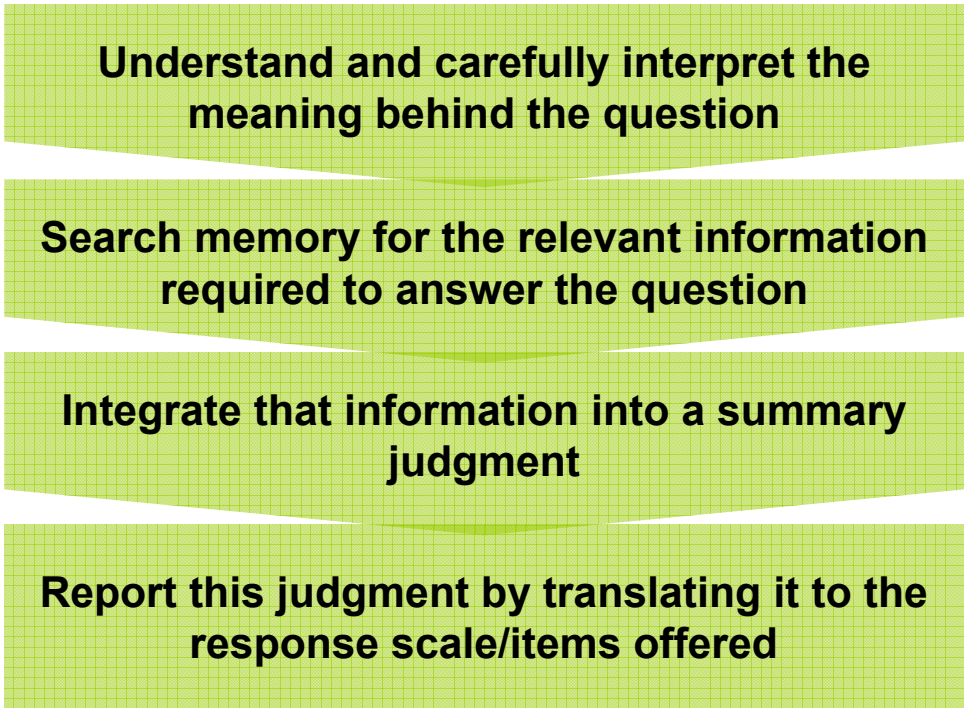
Conclusions

What is satisficing?

The “optimizing” respondent

The ideal respondent “optimizes” as s/he answers every question, conducting a “complete and unbiased search of memory and full integration of retrieved information.”

Optimizing requires a 4-step cognition and retrieval process described by Tourangeau and Rasinski (1988):



What is satisficing?

The “satisficing” respondent

The satisficing theory (Krosnick, 1991) defines satisficers as respondents who cease to follow these steps.

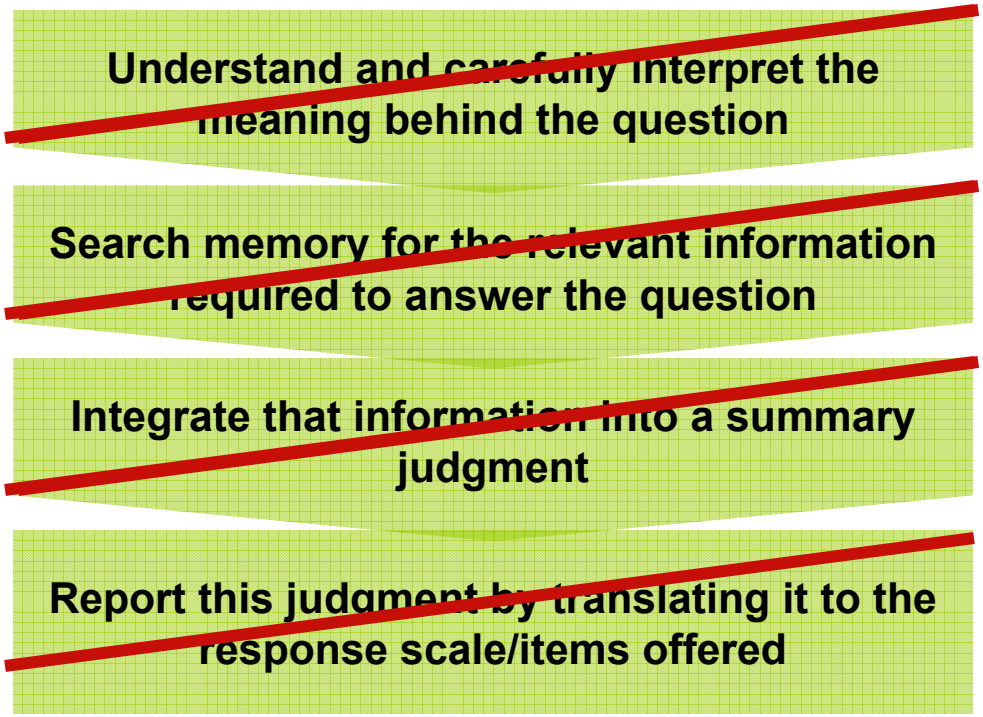
Satisficers’ responses are formulated with “reduced thoughtfulness, careless integration of retrieved information, and a haphazard selection of response choice.”

3 conditions govern the likelihood of satisficing:

- Cognitive sophistication of the respondent.
- Difficulty of the task (not all questions are equally difficult).
- Respondent motivations.

What is satisficing?

Levels of satisficing behavior



Satisficing due to…

Cognitive unsophistication of the respondent: the questionnaire assumed too much about the respondent's abilities and comprehension.

Difficulty of the survey task: the questionnaire was dull and/or difficult relative to the incentive, and the respondent got cranky.

Respondent motivations: the panelist was underqualified, hurried or motivated only by financial gain.

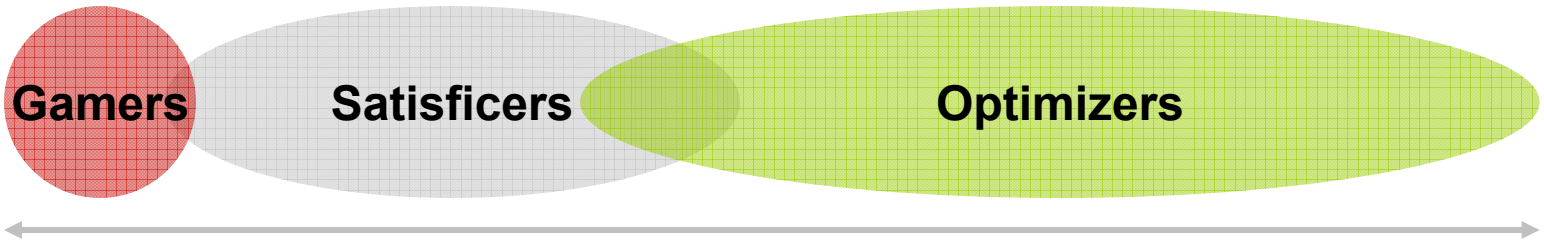
What is satisficing?

Web-based panel opportunities for satisficing

In web-based panels, the combination of means and motive for satisficing can be strong:

- Traditional forms of validation (human monitoring) not available.
- Predictable financial opportunities.
- Ongoing exposure creates considerable opportunity for “learning by doing.” Over time, it becomes easier to predict responses that will not trigger termination in the screener.

Our experience with web survey data allows us to characterize 2 types of satisficers:



Opportunities to reduce satisficing

There are 3 chief opportunities to minimize satisficing and gaming:

Panel building/ management

Survey design

Data cleaning

Best practices in panel management, questionnaire design and web interface design can minimize many forms of satisficing, but are not a panacea.

Opportunities to reduce satisficing

Panel building/management

Panel building/ management

Goals: 1 qualified respondent ▸ 1 complete ▸ optimized responses.
Current best practices:

- Verification:
  - Email address verification: check validity and eliminate duplicate email addresses and sign-ups, with a limit of 1 per panelist.
  - Physical address verification: check validity and eliminate duplicate physical addresses, with a limit of 1 per panelist.
- Mail incentives to a physical address.
- Random validation: periodic, phone-based validation of respondent profile and identity and/or validation of specific survey responses.
- In survey deployment, use PINs that are unique, random and alphanumeric rather than sequential.
- Cap on number of surveys during time period and forced panelist attrition to discourage development of "professional" respondents and minimize satisficing due to panelist fatigue across surveys.

Opportunities to reduce satisficing

The worst case: “Gamers”

Panel building/ management

Weak satisficers are those whose responses to some questions are not logically consistent. Nearly all panelists satisfice at some point.

Gamers, on the other hand, engage in the survey (and possibly the panel) with the intention of minimizing effort and maximizing reward.

Incentive structure is critical in dictating the best way to game the system: Fixed incentives offer panelists a high incentive to qualify, and low incentive to drop out when:

- The survey turns out to be longer than expected
  or
- When exogenous factors kick in that would otherwise prompt the respondent to stop taking the survey.

Opportunities to reduce satisficing

Case study: software reseller research study, Jan 2005

Panel building/ management

Survey of professionals from a “viral recruit” panel.

We suspected a potentially high rate of satisficing due to the length and complexity of the survey.

Mean completion times were suspiciously low, alerting us that there might be significant data quality issues. Looking for bimodal distributions of completion times to help identify speed-clickers.

Examination of verbatim and continuous variable responses showed that 5 identical response patterns were each repeated dozens of times in the dataset…accounting for more than 200 records.

All cases were the result of a single gamer, using keystroke automation software to auto-fill the survey repeatedly.

- The panelist had created more than 200 bogus panel profiles, using “unique” email addresses in a domain he owned.
- The panel provider had no mechanism for ensuring valid panelist physical addresses in order to prevent duplication.

Opportunities to reduce satisficing

Survey design

Survey design

Best practices for offline survey design and for survey web interface design are the best remedy — keep questionnaires short, salient and intuitive.

In addition, several steps can help minimize panelist satisficing in:

- Screening questions.
- “Table-style” questions.
- Verbatim questions.

Opportunities to reduce satisficing

Screening questions

Survey design

Response pattern: Satisficers and gamers select all responses in a multiple response question suspected of leading to termination.

Solution:

- Include low incidence response options (products, brands, qualifying behaviors).
- Avoid terminating respondents at the decoy question.

Which of the following products or services — if any — do you, or do other household members, currently own/use in your home?

- Any type of PC/personal computer
- Any type of printer
- Digital camera
- Portable DVD player
- Any wireless home networking equipment
- Segway Human Transporter
- Camcorder (digital or analog)
- Equipment and a calling plan that allows you to make/receive phone calls using voice over IP (VoIP) on a high-speed internet connection
- TiVo or any other brand of digital video recorder (DVR)
- None of the above

Incidence of “Segway owners” in recent studies:

Case A (N=1,879): 27 panelists (1%) claimed to be Segway owners. Excluding Segway, another 22 respondents chose *all* products mentioned.

Case B (N=728): 13% of panelists claimed to be Segway owners in another consumer study. These respondents completed the survey in an average of 5.5 minutes; 71% of the “Segway owners” reported household incomes of less than $35K in 2004. For the same study, a client-supplied list yielded no Segway owners (N=73) and an interview length of 17 minutes.

Opportunities to reduce satisficing

Screening questions design

Survey design

Response pattern: Satisficers and gamers select all responses in a multiple response question suspected of leading to termination.

Solution: Treat respondents who select every item in a list of products with skepticism.

In a screening question “which of the following technology-related products and services are you involved in purchasing and/or leasing for your business,” we offered 7 response options in total, plus a “none of the above” option (which would lead to immediate termination).

Case study of N=712 professionals:
n=451 from client-provided sample
n=261 from a combination of consumer- and business-focused online panels

Client customer list	Panelists
39% said they had decision-making responsibilities for every listed product/service.	66% said they had decision-making responsibilities for every listed product/service.
Completed survey in mean of 25 minutes (median was 23 minutes).	Completed survey in mean of 14 minutes (median was 12 minutes).

Opportunities to reduce satisficing

Substantive questions design

Survey design

Response pattern: Satisficers and gamers “straightline” (click same response vertically) all responses on a page or use browser auto-fill.

Solutions:

- Include both positive and negative statements, proving respondent inconsistency when straightlining occurs.
- Include verification ratings.
- Use programming logic that prompts respondents to check answers when all responses within a question set are identical.

Convenience is more important to me than low cost	● ● ● ● ● ● ●
Please verify your place in the survey by checking the second box from the left	● ● ● ● ● ● ●

Incidence of failures for verification ratings in table-style questions in recent studies:

Case A: 12% of respondents (N=822) from a consumer panel.

Case B: 34% of respondents (N=234) from a consumer panel.

Opportunities to reduce satisficing

Data cleaning

Data cleaning

Goal: Identify and flag response behaviors and patterns that indicate varying degrees of satisficing, up to and including gaming.

Variations include manual or semi-automated review of:

- Verbatims or continuous numerical responses for gibberish, clear internal inconsistencies and “copy and paste” responses.
- Time stamps (less suspensions) to determine interview length.
- Response patterns for questions that are either cognitively complex or subject to “straightlining.”

Basic process:

- Define criteria upfront unique to study. Each criterion represents a “strike.”
- Fine-tune criteria based on an interim dataset.
- Identify likely satisficers based on the total number of strikes in final dataset.
- Drop and replace cases, or weight results, based on results of data cleaning.

Opportunities to reduce satisficing

Data cleaning *continued*

Data cleaning

Some strikes may be used as absolute filters (delete the case if the condition is not met).

- Example: Interview length cutoff at 10th to 15th percentile, adjusted for those who legitimately follow a shorter survey path due to skip patterns.

Each criterion that is not met is associated with a cleaning rule.

- In practice, we classify as weak satisficers those who may have some strikes against them but are under a minimum number required for case deletion. Such records are cleaned on a question-by-question basis.

Overall, determination of the number of strikes that requires deletion of a case is a matter of judgment unique to each study.

The process identifies strong satisficing behavior but may or may not identify gaming behavior, which may only be discernable across multiple surveys.

Opportunities to reduce satisficing

Case study: hardware rebate study, Feb 2005

Data cleaning

Criteria for review

Interview lengths more than 2 standard deviations LOWER than the trimmed mean.
STRIKE 1: n=125 (length of interview cutoff customized for short/long interview based on skip patterns)

Reported Segway ownership.
STRIKE 2: n=27

Reported owning many types of PCs and all types of peripherals (choosing all responses in a couple of multiple choice questions).
STRIKES 3 and 4: n=8 and n=6, respectively

Selected each item in seemingly diametrically opposed pairs (statement 3 cannot be chosen with statements 1, 2 or 4).
STRIKE 5: n=15

Straightlining in a statement agreement/disagreement series.
STRIKE 6: n=26

Frequency for sum of strikes:

- 0 strikes: n=1,700
- 1 strike: n=158
- 2 strikes: n=20
- 3 strikes: n=1

All cases with 2+ strikes and those failing the interview length filter were deleted from the dataset, reducing the sample size from N=1,879 to N=1,800 (4% of cases deleted).

Which of the following statements, if any, are true about the new [BRAND] [PRODUCT] you purchased for your household? *Please select all that apply.*

1. **The brand/model of [PRODUCT] that was my first choice was not available at the time I wanted to buy, so I ended up getting the [BRAND] [PRODUCT] instead**
2. **I entered the purchase decision-making process expecting to buy another brand of [PRODUCT], but ended up buying a [BRAND] [PRODUCT] instead**
3. **I knew I wanted a [BRAND] [PRODUCT] when I started my purchase decision-making process, and that's the brand I ended up buying**
4. **I really didn't have a strong brand preference when I came into the purchase decision-making process**
5. I knew what I wanted in a [PRODUCT], and basically spent my time shopping for the lowest-priced [PRODUCT] that would meet my needs
6. I set out thinking I wanted to buy a single-function (print-only) printer, but ended up buying an all-in-one printer instead
7. I wasn't necessarily shopping for a [PRODUCT], but I changed my mind during the process of buying a PC
8. None of the above are true for me/accurately describe my situation

Opportunities to reduce satisficing

Case study: BI software study, July 2005

Data cleaning

Criteria for review

Interview length less than 5 minutes (actual median: 14 min) or choice task length less than 2 minutes (actual median: 3 min).
STRIKE 1: n=48

Answering zero to all choice tasks.
STRIKE 2: n=30

Giving the same numeric response for all choice tasks.
STRIKES 3: n=123

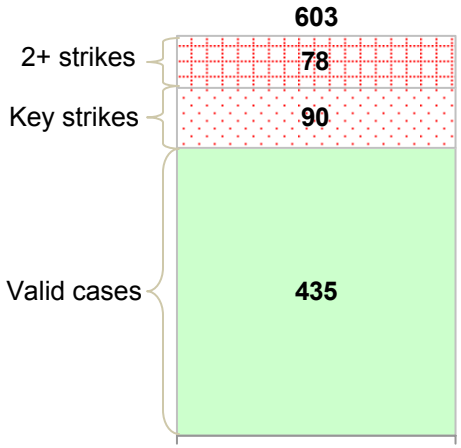
Shotgunning server brands—reports owning all brands.
STRIKE 4: n=59

Shotgunning BI software brands—reports owning all brands.
STRIKE 5: n=3

Frequency for sum of strikes:

0 strikes: n=424
1 strike: n=101
2 strikes: n=72
3 strikes: n=6

All cases with 2+ strikes and those failing the interview length, unresponsive or straightlining DCM filter, or shotgunning BI software filter were deleted from the dataset, reducing the sample size from N=603 to N=435 (28% of cases deleted).



Which of the following BI applications and/or tools is your organization currently using?? *Please select all that apply.*

- ☒ Brio
- ☒ Business Objects
- ☒ Cognos
- ☒ Hyperion
- ☒ IBM
- ☒ JD Edwards
- ☒ MicroStrategy
- ☒ Oracle
- ☒ PeopleSoft
- ☒ SAP
- ☒ SAS
- ☒ A “home-grown” solution
- ☐ Other, please specify
- ☐ Don’t know

Opportunities to reduce satisficing

Can we tell panelists too much?

Panel providers can and should develop relationships with panelists within the bounds of industry ethics guidelines and common sense.

But how much do we want to reveal to panelists about how we ensure the validity of their identities and responses?

From a panel provider's standard survey intro:

The accuracy of data is critical to meeting market research objectives. [Panel] routinely employs a verification process of respondent survey data. Through analysis of respondent data patterns and the overall time that a respondent takes to enter answers to survey questions, [Panel] ensures the integrity of project data. [Panel] reserves the right to reasonably identify surveys as accurate and complete. You, as a respondent to this survey, agree to provide accurate data and agree to abide by [Panel's] membership terms and conditions.

Does such disclosure:

- Create an environment of honesty with panelists?
- Discourage satisficing?
- Give gamers clues as to how to beat the system?

Cost implications

Cleaning of interim and final datasets: 4 to 8 hours per dataset to define criteria, create automated analysis syntax and analyze dataset(s).

Case deletion:

- For all commercial, internet-recruited US web panels, we specify an overage between 10% and 15% above sample size targets per quota (5% in Europe) to compensate for data cleaning (as compared to 1% per quota group for telephone surveys from non-panel sources).
- Our regular panel providers are contractually responsible for costs of replacement, but we:
  - Provide full access to our decision rules so that they participate in defining what constitutes “strikes.”
  - Provide deleted case IDs so that the panel provider can create a feedback loop and determine whether/when a panelist should be attrited.

Conclusions

Adopting some small-scale changes in research design and data cleaning standards can help:

- Minimize the impact of routine satisficing on survey data quality.
- Catch the few “bad apples.”

Suppliers: Minimal extra steps can significantly improve QA.

Clients: Incorporating these steps into proposal requests will move the industry toward de facto standardization.

We need to be judicious, however, in the number of “traps” we use:

- Low incidence response options, ratings verifications and the like increase questionnaire length and may annoy the majority of respondents who are conscientious.
- Obvious traps will, over time, make it wholly obvious to gamers how we are seeking to trap them.

Thank you!

More information:

Theo Downes-Le Guin
Doxus
503 788 7749 x111
theo.downes-leguin@doxus.com