



Designing good closed-ended questions for impact evaluation

Dr. Eric Jensen: e.jensen@warwick.ac.uk

University of Warwick

1

Presentation Roadmap

- Common issues in quantitative survey design (for impact)
- Importing bad methods
- How to improve quantitative survey design



Survey-based Impact Evaluation – Current Approaches

Common problems:

- Relying on post-visit only self-report
- Proxy reporting

Measurement

- A key issue is what will be captured on a particular measure (i.e. ‘what counts?’)
- Measurement error is an issue. (i.e. error due to measurement approach/tool)
- Where feasible, important to directly measure relevant variables such as knowledge, e.g. before/after

Survey Design for Evaluation...

Common Problems



Current Problems

- ‘Industry standard’ survey design often rife with basic errors and poor practice in survey design, sampling and analysis.
- Many institutions are uncritical consumers (and producers) of survey-based evaluation research, quick to believe measuring complex outcomes is simple.



Proxy Reporting: Parents reporting for children

Example: Evaluating California Science Center impacts on children

- Falk and Needham (2011) sought to measure the Science Center's impacts on children by asking parents to report on cognitive and affective outcomes.
- First, parents were asked to indicate whether their children had gained an increased understanding of science and technology after visiting the Science Center.
- Falk and Needham (2011: 5) reported that **'nearly all adults (87%) who indicated that their children had visited the Science Center agreed that the visit increased their children's science or technology understanding, with 45% believing that the experience increased their children's understanding "a lot"'**.

Example: Evaluating California Science Center impacts on children

- This survey item raises obvious issues surrounding the unreliability of expecting different parents within a sample to judge what counts as “a lot” of learning.
- Because this measure was not explicitly operationalized during the design of the survey, respondents will likely interpret “*a lot*” of *learning* in different ways, subsequently undermining the quality of the data being collected.
- In addition, parents are being asked to provide one assessment regardless of the number of children they may have.
 - What if parents feel that one of their children learned “a lot”, while another learned “a little” and a third “nothing” at all? Are parents really likely to be making a considered judgment here?
- Asking parents to provide an off-the-cuff assessment of their child’s learning is highly prone to error.

Proxy Reporting: Teacher Feedback

What can we validly get from teachers?

- Their opinions/satisfaction on a range of topics, for example:
 - ✓ The service they have received.
 - ✓ The education provision on offer and how/if they used it.
 - ✓ Comparisons with competing organisations.
 - ✓ Their professional opinion on any improvements to education provision.

Headline:

**We can never use teacher
opinion as a proxy
indicator of student impact
(*same goes for parents!*)**

Some teacher comments from a zoo evaluation that cannot be taken as statements of fact:

“The kids loved it, and they didn't really think about how much they were learning as they looked around.”

“The cherry on top was when he brought out the snake - the kids loved it and they learnt an exceptional amount of information for a one hour session.”

These are still the perceptions of the teacher, not measures of impact on the learners involved.



Examples courtesy of Andy Moss, Chester Zoo

Some teacher comments from a zoo evaluation that cannot be taken as statements of fact:

“Very enjoyable day for all. Children were very enthusiastic and gained a lot from the experience.”

“I think it's 100% educational as the Zoo is so involved with highlighting the importance of preserving ecosystems (even the cafes); also watching animals invariably increases understanding of them.”

These are still the perceptions of the teacher, not measures of impact on the learners involved.



Examples courtesy of Andy Moss, Chester Zoo

Some 'bad' use of teacher comments (from an MLA museums report)

Section titled 'Powerful learning outcomes for pupils':

- 'Teachers continue to value the inspirational quality of museum work because *it enables their pupils to better understand their subject*'
- '97% of teachers think their *pupils are likely to have been inspired to learn more*'
- '94% of teachers think their *pupils will have gained subject-related facts*'
- '94% of teachers think their *pupils are likely to have developed thinking skills and (89%) communication skills*'

([Greenhill, et al., 2007, p. 28](#))

These are still the perceptions of the teacher, not measures of impact on the learners involved.



To sum up:

...there's no problem with reporting teacher opinion as opinion.

For example:

“Teachers believed that the the visit was a valuable learning experience”.

but important to avoid temptation to confuse opinion with actual impacts on learners...

Question Design:

Not:

“How do you think your visit to the zoo benefitted your students?”

Try:

“Did the session you attended link with the part of the curriculum you teach in school?”

Stick to getting teacher feedback in areas of expertise – much more useful and valid.

Common problems – Self-reporting impact

Impact Evaluation – Common problems

– **AVOID!**

‘Did you learn during your visit to the science museum today?’: Yes or No?.



Science Communication Evaluation – Current Approaches

London's Science Museum's internal guidance for evaluation includes the following flawed survey item:

'To what extent do you agree or disagree with the following statements?' (strongly agree to strongly disagree):

'I have learnt something new today' (National Museum of Science & Industry, 2009).



Evaluation – Current Approaches

Other common problems in evaluation surveys include:

- Skewed response options (e.g. all positive)
- No ‘don’t know’, ‘neutral’ or ‘no opinion’ option
- ‘Other report’ requested (e.g. teacher for students)



Over to you

Would you recommend this workshop to a friend or colleague?

0= Not at all likely

5= Neutral

10= Extremely Likely



1. Give a numerical score for the workshop (no discussion!)
1. Write down a short description of what you mean by this score indicates about your views on this workshop.
2. Be prepared to share your description

Common problems: Self-reporting future behaviour





81%

Net Promoter Score

81% of visitors would recommend
us to friends and family

Net promoter score



$$\text{NPS}^{\text{®}} \text{ (Net Promoter Score)} = \text{Promoters (\%)} \text{ (9s and 10s)} - \text{Detractors (\%)} \text{ (0 through 6s)}$$

Primary concerns about this:

- **Validity** (are you measuring what you are really interested in?)
- **Reliability** (are you gathering data in a consistent way?)

Net promoter score

The NPS Calculation

Calculate your Net Promoter Scores using the answer to a single question, using a 0-10 scale: How likely is it that you would recommend [brand] to a friend or colleague? This is called the Net Promoter Score question or the recommend question.

Respondents are grouped as follows:

- *Promoters* (score 9-10) are loyal enthusiasts who will keep buying and refer others, fueling growth
- *Passives* (score 7-8) are satisfied but unenthusiastic customers who are vulnerable to competitive offerings.
- *Detractors* (score 0-6) are unhappy customers who can damage your brand and impede growth through negative word-of-mouth.

Further examples



Conservation Biology



Conservation Practice and Policy

Knowledge Gain and Behavioral Change in Citizen-Science Programs

REBECCA C. JORDAN,* STEVEN A. GRAY, DAVID V. HOWE, WESLEY R. BROOKS,
AND JOAN G. EHRENFELD†

Ecology, Evolution, and Natural Resources, 14 College Farm Road, Rutgers University, New Brunswick, NJ 08901-8551, U.S.A.,
email jordan@aesop.rutgers.edu

<i>Question</i>	<i>Mean pretest (n = 82)</i>	<i>Mean follow-up (n = 33)</i>
To what extent are you knowledgeable about environmental science?	2.29	2.66

Response options:

- 1 - a lot;
- 2 - a fair amount;
- 3 - only a little;
- 4 - nothing;



What is wrong with this?

To what extent are you knowledgeable about the research and regulatory infrastructure as they relate to environmental issues?

3.54

3.39

Response options:

- 1 - great extent;
- 2 - considerable extent;
- 3 - moderate extent;
- 4 - slight extent;
- 5 - no extent.



What is wrong with this?

Quiz time!

Survey question examples: Test Time!

Q1. What will be the one thing you remember from your museum visit in 3 months time?

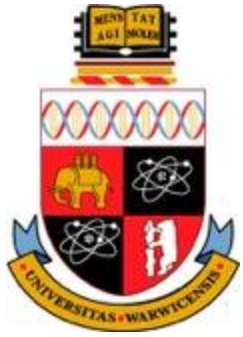
***Q2. How much have you learned about modern art today?
a) nothing at all b) a little c) a moderate amount d) a great deal***

Q4. Are you now aware of London Zoo's conservation work, both in the zoo and in the wild?

Yes / Most of it / No

DIY Survey Design:

In small groups, propose three outcome items for evaluating the effects of this workshop



Dr. Eric Jensen
e.jensen@warwick.ac.uk
University of Warwick