Evaluating quality of education

Zoe Säflund

17 August, 2022

Cambridge, LEANBODY training

Overview

- Course evaluations: your experiences

- Taking course evaluations one step further: the KI experience

- An example: monitoring internationalisation at KI

- Common pitfalls in question design

Let's talk!

1. How do you evaluate quality of education at your university? Examples?

2. What do you think the questions in your course evaluation measure/should measure? Where should the focus be on?

3. As a teacher, what do you think is important to ask your students regarding your course?

4. What other ways could you use or already use to get feedback about the course from your students?

"The purpose of evaluation is to improve, not prove."

Stufflebeam, 1983

Course evaluation versus course analysis

- **Course evaluation** is the summative course participants' feedback given anonymously via a web-based questionnaire at the end of a course.

- The **course analysis** is the summary made by the course director, based on:
    - outcome of the course evaluation
    - teachers' experiences from the course
    - other data sources
    - changes that are planned as a result of the above

Student evaluations at KI

- Newcomers survey

- Course evaluations

- Exit polls

- Alumni survey

- Other thematic surveys (questions about the transition to digital teaching and examinations, equal opportunities survey etc)

- Employer surveys about students

The purpose of a course evaluation

- Student voice should be heard to improve the learning experience and quality

- Feedback on the teaching and learning quality

- Feedback on students' own learning processes

- Opportunity for the student to think about the course as a whole

- Students learn how to give feedback

- An evaluation should <u>not</u> focus on student satisfaction

The course evaluation process at KI

- Standardized procedure to evaluate quality of education at KI.

- A course evaluation should be sent out to students after the completion of each course.

- Questions common to all courses at undergraduate level and doctoral level.

- Possibility to add program-specific and course-specific questions.

How we do it

- Use multiple sources to evaluate quality of education

- Ask feedback on the teaching-learning environment

- Ask feedback on the students own learning processes

- Evaluate students' perceptions of different elements of learning

- Remember: Satisfaction in itself cannot be used to measure quality of teaching

The five general questions

- Compulsory to all courses at undergraduate level.

- Index of student-perceived quality. Scale with high internal consistency.

- Measures five quality aspects of education as experienced by students at course level. These are:
  - Relevance
  - Achievement of learning outcomes
  - Constructive alignment
  - Scientific approach and link to research
  - Student voice

The five general questions

1. In my view, I have developed valuable expertise/skills during the course.

2. In my view, I have achieved all the intended learning outcomes of the course.

3. In my view, there was a common theme running throughout the course – from learning outcomes to examinations.

4. In my view, the course has promoted a scientific way of thinking and reasoning (e.g., analytical and critical thinking, independent search for and evaluation of information).

5. In my view, during the course, the teachers have been open to ideas and opinions about the course's structure and content

   Responses:´1-5 Likert scale (1-to a very small extent, 2-to a small extent, 3-to some extent, 4-to a large extent, 5-to a very large extent)

Question bank

- Includes questions about different aspects of education

- Currently updated to be more in line with the AGENDA 2030 and KI Strategy

- Areas of Intercultural Competence, Global Health perspectives, Gender Equality etc.

- Possibility to add questions in program specific templates

Exit poll: I feel well prepared for my future role's requirements to:

a) co-operate

b) critically review information

c) communicate in writing (present information, problems and solutions in writing)

d) communicate verbally (verbally present information, problems and solutions)

e) communicate in English

f) work independently.

g) search for necessary information

h) solve problems independently

i)

Quotes - exitpolls

- To study together with other international students, with different backgrounds. I have improved my skills by learning from everyone

- I think my favorite part of the study was having an international class. XXX is meant to be a collaborative course. Although we focus on XXXX mostly on the Swedish or European setting, having perspectives from different cultural backgrounds enhances the whole experience.

- Meeting people from different corners of the world.

- Leave bias at the door and be aware of your cultural background when assessing a new piece of information.

Alumni: How well does the content of your education match the following aspects of your current job's demands:

	To what extent did your education contribute to you developing the following knowledge and abilities to:	To what extent do you need at your current job the following knowledge and abilities to:
Cooperate in interprofessional teams	To a very small extent --------To a very large extent	To a very small extent -------To a very large extent
Apply a global health perspective on a variety of issues	To a very small extent --------To a very large extent	To a very small extent -------To a very large extent
Cooperate in diverse cultural environments	To a very small extent --------To a very large extent	To a very small extent -------To a very large extent
Have a broad understanding of international events shaping the world	To a very small extent -------To a very large extent	To a very small extent --------To a very large extent

Usefulness of education

Utbildningens användbarhet
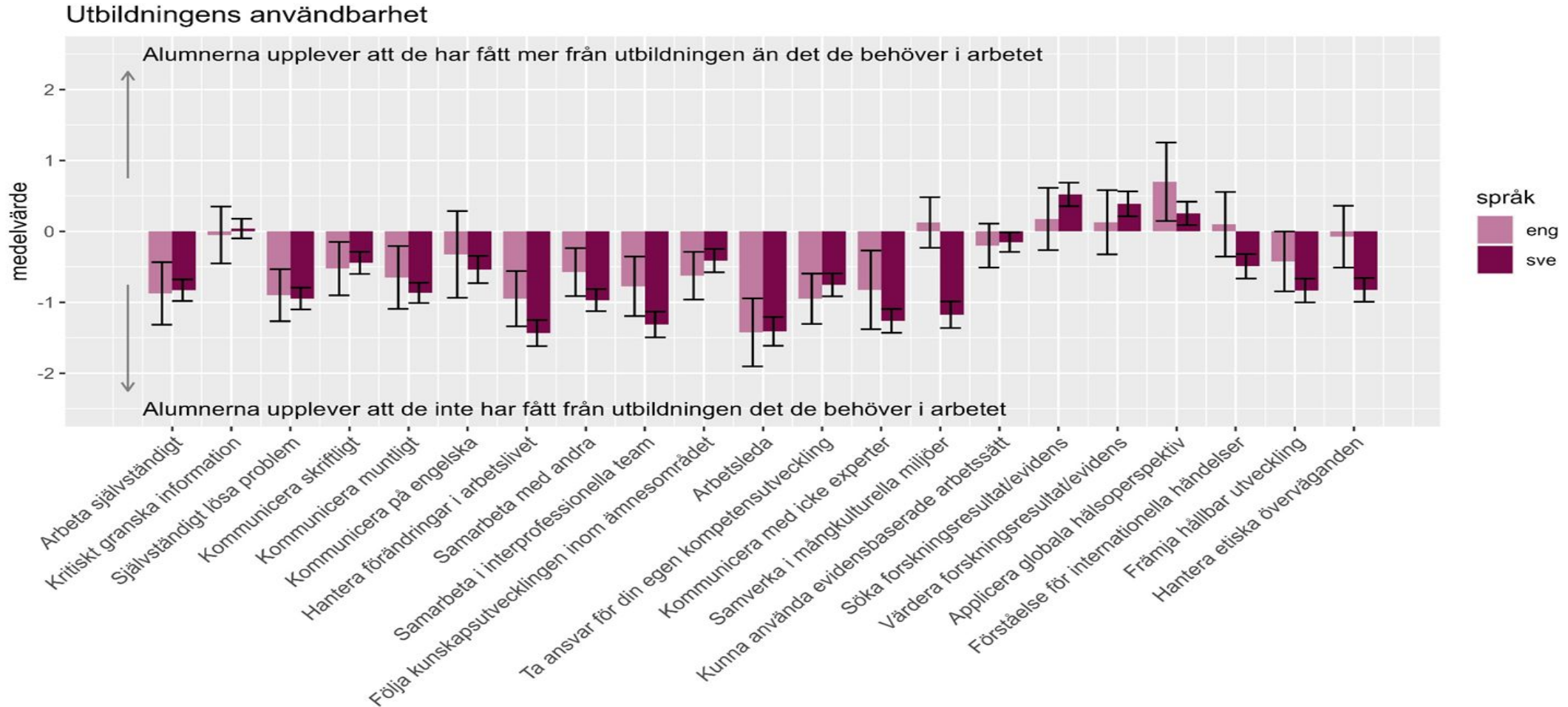
Quotes – Survey to teachers

- I plan to continue to have some parts digital. I think "hybrid" teaching will be very useful in the future. Good to invite teachers without the need to travel to KI (considering CO_2 etc).

- We can now include course lecturers who cannot be on site - they can participate online. Make course scheduling easier and enhances possibilities to get the best speakers!

- Some (lectures) are even better digitally, and we have the opportunity to have lecturers from other universities without traveling.

- I have created short, pre-recorded lectures and provided my students with open online available lectures from world-known experts.

- Tillgång till lärare från andra länder kan bli en viktig kvalitetshöjare. Även internationella alumner som lättare kan nätverka med studenterna.

Common pitfalls in question design

Zoe Säflund

What's wrong with this question?

Question:

„Where was the Declaration of Independence signed?"

Answer...

Question: Where was the Declaration of Independence signed?

Answer: 'At the bottom'.

Validity

Do we measure what we are supposed to measure?

Remember… not measure what we are NOT supposed to measure!

The concept of validity

- A judgement of an estimate of how well a test measures what it purports to measure in a particular context.

- This means that throughout the measurement procedure or process we need to gather evidence about the appropriateness of inferences drawn from the test results.

- A judgement of an instrument's validity shows how useful the instrument is for a particular purpose with a particular population of people.

- What should it mean then that a test is valid?

- Important: Validation process – evaluating evidence about validity.

Threats to validity in question design

Negatively phrased questions

Inappropriate use of language (vague words, jargon etc)

Ambiguous terms, tricky or unnecessarily complicated language

What's wrong?

Question: Would you dare to take medicine developed and quality tested with **AI/computer models**?

Answers:

- ☐ Yes
- ☐ No
- ☐ I don't know

Leading / biased questions

- Not objective questions.
- Phrased or formatted in a way that skew people towards a certain answer.
- They usually include words (subjective adjectives or context-laden words) that frame the question in a positive or negative light.
- It leads to biased results.

What's wrong?

To what extent is your experience as a PhD student problematic?

Responses:
❐ To a very small degree
❐
❐
❐
❐
❐ To a very high degree
❐ I have no opinion

Loaded/ assumptive questions

- Contain an assumption.

- When people answer these questions they iadvertently end up agreeing or disagreeing with an implicit statement.

Ambiguous/ vague language

- Item and option text must be clear and unambiguous. Avoid imprecise phrases such as "is associated with" or "is useful for" or "is important"; words that provide cueing such as "may" or "could be".

- Options should be worded to avoid the use of vague terms (e.g., frequently, occasionally, rarely, usually, commonly) as these terms lack precision and there is seldom agreement on the actual meaning of ''often'' or ''frequently''.

- An item is sometimes complicated by the use of uncommon abbreviations that are not spelled out.

What's wrong?

During the course I developed my ability to critically think and reflect on the facts and knowledge that were presented to me.

- ❒ To a very small degree
- ❒
- ❒
- ❒
- ❒
- ❒ To a very high degree

- ❒ *I have no opinion*

Double-barreled questions
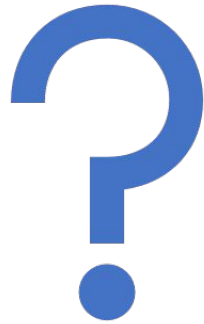
Asking two questions in one.

Proofread your questions for "and" or "or".

Jargon

- Questions that include words or phrases that are difficult to understand or not widely used by the general population.
- Important to test your question with people that are not familiar with the subject.
- eg. "KI" stands for "Karolinska Institut", but not everyone will understand it

What is wrong?

Question: "This would probably not have happened without the program."

Response alternatives:

☐ Yes
☐ No
☐ Do not know

Why avoid Negatively Worded Questions(NWQs)

RISK OF INTRODUCING A 'DOUBLE NEGATIVE'

THE CHALLENGE FOR NON-NATIVE LANGUAGE SPEAKERS IS EVEN GREATER WHEN THE QUESTIONS IS NEGATIVELY FORMULATED

THE COGNITIVE LOAD OF IDEAS EXPRESSED IN NEGATIVE FORM HAS BEEN ESTIMATED AS OCCUPYING 'TWICE AS MUCH SPACE' IN WORKING MEMORY AS THE CORRESPONDING POSITIVE FORM OF THE QUESTION

CONCERN THAT THE NEGATIVE ORIENTATION OF THE QUESTION MAY SIMPLY BE MISSED BY THE RESPONDENTS

Double negatives

A double negative occurs when you use two negatives in the same sentence:

- No/not with "un-" prefix words (also in-, non-, and mis-)

- No/not with negative adverbs (scarcely, barely, or hardly)

- No/not with exceptions (unless + except)

Poor answer scale options

- Answer options are just as important as the questions themselves.

- If they are confusing or unbalanced, the results will be skewed.

- Consider carefully the best way to ask the question (statement or question) and then think about the responce types that are most appropriate for the respondents to provide sincere feedback.

- Proofread for mutually exclusive answer options and remove redundant

- Cover all the likely use cases in multiple choice questions

Mismatched scales

- Sync the answer scale back to the question.
- Use balanced responses (the response options should be spaced-out equally)
- Be consistent with response formatting (be consistent about which end of the scale is positive).

Checklist of dos and dont's

- Avoid leading questions
- Don't make assumptions (avoid loaded questions)
- Avoid double-barreled questions
- Avoid jargon
- Avoid double-negatives
- Sync your answer responses to the question
- Format your questionnaire for all devices
- Think of the data processing methods beforehand

Likert Scales

- A rating scale that measures how people feel about something.

- It often comes with a neutral midpoint and essentially measures the level of agreement.

- They often asked as statements.

- They measure agreement, frequency, quality, importance and likelihood.

Likert Responses

Strenghts

They offer variation in responses than just a simple yes/no answer.

When anonymous they reduce the risk for social desirability bias.

Limitations

The validity of the Likert scale attityde measurement can be compromised due to social desirability.

Even vs. uneven response alternatives

- Likert-responses can be either even (4 or 6 response alternatives) or uneven (5 or 7 response alternatives).

- Each of them have pros and cons.

- Even scales can give wider response variation when the respondents are forced to be positive or negative.

- But if they don't have an opinion about the question they will feel that the response alternatives are limited.

- Choose even scales when you expect your respondents to have an opinion about the subject.

Survey limitations

- The survey is only as good as your questions
- Often "unable" to answer the "why of behavior, only the "what"

More issues to consider when designing a questionnaire

- Assuming prior knowledge or understanding (Please don't)
- Consider what demographics are necessary for your project and think carefully when in the questionnaire they should be included.
- Inadequate response options
- Grammatical inconsistencies
- Too many open-ended questions – Do not include a lot of "please explain" options after each question. Choose wisely when to use open-ended questions.
- Asking too much of respondents (Too long, too complex and too confusing should be avoided)
- Unnecessary questions

Questions to ask yourself when designing a questionnaire:

- What do you want to know?

- Who do you want to ask?

- How will you use the responses you collect?

Thank you!