

Mixed method approaches: design of authentic assessment tasks

Carla van Boxtel
Anne Hemker
Thomas Klijnstra
Gerard Ruijs

University of Amsterdam

Geschichtsdidaktik-Tagung Eichstätt
17 november 2017



Study

Practice oriented

16 months

Team

- 8 social science teachers
- 2 teacher educators
- 2 researchers

Assessment of causal reasoning in social science

New social science curriculum in upper secondary education:

- ❖ *key concepts*, e.g. socialisation, social cohesion
- ❖ *thinking skills*, e.g. causal reasoning

Central examination: short answer questions

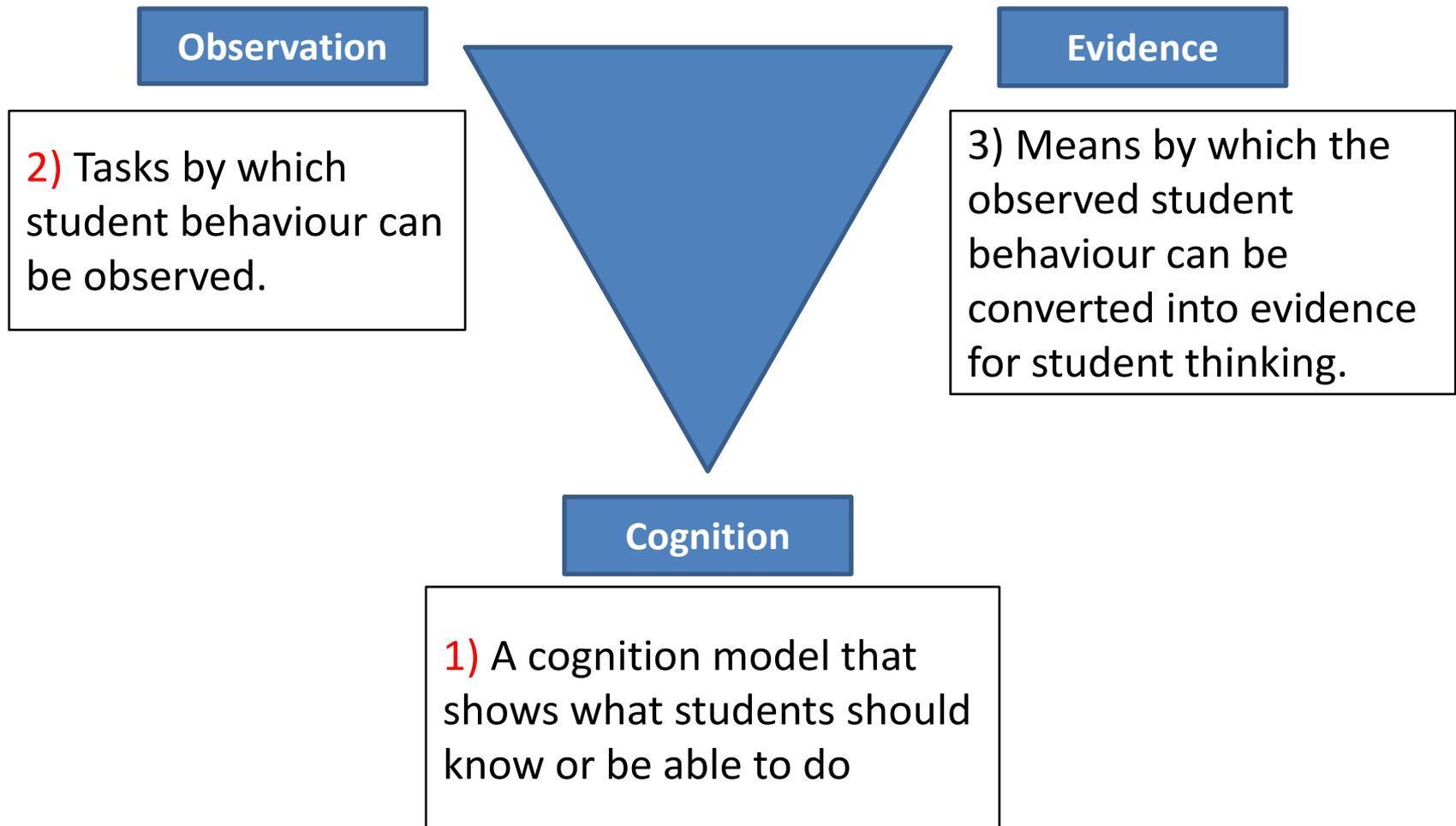
School examination: also **authentic assessment tasks**

Research questions

1. How can we operationalize 'causal reasoning with main and key social science concepts' in terms of concrete student behaviour?
2. Which design principles can be used to design authentic assessment tasks to assess this skill?
3. To what extent are tasks that are developed according these design principles suitable to assess this skill?

The assessment triangle

(Pellegrino *et al*, 2001)



Mixed method

Qualitative

- Literature study
- **Expert** consultation
- Interviews with **students**
- Evaluations from **teachers**

Quantitative

- Questionnaire for students
- Content analysis of tasks and **student products**

valid
operationa-
lization of the
skill?

Suitable
tasks?

Cognitive model

- Version 1 after consultation of the literature
- Version 2 after expert meeting
 - Teacher educators, researchers (N = 7)
 - Clarity, validity, usability
- Version 3 after implementation and evaluation of authentic assessment tasks
- Version 4 after second expert meeting (N = 7)

COGNITIEF MODEL VAN DE VAARDIGHEID 'VERBANDEN LEGGEN BIJ MAATSCHAPPIJWETENSCHAPPEN'

Wat leerlingen moeten kunnen *doen*

De kennis en inzichten die leerlingen hierbij moeten kunnen toepassen

1. LEERLINGEN KUNNEN OORZAKEN, GEVOLGEN EN/OF OPLOSSINGEN VAN EEN MAATSCHAPPELIJK VERSCHIJNSEL/PROBLEEM BESCHRIJVEN:

- het verschijnsel of probleem definiëren
 - een onderzoeksvraag en/of hypothese formuleren
 - mogelijke oorzaken, gevolgen en oplossingen beschrijven
 - relaties tussen problemen/verschijnselen, oorzaken, gevolgen en oplossingen specificeren
 - hierbij gebruikmaken van maatschappijwetenschappelijke methodologische en inhoudelijke begrippen, theorieën, paradigma's en/of politieke ideologieën
- kennis van hoe je een probleem of verschijnsel beschrijft: afbakening naar onderwerp, actoren, tijd en plaats
 - kennis van het onderscheid tussen een onderzoeksvraag en een hypothese, operationaliseren en het verschil tussen correlatie en een causaal verband
 - het inzicht dat er doorgaans meerdere oorzaken, gevolgen en oplossingen zijn
 - kennis van soorten oorzaken, gevolgen en oplossingen die je kunt onderscheiden
 - kennis van hoe je een verband beschrijft: richting en sterkte, onderscheid tussen samenhang (correlatie) en causaal verband, mogelijke wisselwerking
 - kennis van hoofd- en kernconcepten, theorieën, paradigma's, ideologieën
 - kennis van de methodologische termen onafhankelijke, afhankelijke en interveniërende variabelen, indicatoren, correlatie, causaal verband, wisselwerking

2. LEERLINGEN KUNNEN UITSPRAKEN OVER OORZAKEN, GEVOLGEN EN/OF OPLOSSINGEN MET GEGEVENS ONDERBOUWEN:

- relevante informatie(bronnen) zoeken / selecteren
 - passende voorbeelden geven / illustreren met kwalitatieve gegevens
 - passende statistische / kwantitatieve gegevens gebruiken
 - vergelijkingen maken tussen effecten van variaties in een variabele, tijdstippen, plaatsen, situaties, groepen of actoren
 - een mechanisme uitleggen m.b.v. een concept of theorie
 - rekening houden met de validiteit, betrouwbaarheid, representativiteit en/of generaliseerbaarheid van gegevens
- kennis van bruikbare bronnen voor informatie
 - kennis van specifieke contexten die als voorbeeld kunnen dienen en van kwalitatief onderzoek: bv. interviews, observaties
 - kennis van kwantitatief onderzoek: bv. enquête, experiment, statistisch bewijs
 - kennis van hoe je effecten van variaties in variabelen, plaatsen, tijden of groepen/actoren kunt vergelijken om oorzaken, gevolgen of oplossingen te identificeren of uit te sluiten
 - kennis van de methodologische term mechanisme, hoofd/kernconcepten, theorieën
 - kennis van hoe je betrouwbaarheid, validiteit, representativiteit en generaliseerbaarheid van gegevens kunt beoordelen

3. LEERLINGEN KUNNEN TOT EEN BEARGUMENTEERD(E) STANDPUNT / CONCLUSIE KOMEN OVER OORZAKEN, GEVOLGEN EN/OF OPLOSSINGEN:

- een conclusie / standpunt formuleren
 - deze onderbouwen met argumenten
 - mogelijke tegenargumenten benoemen en/of ontkrachten
 - uitspraken nuanceren
 - reflecteren op (eigen) vooronderstellingen
- kennis van hoe je door vergelijken en waarden van mogelijke oorzaken, gevolgen, doelen en/of oplossingen tot een beargumenteerde conclusie/standpunt komt
 - kennis van hoe je reflecteert op de kwaliteit van bewijs, verbanden met gepaste nuance beschrijft (zien als mogelijk/waarschijnlijk, aandacht voor de sterkte)
 - het inzicht dat wat je vooraf voor waar aannam misschien niet klopt

An assessment task that is meant to assess to what extent students are able to reason causally in social science:

1. concerns a social problem or phenomenon that serves as a context for making connections while using key concepts of social science.

2. must be enough authentic. Which means:

- The assessment task has an open question or problem;
- The assessment task is a realistic task (a task that fits participation as a citizen or a professional practice in which people analyze social issues);
- There is more time for the task compared with a regular – paper and pencil - test;
- The task demands application of complex skills, such as analyzing, argumentation
- The assessment criteria are provided. (Frey, Schmitt & Allen, 2012; Kuhlemeier, 2013)

3. provides students with clear information about what they are supposed to do concerning the problem or phenomenon: investigating, providing an advise or providing a substantiated opinion.

4. elicits behaviour that is described in the cognitive model of causal reasoning in social science.

5. corresponds with the level that can be expected of the students. When skills are not so much practiced yet, the task must provide more structure and support.

Factors that affect complexity

Campbell, 1988; Van Merriënboer & Kirschner, 2012

- Substantive and methodological concepts: *number, complexity, provided or not, more or less practiced with*
- Context: *more or less familiar*
- Sources: *number, complexity, provided or not*
- Structure: *subtasks that guide students*
- Support: *hints how to do certain things, instruction by the teacher*

Authentic assessment tasks

- Investigate as a member of the Junior Scientific Council for Government Policy – To what extent are the conclusions in the ‘Spirit Level’ valid?
- Write an advise to the mayor of your city – Which form of shelter for refugees is the best?
- Participate in a round table about social cohesion – Give your opinion reasoning from a particular paradigm

Evaluation: student questionnaire

8 tasks, students from 15 classes from 8 schools

Congruency – 7 items, 5 point Likert scale (alpha .76; N = 456)

This task fits well with what we learn in the lessons.

With this task, I was able to show that I could apply the key concepts.

Authenticity – 7 items (alpha = .85; N = 422)

In this task I had to apply my knowledge to a real social issue.

What I learned from this task is also relevant outside school.

Transparency – 6 items (alpha = .81; N = 454)

It was clear what I had to know and be able to do for this task.

I knew in advance what would be assessed in this task.

Difficulty – 3 items (alpha = .85; N = 460)

I thought this task was doable.

I found this a difficult task.

Evaluation: interviews with students

6 tasks: 9 students each; 2 tasks: 6 students each

- **Authenticity**
 - Did you like the task?
 - Do you think the task is relevant?
- **Transparency**
 - Can you tell in your own words the purpose of the task?
- **Congruency**
 - You had to identify causes and consequences. Do you recognize this? How did you do that?
- **Difficulty**
 - Was the task easy or difficult?
 - What would you suggest to improve the task?

Results student perceptions

- Authenticity
 - Range for 8 tasks: $m = 2.47$ to 3.90
- Transparency
 - Range for 8 tasks: $m = 2.31$ to 3.89
- Congruency
 - Range for 8 tasks: $m = 3.11$ to 3.77
- Difficulty
 - Range for 8 tasks: $m = 2.52$ to 3.69

Results student perceptions *interviews*

Students considered the task relevant / meaningful, because ...

- “it is about a present-day issue”
- “more knowledge changes your opinion” (values)
- “you develop research skills, which are useful outside school”
- “the task connects theory to practice”
- “it makes you think”
- “you are better informed when you are going to vote”
- ...

Results content analysis

- Coding of the 8 tasks: which components of the cognitive model are asked for in the task?
- Coding of the student products (N = 388): which components of the cognitive model can be observed (and assessed)?
 - Percentage of agreement (N = 77 products): 79% to 100%

Qualitative analysis of students' reasoning in the products

- To define levels of causal reasoning
- To construct better assessment instruments (rubrics)
 - Examples of naive reasoning versus (more) social scientific reasoning

Conclusions

- The tasks were considered meaningful, doable (not too simple or difficult) and congruent with the aims and lessons of social science
- Several tasks were insufficiently clear for students
- The tasks elicited the identification of causes, effects, solutions, specification of relations and the application of main and key concepts
 - But (even when explicitly asked for) much less: application of theories, substantiation with quantitative data, comparing, assessment of reliability/validity

Discussion

- Combination of qualitative and quantitative data gave us a fuller picture of the validity and usability of the tasks
- Qualitative data can provide a different view on the results of the questionnaire (e.g. perception of authenticity)
- We didn't define any levels of reasoning / knowledge yet – more qualitative (and quantitative) research is needed
- When we didn't find the behaviour we wanted to assess, was it because of the task or because students hardly developed the skill?