

Unpacking decision making in comparative judgement: A stimulated think-aloud methodology to gain insight into young peoples' decision making

Eva Hartell^{1*}, Jeffrey Buckley^{1,2}

¹Department of Learning, KTH Royal Institute of Technology, Stockholm, Sweden

²Faculty of Engineering and Informatics, Technological University of the Shannon: Midlands Midwest, Athlone, Ireland

*** Correspondence:**

Dr Eva Hartell
ehartell@kth.se

Keywords: Comparative judgement, assessment, stimulated recall, think aloud protocol, formative assessment, stem education, technology education

While reliability has been the subject of much Comparative Judgement (CJ) research, understanding its validity, which and relates directly to the included judges and their decision-making, is paramount. Understanding this decision-making better would add significantly to the formative use of CJ and its use in educational task-design.

This paper reports on a pilot study exploring a novel methodology aiming to unpack judges' decision-making. One 11-year-old student completed a CJ session on a selection of portfolios developed in response to an authentic design-task in STEM education. During this, a novel "stimulated think aloud protocol" was implemented, which was developed by synthesising aspects of traditional think-aloud-protocols with stimulated recall interviews.

The approach is considered to have worked well as it was immediately evident that prompts were required to keep the participant on task and to continue verbalising their thoughts. As in this case the participant was younger, giving support in what to verbalise appeared necessary. The approach was possibly more useful due to the age of the participant. Limitations exist in that the stimulating prompts could influence participant decision-making if they provoke reflection which otherwise would not have occurred in an uninterrupted CJ judging session.

Introduction

Many studies have shown that comparative judgement (CJ) can be used to generate highly reliable rank orders of student work generated in response to open ended tasks. There is now a need to move beyond studies of reliability and to instead focus on validity. The validity of CJ relates directly to the included judges and their decision making (Hartell & Buckley, 2020).

Software solutions for CJ typically include comment features which allow judges to note their thoughts relating to pieces of work and to explain why they made their decision. This data can be quantified in attempts to gain insight into influences on judges decisions (e.g. Buckley et al., 2020), however this approach is limited in that it is retrospective. An alternative approach which has been trialled is the use of a think-aloud protocol analysis to capture judges thinking during CJ sessions (Hartell & Skogh, 2015). The underlying theory of think-aloud protocols is that the elicited thoughts are a “valid reflection of at least a subset of the thoughts involved in the mediation of the task being performed” (Eccles & Arsal, 2017, p. 514). In think-aloud protocols real-time thought is captured which is of significant use however there is a potential limitation of lack of depth. Where methodological implementation is strict prompts should only be used to instruct participants to verbalise their thoughts (Desoete, 2008).

With respect to limitations of think aloud protocols, it has been observed that students may not know what kinds of thoughts to articulate and that in cognitively demanding tasks it can be difficult for them to devote attention to the continuous verbalisation of thoughts (Cotton & Gresty, 2006). There appears a need for a methodology which involves the augmentation of the traditional think aloud protocol analysis with supplementary concurrent stimulating prompts to gather deeper insight, support participants in staying on task, and offer guidance on the nature of thoughts to verbalise. Under the assumption that CJ activity is cognitively demanding on judges and that verbalisation may remain at a descriptive level of portfolios rather than of the thought processes underpinning decision making, this study explored a novel methodological adaption to the think aloud protocol, a “stimulated think aloud protocol”.

Method

A group of middle school students engaged with an online CJ session to assess responses to an authentic STEM education open-ended design task. All but one student wrote their thoughts on why they made specific decisions using the CJ comment system. To gain deeper

insights in terms of the reasoning behind the decision-making, one student was asked to think out loud while undertaking the CJ session. Specifically, the efficacy of a stimulated think aloud protocol was explored with this student. Due to the Covid-19 situation observations had to be made entirely online which afforded the possibility of closely following the process in real time.

The student was initially instructed to verbalise their thoughts continuously during the CJ session, akin to a traditional think aloud methodology. In response to the student and the situation, an observer, who was an expert researcher in technology education assessment and in the use of CJ, offered prompts if the student fell silent or if their thoughts deviated too far from the research question at hand, i.e., the justification of why a decision was being made during the CJ session.

Researcher insight and reflection

Based on this experience, it appears that the combination of the think aloud protocol methodology with stimulated recall prompts has the potential to support understanding reasoning behind judges' decision-making during CJ sessions. As in this study the participant was younger, it became immediately evident that prompts were required to keep them on task, to support them in continually verbalising their thoughts, and most importantly to move them beyond offering only surface level descriptions of portfolios. Further, the insight gained into the thoughts of this student was deeper and richer in comparison to the written justifications provided by students in the larger group. Coupled with this being an online activity enacted through video-conferencing software, the stimulated think aloud protocol made it possible to follow the students decision-making thought process in real time. This opens the possibility to follow the process more in-depth compared to analysing data retrospectively.

Limitations exist in terms this being a one participant pilot in that the design of the stimulating prompts was reactive and unstructured. It is also important to consider that the protocol could have influenced the participants decision making by provoking reflection which otherwise may not have occurred in an undisrupted CJ session. Still, the insight gained from this approach was found more informative in terms of understanding the reasoning and decision-making of the student compared to the written comments given by the other participants who undertook CJ on the same cohorts of student work.

References

- Buckley, J., Canty, D., & Seery, N. (2020). An exploration into the criteria used in assessing design activities with adaptive comparative judgment in technology education. *Irish Educational Studies*. <https://doi.org/10.1080/03323315.2020.1814838>.
- Cotton, D., & Gresty, K. (2006). Reflecting on the think-aloud method for evaluating e-learning. *British Journal of Educational Technology*, 37(1), 45–54.
<https://doi.org/10.1111/j.1467-8535.2005.00521.x>
- Desoete, A. (2008). Multi-method assessment of metacognitive skills in elementary school children: How you test is what you get. *Metacognition and Learning*, 3(3), 189.
<https://doi.org/10.1007/s11409-008-9026-0>
- Eccles, D. W., & Arsal, G. (2017). The think aloud method: What is it and how do I use it? *Qualitative Research in Sport, Exercise and Health*, 9(4), 514–531.
<https://doi.org/10.1080/2159676X.2017.1331501>
- Hartell, E., & Buckley, J. (2021). Comparative Judgment: An Overview. In A. Marcus-Quinn & T. Hourigan (Eds.), *Handbook for Online Learning Contexts: Digital, Mobile and Open: Policy and Practice* (pp. 289–307). Springer International Publishing.
https://doi.org/10.1007/978-3-030-67349-9_20
- Hartell, E., & Skogh, I.-B. (2015). Criteria for success: A study of primary technology teachers' assessment of digital portfolios. *Australasian Journal of Technology Education*, 2(1), 1–17. <https://doi.org/10.15663/ajte.v2i1.27>