Exploring Science Teacher Educators' Evaluation of a Score Report to Support Content Knowledge for Teaching

Dante Cisterna, Jamie N. Mikeska, Katherine Castellano & Jennifer Lentini
Educational Testing Service

Abstract

We conducted a usability study to explore how elementary science teacher educators: (1) understand and interpret a score report prototype providing information on one class of preservice teachers (PSTs)' content knowledge for teaching (CKT) about matter and its interactions and (2) evaluate its utility for use within elementary science teacher education courses.

Conceptual Framework

- The score report prototype was developed along with a new CKT assessment to measure PST's CKT about matter and its interactions. The score report was designed to support preservice teacher educators in identifying strengths and areas for improvement in PSTs' CKT performance and making formative decisions to help develop PSTs' CKT.
- *Score report characteristics*. When designing score reports research has suggested the importance of defining a purpose, identifying target audience and needs, developing a prototype, and conduct usability testing (Zenisky and Hambleton, 2012). Displays that combine enhanced text and enhanced graphics tend to be more helpful to facilitate users' understanding (Zwick, Zapata-Rivera, & Hegarty, 2014), so they can take actionable steps.

Research Questions

- *Research Question 1*. How does a group of elementary teacher educators understand and interpret information on a score report about PSTs' CKT about matter and its interactions?
- *Research Question 2*. How does a group of elementary teacher educators evaluate the utility of the score report for use within science teacher education courses?

Score Report Sections

- *Read me*. Introduces the score report and explains how to navigate through the sections.
- *Scale scores*. Describes individual (i.e., PSTs) and aggregated (i.e., class) scores on the CKT matter assessment, using various tables and graphs (see Figure 1).
- *Item map*. Ranks a subset of the CKT assessment items by difficulty and summarizes the skills assessed by each item.
- *Resources*. Describes the content categories of the assessment and provides links to instructional materials and assessment tasks.

Methods

We interviewed seven elementary science teacher educators regarding their understanding and perceptions of the score report prototype. Our analysis included computing descriptive statistics for usability categories and examining patterns across responses to the interview questions and summarized them using a thematic analysis (Clarke & Brown, 2014).

Findings: Research Question 1

- Most teacher educators were able to correctly understand the information presented about individual and class performance and use the functions included in the score report, such as sorting and filtering.
- Their understanding of how to interpret score certainty, however, tended to be less accurate (64%) in comparison to the other item categories.
- This finding suggests that several teacher educators, although able to recognize the importance of the standard error of measurement, struggled to explain what it meant and how to use that information to interpret and understand the PSTs' CKT scores, especially when comparing two or more PSTs' scores.

Note. Pretest and posttest scores are provided as the PSTs in this class took the CKT matter assessment at the beginning and at the end of their teacher education course.

Figure 1. Example of Individual Score Table and Class Chart Included in the Scale Scores Section.

Table 1. Summary of Overall Understanding of Score Report (n=7).

Item Category	Example Items	Number of Items	Correct	SD
Interpretation of class averages, ranges, categorization	• What percentage of preservice teachers are in the medium level for the pretest? • How many preservice teachers received a score in the 275 to 285 range on the posttest?	6	98%	0.38
Functionality	• Sort the scores from smallest to largest. • There are some blank cells for one PST. What do these blank cells mean?	5*	89%	0.20
Interpretation of score certainty	• Are the posttest scores for PST 1 and PST 2 about the same? How do you know? • What information gives you a sense of certainty of their scores?	2	64%	0.49
Interpretation of item map	• The class average on the posttest is 208. What information can you use to help you understand the class's average score? • Which CKT matter items are high difficulty? How do you know?	5	91%	0.79

Findings: Research Question 2

Table 2. Emerging themes regarding score report utility for use.

Score Report Theme	Examples
Score report includes relevant information to understand and interpret PSTs' scores and class performance	"…most students showed improvement, the gains are mostly positive…[but] two students have low gains or negative gains. I recognize the class is above the field sample. The score looks good, aggregated at the individual level."
Appreciation of the score report organization and layout -- including multiple representations to represent PSTs' performance	"…the histograms and bar charts provide quick summaries and show how the class moved." "…right amount of information that I can manage it…, what the columns are, maps and diagrams about PSTs' performance. It is good."
Score report needs to better connect the information about PST performance, assessment content categories, and instructional resources	"…the content category is very broad, but the question is pretty specific…, the content is models of matter, but is there more info that could be given to her about what is hard about this question?"
Teacher educators need additional support for taking actionable steps based on information provided in the report	"I am not sure what to do with this…, so what info does this give her about what to prioritize and what not to prioritize?"

Discussion and Implications

Discussion

- Teacher educators were able to understand the information provided in the score report about PSTs' performance at the individual and class levels and appreciated the inclusion of different representations for scores as well as supporting materials for interpretation and instructional support.
- However, teacher educators tended to struggle to determine how to use this information to provide specific help to PSTs and tended to show limited understanding of some information about student and class performance, such as the standard error of measurement.
- Adding guidance about how to connect the item descriptors, the information of the content categories and teacher practices, and the use of the resources provided would be important for improving the usability of the score report (Zapata-Rivera, van Winkle & Zwick, 2012), especially to identify specific actions to be implemented in their science content and methods courses to better support PSTs' CKT development.

Study Implications

- Results of this study can provide insights on how to better support teacher educators' use of assessment results within elementary science teacher education courses.
- We used the results of this study to refine the score report prototype for teacher educators' use. For example, we added a new section in the score report that summarizes PST performance linked to the content categories and teaching practices included in the assessment.

Contact

Dante Cisterna
Associate Research Developer
Educational Testing Service
dcisterna@ets.org

References

1. Clarke, V., & Braun, V. (2014). Thematic analysis. In: *Encyclopedia of critical psychology* (pp. 1947-1952). Springer, New York. https://doi.org/10.1007/978-1-4614-5583-7_311
2. Zapata-Rivera, D., Van Winkle, W., & Zwick, R. (2012). Applying Score Design Principles in the Design of Score Reports for CBAL Teachers. *ETS Research Memorandum*, 12-20.
3. Zenisky, A. L., & Hambleton, R. K. (2012). Developing test score reports that work: The process and best practices for effective communication. *Educational Measurement: Issues and Practice, 31*(2), 21-26. https://doi.org/10.1111/j.1745-3992.2012.00231.x
4. Zwick, R., Zapata-Rivera, D., & Hegarty, M. (2014). Comparing graphical and verbal representations of measurement error in test score reports. *Educational Assessment, 19*(2), 116-138. https://doi.org/10.1080/10627197.2014.903653

For more information, visit
http://cktscience.org/

or scan this QR code