

DEVELOPMENT AND MEASUREMENT OF STATISTICAL KNOWLEDGE FOR TEACHING

Stephanie Casey
Eastern Michigan University
scasey1@emich.edu

Andrew Ross
Eastern Michigan University
aross15@emich.edu

Jeremy Strayer
Middle Tennessee State
University
jeremy.strayer@mtsu.edu

We describe novel teacher education curriculum materials designed to develop secondary PSTs' Statistical Knowledge for Teaching (SKT) along with a new test for measuring teachers' SKT. We report PSTs' changes in SKT from learning with the materials in a preliminary study.

Keywords: Teacher Knowledge; Data Analysis and Statistics

Teachers create meaningful student learning opportunities by drawing on a variety of knowledge domains—encompassing not only knowledge of the content to be learned but also additional knowledge connected to the work of teaching (Hill et al., 2008; Shulman, 1986; Thwaites et al., 2011). Assisting preservice teachers (PSTs) in developing this sort of mathematical and statistical knowledge (termed mathematical knowledge for teaching (MKT) and statistical knowledge for teaching (SKT)) is of particular interest to those who teach mathematics and statistics content courses in secondary mathematics teacher preparation programs because we know that there is an association between teachers' MKT/SKT and high quality instruction (Hill et al., 2007). However, there is dissonance between the preparation PSTs are receiving in these courses and the preparation they need to develop MKT/SKT (Goulding et al., 2003). Additionally, understanding how PSTs develop MKT/SKT and how researchers observe and/or measure MKT/SKT is still an open question (Groth, 2017; Silverman & Thompson, 2008; Thwaites et al., 2011). The MODULE(S²) project works in this space and has created curricular materials that provide opportunities for secondary PSTs to develop MKT and SKT in university algebra, geometry, modeling, and statistics courses. With regard to measures, some tests exist for measuring MKT (e.g., Hill et al., 2007; Mohr-Schroeder et al., 2017), but none exist for measuring secondary (grades 6-12) teachers' SKT. In this brief report, we describe the development of a test we created to assess secondary PSTs' SKT prior to and after learning with MODULE(S²) materials, and report PSTs' changes in SKT from learning with the statistics MODULE(S²) materials in a preliminary study.

Just as mathematics and statistics are related yet distinct fields (Cobb & Moore, 1997), MKT and SKT are related yet distinct domains of knowledge (Groth, 2007, 2013). Both MKT and SKT include subject matter knowledge (SMK: understanding the logical necessity of aspects of the content to be taught) and pedagogical content knowledge (PCK: understanding how to make content understandable to students) (Groth, 2013; Hill et al., 2008; Simon, 2006). However, SKT differs from MKT because it contains nonmathematical elements, such as assessing the suitability of statistical questions for the context of a statistical investigation (Groth, 2007). The *MODULE(S²): Statistics for Secondary Mathematics Teaching* curriculum (Casey et al., 2019) was created to develop PSTs' SKT regarding study design, exploratory data analysis, statistical inference, and association, through enactment of the statistical investigation cycle. The materials align with the content of a modern college introductory statistics course while also preparing PSTs to address the CCSS-M (CCSSI, 2010) statistics standards in their teaching and meeting the recommendations from professional societies (AMTE, 2017; Bargagliotti et al., 2020; Franklin, 2015). Attention is given to developing PSTs' critical statistical literacy (Weiland,

2017) and abilities to implement equitable teaching practices. This is accomplished through a sustained focus on data sets addressing inequities in educational and social contexts, asking PSTs to wrestle with the dissonances that analyses of these data sets illuminate, and prompting PSTs to reflect on how they might teach lessons with their future students to leverage these contexts for learning. Opportunities for PSTs to develop their PCK are integrated throughout the materials, including learning common conceptions of students when learning statistics, responding to student work, constructing lessons, analyzing curriculum standards, and learning how to use statistics to empower action for social change. The materials are organized into three modules that could collectively be taught in a semester course.

SKT Test

We developed a 7-item SKT test along with an accompanying scoring rubric for the MODULE(S²) project. It is designed to assess secondary teachers' knowledge for teaching the statistics standards in the CCSS-M (CCSSI, 2010). It is aligned with Groth's (2013) conceptualization of SKT, comprehensively assessing SMK and PCK. Its design drew upon previous work to develop SKT assessment items (e.g., Casey et al., 2019; Groth, 2014). Items assessing PCK utilized released student work from LOCUS (n.d.) to document student conceptions of statistics. We piloted and revised the instrument and its accompanying rubric for two years. Revisions focused on improving the clarity of the language as well as the reliability of the scoring. Thus, we used multiple methods to work towards test content validity (AERA et al., 2014) for the use of this SKT test for this study.

The test contains 12 selected response prompts (e.g., multiple choice) and 20 open-ended prompts across the 7 items. The maximum number of points possible on the test is 45. Three of the items, worth a total of 8 points, address SMK. An additional 3 items, totaling 29 points, address PCK. Finally, a single item addresses both SMK (parts i and iii, worth 3 points each) and PCK (part ii, worth 2 points). This item is shown in Figure 1 as a sample.

Imagine you were teaching a high school (non-AP) class and presented the following task to your students:

A simple random sample of 100 high school seniors was selected from a large school district. The gender of each student was recorded, and each student was asked the following questions: 1. Have you ever had a part-time job? 2. If you answered yes to the previous question, was your part-time job in the summer only? Their responses are shown in the table to the right.

Construct a graphical display that represents the association between gender and job experience for the students in the sample.

As your students work on the task, you notice a student named Juan has created this graph (shown at the right).

- Comment on the strengths and weaknesses of this graph.
- How would you respond to Juan in order to advance his thinking concerning graphical displays to show association between categorical variables?
- Complete the task yourself, creating the best graphical representation for representing the association between gender and job experience.

Job Experience	Gender		Total
	Male	Female	
Never had a part-time job	21	31	52
Had a part-time job during summer only	15	13	28
Had a part-time job but not only during summer	12	8	20
Total	48	52	100

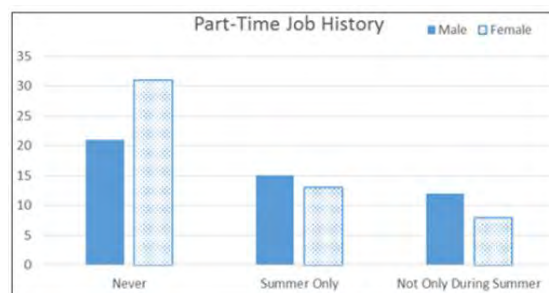


Figure 1: Sample item; parts (i) and (iii) focus on SMK while (ii) focuses on PCK

The scoring rubric for the item in Figure 1 delineates how the points should be awarded for this item. To earn 3 points for part (i), the response needs to include at least one strength and the following two notable weaknesses: (1) there is no label on the vertical axis; and (2) it displays frequencies instead of conditional relative frequencies. Relatedly, when creating the best graphical representation in part (iii), a representation received all three points if it was a segmented bar graph showing percentages conditioned by either categorical variable and there were no flaws in the execution of this representation. The rubric tiered down the point values when some elements were missing. For part (ii), a response received 2 points if it included questions or proposed activities for Juan that addressed either of the notable weaknesses from part (i). A response to part (ii) received 1 point if it accurately gave an explanation about what is incorrect or correct about Juan's graph but did not prompt Juan to think about it himself.

Data Collection and Results

The SKT test was given to PSTs in two statistics-for-teachers classes at separate universities that used the MODULE(S²) materials in the 2020-21 academic year. It was given at the start of the semester as a pre-test and at the end of the semester as a post-test. It was administered as a take-home test without a time limit. The tests were independently scored by 2 graduate assistants using the scoring rubric, and they reconciled any scoring discrepancies for each item. The sample sizes were 11 on the pre-test and 11 on the post-test; we present results here for the 10 PSTs who submitted both tests.

Figure 2a shows the pre- and post-test scores for PSTs' overall SKT, with line segments linking each PST's pre- and post-test scores.

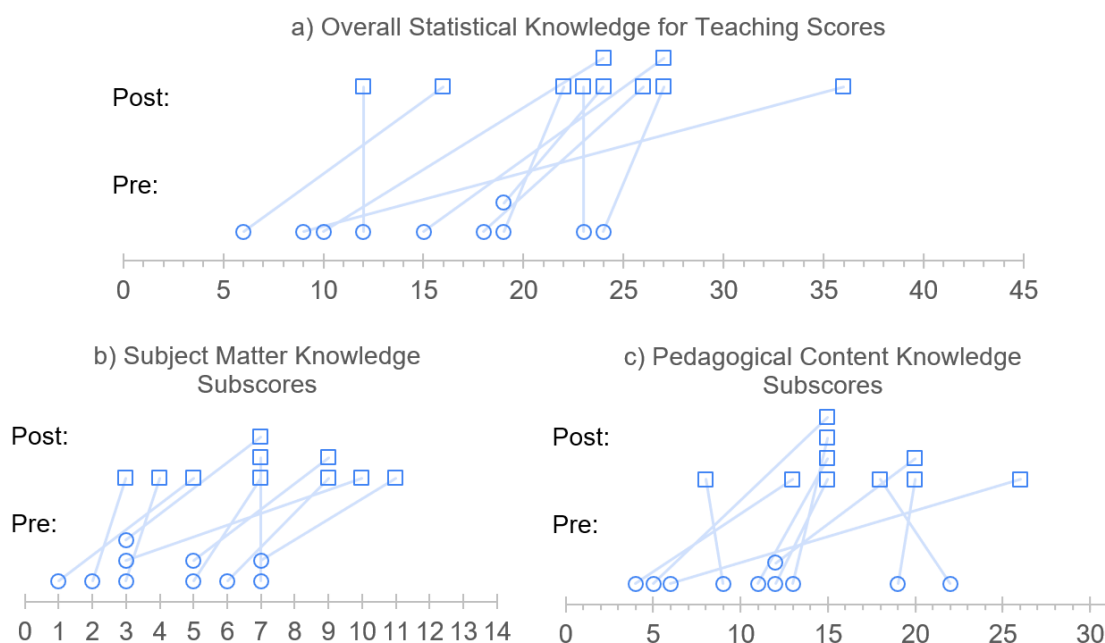


Figure 2: Paired dotplots of pre-test (circles) and post-test (squares) scores.
(a) Overall scores, (b) SMK subscores, (c) PCK subscores

We observe a general pattern of increases from pre to post; the mean pre-test and post-test scores were 15.5/45 (34.4%) and 23.7/45 (52.7%), for an average increase of 18.22 percentage points (standard deviation 18.20). This results in a strong effect size, measured as 0.92 by Cohen's d

with Hedge's correction. Similar graphs are provided in Figure 2 for SMK (2b) and PCK (2c). These also show that PSTs generally improved from pre- to post-test in both of these areas, some quite substantially. The mean SMK pre-test and post-test subscores were 4.2/14 (30.0%) and 7.2/14 (51.4%), a mean increase of 21.4. The mean PCK pre-test and post-test subscores were 11.3/31 (36.5%) and 16.5/31 (53.2%), a mean increase of 16.8. We concluded that PSTs made similar gains on the SMK and PCK portions of the test. An item-by-item analysis revealed that, on average, the PSTs considerably improved their performance on all items from pre- to post-test and that their mean change was roughly the same for SMK- and PCK-related items overall.

Discussion

The MODULE(S²) project works to resolve the conflict between the preparation PSTs have historically received in their university content courses and the preparation they need to develop their MKT/SKT. The statistics teacher education curriculum materials made by the MODULE(S²) project were designed to address this by attending to the development of PSTs' SKT, including SMK and PCK development simultaneously. This preliminary study has provided initial evidence that harmony in developing these aspects of SKT concurrently is possible and productive. Use of the MODULE(S²) project's materials considerably enhanced PSTs' SKT, including both their SMK and PCK. Additionally, we believe that strong SKT supports teachers to effectively implement data-driven lessons focused on equity. The design of the MODULE(S²) project's materials to improve PSTs' critical statistical literacy along with evidence from this study that PSTs' SKT improved through their use of the MODULE(S²) project's materials is promising for the development of teachers who can teach equity-focused statistical lessons. Further research on a larger scale to study the effects of the MODULE(S²) project's materials on SKT and critical statistical literacy is needed to move to more substantial findings. In addition, further research that explores how PSTs' SKT evolves through use of the MODULE(S²) project's materials—including features of the materials' implementation that support or inhibit PSTs' SKT development—is recommended.

An additional contribution of this preliminary study is the development of a SKT test for secondary teachers, addressing a longstanding lack of teacher instruments in statistics education (Groth & Meletious-Mavrotheris, 2018). Future projects can conduct large-scale validation studies of the test to work towards establishing more validity evidence for its use in measuring teachers' SKT.

Acknowledgements

The work reported here was supported by the National Science Foundation (Awards 1726252 and 1726723). The content of this manuscript does not necessarily reflect the views of the National Science Foundation.

References

- AERA, APA, & NCME. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Association. <https://www.testingstandards.net/open-access-files.html>
- Association of Mathematics Teacher Educators (AMTE). (2017). *Standards for Preparing Teachers of Mathematics*. Retrieved from: amte.net/standards.
- Bargagliotti, A., Franklin, C., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. (2020). *Pre-K-12 Guidelines for Assessment and Instruction in Statistics Education (GAISE) report II*. American Statistical Association and National Council of Teachers of Mathematics.
- Casey, S.A., Albert, J., & Ross, A. (2018). Developing knowledge for teaching graphing of bivariate categorical data, *Journal of Statistics Education*, 26(3), 197-213. <https://doi.org/10.1080/10691898.2018.1540915>

- Casey, S., Ross, A., Maddox, S., & Wilson, M. (2019). *MODULE(S2): Statistics for Secondary Mathematics Teaching*. Mathematics of doing, understanding, learning, and educating for secondary schools. APLU. <http://modules2.com>
- Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, 104(9), 801–823. <https://doi.org/10.2307/2975286>
- Common Core State Standards Initiative (CCSSI). (2010). *Common core state standards for mathematics*. Retrieved from <http://www.corestandards.org/the-standards/mathematics>
- Franklin, C.A., Bargagliotti, A.E., Case, C.A., Kader, G.D., Schaeffer, R.L., & Spangler, D. A. (2015). *Statistical Education of Teachers*. American Statistical Association. <https://www.amstat.org/asa/files/pdfs/EDU-SET.pdf>
- Goulding, M., Hatch, G., & Rodd, M. (2003). Undergraduate mathematics experience: its significance in secondary mathematics teacher preparation. *Journal of Mathematics Teacher Education*, 6(4), 361–393.
- Groth, R. E. (2017). Developing statistical knowledge for teaching during design-based research. *Statistics Education Research Journal*, 16(2), 376.
- Groth, R. E. (2014). Using work samples from the National Assessment of Educational Progress (NAEP) to design tasks that assess statistical knowledge for teaching. *Journal of Statistics Education*, 22(3), Available: <http://www.amstat.org/publications/jse/v22n3/groth.pdf>
- Groth, R. E. (2013). Characterizing key developmental understandings and pedagogically powerful ideas within a statistical knowledge for teaching framework. *Mathematical Thinking and Learning*, 15(2), 121–145. <https://doi.org/10.1080/10986065.2013.770718>
- Groth, R. E. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for Research in Mathematics Education*, 38(5), 427–437. <https://doi.org/10.2307/30034960>
- Groth, R. E. & Meletious-Mavrotheris, M. (2018). Research on statistics teachers' cognitive and affective characteristics. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 327–355). Springer.
- Hill, H. C., Ball, D. L., Blunk, M., Goffney, I. M., & Rowan, B. (2007). Validating the ecological assumption: The relationship of measure scores to classroom teaching and student learning. *Measurement: Interdisciplinary Research and Perspectives*, 5(2–3), 107–118. <https://doi.org/10.1080/15366360701487138>
- Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26, 430–511.
- LOCUS (n.d.) *Levels of Conceptual Understanding in Statistics*. <https://locus.statisticseducation.org/>
- Mohr-Schroeder, M., Ronau, R. N., Peters, S., Lee, C. W., & Bush, W. S. (2017). Predicting student achievement using measures of teachers' knowledge for teaching geometry. *Journal for Research in Mathematics Education*, 48(5), 520–566.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4–14.
- Silverman, J., & Thompson, P. W. (2008). Toward a framework for the development of mathematical knowledge for teaching. *Journal of Mathematics Teacher Education*, 11, 499–511. <https://doi.org/10.1007/s10857-008-9089-5>
- Simon, M. A. (2006). Key developmental understandings in mathematics: A direction for investigating and establishing learning goals. *Mathematical Thinking and Learning*, 8(4), 359–371. https://doi.org/10.1207/s15327833mtl0804_1
- Thwaites, A., Jared, L., & Rowland, T. (2011). Analysing secondary mathematics teaching with the knowledge quartet. *Research in Mathematics Education*, 13(2), 227–228. <https://doi.org/10.1080/14794802.2011.585834>
- Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics*, 96(1), 33–47.