Journal of Educational Research and Innovation

Volume 8
Number 1 *Rising to Global Challenges in Education and Research*

Article 4

2020

Accreditation by Design: Construction of an Instrument to Measure Teacher Candidates' Perceptions of Preparedness to Meet InTASC Standards

Michael Floren
Misericordia University, mfloren@misericordia.edu

Chelsie Hess
Colorado Mesa University, chess@coloradomesa.edu

Valerie J.H. Sherman
Colorado Center for Rural Education, Valerie.Sherman@unco.edu

Nancy M. Sileo
University of Northern Colorado, Nancy.Sileo@unco.edu

Journal of Educational Research and Innovation
2020, Vol. 8, No. 1

Accreditation by Design: Construction of an Instrument to Measure Teacher Candidates' Perceptions of Preparedness to Meet InTASC Standards

Michael Floren
Misericorida University

Chelsie A. Hess
Colorado Mesa University

Valerie JH Sherman
University of Northern Colorado

Nancy M. Sileo
University of Northern Colorado

The Council for the Accreditation of Educator Preparation (CAEP) was established in 2013 as an agency responsible for the accreditation of educator preparation programs (EPPs). Within this framework, CAEP requires EPPs to submit "solid evidence" for the competency of their graduates and quality of their programs (CAEP, 2013; Immekus, 2016). In other words, CAEP requires EPPs to provide high-quality learning experiences and to utilize measurements that yield valid and reliable data demonstrating the EPPs ability to prepare high-quality teachers (Immekus, 2016).

High-quality teacher performance in EPPs is best accomplished when there is a set of governing accreditation and performance standards (Schacter & Thum, 2004). These governing standards encourage EPPs to adopt best practices which promote teacher effectiveness (Darling-Hammond, 2020). To this end, the Interstate Teacher Assessment and Support Consortium (InTASC) developed a set of model standards outlining what teachers should know and be able to do in order to improve student outcomes and achievement (CCSSO, 2013). The InTASC standards identify the content knowledge, pedagogical skills, and professional dispositions that teacher candidates need to learn and master in order to advance the learning needs of preschool through 12th grade (P-12) students (CAEP, 2013; CCSSO, 2013). Indeed, CAEP standard I.1 requires EPPs to show "candidates demonstrate an understanding of the 10 InTASC standards…" (CAEP, 2013), and the InTASC standards have been integrated into licensing and accreditation in more than 40 states (Darling-Hammond, 2020).

A myriad of research has investigated teacher candidates' perceptions of the skills, concepts, and dispositions acquired throughout their preparation programs (Hoffman et al., 2005; Pajares, 1992; Wolsey et al., 2013; Zeichner et al., 2008). However, Darling-Hammond (2006) has

observed that collecting P-12 learning data is both difficult and time-intensive for EPPs. As a result, many EPPs must decide between committing considerable time to search for appropriate measures or creating and utilizing internal measures to gather data from teacher candidates. Though this seems straightforward, finding or creating instruments that yield valid and reliable data is a complex task, limiting the quality of data obtained by many EPPs (Darling-Hammond, 2006; Immekus, 2016). Immekus (2016) further describes how the vague nature of accreditation standards creates additional difficulties for programs seeking to implement high quality measures and determine their adequacy. For this reason, recent calls have been made to provide funding to create robust measurement tools in teacher education (Richmond, Salazar, & Jones, 2019).

Previous research has suggested ways that programs can address CAEP and InTASC standards through a variety of measures (Heafner, McIntyre, & Spooner, 2014; Wentworth, Erickson, Lawrence, Popham, & Korth, 2009). Though valuable, research in creating instruments to measure program alignment with the InTASC standards use the former set of standards and are focused on direct observation of candidates instead of candidate perceptions of their own experiences (Wentworth et al., 2009). As the updated standards have been better aligned with state and national accreditation standards, instruments based upon these standards will also better reflect that scope and offer a common benchmark for competence (CCSSO, 2013; Darling-Hammond, 2020).

This study proposes the construction and evaluation of an instrument designed to meet these requirements. The instrument, called the InTASC Candidate Self-Perception Instrument (ICSPI), is designed to obtain feedback from candidates on how well their EPP prepared them to meet a variety of elements indicated in the InTASC standards.

Instrument Creation, Validation, and Pilot Reliability

Requisition and Construction of the InTASC Candidate Self-Perception Instrument

Background. The ICSPI was designed to be distributed to teacher candidates at multiple points during the EPP. In a pilot distribution, the ICSPI was used to gather candidate perceptions after completion of a final methods/strategies course and the final field experience (i.e., student teaching). In this pilot, the delivery of the ICSPI at multiple points in the program was used to isolate the effects that the final field experience had on candidate self-perceptions of preparation to meet the InTASC standards. Additionally, the final field experience marked the end of the EPP for the majority of candidates and was an ideal point to assess candidates' perceptions of the culmination of their preparation in the program.

Item creation. Item creation began by forming an instrument Research and Development (R&D) team. Members of the eight-person team had a diverse set of education-related experiences, knowledge, and areas of expertise. These included training and experience in elementary education, secondary education, special education, school psychology, educational psychology, research methodology, and statistics. The diverse backgrounds of the R&D team provided a wide range of perspectives which were utilized throughout the instrument creation and validation processes.

Due to the breadth of information covered in the InTASC standards and the goal of creating an instrument that teacher candidates would be willing to complete, all indicators of the standards were not included in the instrument. However, each of the indicators contributed important and meaningful information to each of the ten InTASC standards. Therefore, as covering the broad scope of each standard was essential to the development of a valid instrument, it was determined that all indicators (*n* = 216) would be evaluated for inclusion in the instrument.

The R&D team met on multiple occasions to translate each indicator into potential items. When possible, indicators were simply restated in question form to maintain the direct relationship between the survey items and the InTASC indicators. In many cases, however, this process created double-barreled items, which are difficult for participants to interpret (Groves et al., 2009). In these situations, the R&D team created items for each unique component of the indicator. After this step, the R&D team discussed each item to ensure there was direct alignment between the item, the intent of the indicator, and the corresponding InTASC standard. Due to the changes in meaning that minor differences in wording can cause, a special effort was made to preserve the wording of the indicator in each of the items created (Goodman, Iervolino, Collishaw, Pickles, & Maughan, 2007; Thorndike & Thorndike-Christ, 2010). For an example of the process, consider the indicator 2(n): "The teacher makes learners feel valued and helps them learn to value each other." This indicator includes two facets: the teacher valuing learners, and teaching learners to value each other. These facets are related, but address two potentially independent candidate dispositions. Thus, this indicator was considered as two separate items: "The teacher makes learners feel valued" and "The teacher helps learners to value each other." Note that indicators that included lists or examples to clarify the intent of the indicator, such as indicator 5(b) "The teacher engages learners in applying content knowledge to real world problems through the lens of interdisciplinary themes (e.g., financial literacy, environmental literacy)" were not considered as separate items (Groves et al., 2009). To promote transparency in the relationship between potential survey items and indicators, all potential items listed the indicator source. Items were constructed to be combined into an instrument that would be distributed to and completed by teacher candidates. Therefore, items were not considered if the R&D team felt the items were not applicable and/or failed to relate to the typical range of experiences, types of knowledge, and skills normally gained by teacher candidates while enrolled in an EPP. For instance, the R&D team agreed to exclude potential items related to InTASC Standard 10(k) which states "The teacher takes on leadership roles at the school, district, state, and/or national level" from the consideration because the team believed it would be unreasonable to expect teacher candidates to have a leadership role of such magnitude before or during their final field experience (CCSSO, 2013).

Item consolidation. For each InTASC standard, the team reviewed the contributions of all potential items to the intent and purpose of each standard. After all items were reviewed, each team member selected the minimum number of items that they believed best captured the purpose and breadth of the standard. These

selections were then compared with potential items identified by other team members. After extensive discussion, decisions were made to retain items based on each item's unique contribution to the instrument. To promote content validity, the R&D team ensured the breadth of each standard was covered by including all items that were viewed as having a necessary contribution by one or more team member.

In addition to identifying item alignment with InTASC standards, cross-standard themes were also considered when determining which items to include in the ICSPI (CCSSO, 2013). For example, the themes "technology" and "cultural competence" are present in indicators from multiple standards. After producing the consolidated list of items, the R&D team re-examined the contribution of each item with respect to its associated standard and the cross-standard themes.

After the item consolidation process, all items were formatted to fit the common stem "My program prepared me to…" to enhance comprehension and shorten reading for participants. For the vast majority of items, reformatting to accommodate the uniform question stem was not required. For items that required reformatting, every effort was made to maintain the original wording of the indicator. Extending the example from indicator 2(n), the re-formatted items read "My program prepared me to make learners feel valued" and "My program prepared me to help learners to value each other."

After the items were fit to a common stem, the instrument was formatted to facilitate online distribution. A five-point Likert scale was used to measure the candidates' responses. In an effort to increase the overall readability of the instrument, small groups of items (i.e., items from two standards at a time) were used so that they would be displayed on a single page. The "My program prepared me to…" question stems were placed near the top of each page in large font and bold letters. The team used alternating background colors for each item so that candidates could easily match the response options with their corresponding question. To minimize missing data, the electronic distribution system delivered a pop-up message to candidates who attempted to advance to the next set of questions prior to providing an answer to all items. This message did not prevent candidates' progress, but notified the candidates that they would be continuing before answering all questions on the page.

Review of the ICSPI

Prior to distribution of the survey, several additional validation procedures were conducted to ensure adequate coverage, wording, and formatting. To this end, a panel of experts was convened to evaluate the procedures used by the R&D team during the creation process. This panel also reviewed item wording and coverage to ensure the instrument was aligned with the InTASC standards. After the expert panel review, a small pilot distribution was conducted. A focus group of the pilot participants met to provide suggestions and feedback to the R&D team. After feedback was incorporated, a larger pilot distribution was conducted to provide reliability estimates for each subscale. This distribution utilized candidates from a variety of majors and included candidates in either their final methods course or their final field experience.

Expert review. The panel of experts was gathered to review the instrument creation process, procedures, and the

instrument itself. This panel consisted of the dean and assistant dean of a large and robust EPP, a survey methodologist, and a statistician. Though all panel members were familiar with the project prior to the review, none were part of the research team, participated in item creation, or were involved in the formatting process. Several members of the expert panel met multiple times during the instrument creation process to ensure fidelity and that the instrument covered the breadth of each of the InTASC standards. After the initial draft of the instrument was created, it was distributed to all members of the expert panel for feedback related to item wording, formatting, overall structure, and to ensure sufficient content coverage. After review, the suggestions made by the expert panel were discussed by the R&D team and incorporated into the instrument.

Initial pilot and focus group. After the expert review, the R&D team identified a small sample of teacher candidates who were seeking an elementary education teaching license and were enrolled in an EPP course. The R&D team invited these candidates to take a pilot version of the instrument, which was modified to include a "Don't Know" (DK) response option in addition to the original 5-point Likert options. Pilot participants were asked to take the instrument and answer DK to any questions they found confusing or difficult to understand. Of the 28 teacher candidates, 17 completed the 50-item survey and nine agreed to participate in the follow up focus group. Of the 50 items, six items were marked DK by one participant, and one item was marked DK by two participants. Questions with one or more DK responses (seven total) were used as prompts to guide the focus group conversation between participants regarding item wording and intent. The focus group was facilitated by one moderator and two assistant moderators, all of whom were members of the R&D team. The focus group provided a variety of ideas that were used to improve the items marked with DK. For their efforts, these initial pilot participants received extra credit points in their course.

Reliability. Following the initial pilot, the ICSPI was disseminated in three consecutive semesters to all candidates who were completing either a final methods course or a final field experience in that semester (number of respondents listed in Table 1). Reliability estimates were calculated using the "psych" package in R (R Core Team, 2016; Revelle, 2016). Cronbach's alpha was used to estimate reliability of the data from each subscale for a variety of candidate majors and for the methods and field distributions. In addition to traditional estimates of alpha, ordinal alpha reliability estimates were also calculated and provided (Gadermann, Guhn, & Zumbo, 2012; Zumbo, Gadermann, & Zeisser, 2007). As the instrument used 5-point Likert response options for all subscales, the ordinal alpha was viewed as the more appropriate of these two measures of reliability. Table 1 shows that reliabilities estimates based on data from these three semesters are at acceptable levels (i.e., $\alpha > .70$) for all subscales in all majors and for the two distributions in the program, with reliability estimates at desirable levels for basic research (i.e., $\alpha > .80$) for all subscales (Nunnally, 1978).

Instrument description. The final instrument consisted of a total of 48 items. Each item used a 5-point Likert response ranging from Strongly Disagree to Strongly Agree with a neutral category. The main item stem was maintained at the beginning

of each page, though, based on focus group feedback, the size of the font was increased

Discussion

This study has provided details on the creation of the ICSPI and demonstrated that this instrument is an appropriate tool to complement an EPP's evaluation framework, allowing them to evaluate candidates' perceptions of the preparation provided by their programs to meet the InTASC standards. Reliability estimates of pilot data were found to be at acceptable levels for all candidate majors and for candidates at multiple points in their program. Evidence towards validity was established through methods designed to utilize input from individuals with a wide range of backgrounds on the alignment and coverage of the items with the InTASC standards. Additionally, multiple checks of item quality, coverage, and formatting by experts and members of the target population were conducted.

Reliability

Reliabilities estimates were based on data from a pilot distribution disseminated in three consecutive semesters. Both traditional and ordinal reliability estimates, given in Table 1, show adequate reliability for the subscales based on each of the InTASC standards. Additionally, reliability estimates were calculated by candidates' majors and candidates' levels in the program (i.e., whether candidates were in methods or fieldwork). This was to ensure adequate reliability was observed across candidate majors and levels. Though the sample sizes within some groups were small, adequate reliability was observed for all subscales within all programs and candidate levels. These results compare favorably to previous research of instruments based on the InTASC standards, which showed low reliabilities for some subscales and student majors (Wentworth et al., 2009). This discrepancy may be due to the self-perception nature of the ICSPI, as former research utilized field observation.

Validity

The R&D team used a variety of different methods to increase and improve the content and construct validity of the ICSPI. Evidence towards content validity was established using a test blueprint that was derived from the InTASC standards. Considered items were aligned with each standard and with the cross-standard themes. Evidence towards construct validity was established using feedback from the panel of experts during multiple stages of the instrument development process. Experts examined the coverage of the content, the alignment of the items with the InTASC standards, and provided input on clarity of item wording to establish face validity. Additionally, construct validity was further refined by recommendations provided by the focus group related to item wording and comprehensibility. Adequate reliabilities of all standards also provide evidence towards consistent item alignment within each InTASC standard.

Limitations

There are several limitations for the present form of the instrument. First, item selection was geared towards EPPs rather than practicing teachers, limiting the scope of the instrument. Additionally, there is an inherent limitation to item selection processes. Though a relatively large team with diverse backgrounds was used to offset selection bias, a certain amount of bias in item selection may remain. Also, the present instrument is limited to evaluating candidate self-perceptions of preparation throughout their EPP experience.

Therefore, use of this instrument for other purposes, such as to evaluate teaching performance, is not supported by this study. Finally, this study is limited by the scope of the available sample. Inclusion of multiple universities and larger sample sizes within major would allow for a formal study of measurement invariance across these factors. Additionally, larger sample sizes would allow testing of the underlying factor structure across major and university.

Future Directions

It is important to acknowledge that validation is a continuing process (Messick, 1995). Criterion validity of the new measurement will continue to be developed over time as the measurement procedure continues to be examined. Future researchers could pursue a large enough sample size to allow the identification of the underlying factor structure of data produced by the ICSPI (Comrey & Lee, 1992). Additionally, future partnerships between EPPs would facilitate continued instrument validation, further extending the generalizability of results to teacher candidates enrolled in a more diverse sample of EPPs. Finally, larger sample sizes may allow analytic options (e.g., structural equation modeling) to be used to test the fit of the data on an a-priori theoretical model (Kline, 2016).

Continued validation efforts should also incorporate qualitative data from teacher candidates, cooperating/mentor teachers, and university supervisors. Focus groups to collect rich and descriptive information from additional populations of teacher candidates would be a valuable addition to build the usability of the instrument. For example, faculty members, program coordinators, and teacher candidates majoring in content areas of interests could provide alternative suggestions that would be informative and serve to increase the overall reliability and validity of data the instrument generates.

Implications

The present study's findings have direct implications for how EPPs can collect data in support of accreditation efforts and inform programmatic improvements utilizing their teacher candidates' perceptions of preparedness. The ICSPI provides data which can be used to inform EPP practices, refine and revise program requirements, align course outcomes, and assist program coordinators and administrators in identifying the strengths and weaknesses of their licensure programs.

***Michael Floren**, Ph.D., is Assistant Professor of Mathematics: Statistics in the Department of Mathematics at Misericorida University. He can be contacted at mfloren@misericordia.edu.*

***Chelsie A. Hess**, Ph.D., is Assistant Professor of Psychology in the Department of Social and Behavioral Sciences at Colorado Mesa University. She can be contacted at chess@coloradomesa.edu.*

***Valerie JH Sherman**, Ph.D., is Rural Education Coordinator of the Colorado Center for Rural Education at the University of Northern Colorado. She can be contacted at valerie.sherman@unco.edu.*

***Nancy M. Sileo**, Ed.D., is Professor of Special Education: Early Childhood Special Education in the School of Special Education at the University of Northern Colorado. She can be contacted at nancy.sileo@unco.edu.*

References

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis*. New York: Academic Press.

Council for the Accreditation of Educator Preparation [CAEP] (2013). *CAEP accreditation standards.*

Council of Chief State School Officers [CCSSO] (2013, April). *Interstate Teacher Assessment and Support Consortium InTASC Model Core Teaching Standards and Learning Progressions for Teachers 1.0: A resource for ongoing teacher development*. Washington, DC: Author.

Darling-Hammond, L. (2006). Assessing teacher education, the usefulness of multiple measures for assessing program outcomes. *Journal of Teacher Education*, *57*(2), 120-138. https://doi.org/10.1177/0022487105283796

Darling-Hammond, L. (2020). Accountability in teacher education. *Action in Teacher Education, 42*(1), 60-71. https://www.tandfonline.com/doi/full/10.1080/01626620.2019.1704464

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation, 17*(3), 1–13.

Goodman, R., Iervolino, A. C., Collishaw, S., Pickles, A., & Maughan, B. (2007). Seemingly minor changes to a questionnaire can make a big difference to mean scores: A cautionary tale. *Social Psychiatry and Psychiatric Epidemiology*, *42*(4), 322–327. doi:10.1007/s00127-007-0169-0.

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Hoboken, NJ: Wiley & Sons, Inc.

Heafner, T., McIntyre E., & Spooner, M. (2014). The CAEP standards and research on educator preparation programs: Linking clinical partnerships with program impact. *Peabody Journal of Education, 89*(4), 516-532. https://doi.org/10.1080/0161956X.2014.938998

Hoffman, J. V., Roller, C., Maloch, B., Sailors, M., Duffy, G., & Beretvas, S. N. (2005). Teachers' preparation to teach reading and their experiences and practices in the first three years of teaching. *The Elementary School Journal*, *105*(3), 267-287.

Immekus, J. C. (2016). The use of surveys in teacher education programs to meet accreditation standards: Preservice teachers' culturally responsive beliefs and practices. *Research & Practice in Assessment*, *11*, 18–28.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York: Guilford Press.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741. doi: 10.1037/0003-066X.50.9.741

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*(3), 307-332.

R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Revelle, W. (2016) *Psych: Procedures for personality and psychological research*. Evanston, Illinois: Northwestern University. http://CRAN.R-project.org/package=psych Version = 1.6.4.

Richmond, G., Salazar, M.C., & Jones, N. (2019). Assessment and the Future of Teacher Education. *Journal of Teacher Education, 70*(2), 86-89. https://doi.org/10.1177/0022487118824331

Schacter, J., & Thum, Y. M. (2004). Paying for high- and low-quality teaching. *Economics of Education Review, 23*(4), 411-430. https://doi.org/10.1016/j.econedurev.2003.08.002

Thorndike, R. M., & Thorndike-Christ, T. M. (2010). *Measurement and evaluation in psychology and education* (8th ed.). New York: Pearson Education.

Wentworth, N., Erickson, L. B., Lawrence, B., Popham, J. A., & Korth, B. (2009). A paradigm shift toward evidence-based clinical practice: Developing a performance assessment. *Studies in Educational Evaluation*, *35*(1), 16-20. https://doi.org/10.1016/j.stueduc.2009.01.006

Wolsey, T. D., Young, J. R., Scales, R. Q., Scales, W. D., Lenski, S., Yoder, K. K., ... & Dobler, E. (2013). An examination of teacher education in literacy instruction and candidate perceptions of their learned literacy practices. *Action in Teacher Education*, *35*(3), 204-222. https://doi.org/10.1080/01626620.2013.806230

Zeichner, K., Conklin, H. G., Cochran-Smith, M., Feiman-Nemser, S., McIntyre, D. J., & Demers, K. E. (Ed.). (2008). *Handbook of research on teacher education: Enduring questions in changing contexts.* New York, NY: Routledge.

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods, 6*(1), 21–29. doi:10.1107/S0907444909031205

Table 1
Reliability Estimates for ICSPI Subscales

Standard (# of Items)		EPP					Distribution		
		ECE	ELEM	K-12	SEC	SPED	Methods	Field	Overall
Learner Development (5 items)	α_0	0.87	0.83	0.79	0.83	0.86	0.87	0.86	0.88
	α	0.97	0.79	0.77	0.89	0.85	0.90	0.82	0.85
	n	13	104	26	51	63	79	178	257
Learning Differences (6 items)	α_0	0.87	0.84	0.82	0.85	0.87	0.87	0.87	0.89
	α	0.96	0.81	0.83	0.90	0.89	0.87	0.85	0.86
	n	13	104	26	51	63	79	178	257
Learning Environments (4 items)	α_0	0.84	0.88	0.76	0.85	0.85	0.82	0.91	0.90
	α	0.92	0.86	0.81	0.92	0.90	0.84	0.89	0.88
	n	13	102	26	49	59	75	174	249
Content Knowledge (5 items)	α_0	0.88	0.87	0.85	0.88	0.84	0.89	0.90	0.91
	α	0.96	0.86	0.80	0.93	0.90	0.90	0.88	0.89
	n	13	102	26	49	59	75	174	249
Application of Content (4 items)	α_0	0.71	0.84	0.72	0.83	0.83	0.80	0.87	0.87
	α	0.76	0.80	0.76	0.86	0.90	0.79	0.85	0.83
	n	13	100	25	49	57	73	171	244
Assessment (5 items)	α_0	0.86	0.87	0.79	0.85	0.86	0.85	0.90	0.90
	α	0.91	0.85	0.78	0.89	0.91	0.86	0.88	0.87
	n	13	100	25	49	57	73	171	244
Planning for Instruction (4 items)	α_0	0.89	0.86	0.78	0.87	0.86	0.84	0.90	0.90
	α	0.88	0.84	0.77	0.93	0.91	0.89	0.87	0.88
	n	12	99	25	47	55	71	167	238
Instructional Strategies (6 items)	α_0	0.91	0.90	0.88	0.89	0.89	0.88	0.93	0.93
	α	0.86	0.89	0.91	0.95	0.93	0.92	0.93	0.92
	n	12	99	25	46	55	71	166	237
Professional Learning and Ethical Practice (6 items)	α_0	0.92	0.88	0.80	0.88	0.88	0.86	0.91	0.91
	α	0.82	0.86	0.76	0.92	0.89	0.86	0.88	0.88
	n	11	98	23	46	55	68	165	233
Leadership and Collaboration (3 items)	α_0	0.92	0.78	0.70	0.78	0.80	0.77	0.84	0.84
	α	0.87	0.76	0.76	0.84	0.78	0.78	0.79	0.79
	n	11	98	23	46	55	68	165	233

Note. Standards are ordered to match presentation of InTASC standards.