AUTHORS
Michael Floren Ph.D.
University of North Alabama

N. M. Sileo Ed.D.
University of Northern Colorado

V. J. H. Sherman Ph.D.
University of Colorado Denver

Kiley Floren, M.P.H.
Practical Statistics LLC

Abstract

Educator preparation programs (EPPs) provide a foundation for preservice teachers to gain the knowledge, skills, and dispositions needed by classroom teachers. Moreover, EPPs also provide educational foundations and professional development for inservice teachers, school counsellors, school psychologists, educational administrators such as principals and district superintendents, as well as other direct service providers such as physical therapists, occupational therapists, and speech language therapists. The Council for the Accreditation of Educator Preparation (CAEP), the Association for Advancing Quality in Educator Preparation (AAQEP), and state-level agencies have identified standards and structures to ensure EPPs address the needs of individuals in educator preparation programs to ensure these individuals have the skills needed to maximize student learning outcomes in P-12 systems. The Interstate Teacher Assessment and Support Consortium (InTASC) Model Standards (CCSSO, 2013) are often used to define candidates' skills. This study focused on the structural validation of the InTASC Candidate Self-Perception Instrument (ICSPI; Floren et al., 2020), designed to obtain feedback of teacher candidates' perceptions of how well their EPP prepared them. A confirmatory factor analysis (CFA) was conducted using item placement within a structure based on the InTASC Standards. The CFA demonstrated the ICSPI's adherence to InTASC Standards, providing structural evidence for the content and construct validity and reliability of instrument data in reflecting candidate self-perceptions on their preparedness.

Validation of an Instrument to Measure Teacher Candidates' Perceptions of Preparedness to Meet InTASC Standards

CORRESPONDENCE
Email
mfloren@una.edu

Training, practice, and pedagogical reflection are necessary for teacher candidates to become proficient educators. The training teacher candidates receive while enrolled in educator preparation programs (EPPs) helps develop teacher candidates' professional competencies and ensure they have the necessary skills and dispositions required to have a positive impact on student learning (Struyf et al., 2011). Because of this, it is essential EPPs reflect on the training provided to teacher candidates. Previous research examined teacher candidates' perceptions of the skills, content knowledge, and dispositions acquired throughout their preparation programs (Hoffman et al., 2005; Pajares 1992; Wolsey et al., 2013; Zeichner & Conklin, 2008). These self-reflections are critical to both program evaluation and student growth. For instance, according to Chan and Luk (2021), it is essential for teacher candidates to understand and evaluate their own pedagogical capacities as this awareness provides a foundation for future development and acts as a source of personal motivation as they develop, refine, and improve their teaching skills.

In support of programs determining the proper concepts and skills to cover, the Council of Chief State School Officers (CCSSO, 2013) developed a series of competencies designed for beginning teachers in the United States. These standards identify what teachers should know and be able to do and are called the Interstate Teacher Assessment and Support Consortium (InTASC; CCSSO, 2013). Throughout their programs, teacher candidates need to learn and master content knowledge, pedagogical skills, and professional dispositions to

support the learning needs of P-12 (CAEP, 2013; CCSSO, 2013; Floren et al., 2020). The Council for the Accreditation of Education Preparation (CAEP), one of the major educator preparation accrediting bodies, required EPPs to show "candidates demonstrate an understanding of the 10 InTASC standards…" (CAEP, 2013, p. 2). Further, according to Darling-Hammond (2020), InTASC standards have been integrated into licensing and accreditation in more than 40 states. While these InTASC Standards can help programmatic alignment, they also serve a variety of other purposes. For example, the InTASC Standards detail the skills preservice teaching candidates should possess in order to have a positive impact on student achievement. The InTASC Standards also serve as a source of information used to solidify expectations for early career teachers across states and within school districts (CCSSO, 2013).

Many EPPs seek recognition from accreditation organizations such as the CAEP, the Association for Advancing Quality in Educator Preparation (AAQEP), and state-level accreditation agencies to endorse the quality of their programs and elevate the accomplishments of their teacher candidates. In all cases, these accreditation organizations require programs to demonstrate alignment with professional standards. Darling-Hammond (2020), Floren et al. (2020), and Schacter and Thum (2004) reinforced the concept that high quality EPPs should adopt and use high-quality educator preparation standards, such as those developed by state and professional organizations. The InTASC Standards are examples of foundational elements that can be incorporated into educator preparation programming and course structures used to support accreditation efforts.

To demonstrate programs are offering training in accordance with professional standards, EPPs are expected to report on measures that exhibit the alignment. Heafner et al. (2014) and Wentworth et al. (2009) identified a variety of methods programs can utilize to demonstrate alignment with CAEP and/or AAQEP and InTASC Standards within their frameworks. While valuable, research focused on creating instruments to assess program alignment with the InTASC Standards predominantly incorporated direct observation of teacher candidates instead of providing an instrument capable of assessing candidate perceptions of their own experiences (Wentworth et al., 2009). Further accrediting bodies require EPPs to collect data and demonstrate candidate abilities, knowledge, and skills using a variety of measures, such as observations, data collected on candidate academic performance, and candidate self-assessment measures (CAEP, 2013). This can be a daunting task, as many of the indicators included within the InTASC Standards may not be present even in high-quality lessons and programs, as are typically observed of candidates during lesson demonstrations and field experiences. More broadly, Immekus (2016) described how the vague nature of accreditation standards is one of the challenges for programs seeking to implement high-quality measures and determine the adequacy of their programs. Additionally, instruments created for this purpose are not always evaluated utilizing an assessment of a pre-hypothesized factor structure (e.g., Struyf et al., 2011; Wentworth et al., 2009). Thus, EPPs must decide between committing considerable time to search for appropriate measures or creating instruments.

The ICSPI is a recently-created instrument designed to obtain feedback from teacher candidates on how well their educational preparation program prepared them to meet a variety of elements indicated in the InTASC Standards

One instrument designed to fill this gap is the InTASC Candidate Self-Perception Instrument (ICSPI; Floren et al., 2020). The ICSPI is a recently-created instrument designed to obtain feedback from teacher candidates on how well their educational preparation program prepared them to meet a variety of elements indicated in the InTASC Standards (Floren et al., 2020). The ICSPI can be used to determine how well an EPP prepares teacher candidates to meet a variety of elements indicated in the InTASC Standards (Floren et al.). Instruments such as the ICSPI can be used as a source of evidence for EPPs to demonstrate and elevate program quality (Chan & Luk, 2021). Moreover, Chan and Luk (2021) indicated that data from a singular validated source or instrument can be used across disciplines to help leaders of EPPs come to a common understanding of the skills their candidates possess.

Current Study

Floren et al. (2020) provided evidence for the reliability and validity of data obtained from the ICSPI when gathered from teacher candidates enrolled in final methods courses or during final field experiences. While Floren et al. (2020) present important procedural

validation and reliability information of the ICSPI, confirmatory analysis of the factor structure hypothesized in their paper has yet to be conducted (their paper describes limitations of sample size as the reason this was not done in the initial study). The current study addresses the gap by assessing the fit of the hypothesized factor structure using confirmatory factor analysis (CFA) in addition to providing further information on the reliability of this relatively new instrument.

Method

Sample

Overall, 625 undergraduate teacher candidates agreed to participate in this study. Of these, 578 began the instrument and 522 completed the demographic information and all items (90.3% completion rate). Teacher candidates within the sample represented a variety of different content areas. Demographic information for the sample is presented in Table 1. As illustrated in the table, the sample leans towards female participants who were seniors at the point they completed the ICSPI and has broad representation regarding program and emphasis. Note that not all programs have an emphasis. Instrumentation

The ICSPI is an instrument designed to obtain feedback from teacher candidates regarding how well their program prepared them to meet the benchmarks described in the InTASC Standards. The ICSPI subscales follow each of the InTASC Standards: learner development, learning differences, learning environments, content knowledge, application of content, assessment, planning for instruction, instructional strategies, professional learning and ethical practice, and leadership and collaboration. Floren et al. (2020) describe the process used during the item creation, piloting, and initial assessment of the instrument. Designed to be given to candidates at multiple points throughout their EPP, the ICSPI consists of 48 items distributed across all standards. Each item utilized a 5-point Likert scale ranging from Strongly Disagree (1) to Strongly Agree (5) with a neutral category. Floren et al. (2020) present evidence of the validity of the instrument and include reliability information for each subscale broken down by candidates' program and program year (i.e., methods and field) from their sample of 257 undergraduate students. Floren et al. (2020) reported Cronbach's alphas for the subscales ranging from 0.84 to 0.93.

Procedures

Ethical approval was obtained from an institutional review board before data were collected or analysed. In conjunction with EPP administration, the final methods course and final field experience from within each program were identified. University-based email addresses for teacher candidates within the identified courses were obtained through the university student information system. For the purposes of program evaluation, college administrators requested a program-related demographic survey and a pre-existing field experience follow-up questionnaire be distributed with the instrument. The ICSPI required roughly 15 minutes to complete.

Teacher candidates in the sample pool received an email with a link to the instrument after completing their respective target courses. Reminder emails were sent to candidates who had not started the instrument at two and four weeks after the initial dissemination. As part of the consent process, all teacher candidates were informed that participation was completely voluntary and that responses and data would be kept confidential.

The current study addresses the gap by assessing the fit of the hypothesized factor structure using confirmatory factor analysis (CFA) in addition to providing further information on the reliability of this relatively new instrument.

Analysis

To assess the intended fit of the ICSPI to the InTASC Standards, a CFA was conducted using item placement within the structure based on the InTASC Standards described by Floren et. al. (2020). Statistics regarding item alignment with scale (i.e., item-total correlation), as well as scale reliability (i.e., Cronbach's alpha), were also calculated. Data analysis for the CFA was fit utilizing the robust maximum likelihood estimation method available in the *lavaan* package, electing for the tradeoff (due to the ordinal nature of the response) of a slight negative bias in factor loadings for the beneficial statistical properties of maximum likelihood with the trivial level of bias in standard errors and factor correlations with sample sizes

Table 1
Sample Demographics

Characteristics[a]	*n* (students)	Percentage[b]
Sex *(n=522)*		
Female	453	86.8
Male	69	13.2
Location *(n=522)*		
Main Campus	419	80.3
Online	42	8.0
Remote Campus	61	11.7
Year *(n=522)*		
Sophomore	5	1.0
Junior	39	7.5
Senior	289	55.5
Graduate	187	36.0
Program *(n=520)*		
Early Childhood Education	40	7.7
Elementary Education	208	39.8
K-12 Education	49	9.4
Secondary Education	98	18.8
Special Education	127	24.3
Emphasis *(n=273)*		
Art[c]	17	6.2
Deaf and Hard of Hearing[e]	6	2.2
Early Childhood Special Education[e]	25	9.2
English[d]	21	7.7
Generalist[e]	78	28.6
History[d]	12	4.4
Mathematics[d]	19	7
Modern Languages[d]	4	1.5
Music[c]	20	7.3
Science[d]	23	8.4
Social Science[d]	3	1.1
Spanish[d]	6	2.2
Sports and Exercise Science[c]	12	4.4
Theater[c]	9	3.3
Visually Impaired[e]	18	6.6

Note: [a]Sample sizes provided next to demographic indicate the number of respondents per item. [b]Percentages are calculated with respect to non-missing responses. [c]Emphasis in K-12, [d]Emphasis in Secondary Education, eEmphasis in Special Education.

The ICSPI is an instrument designed to obtain feedback from teacher candidates regarding how well their program prepared them to meet the benchmarks described in the InTASC Standards. The ICSPI subscales follow each of the InTASC Standards: learner development, learning differences, learning environments, content knowledge, application of content, assessment, planning for instruction, instructional strategies, professional learning and ethical practice, and leadership and collaboration.

observed here, even when underlying distributions were moderately nonnormal (Li, 2016). Reliability and item-total correlations were conducted using the *psych* package (Revelle, 2022; Rosseel, 2012). Both packages were loaded from R and implemented in RStudio (R Core Team, 2023; RStudio Team, 2020).

To evaluate the fit of the data to the hypothesized model, several CFA indices were calculated: Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). While classic rules-of-thumb are available regarding acceptable fit on these indices (e.g., CFI, TLI > 0.90; SRMR, RMSEA < 0.10), it has been suggested that interpreting a combination of fit statistics provides lower Type 1 and Type 2 error rates, and that the thresholds for acceptable fit change with this approach (Hu & Bentler, 1999). Thus, model fit was interpreted using a combination of these fit statistics. Additionally, robust fit index estimates were calculated.

Results

We first present the results of the CFA used to validate the ICSPI items on the 10-factor structure of InTASC Standards. All fit indices were within acceptable ranges and the model overall was determined to have good model fit, with the CFI and TLI values within acceptable ranges and a combination of RMSEA and SRMR values satisfying the ranges identified by Hu and Bentler (1999), with (1035) = 3084.07, $p < 0.001$, CFI = 0.92, TLI = 0.91, Robust RMSEA = 0.053, 90% CI [0.050, 0.056], SRMR = 0.042. Item factor loadings were all statistically significantly different from zero ($p < 0.01$), with values ranging from 0.60 to 0.86 as shown in Table 2.

Table 2
Factor Loadings Of Items In The ICSPI From Confirmatory Factor Analysis

InTASC Standard	*α*	Loading	***InTASC Standard***	*α*	Loading
Learner Development	0.87		***Assessment***	0.87	
Item 1		0.78	Item 25		0.75
Item 2		0.80	Item 26		0.75
Item 3		0.75	Item 27		0.82
Item 4		0.77	Item 28		0.75
Item 5		0.72	Item 29		0.72
Learning Differences	0.86		***Planning for Instruction***	0.86	
Item 6		0.62	Item 30		0.81
Item 7		0.78	Item 31		0.79
Item 8		0.76	Item 32		0.78
Item 9		0.67	Item 33		0.76
Item 10		0.77	***Instructional Strategies***	0.92	
Item 11		0.72	Item 34		0.79
Learning Environments	0.87		Item 35		0.80
Item 12		0.79	Item 36		0.85
Item 13		0.80	Item 37		0.82
Item 14		0.80	Item 38		0.80
Item 15		0.77	Item 39		0.82
			Professional Learning		
Content Knowledge	0.88		***and Ethical Practice***	0.88	
Item 16		0.78	Item 40		0.74
Item 17		0.83	Item 41		0.72
Item 18		0.78	Item 42		0.76
Item 19		0.81	Item 43		0.73
Item 20		0.68	Item 44		0.76
Application of Content	0.82		Item 45		0.77
			Leadership and		
Item 21		0.74	***Collaboration***	0.83	
Item 22		0.86	Item 46		0.77
Item 23		0.60	Item 47		0.81
Item 24		0.77	Item 48		0.79

Note: As ICSPI items are protected via non-disclosure, items numbers are reported instead.

Item factor loadings were all statistically significantly different from zero (p < 0.01), with values ranging from 0.60 to 0.86 as shown in Table 2.

Table 3 shows Cronbach's alpha for the subscales ranged from 0.83 to 0.92, exemplifying how reliable data can be produced by the subscales. The 95% confidence interval for alpha values within each subscale are also provided. Item-total correlations for items within each subscale ranged from 0.76 to 0.89, demonstrating a strong positive correlation, which indicates the survey items were consistently associated with one another. The high alpha and item-total correlation values suggest the items within each scale measured a consistent construct.

Table 3
Reliability Evidence of ICSPI Items Broken Down by InTASC Standard

Sub-Scale	# of Items	*n*	*M*	*SD*	Item-Total Correlation	Cronbach's Alpha (95% CI)
Learner Development	5	578	4.07	0.70	0.77 - 0.84	0.87 (0.86, 0.89)
Learning Differences	6	578	4.02	0.71	0.73 - 0.82	0.86 (0.85, 0.88)
Learning Environments	4	563	4.21	0.70	0.79 - 0.88	0.87 (0.85, 0.89)
Content Knowledge	5	563	4.12	0.70	0.78 - 0.85	0.88 (0.86, 0.89)
Application of Content	4	553	3.84	0.79	0.74 - 0.86	0.82 (0.80, 0.84)
Assessment	5	553	4.05	0.75	0.78 - 0.83	0.87 (0.86, 0.89)
Planning for Instruction	4	538	4.19	0.68	0.80 - 0.86	0.86 (0.85, 0.88)
Instructional Strategies	6	537	4.19	0.68	0.81 - 0.87	0.92 (0.91, 0.93)
Professional Learning and Ethical Practice	6	528	4.44	1.02	0.76 - 0.82	0.88 (0.87, 0.90)
Leadership and Collaboration	3	528	4.58	1.06	0.85 - 0.89	0.83 (0.80, 0.85)

Discussion

This study describes results of an assessment of structural validity and reliability of data from the ICSPI, a survey instrument designed to examine teacher candidates' perceived preparation in the InTASC Standards. A CFA was utilized to assess the fit of the ICSPI data to the InTASC Standards structure. The analysis provided strong evidence supporting the structural validity of the ICSPI when completed by preservice teacher candidates. Cronbach's alpha scores and confidence intervals were calculated using data from each of the subscales and indicated reliable measurement.

The analysis provided strong evidence supporting the structural validity of the ICSPI when completed by preservice teacher candidates….This alignment strengthens the evidence of structural validity for using the ICSPI scores for both EPP program assessment and accreditation in reference to the InTASC Standards.

Evidence of this structural alignment both complements and aligns with the content and construct validity evidence originally provided in Floren et al. (2020). This alignment strengthens the evidence of structural validity for using the ICSPI scores for both EPP program assessment and accreditation in reference to the InTASC Standards. Reliability results were strong across all subscales, which also aligns with results reported in Floren et al. (2020).

Psychometric results of this study also align with those available for other scales regarding teacher candidate self-perceptions on knowledge and dispositions for teaching (e.g., Struyf et al., 2011; Wentworth et al., 2009). While fit statistics for CFA are not always presented for these scales, measures of reliability are more consistently reported. For example, Cronbach's alpha reliability scores for the BeTaBas (60 item, 12 subscale instrument) presented by Struyf et al. (2011) are between 0.63 - 0.90, while Cronbach's alpha reliability scores for the InTASC-based CPAS (34 item, 10 subscale instrument) presented by Wentworth et al. (2009) are between 0.49 - 0.98 for Elementary Education, and between 0.32 - 0.71 in Secondary Education. Wentworth et al. (2009) acknowledges the low reliability of scale use within Secondary Education citing issues with insufficient training of raters on InTASC Standards. In comparison, Cronbach's alpha reliability scores for the ICSPI ranged from 0.82 - 0.92, demonstrating comparable or stronger reliability when measuring similar constructs within a similar population. Combined with the comparative ease of data collection using the ICSPI, this makes the ICSPI an attractive option for researchers and administrators seeking to provide reliable evidence regarding programmatic contribution to teacher candidates' ability to have a positive impact on student learning and achievement as advanced by the InTASC Standards.

Future Research

One limitation of this study is that teacher candidates self-reported their competencies on the ICSPI scale and may have over- or underestimated their level of competency regarding the InTASC Standards (Dunning et al., 2003; Zlatkin-Troitschanskaia et al., 2015). In the future, researchers could consider triangulating data with alternate sources of information related to teacher candidate preparation (e.g., teaching artifacts, classroom observations, etc.) to reduce

the potential bias associated with self-report measures and establish evidence for predictive validity of the ICSPI.

The sample in Floren et al. (2020) was predominantly white women as is typical of the demographics of preservice teacher educator preparation programs throughout the United States (Causey et al., 2000), and did not specifically analyze subgroup data such as gender, race, ethnicity, and/or language within the larger sample, Therefore, the CFA results do not discuss or break down the sample using these subgroups. Previous research (e.g., Causey et al., 2000; Whitaker & Valtierra, 2018) indicates these factors can impact self-reporting and self-efficacy perceptions. Researchers who seek to replicate Floren et al.(2020) may wish to consider analyzing subgroup data associated with gender, race, ethnicity, and/or language to explore how those factors may affect outcomes on the ICSPI.

To continue to build evidence for generalizability of the ICSPI, future researchers could consider assessing the reliability and structural validity within samples from multiple institutions and geographical regions.

Additionally, this measure was validated based on data collected from primarily undergraduate teacher candidates from a single institution in the western United States. To continue to build evidence for generalizability of the ICSPI, future researchers could consider assessing the reliability and structural validity within samples from multiple institutions and geographical regions. Concomitantly, future studies could consider expanding the sample to further include graduate students and early-career program completers.

Finally, research to date has focused on content and construct validity and reliability of the ICSPI. Concurrent and divergent validity have not been established between this measure and other measures of candidate perceptions of preparedness. To continue to build evidence for validity of the ICSPI in relation to other measures, future researchers may consider a joint distribution of the ICPSI with other instruments designed to assess candidate perceptions of preparedness to teach.

This study demonstrated the structural validity of the ICSPI when used as a measurement instrument to capture teacher candidates' self-perceptions of their preparedness to meet the InTASC Standards. Confirmatory factor analysis and reliability results demonstrate all sub-scales of the ICSPI have produced reliable and valid data, adding to evidence from previous research. We believe that this growing body of evidence will allow researchers and administrators to utilize the ICSPI for program review and accreditation with confidence.

References

Causey, V. E., Thomas, C. D., & Armento, B. J. (2000). Cultural diversity is basically a foreign term to me: The challenges of diversity for preservice teacher education. *Teaching and teacher education*, *16*(1), 33-45. https://doi.org/10.1016/S0742-051X(99)00039-6

Chan, C. K. Y., & Luk, L. Y. Y. (2021). Development and validation of an instrument measuring undergraduate students' perceived holistic competencies. *Assessment & Evaluation in Higher Education*, *46*(3), 467-482. https://doi.org/10.1080/02602938.2020.1784392

Council for the Accreditation of Educator Preparation [CAEP]. (2013). *CAEP Accreditation Standards*. Washington, DC: Council for the Accreditation of Educator Preparation.

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: Council of Chief State School Officers.

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

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*(3), 83-87. https://doi.org/10.1111/1467-8721.01235

Floren, M., Hess, C., Sherman, V. J. H., & Sileo, N. M. (2020). Accreditation by design: Construction of an instrument to measure teacher candidates' perceptions of preparedness to meet InTASC standards. *Journal of Educational Research and Innovation*, *8*(1), Article 4. https://digscholarship.unco.edu/jeri/vol8/iss1/4

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 https://doi.org/10.1086/428744

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling 6*, 1-55. http://dx.doi.org/10.1080/10705519909540118

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. https://eric.ed.gov/?id=EJ1137992

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936-949. https://doi.org/10.3758/s13428-015-0619-7

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*(3), 307-332. https://doi.org/10.3102/00346543062003307

R Core Team. (2023). "R: A language and environment for statistical computing. *R Foundation for Statistical Computing. Vienna*. Austria: R Core Team. https://www.R-project.org/

Revelle, W. (2022). psych: Procedures for personality and psychological research. Northwestern University. Evanston, IL. https://CRAN.R-project.org/package=psych Version = 2.2.5.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

RStudio Team. (2020). RStudio: Integrated development for R. *RStudio*. PBC. Boston, MA: RStudio Team. http://www.rstudio.com/.

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

Struyf, E., Adriaensens, S., & Meynen, K. (2011). Are beginning teachers ready for the job? The development and validation of an instrument to measure the basic skills of beginning secondary teachers. *Assessment & Evaluation in Higher Education*, *36*(4), 429-449. https://doi.org/10.1080/02602938.2011.581748

Wentworth, N., Erickson, L. B., Lawrence, B., Popham, J. A., & Korth, J. A. (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

Whitaker, M. C., & Valtierra, K. M. (2018). Enhancing preservice teachers' motivation to teach diverse learners. *Teaching and Teacher Education*, *73*, 171-182. https://doi.org/10.1016/j.tate.2018.04.004

Wolsey, T. D., Young, J. R., Scales, R. Q., Scales, W. D., Lenski, S., Yoder, K. K., Wold, L., Smetana, L., Grisham, D. L., Ganske, K., Dobler, E. & Chambers, S. A. (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. (2008). Teacher education programs as sites for teacher preparation. In M. Cochran-Smith, S. Feiman-Nemser, D. J. McIntyre, & K. E. Demers (Eds.), *Handbook of research on teacher education: Enduring questions in changing contexts* (3rd ed., pp. 269-289). Routledge.

Zlatkin-Troitschanskaia, O., Shavelson, R. J., & Kuhn, C. (2015). The international state of research on measurement of competency in higher education. *Studies in Higher Education*, *40*(3), 393-411. https://doi.org/10.1080/03075079.2015.1004241