FACTOR ANALYTIC INSIGHTS INTO MICRO-TEACHING PERFORMANCE OF TEACHER CANDIDATES

Atakan Ata [iD]
Bahcesehir University
atakan.ata@es.bau.edu.tr

Kadir Kozan [iD]
Istanbul Bilgi University
kadir.kozan@bilgi.edu.tr

Atakan Ata, PhD-University of Georgia (Athens, GA), is an assistant professor at Bahcesehir University in Istanbul. Dr. Ata currently teaches education law, educational sciences, and social research methods. His research is focused on promoting student civic engagement, and teachers' professional development.

Kadir Kozan, PhD-Purdue University (West Lafayette, IN), is an assistant professor at Istanbul Bilgi University. His main research interest is pedagogy-rich technology-enabled learning environments. He also has an M.A. and a B.A. degree in English Language Education.

FACTOR ANALYTIC INSIGHTS INTO MICRO-TEACHING PERFORMANCE OF TEACHER CANDIDATES

Atakan Ata

atakan.ata@es.bau.edu.tr

Kadir Kozan

kadir.kozan@bilgi.edu.tr

Abstract

This study investigated the factor structure of the intern keys teacher candidate assessment instrument used to assess micro-teaching performance of teacher candidates. In other words, the purpose of the current study was to explore the construct validity of the assessment tool which is considered to consist of one overarching construct: teaching effectiveness. To this end, an exploratory factor analysis in addition to a parallel analysis were employed using the ratings of 116 faculty members at various U.S. universities who supervised the teacher candidates' performance in their final semester. The results indicated a one-factor structure and high reliability indices. All these findings align with theoretical assumptions and call for further factor analytic studies on the instrument using different samples.

Keywords: practicum, micro-teaching, teacher candidate, factor analysis, effectiveness

1. Introduction

In recent years, teacher education programs at higher education institutions have begun adapting a practice-based approach (McDonald, Kazemi, & Kavanagh, 2013) because learning to teach is not easy and it is cultivated through high-quality opportunities to practice, along with support and feedback (Benedict, Holdheide, Brownell, & Foley, 2016). For instance, most states in the United States have adapted some teaching standards for teacher performance assessment recommended by the Interstate Teacher Assessment and Support Consortium, and these standards are also being used for pre-service teachers' performance evaluation. In 2016, Turkish Department of Education also created some teacher observation criteria for teacher performance evaluation purposes. Teacher education programs at universities are closely following the developments, adjusting their curricula accordingly, and extending pre-service teachers' practicum hours in the field.

Practice-based approaches are not a new idea (Zeichner, 2012), nevertheless, because they provide teacher candidates with opportunities to integrate both content and pedagogy acquired through coursework into instruction (Ericsson, 2014), they have recently drawn attention again. Another reason is because some novice teachers feel that they are not ready to teach even after graduation, or first-year teachers face with more problems compared to experienced teachers (Rust, 1994). Teacher educators often experience that 'theory first', and

applying theories they had learnt in class in practicum later, is not very productive (Emsheimer & Ljunggren de Silva, 2011; Hennissen, Beckers, & Moerkerke, 2017). Educator preparation programs have developed alternative ways for pre-service teachers to bridge the gap between theory and practice.

One of the approaches helping teacher candidates become *effective* teachers is microteaching (Arsal, 2015; Lenihan, 2016). Based on various studies (Amobi, 2005; Benton-Kupper, 2001; Gorgen, 2003), Yangin-Eksi and Asik (2015, p. 28) stated that microteaching "help[s] prospective teachers transfer their knowledge and skills into action, having reflective teaching practices and experiencing teaching profession." This is, however, possible when microteaching is done properly. For instance, Allen and Eve introduced a microteaching model in 1960s which included six interrelated stages: Planning, teaching, observation (criticism), re-planning, re-teaching, and re-observation (Arsal, 2015). Yet today, feedback, a valuable part of microteaching, is mostly absent (Lenihan, 2016). Even though microteaching has one of the highest teacher-controlled effects on student achievement (Hattie, 2009), educator preparation programs seem to miss the potential this method presents (Lenihan, 2016).

Another problem with practice-based approach is the design. In a study among Flemish pre-service secondary teachers studying, Schelfhout et al. (2006) found that pre-service teachers did not change their behavior when the curriculum only offered theoretical topics that were unrelated to learning in practice (Hennissen et al., 2017). Brouwer and Korthagen (2005) found that practicum experiences were greatly enhanced when theory and practice are integrated into the curriculum, and theoretical elements are perceived by pre-service teachers as useful for practice. Because most pre-service teachers are not getting the most out of microteaching, adapting teacher performance evaluation systems that are designed for in-service teachers is becoming common. This approach encourages feedback and prepares student for 'real-life' while assessing teacher candidates' teaching performance.

As teacher education programs are aligning their regulations with teacher performance evaluation systems, and basing their courses on instructional performance, teacher candidate observations are taking a new shape and new tools are needed for assessing candidates' performance. Assessing pre-service teachers' instructional performance is not easy, mostly because they are not teachers yet, but most of the evaluation criteria are adapted from teacher evaluation forms which include items like communication with students' parents. It is also not easy because number of observations a pre-service teacher would get might not reflect the reality. Teacher education programs must take equity into account considering different observers involved in the process as well.

In this movement of shifting initial teacher education to a point after teachers assume full responsibility, Zeichner (2012) warns us that there is a danger of narrowing the role of teachers to that of technicians who implement a particular set of teaching strategies, but who do not develop broad professional vision, and relational skills. For this very reason, and also because we cannot expect teacher candidates to assume full responsibility of teaching, teacher candidate effectiveness and its assessment should be well thought and established.

Therefore, adopting a comprehensive assessment tool that truly assess instructional performance, and that minimizes personal bias is essential for faculties of education.

1.1. Research Question and Objective

To address the purpose above, the present study was to test the construct validity of the teacher candidate effectiveness instrument thereby addressing the following research question:

- What is the factor structure of the teacher candidate effectiveness instrument?

In other words, the present study focused on whether the teacher candidate effectiveness assessment tool would have an interpretable factor structure that aligns with theoretical assumptions.

1.2. Theoretical Framework

A comprehensive report on practice-based preparation in teacher education by the Collaboration for Effective Educator Development Accountability and Reform, U.S. Office of Special Education Programs, and the Center on Great Teachers and Leaders (Benedict et al., 2016) suggests that practice-based opportunities "teach novices to integrate critical knowledge and skills they need to teach effectively while receiving valuable feedback," (p. 2). However, the power of practice can be weakened unless they are strategically organized and delivered. The report (Benedict et al., 2016, p. 4-6), therefore, offers six features of high-quality practice based opportunities:

1) Modeling: How teacher educators provide candidates examples of what expert performance looks like in practice.
2) Spaced learning: It offers candidates opportunities to practice the knowledge and skills acquired in coursework over a period of time that are sustained and repeated.
3) Varied learning: It offers candidates opportunities to practice across varying contexts, with a diverse range of student learners, and with differing degrees of support.
4) Coaching and feedback: The focus of the coaching and feedback is on improving candidates' practice and expertise.
5) Analyzing and reflecting: Engagement in analysis and reflection upon both their practice and their impact on student learning.
6) Scaffolded practice: Teaching experiences that gradually increase in complexity over time with fading support from teacher educators.

Based on research and the suggestions above, teacher candidates should be provided with opportunities to analyze and reflect upon their practice before, during, and after instruction, and they are expected to employ metacognitive skills to both reflect upon and improve their practice (Benedict et al., 2016; Berliner, 1986; Nagro, deBettencourt, Rosenberg, Carran, & Weiss, 2017).

1.3. The Intern Keys Teacher Candidate Assessment

Intern Keys is one of the assessment tools that is used for assessing pre-service teachers' instructional performance. Intern Keys is described to be "not an observation checklist," but a collection of standards that a pre-service teacher should meet (Elder, Ata, & Cramer, 2016, p.

2). The Intern Keys validation project is being conducted in the State of Georgia (USA) and was funded through a grant from the Georgia Network for Transforming Educator Preparation and the Council of Chief State School Officers.

Intern Keys Teacher Candidate Assessment tool is adapted from Teacher Assessment on Performance Standards. Teacher Assessment on Performance Standards are similar to the Stronge Teacher Effectiveness Performance Evaluation (Stronge & Associates, 2016), and, as an assessment tool, it is one of the components of the Teacher Keys Effectiveness System in Georgia - an evaluation system designed for building teacher effectiveness and ensuring consistency and comparability throughout the state (Georgia Department of Education, 2016).

2. Methodology

2.1. Participants

The research data were collected from 116 faculty members working at different universities across the US. All the faculty members were working as teacher educators at the time of data collection. Specifically, the participants were working as faculty supervisors of more than 500 different pre-service teacher candidates. MacCallum, Widaman, Zhang, and Hong (1999) stated that a sample of 100 to 200 participants is enough with communalities of ≥. 5 and well-determined factors, which held true in the current study: Both initial and extraction communalities were higher than .5 and they ranged from .555 to .738. After all, MacCallum et al. (1999) also stated that "common rules of thumb regarding sample size in factor analysis are not valid or useful" (p. 96).

2.2. Design

The present study basically had a correlational research design since factor analysis is one of the statistical techniques used to examine relationships (e.g., Pallant, 2007). According to Pallant (2007), factor analysis is used when there is "a large number of related variables (e.g. the items that make up a scale)" in order to "explore the underlying structure of this set of variables. It is useful in reducing a large number of related variables to a smaller, more manageable, number of dimensions or components" (p. 120). As such, the current study checked whether the teacher candidate assessment tool has a one-factor structure thereby aligning with the theoretical assumptions.

2.3. Data Collection

The teacher candidate effectiveness assessment tool used to collect data included 10 main standards having 6-10 criteria to measure under each main standard. The instrument is very similar to Stronge Teacher Effectiveness Performance Evaluation System (2012) which is widely used by many school districts in the United States, and used in different effectiveness projects before. The instrument had a scale ranging from 4 (exemplary) to 1 (ineffective).

The faculty supervisors were emailed to complete the survey after the candidates were finished with their school experiences. Originally, there were 296 numbers of surveys collected but due to the lack of names or other identifying information and a random elimination of duplicate cases (each supervisor completed the survey for more than one candidate), 116 of them were analyzed here.

2.4. Data Analysis

Because the research design was correlational or factor analytical, the present research data were analyzed through factor analysis. Specifically speaking, exploratory factor analysis (e.g., Field, 2009; Pallant, 2007; Tabachnick & Fidell, 2013) was employed to establish whether the teacher candidate assessment tool has an interpretable factor structure or not. The data analysis started with checking for univariate and multivariate outliers. The 5 % trimmed mean values did not reveal any high-potential problems either in addition to no missing data points. The item ratings violated the normality assumption based on both Kolmogorov-Smirnov and Shapio-Wilk statistics (p's < .001). The relevant transformations employed resulted in no significant improvements. Therefore, the data were left as they were for the factor analysis. According to Tabachnick and Fidell (2013), as long as factor analysis is used for descriptive purposes, assumptions pertaining to data distribution may not be important. Tabachnick and Fidell (2013) further stated that the violation of the normality assumptions weakens the factorial solution, it "may still be worthwhile" (p. 618). Likewise, Field (2009) claimed that as the sample size gets larger the data distribution becomes closer to normality. Accordingly, given the more than 100 participants enrolled, violation of the normality assumption may not affect the results dramatically. Further, bivariate Spearman's rho correlations did not indicate any multicollinearity or singularity problems (r_s's $\leq$ .775). Finally, being higher than Tabachnick and Fidell's (2013) benchmark of >.30, the positive and significant correlations among the variables referred to a reasonable factorability level.

3. Results

The initial descriptive analyses showed that, on average, participants' ratings were quite high given that the scale had a maximum point of 4. Table 1 displays the descriptive statistics for each item on the survey:

Table 1. *Descriptive statistics (N = 116)*

	Possible Min.	Min.	Possible Max.	Max.	*M*	*SD*
Item 1: Professional Knowledge	1	1	4	4	3.06	.48
Item 2: Instructional Planning	1	1	4	4	3.04	.53
Item 3: Instructional Strategies	1	1	4	4	3.02	.51
Item 4: Differentiated Instruction	1	1	4	4	2.96	.53
Item 5: Assessment Strategies	1	1	4	4	2.96	.50
Item 6: Assessment Uses	1	1	4	4	2.90	.52
Item 7: Positive Learning Environment	1	1	4	4	3.15	.57
Item 8: Academically Challenging Environment	1	1	4	4	3.05	.54

Item 9: Professionalism	1	1	4	4	3.22	.54
Item 10: Communication	1	1	4	4	3.12	.54

The overall descriptive insights above suggest that the candidates were successful at their teaching practice since their mean ratings were either above three or very close to three (i.e., items 4, 5 and 6). An initial exploratory factor analysis (EFA) with direct oblimin using principal axis factoring was run to check sampling and data adequacy. No number of factors were extracted during this initial EFA. The correlation matrix showed large correlation values bigger than .30, which was suitable for factor analytic purposes. Bartlett's test of sphericity, $\chi 2$ (45.00) = 907.66, p < .001, revealed that correlations among the variables were large enough for an EFA. The Kaiser–Meyer–Olkin (KMO) indicated that sampling was suitable for the present analyses, KMO = .91 which is bigger than the suggested minimum values of .5 (Field, 2009) and .6 (Tabachnick & Fidell, 2013). All KMO values for each item was bigger than .875 as well.

The initial EFA above suggested one factor with one eigenvalue higher than 1 thereby referring to a one-factor structure for the ten items. Likewise, the scree plot also suggested one very strong factor leveling off at the second factor:
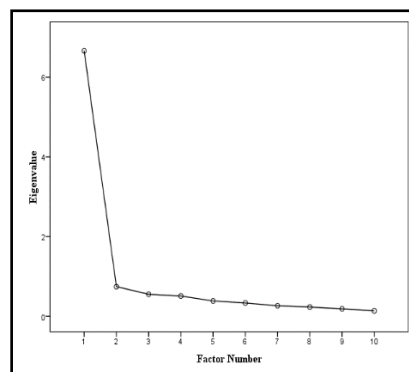

Figure 1. *Factor Structure*

Consequently, a second and final EFA with one factor extracted and no rotations was implemented. The factor matrix also suggested one factor with all the ten survey items having strong loadings on it (.726 to .834), and it was able to explain approximately 67 % of variance in the data. A parallel analysis (Watkins, 2000) including 10 variables, 116 participants and 100 replications produced one factor having an eigenvalue (i.e., 6.66) bigger than its random eigenvalue (1.48). All the upcoming potential nine factors, however, had a smaller measured eigenvalue (.137 to .743) than their corresponding possible random eigenvalues (.586 to 1.33). Further, the overall Cronbach's Alpha value was .944 and if-item-deleted statistics did not indicate any problematic items. Table 2 presents the final factor loadings.

Table 2. *Factor loadings*

Item	Factor loadings
# 2: Instructional Planning	.834
# 6: Assessment Uses	.821
# 7: Positive Learning Environment	.816
# 5: Assessment Strategies	.808
# 8: Academically Challenging Environment	.804
# 10: Communication	.798
# 1: Professional Knowledge	.781
# 9: Professionalism	.774
# 3: Instructional Strategies	.766
# 4: Differentiated Instruction	.726

Table 2 shows that all factors loaded positively and significantly on the factor which can be called "teaching effectiveness". In other words, these findings suggest that the construct validity of the instrument is high measuring one factor which is what is to be measured. Moreover, the second item (Instructional Planning) showed the highest loading in magnitude while the fourth item (Differentiated Instruction) showed the smallest one.

4. Discussion

The purpose of the present paper was to examine the construct validity of the interns key assessment instrument through an exploratory factor analysis (EFA). Even though the mean ratings of the items on the instrument were quite high (bigger than three on a 1-4 scale), there were also three items the mean ratings of which were quite close to but lower than three. Interestingly enough, these three items pertained to providing differentiated instruction, choosing appropriate assessment strategies, and appropriately using the insights gained through assessment. Needless to say, all these three perspectives seem to be vulnerable to experience: Teacher candidates may need to gain a certain level of experience or practice before mastering assessment and differentiated instruction.

Furthermore, the results above revealed significant theoretical and practical conclusions. First, the results strongly suggested that the teacher candidate effectiveness instrument is highly reliable. In other words, the ten items used in the instrument are based on a meaningful whole construct of "effectiveness" as it is measured through performance standards. This insight further suggests that the instrument can be used to during both pre-service and early in-service teacher training to assess whether teacher candidates perform effectively. Second, the scale item with the strongest factor loading (i.e., # 2) implies that instructional planning is of great importance for teacher educator faculty members. Similarly, the next two strongest

items (i.e., 6 & 7) focus on assessment use and creating a positive learning environment that are complementary to provide a quality and meaningful learning experience. However, the fourth item focusing on differentiated instruction had the smallest factor loading, which is in line with the findings that it was one of the items with the lowest mean rating. As suggested before, such a finding seems to suggest that faculty members or teacher educators may regard differentiated instruction as a high-level skill that comes with practice. Still though, it is a significant measure of teaching effectiveness. Consequently, further research can also run the instrument on different occasions over time to detect which aspect of teaching becomes important during teaching.

All these insights suggest that it is important to provide teacher candidates with an understanding of how to design, run, and evaluate meaningful learning experiences in teacher preparation programs. Even though educator preparation programs strive to meet elaborate accountability criteria to monitor their compliance with state requirements, much of this monitoring activity does not address or contribute to improving the quality of programs (Johnson, Johnson, Farenga, & Ness, 2005; Sleeter, 2008; Zeichner, 2010). Moreover, the effort and resources to produce detailed and extensive reports to states and accreditation agencies have diverted the attention of teacher educators away from creating innovative practices (Kornfeld, Grady, Marker, & Ruddell, 2007; Rennett-Ariev, 2008; Zeichner, 2010). Accordingly, the current research and future similar research on the development of valid, reliable and practical teaching effectiveness measures can provide us with more insights into how to efficiently foster teaching to actualize our ultimate goal of enhancing student learning.

References

Amobi, F. A. (2005). Preservice teachers' reflectivity on the sequence and consequences of teaching actions in a microteaching experience. *Teacher Education Quarterly, 32*(1), 115-130.

Arsal, Z. (2015). The effects of microteaching on the critical thinking dispositions of pre-service teachers. *Australian Journal of Teacher Education, 40*(3), 140-153.

Benedict, A., Holdheide, L., Brownell, M., & Foley, A. M. (2016). *Learning to Teach: Practice-Based Preparation in Teacher Education.* Special Issues Brief. Washington. DC: American Institutes for Research.

Benton-Kupper, J. (2001). The microteaching experience: Student perspectives. *Education, 121*(4), 830-830.

Berliner, D. C. (1986). In pursuit of the expert pedagogue. *Educational Researcher, 15*(7), 5-13. doi:128.227.169.113

Brouwer, N., & Korthagen, F. (2005). Can teacher education make a difference? *American Educational Research Journal, 42*(1), 153-224.

Eksi, G. Y., & Asik, A. (2015). Enriching microteaching in teaching English to young learners (TEYL): an action research. *International Online Journal of Education and Teaching, 2*(1), 26-41.

Elder, T, Ata, A., & Cramer, S.E. (2016). *An evaluation of the validity and reliability of the Intern Keys Assessment – Phase two.* (Report published by the GA Department of ED).

Ericsson, K. A. (2014). *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games.* Florence, KY: Psychology Press.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: SAGE Publications.

Georgia Department of Education. (2016). *Georgia's Teacher Keys Effectiveness System.* Atlanta, GA: Office of School Improvement.

Gorgen, I. (2003). The effect of microteaching practises on student teachers' views of giving lessons in the classroom. *Journal of Hacettepe University Education Faculty, 24*, 56-63.

Hennissen, P., Beckers, H., & Moerkerke, G. (2017). Linking practice to theory in teacher education: A growth in cognitive structures. *Teaching and Teacher Education, 63*, 314-325.

Johnson, D., Johnson, B., Farenga, S., & Ness, D. (2005). *Trivializing teacher education: The accreditation squeeze.* Lanham, MD: Rowman & Littlefield.

Kornfeld, J., Grady, K., Marker, P., & Ruddell, M. (2007). Caught in the current: A self-study of state-mandated compliance in a teacher education program. *Teachers College Record, 109*, 1902-1930.

Lenihan, E. (2016). Developing a culture of feedback through microteaching. *International Schools Journal, 35*(2), 82-85.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). *Sample size in factor analysis. Psychological Methods, 4*(1), 84–99.

McDonald, M., Kazemi, E., & Kavanagh, S. S. (2013). Core practices and pedagogies of teacher education a call for a common language and collective activity. *Journal of Teacher Education, 64*(5), 378-386.

Nagro, S. A., deBettencourt, L. U., Rosenberg, M. S., Carran, D. T., & Weiss, M. P. (2017). The Effects of Guided Video Analysis on Teacher Candidates' Reflective Ability and Instructional Skills. *Teacher Education and Special Education*, *40*(1), 7-25.

Pallant, J. (2007). *SPSS survival manual: A step by step guide to data analysis using SPSS for Windows* (3rd ed.). New York, NY: Open University Press.

Rennert-Ariev, P. (2008). The hidden curriculum of performance based teacher education. *Teachers College Record, 110*, 105-138.

Rust, F.O. (1994). The first year of teaching: It's not what they expected. *Teaching & Teacher Education, 10*(2), 205-217.

Schelfhout, W., Dochy, F., Janssens, S., Struyven, K., Gielen, S., & Sierens, E. (2006). Educating for learning-focused teaching in teacher training: The need to link learning content with practice experiences within an inductive approach. *Teaching and Teacher Education, 22*(7), 874-897.

Sleeter, C. (2008). Equity, democracy, and neo-liberal assaults on teacher education. *Teaching and Teacher Education, 24*, 1947-1957. State College, PA: Ed & Psych Associates.

Stronge, J. H. (2012). *Teacher Effectiveness Performance Evaluation System*. Stronge & Associates.

Stronge & Associates. (2016). *A validation report of Stronge evaluation system.* Stronge & Associates.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson.

Watkins, M. W. (2000). *Monte Carlo PCA for parallel analysis* [computer software].

Zeichner, K. (2010). Rethinking the connections between campus courses and field experiences in college-and university-based teacher education. *Journal of teacher education, 61*(1-2), 89-99.

Zeichner, K. (2012). The turn once again toward practice-based teacher education. *Journal of Teacher Education, 63*(5), 376–382. doi:10.1177/0022487112445789