



Research Memorandum
ETS RM-16-11

**Investigating the Relevance and
Importance of High-Leverage
Practices for Beginning Elementary
School Teachers**

Michelle P. Martin-Raugh

Jonathan H. Steinberg

Clyde M. Reese

Jun Xu

Richard J. Tannenbaum

September 2016

ETS Research Memorandum Series

EIGNOR EXECUTIVE EDITOR

James Carlson
Principal Psychometrician

ASSOCIATE EDITORS

Beata Beigman Klebanov
Senior Research Scientist

Heather Buzick
Research Scientist

Brent Bridgeman
Distinguished Presidential Appointee

Keelan Evanini
Research Director

Marna Golub-Smith
Principal Psychometrician

Shelby Haberman
Distinguished Presidential Appointee

Anastassia Loukina
Research Scientist

Donald Powers
Managing Principal Research Scientist

Gautam Puhan
Principal Psychometrician

John Sabatini
Managing Principal Research Scientist

Matthias von Davier
Senior Research Director

Rebecca Zwick
Distinguished Presidential Appointee

PRODUCTION EDITORS

Kim Fryer
Manager, Editing Services

Ayleen Gontz
Senior Editor

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Memorandum series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Memorandum series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

**Investigating the Relevance and Importance of High-Leverage Practices
for Beginning Elementary School Teachers**

Michelle P. Martin-Raugh, Clyde M. Reese, Richard J. Tannenbaum,
Jonathan H. Steinberg, and Jun Xu
Educational Testing Service, Princeton, New Jersey

September 2016

Corresponding author: Michelle P. Martin-Raugh, E-mail: mmartin-raugh@ets.org

Suggested citation: Martin-Raugh, M. P., Reese, C. M., Howell, H., Tannenbaum, R. J., Steinberg, J. H., & Xu, J. (2016). *Investigating the relevance and importance of high-leverage practices for beginning elementary school teachers* (Research Memorandum No. RM-16-11). Princeton, NJ: Educational Testing Service.

Find other ETS-published reports by searching the ETS ReSEARCHER
database at <http://search.ets.org/researcher/>

To obtain a copy of an ETS research report, please visit
<http://www.ets.org/research/contact.html>

Action Editor: Heather Buzick

Reviewers: Patrick Kyllonen and Gary Sykes

Copyright © 2016 by Educational Testing Service. All rights reserved.

ETS and the ETS logo are registered trademarks of Educational Testing Service (ETS).

MEASURING THE POWER OF LEARNING is a trademark of ETS. All other trademarks are the property of their
respective owners.



Abstract

The purpose of this study is to explore the validity evidence supporting the high-leverage practices (HLPs) of the *ETS*[®] National Observational Teaching Exam (NOTE) assessment series, a kindergarten through 6th grade teacher licensure assessment. HLPs include “tasks and activities that are essential for skillful beginning teachers to understand, take responsibility for, and be prepared to carry out in order to enact their core instructional responsibilities” (Ball & Forzani, 2009, p. 504). We accumulated relevance, importance, and frequency judgments of 20 HLPs from 569 practitioners in the field (385 teachers and 184 college faculty) verifying the necessity of these competencies for elementary school teachers first entering the teaching profession. Implications for the performance components of the NOTE assessment series are discussed.

Key words: high-leverage practices, content knowledge for teaching, elementary school, teacher licensure

Table of Contents

	Page
Licensure and Licensure Assessments.....	2
Content-Related Validity Evidence.....	3
The Current Study.....	4
Method.....	5
Survey.....	6
Sample.....	6
Analysis.....	10
Results.....	10
Relevance Judgments.....	10
Importance Judgments.....	10
Frequency.....	18
Relative Importance.....	26
Discussion.....	28
Limitations.....	30
References.....	31
Appendix. High-Leverage Practices (HLPs).....	35
Notes.....	39

List of Tables

	Page
Table 1. Background Information—Overall Samples—for English Language Arts (ELA) and Mathematics	8
Table 2. Teacher-Specific Background Information for English Language Arts (ELA) and Mathematics	9
Table 3. Faculty-Specific Background Information for English Language Arts (ELA) and Mathematics	9
Table 4. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Subject Area for English Language Arts (ELA) and Mathematics	11
Table 5. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Current Grade Level Taught for Mathematics	12
Table 6. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Current Grade Level Taught for English Language Arts (ELA)	13
Table 7. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Race/Ethnicity for Mathematics	14
Table 8. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Race/Ethnicity for English Language Arts (ELA)	15
Table 9. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Geographic Region for Mathematics	16
Table 10. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Geographic Region for English Language Arts (ELA)	17
Table 11. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Subject Area for English Language Arts (ELA) and Mathematics	18
Table 12. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Current Grade Level Taught for Mathematics	19
Table 13. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Current Grade Level Taught for English Language Arts (ELA)	20
Table 14. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Race/Ethnicity for Mathematics	21
Table 15. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Race/Ethnicity for English Language Arts (ELA)	22
Table 16. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Geographic Region for Mathematics	23
Table 17. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Geographic Region for English Language Arts (ELA)	24
Table 18. Summary (Overall Sample) of High Importance and Frequency Ratings for High-Leverage Practices (HLPs) for English Language Arts (ELA) and Mathematics	25

Table 19. Percentage of Teacher/Faculty Categorizations for High-Leverage Practices (HLPs) as Five Least and Most Important Practices for English Language Arts (ELA) 26

Table 20. Percentage of Teacher/Faculty Categorizations for High-Leverage Practices (HLPs) as Five Least and Most Important Practices for Mathematics..... 27

Table 21. Relative Importance Rankings for English Language Arts (ELAs) and Mathematics High-Leverage Practices (HLPs)..... 28

Teacher quality is of great interest to educators, policy makers, and the public at large. Concern over how teachers are selected into the profession and subsequently evaluated has existed nearly as long as the public school system (Wilson, 2009). Empirical studies have found evidence supporting the notion that the quality of a student’s teacher is the most important in-school factor predicting student outcomes (e.g., Ferguson, 1998; Goldhaber, 2002; Hanushek, Kain, & Rivkin, 1999). Consequently, there is agreement among both researchers and educators that teachers have a large effect on the lives of students (Harris & Rutledge, 2010) and that teacher quality is an important area of research.

Deborah Ball and her colleagues have argued that “any examination of teacher quality must, necessarily, also grapple with issues of *teaching* quality” (Ball & Hill, 2008, p. 81). Efforts to improve teaching quality can start with examining teacher quality at the point of entry into the profession and the licensure processes that are intended to safeguard the public.

Teacher licensure focusing on the beginning teacher is a process established by state boards of education, departments of education, or designated teacher licensure agencies. Licensure is intended to ensure that any candidate permitted to enter a profession—in this case, teaching— “possesses knowledge and skills in sufficient degree to perform important occupational activities safely and effectively” (American Educational Research Association [AERA], American Psychological Association [APA], National Council on Measurement in Education [NCME], 2014, p. 174). Therefore, in addressing the quality of teaching, a first step should be examining the gateway into the profession—the licensure process. A rationale used by states to first institute teacher licensure assessments was to upgrade the quality of teacher preparation programs (M. T. Kane, Kingsbury, Colton, & Estes, 1989) and, as a result, the quality of teachers.

The *ETS*[®] National Observational Teaching Exam (NOTE) assessment series is a new licensure/certification assessment designed to evaluate a prospective elementary school (kindergarten through 6th grade) teacher’s ability to translate knowledge of content and of teaching into effective practice in the classroom. The NOTE assessment series consists of (a) a performance component that addresses a teacher candidate’s ability to model and explain content, lead group discussions, and elicit and interpret student thinking, and (b) computer-delivered components that address a teacher candidate’s content knowledge for teaching core subjects (English language arts [ELA], mathematics, science, and social studies) at the

elementary grades. Created by Educational Testing Service (ETS) in collaboration with TeachingWorks,¹ NOTE is intended to be a component of a state's initial licensure process for beginning elementary school teachers. As NOTE is a licensure assessment designed specifically for elementary school teachers, it is critical to identify knowledge and skills necessary for effective practice and to demonstrate that the knowledge and skills assessed by NOTE are essential to beginning teaching (Clauser, Margolis, & Case, 2006; AERA et al., 2014). This report discusses the collection of content-related validity evidence to support one feature, the on-demand performance component, of NOTE. Companion reports² address the content-related validity evidence for ELA and mathematics content knowledge for teaching components of the NOTE assessment series.

Licensure and Licensure Assessments

As it refers to public policy, licensure historically denoted the issuance of permission to engage in a particular occupation or profession by a governmental agency for the purpose of protecting the public (Boulet & Zanten, 2014; Schmitt, 1995; Shimberg, 1981). Licensure is a legal requirement to practice and, thus, is a gatekeeper to the profession. Licensure assessments are primarily intended to differentiate between candidates who possess the knowledge and skills required for practice at the time of entry into the profession and those who do not (Clauser et al., 2006; Smith & Hambleton, 1990). A passing score on a state licensure assessment is meant to signify that a candidate has the ability to perform the job of a teacher in a way that ensures the welfare of the public (AERA et al., 2014; Raymond & Luecht, 2013). Laws pertaining to licensure define the range of activities a licensed person may engage in and those activities that are prohibited (Raymond & Luecht, 2013). In most states, licenses are not permanent but rather are granted to individuals for a fixed term and may be renewed pending evidence of professional development work (Danielson, 2006).

Credentialing³ assessments are distinct from assessments meant to screen and select applicants for placement into specific jobs (Mehrens, 1995). Importantly, scores that suggest a candidate is competent do not necessarily mean that a candidate will perform his or her job effectively, and higher scores on a licensure assessment do not necessarily translate to better job performance (Mehrens & Lehmann, 1984). Although personnel selection assessments may measure any knowledge, skill, or ability (KSA) that predicts job success, licensure assessments measure only those KSAs that protect the public (M. T. Kane, 1982; M. Kane, 2004; Raymond

& Luecht, 2013; Shimberg, 1981). As noted by M. Kane (2004), licensure assessments are intended to measure KSAs that may be “necessary but not sufficient for effective practice” (p. 142). Assessments used for selection often rely on predictive validity evidence, whereas licensure assessments often rely primarily on practice analysis or other content-related validation strategies (AERA et al., 2014; M. Kane, 2004; Raymond, 2001; Raymond & Luecht, 2013; AERA et al., 2014).

Content-Related Validity Evidence

Typically, content-related validity evidence is accumulated via expert judgments indicating that the assessment adequately represents the content domain of the occupation or specialty of interest (AERA et al., 2014; M. Kane, 2004; Tannenbaum, Robustelli, & Baron, 2008). A common method for collecting data to support content-related validity is via administering a survey to subject matter experts (SMEs). SMEs are selected based on their knowledge of and experience in the profession (Gael, 1983; Raymond, 2005; Raymond & Luecht, 2013; Rosenfeld & Tannenbaum, 1991) and often include both faculty members and practitioners (Clauser et al., 2006). The survey contains a list of knowledge or skills necessary to perform the activities or responsibilities of the particular job effectively (Raymond, 2001, 2005; Raymond & Neustel, 2006). Surveys are efficient, as they allow for a large number of experts to provide information regarding a large number of knowledge or skill statements in an effective manner across multiple locations (Cascio, 1982; Raymond, 2001, 2005; Raymond & Neustel, 2006). Surveys also provide the opportunity to increase the representation and diversity of professional perspectives into the process of identifying KSAs most important for professional practice (Tannenbaum & Wesley, 1993).

Using SME judgments, test specifications may be derived to represent the occupational domain (Sireci & Sukin, 2013). These analyses are valuable because they represent independent evaluations of what is intended to be measured by the assessment (Sireci & Sukin, 2013). Content-related validity evidence for credentialing should rely on a diverse sample of SMEs who span a variety of work environments and job positions (Raymond, 2001; Raymond & Luecht, 2013). Among other characteristics, factors such as race, ethnicity, gender, urban/rural setting, and geographic location should be considered in selecting an appropriate group of experts for this purpose (Clauser et al., 2006; Tannenbaum & Wesley, 1993).

The Current Study

The purpose of this paper is to present evidence addressing the content-related validity of the performance component of the NOTE assessment series. One viable framework for selecting teaching practices is that of high-leverage practices (HLPs).⁴ HLPs include “tasks and activities that are essential for skillful beginning teachers to understand, take responsibility for, and be prepared to carry out in order to enact their core instructional responsibilities” (Ball & Forzani, 2009, p. 504). Deborah Ball and colleagues at the University of Michigan have proposed a set of 20 HLPs⁵ that forms the basis for the teaching practices examined in the content-related validity study presented here. The list of HLPs appears below.

1. Making content and practices (e.g., specific texts, problems, ideas, theories, processes) explicit through explanation, modeling, representations, and examples
2. Leading a group discussion
3. Eliciting and interpreting individual students’ thinking
4. Establishing norms and routines for classroom discourse and work that are central to the subject matter domain
5. Recognizing particular common patterns of student thinking and development in a subject matter domain
6. Identifying and implementing an instructional response or strategy in response to common patterns of student thinking
7. Teaching a lesson or segment of instruction
8. Implementing organizational routines, procedures, and strategies to support a learning environment
9. Setting up and managing small-group work
10. Engaging in strategic relationship-building conversations with students
11. Learning about students’ cultural, family, intellectual, and personal experiences and resources
12. Setting long- and short-term learning goals for students referenced to external benchmarks
13. Appraising, choosing, and modifying tasks and texts for a specific learning goal
14. Designing a sequence of lessons toward a specific learning goal

15. Selecting and using particular methods to check understanding and monitor student learning during and across lessons
16. Composing, selecting, and interpreting and using information from quizzes, tests, and other methods of summative assessment
17. Providing oral and written feedback to students on their work
18. Communicating about a student with a parent or guardian
19. Analyzing instruction for the purpose of improving it
20. Communicating with other professionals

Consistent with the primary aim of licensure in protecting the public from unqualified practitioners (Clauser et al., 2006; Raymond & Luecht, 2013), the content domain of a licensure assessment should emphasize the knowledge and skills that enable aspiring practitioners to be ready for professional entry and effective practice. Ball and colleagues (Ball & Forzani, 2009; Ball & Hill, 2008) argued that a set of HLPs defines teaching competencies germane to effective teaching, forming the content domain of relevance for this assessment. The current investigation is designed to accumulate evidence from practitioners in the field verifying the necessity of these competencies for elementary school teachers first entering the field. Although the NOTE assessment series only assesses a subset of the HLPs described by TeachingWorks, we will examine the full, longer list of 20 HLPs in this investigation to empirically examine the relevance and importance of each of these practices. Teaching is a complex profession, and no one assessment, or even battery of assessments, can be assumed to address all important knowledge and skills.

Method

Two versions of an online survey were constructed. Each version included two sections. The first asked respondents about HLPs (common across content areas), and the second asked about high-leverage content (specific to a content area). The mathematics version couched judgments about the HLPs in terms of teaching elementary school mathematics and included high-leverage content for teaching mathematics. The ELA version couched judgments about the HLPs in terms of teaching elementary school ELA and included high-leverage content for teaching ELA. Elementary school teachers and teacher preparation faculty were assigned to one of the two versions of the survey. The components of the survey that focused on judgments regarding the HLPs will be described later in this paper.

Survey

For each of the 20 HLPs, we posed these three content-related validity questions to participants:

1. Is the practice relevant to a beginning elementary school teacher's ability to be a safe and effective educator?
2. If this practice is relevant, how important is it to a beginning elementary school teacher's ability to be a safe and effective educator?
3. How frequently is this practice applied by beginning elementary school teachers when teaching?

If educators indicated an HLP is relevant, they then rated the importance and frequency of the HLP (using a 6-point judgment scale). Therefore, importance and frequency ratings were only collected from respondents who judged the HLP as relevant.

Following the relevance, importance, and frequency judgments for the separate HLPs, participants judged the five least and five most important HLPs; a participant could not select the same HLP for each category.

Sample

Working with a mailing list of 8,841 educators obtained from a national educational marketing firm, a multiphase outreach effort was conducted. The mailing list was sampled from a much larger, national database of teachers and teacher preparation faculty. The intent was to sample elementary school teachers and college faculty who prepare elementary school teachers. Sampling twice as many teachers was intended to result in a significant number of teachers currently teaching lower (kindergarten to Grade 3) and upper (Grades 4 through 6) elementary classes as well as oversampling Black/African American and Hispanic/Latino teachers. The sample included an approximately equal number of teachers from each of the four United States Census regions. We also sought to oversample faculty from minority-serving institutions.

Two versions of an online survey were constructed and respondents were contacted, via e-mail and letter, to invite them to complete one of the surveys. Participants were paid \$25 in exchange for their participation in the study. Three rounds of e-mail follow-ups occurred during the data collection period. The assignment of teachers to one of the two versions of the survey was dependent on their current teaching assignments. Teachers who only taught mathematics

were assigned to the mathematics version, and teachers who only taught ELA were assigned to the ELA version. Teachers who indicated they taught ELA and mathematics (more than 75% of the sample) or neither (approximately 4% of the sample) were randomly assigned to one of the two versions. Faculty also were randomly assigned to one of the two versions.

Of the original 8,841 educators contacted, 700 e-mails were not deliverable. Therefore, the number of educators successfully contacted was 8,141. Of these, 607 (or 7.5%) completed one of the two versions of the survey. An additional 31 educators were forwarded the survey by colleagues and completed it. In total, 638 educators completed either the mathematics or ELA version of the online survey.

Of the respondents, 387 (or 61%) indicated they were teachers (pre-K to Grade 12) and 202 (or 32%) indicated they were college faculty.⁶ The remaining 49 (or 8%) respondents indicated they were administrators, held other education-related positions, or preferred not to provide information regarding their current position. Given the purpose of the survey, the 49 respondents who did not indicate they were teachers or faculty were removed from the sample.

The resulting sample—currently licensed teachers and college faculty currently preparing elementary school teacher candidates—included 569 respondents (385 teachers and 184 college faculty) across the two versions of the survey.

While the overall response rate⁷ for the survey was 7.5%, the resulting sample of elementary school teachers does reflect the composition of the national population when compared to the National Center for Education Statistics (NCES) 2011-12 School and Staffing Survey (SASS; Goldring, Gray, & Bitterman, 2013). The sample of teachers slightly overrepresents the percent of Black/African American (11.5% in the sample compared to 7.1% nationally) and Hispanic/Latino (9.3% in the sample compared to 8.7% nationally) teachers compared to the latest SASS results (Goldring et al., 2013). The sample approximately mirrors elementary school teachers nationally in terms of years of experience, with approximately 40% of teachers with less than 10 years of experience and another 40% with more than 15 years of experience, and gender, with approximately 90% of elementary school teachers being female (Goldring et al., 2013).

For the following analyses, samples of 279 participants completed the ELA version (197 teachers and 82 faculty) and 290 participants completed the mathematics version (188 teachers and 102 faculty). Table 1 provides a summary of background information for the samples

overall. Tables 2 and 3 provide specific background information for teachers and faculty, respectively. Given the descriptive nature of the results presented for this study, the sample sizes for both versions are within the 200 to 400 respondents viewed as adequate for generalizable findings (M. T. Kane, Miller, Trine, Becker, & Carson, 1995).

Table 1. Background Information—Overall Samples—for English Language Arts (ELA) and Mathematics

Background information	ELA	Mathematics
Current position		
Teacher	197 (71%)	188 (65%)
Faculty	82 (29%)	102 (35%)
Gender		
Female	237 (85%)	243 (84%)
Male	34 (12%)	39 (13%)
Other/prefer not to answer	8 (3%)	8 (3%)
Race/ethnicity		
American Indian or Alaska Native	2 (1%)	3 (1%)
Asian	3 (1%)	7 (2%)
Black/African American	29 (10%)	35 (12%)
Hispanic/Latino	27 (10%)	25 (9%)
Native Hawaiian or other Pacific Islander	1 (0%)	0 (0%)
White	199 (71%)	196 (68%)
Two or more races	4 (1%)	8 (3%)
Other/prefer not to answer/missing	14 (5%)	16 (6%)
Geographic region		
Northeast	51 (18%)	55 (19%)
Midwest	77 (28%)	81 (28%)
South	94 (34%)	97 (33%)
West	57 (20%)	57 (20%)

Table 2. Teacher-Specific Background Information for English Language Arts (ELA) and Mathematics

Background information	ELA	Mathematics
Current teaching assignment		
Lower elementary (Grades K–3)	103 (52%)	108 (57%)
Upper elementary (Grades 4–6)	75 (38%)	59 (31%)
Other ^a	19 (10%)	21 (11%)
Years of experience		
3 years or less	10 (5%)	26 (14%)
4 to 9 years	40 (20%)	75 (40%)
10 to 14 years	44 (22%)	36 (19%)
15 years or more	102 (52%)	50 (27%)
Other/missing	1 (1%)	1 (1%)
Mentored student teachers		
Yes	96 (49%)	85 (45%)
No	99 (50%)	102 (54%)
Missing	2 (1%)	1 (1%)
Type of School		
Public (noncharter)	173 (88%)	170 (90%)
Public (charter)	12 (6%)	5 (3%)
Private	11 (6%)	13 (7%)
Missing	1 (1%)	0 (0%)
School location		
Urban	67 (34%)	68 (36%)
Suburban	82 (42%)	79 (42%)
Rural	48 (24%)	41 (22%)

^aTeachers who taught across the elementary level (lower and upper) or across elementary and secondary grades.

Table 3. Faculty-Specific Background Information for English Language Arts (ELA) and Mathematics

Background information	ELA	Mathematics
Years of experience		
3 years or less	6 (7%)	9 (9%)
4 to 9 years	24 (29%)	29 (28%)
10 to 14 years	24 (29%)	19 (19%)
15 years or more	28 (34%)	45 (44%)
Supervised student teachers		
Yes	49 (60%)	81 (79%)
No	32 (39%)	21 (21%)
Missing	1 (1%)	0 (0%)
Minority-serving institution		
Yes	25 (30%)	27 (26%)
No	53 (65%)	72 (71%)
Designation not available	4 (5%)	3 (3%)
Institution location		
Urban	26 (32%)	37 (36%)
Suburban	31 (38%)	34 (33%)
Rural	25 (30%)	31 (30%)

Analysis

The analyses in this report were focused in two main categories regarding the content-related validity evidence collected from the online survey. The first category was to describe patterns in average relevance and importance judgment ratings in various ways, starting with all participants (teachers and faculty), and then making comparisons between teachers and faculty, lower elementary and upper elementary teachers, teachers across race/ethnicity groups, and teachers across geographic regions. The sample sizes for Black/African American and Hispanic/Latino participants (across content areas, 64 and 52, respectively) are relatively small, but whether the importance judgments of these subgroups support or run counter to the finding for the overall sample will be highlighted. Any differences across the four census regions will also be highlighted.

Comparisons were made using effect sizes (Cohen, 1988) where the mean difference between two groups was divided by a combination of group sample sizes and standard deviations. In the case of race/ethnicity, White teachers were used as the reference group, and in the case of geographic region, Northeast teachers were used as the reference group. The second category of analyses was to index agreement between relevance and importance using intraclass correlations (ICC[2]; Shrout & Fleiss, 1979).

Results

Relevance Judgments

Across the 20 HLPs and for both ELA and mathematics, the overwhelming majority of educators agreed that the HLPs are relevant for effective practice for beginning elementary school teachers. The percentage of educators judging each HLP as relevant ranged from 93% to 100%; the judgments were similar for teachers and faculty, and for ELA and mathematics.

The ICC(2) (Shrout & Fleiss, 1979) indexing agreement among educators regarding their relevance ratings across the 20 HLPs is 0.71 (95% CI [.49, .86]) for ELA. For mathematics, after removing 14 individuals who were missing ratings for over half of the 20 HLPs, the ICC(2) for the remaining 276 respondents is 0.60 (95% CI [.14, .88]).

Importance Judgments

Table 4 summarizes educators' judgments regarding the importance of each HLP by subject area. Results are presented for teachers, faculty, and the total sample. Comparing

teachers and faculty, the differences in importance judgments were minor; the average absolute difference was approximately 0.15 for both ELA and mathematics (on the 6-point judgment scale). Therefore, summaries across HLPs and between subject areas will focus on the total sample.⁸

Table 4. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Subject Area for English Language Arts (ELA) and Mathematics

HLP	ELA			Mathematics		
	Teachers	Faculty	Overall	Teachers	Faculty	Overall
1 ^a	5.36 (0.75)	5.23 (0.73)	5.32 (0.74)	5.27 (0.65)	5.30 (0.71)	5.28 (0.67)
2 ^a	5.10 (0.80)	5.22 (0.72)	5.14 (0.78)	5.14 (0.71)	5.00 (0.82)	5.09 (0.76)
3 ^a	5.36 (0.70)	5.55 (0.52)	5.41 (0.66)	5.22 (0.75)	5.37 (0.74)	5.27 (0.75)
4	5.35 (0.75)	5.10 (0.80)	5.27 (0.77)	5.33 (0.77)	4.95 (0.87)	5.19 (0.83)
5	4.87 (0.74)	5.00 (0.86)	4.91 (0.78)	5.01 (0.77)	5.04 (0.97)	5.02 (0.85)
6	5.03 (0.74)	5.11 (0.68)	5.05 (0.73)	4.88 (0.80)	5.16 (0.84)	4.99 (0.82)
7	5.44 (0.71)	5.52 (0.55)	5.47 (0.67)	5.36 (0.72)	5.54 (0.67)	5.43 (0.71)
8	5.61 (0.65)	5.48 (0.63)	5.57 (0.65)	5.58 (0.63)	5.21 (0.92)	5.46 (0.75)
9	5.13 (0.88)	5.21 (0.72)	5.16 (0.84)	5.15 (0.79)	5.10 (0.79)	5.13 (0.79)
10	5.31 (0.79)	5.43 (0.71)	5.34 (0.77)	5.22 (0.76)	5.32 (0.72)	5.25 (0.75)
11	4.94 (0.87)	5.31 (0.77)	5.05 (0.86)	5.01 (0.92)	5.16 (0.88)	5.06 (0.91)
12	5.04 (0.86)	5.04 (0.77)	5.04 (0.84)	5.16 (0.77)	5.07 (0.90)	5.13 (0.82)
13	5.11 (0.76)	5.26 (0.75)	5.15 (0.76)	5.18 (0.72)	5.30 (0.70)	5.22 (0.72)
14	5.27 (0.73)	5.36 (0.62)	5.29 (0.70)	5.26 (0.73)	5.44 (0.67)	5.32 (0.72)
15	5.36 (0.75)	5.59 (0.57)	5.43 (0.71)	5.34 (0.75)	5.40 (0.72)	5.36 (0.74)
16	5.16 (0.93)	5.24 (0.71)	5.19 (0.87)	5.26 (0.75)	5.24 (0.90)	5.25 (0.80)
17	5.20 (0.79)	5.41 (0.67)	5.27 (0.76)	5.17 (0.82)	5.39 (0.78)	5.24 (0.81)
18	5.39 (0.78)	5.30 (0.66)	5.36 (0.74)	5.45 (0.74)	5.20 (0.80)	5.37 (0.76)
19	5.10 (0.85)	5.50 (0.64)	5.22 (0.81)	5.22 (0.78)	5.34 (0.66)	5.26 (0.74)
20	5.18 (0.81)	5.04 (0.86)	5.14 (0.82)	5.11 (0.78)	4.98 (0.90)	5.07 (0.82)
Minimum	4.87	5.00	4.91	4.88	4.95	4.99
Maximum	5.61	5.59	5.57	5.58	5.54	5.46
Sample size	182–196	76–82	261–278	169–179	85–102	255–281

Note. Importance scale: 1 (*not at all important*), 2 (*of little importance*), 3 (*of some importance*), 4 (*moderately important*), 5 (*very important*), 6 (*extremely important*). Respondents who judged the practice not relevant are not included in the calculation of the average importance judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

In defining HLPs, we are identifying practices that cut across content areas. Therefore, in examining the relevance and importance of the 20 HLPs identified in the literature, we examined the consistency across subject areas (ELA and mathematics) at the elementary grade level. The average importance judgment for all 20 HLPs and for both subject areas was approximately 5.0 or higher (4.91–5.57 for ELA and 4.99–5.46 for mathematics), well above 4.2 (on the 6-point judgment scale). Research by Tannenbaum and Rosenfeld (1994) recommended that an average importance judgment of 3.5 on a 5-point scale was sufficient to determine importance for

licensure. Translating this finding to a 6-point scale would result in a threshold of 4.2. The absolute difference in average importance judgments between the two subject areas was 0.15 or less on the 6-point scale.

Tables 5 and 6 compare HLP importance ratings between lower elementary teachers (K to 3) and upper elementary teachers (Grades 4 to 6) in mathematics and ELA, respectively. In both content areas, for 13 of the 20 HLPs, effect sizes were below 0.20. Among the other seven statements with effect sizes above 0.20 in mathematics and all but one in ELA, lower elementary teachers on average assigned higher importance ratings to these statements than upper elementary teachers. The three HLPs with effect sizes above 0.20 in both subject areas were HLP 2 (Leading a group discussion), HLP 8 (Implementing organizational routines, procedures, and strategies to support a learning environment), and HLP 9 (Setting up and managing small-group work).

Table 5. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Current Grade Level Taught for Mathematics

HLP	Lower (K–3)	Upper (4–6)	Difference
1 ^a	5.27 (0.60)	5.24 (0.73)	0.03 (0.05)
2 ^a	5.23 (0.66)	5.00 (0.76)	0.23 (0.34)
3 ^a	5.23 (0.78)	5.24 (0.70)	0.02 (0.02)
4	5.34 (0.72)	5.40 (0.77)	0.05 (0.07)
5	5.03 (0.74)	4.93 (0.86)	0.10 (0.13)
6	4.81 (0.80)	4.96 (0.81)	0.15 (0.19)
7	5.35 (0.70)	5.30 (0.80)	0.05 (0.07)
8	5.64 (0.61)	5.52 (0.63)	0.13 (0.21)
9	5.19 (0.72)	5.15 (0.86)	0.04 (0.05)
10	5.31 (0.69)	5.15 (0.86)	0.16 (0.21)
11	5.02 (0.97)	5.02 (0.84)	0.00 (0.00)
12	5.21 (0.74)	5.04 (0.82)	0.17 (0.22)
13	5.22 (0.73)	5.17 (0.71)	0.05 (0.07)
14	5.27 (0.75)	5.25 (0.72)	0.02 (0.02)
15	5.40 (0.68)	5.27 (0.84)	0.13 (0.17)
16	5.26 (0.78)	5.30 (0.72)	0.04 (0.05)
17	5.17 (0.77)	5.22 (0.96)	0.05 (0.06)
18	5.53 (0.71)	5.33 (0.79)	0.20 (0.27)
19	5.28 (0.80)	5.09 (0.83)	0.19 (0.23)
20	5.21 (0.72)	4.91 (0.84)	0.30 (0.39)
Minimum	4.81	4.91	0.00
Maximum	5.64	5.52	0.30
Sample size	97–105	52–58	

Note. Importance scale: 1 (*not at all important*), 2 (*of little importance*), 3 (*of some importance*), 4 (*moderately important*), 5 (*very important*), 6 (*extremely important*). Respondents who judged the practice not relevant are not included in the calculation of the average importance judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

Table 6. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Current Grade Level Taught for English Language Arts (ELA)

HLP	Lower (K–3)	Upper (4–6)	Difference
1 ^a	5.41 (0.62)	5.33 (0.90)	0.08 (0.11)
2 ^a	5.20 (0.73)	5.00 (0.90)	0.20 (0.25)
3 ^a	5.34 (0.70)	5.42 (0.75)	0.09 (0.12)
4	5.24 (0.75)	5.44 (0.76)	0.19 (0.26)
5	4.93 (0.68)	4.78 (0.83)	0.15 (0.21)
6	5.03 (0.74)	5.01 (0.74)	0.02 (0.02)
7	5.46 (0.67)	5.42 (0.79)	0.04 (0.06)
8	5.67 (0.51)	5.49 (0.83)	0.18 (0.28)
9	5.24 (0.80)	5.01 (0.97)	0.22 (0.25)
10	5.29 (0.74)	5.29 (0.89)	0.00 (0.00)
11	4.95 (0.78)	4.88 (1.02)	0.07 (0.08)
12	5.06 (0.81)	5.00 (0.96)	0.06 (0.07)
13	5.21 (0.67)	4.96 (0.90)	0.25 (0.33)
14	5.27 (0.66)	5.25 (0.83)	0.02 (0.03)
15	5.38 (0.69)	5.35 (0.83)	0.02 (0.03)
16	5.09 (0.96)	5.22 (0.95)	0.13 (0.14)
17	5.18 (0.72)	5.21 (0.92)	0.03 (0.04)
18	5.47 (0.68)	5.25 (0.90)	0.22 (0.28)
19	5.07 (0.84)	5.07 (0.92)	0.00 (0.00)
20	5.17 (0.80)	5.19 (0.84)	0.02 (0.02)
Minimum	4.93	4.78	0.00
Maximum	5.67	5.49	0.25
Sample size	95–103	67–74	

Note. Importance scale: 1 (*not at all important*), 2 (*of little importance*), 3 (*of some importance*), 4 (*moderately important*), 5 (*very important*), 6 (*extremely important*). Respondents who judged the practice not relevant are not included in the calculation of the average importance judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

Importance ratings by teacher race/ethnicity are disaggregated in Tables 7 and 8 for mathematics and ELA, respectively. The average importance judgments for Black/African American teachers were above 5.0 for all HLPs except ELA HLP 6 (average = 4.95); for Hispanic/Latino teachers, the average was above 5.0 for all HLPs except ELA HLP 11 (average = 4.79), Math HLP 6 (average = 4.83), and Math HLP 11 (average = 4.89). Using White teachers as a reference group, importance ratings were lower for HLP 12 (Setting long- and short-term learning goals for students referenced to external benchmarks) compared to Black/African American and Hispanic/Latino teachers in both content areas with effect sizes above 0.20. In mathematics, relative to Black/African American teachers, the difference was 0.17 (ES = 0.22), and relative to Hispanic/Latino teachers, the difference was 0.47 (ES = 0.59). In ELA, the relative differences were 0.47 (ES = 0.53) and 0.19 (ES = 0.22), compared to Black/African American and Hispanic/Latino teachers, respectively.

Table 7. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Race/Ethnicity for Mathematics

HLP	Black/African American	Hispanic/Latino	White	Overall
1 ^a	5.14 (0.64)	5.40 (0.60)	5.24 (0.68)	5.24 (0.66)
2 ^a	5.29 (0.56)	5.29 (0.78)	5.09 (0.73)	5.14 (0.72)
3 ^a	5.42 (0.51)	5.19 (0.75)	5.18 (0.81)	5.21 (0.77)
4	5.41 (0.67)	5.45 (0.67)	5.29 (0.83)	5.33 (0.79)
5	5.05 (0.50)	5.20 (0.62)	5.00 (0.83)	5.03 (0.77)
6	5.24 (0.54)	4.83 (0.99)	4.83 (0.82)	4.88 (0.82)
7	5.45 (0.51)	5.09 (0.87)	5.39 (0.75)	5.36 (0.74)
8	5.52 (0.60)	5.75 (0.44)	5.60 (0.62)	5.61 (0.59)
9	5.19 (0.68)	5.28 (0.75)	5.10 (0.84)	5.13 (0.81)
10	5.38 (0.50)	5.25 (0.79)	5.17 (0.83)	5.21 (0.78)
11	5.48 (0.51)	4.89 (1.18)	4.92 (0.94)	4.99 (0.94)
12	5.26 (0.45)	5.56 (0.62)	5.09 (0.80)	5.17 (0.76)
13	5.24 (0.62)	5.22 (0.81)	5.13 (0.74)	5.15 (0.73)
14	5.53 (0.51)	5.37 (0.68)	5.20 (0.75)	5.26 (0.73)
15	5.30 (0.57)	5.32 (0.67)	5.32 (0.81)	5.32 (0.77)
16	5.47 (0.61)	5.26 (0.73)	5.17 (0.79)	5.22 (0.76)
17	5.29 (0.72)	5.21 (0.63)	5.09 (0.88)	5.13 (0.83)
18	5.45 (0.60)	5.40 (0.99)	5.45 (0.71)	5.44 (0.73)
19	5.40 (0.60)	5.25 (0.91)	5.17 (0.80)	5.21 (0.79)
20	5.05 (0.67)	5.16 (0.69)	5.19 (0.74)	5.17 (0.72)
Minimum	5.05	4.83	4.83	4.88
Maximum	5.53	5.75	5.60	5.61
Sample size	19–22	18–22	111–119	150–161

Note. Importance scale: 1 (*not at all important*), 2 (*of little importance*), 3 (*of some importance*), 4 (*moderately important*), 5 (*very important*), 6 (*extremely important*). Respondents who judged the practice not relevant are not included in the calculation of the average importance judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

Table 8. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Race/Ethnicity for English Language Arts (ELA)

HLP	Black/African American	Hispanic/Latino	White	Overall
1 ^a	5.75 (0.44)	5.30 (0.82)	5.34 (0.71)	5.38 (0.71)
2 ^a	5.45 (0.60)	5.13 (0.76)	5.10 (0.80)	5.14 (0.78)
3 ^a	5.67 (0.48)	5.17 (0.83)	5.35 (0.71)	5.37 (0.72)
4	5.52 (0.87)	5.38 (0.67)	5.32 (0.76)	5.35 (0.76)
5	5.11 (0.74)	5.17 (0.72)	4.78 (0.75)	4.87 (0.76)
6	4.95 (0.85)	5.09 (0.60)	5.05 (0.78)	5.04 (0.76)
7	5.52 (0.60)	5.42 (0.50)	5.45 (0.74)	5.45 (0.70)
8	5.81 (0.51)	5.67 (0.70)	5.60 (0.63)	5.64 (0.62)
9	5.30 (0.92)	5.17 (0.78)	5.09 (0.90)	5.12 (0.88)
10	5.26 (0.99)	5.29 (0.69)	5.32 (0.78)	5.31 (0.79)
11	5.14 (1.06)	4.79 (0.66)	4.93 (0.87)	4.94 (0.87)
12	5.45 (0.76)	5.17 (0.70)	4.98 (0.90)	5.06 (0.87)
13	5.40 (0.68)	5.23 (0.69)	5.04 (0.76)	5.10 (0.75)
14	5.38 (0.67)	5.43 (0.59)	5.25 (0.73)	5.29 (0.70)
15	5.52 (0.75)	5.48 (0.79)	5.33 (0.74)	5.37 (0.74)
16	5.62 (0.74)	5.29 (0.69)	5.11 (0.95)	5.19 (0.91)
17	5.48 (0.93)	5.17 (0.70)	5.19 (0.77)	5.22 (0.78)
18	5.57 (0.98)	5.63 (0.49)	5.32 (0.78)	5.39 (0.78)
19	5.33 (0.97)	5.21 (0.72)	5.06 (0.84)	5.11 (0.84)
20	5.35 (0.75)	5.25 (0.68)	5.12 (0.84)	5.17 (0.81)
Minimum	4.95	4.79	4.78	4.87
Maximum	5.81	5.67	5.60	5.64
Sample size	19–21	21–24	123–134	165–179

Note. Importance scale: 1 (*not at all important*), 2 (*of little importance*), 3 (*of some importance*), 4 (*moderately important*), 5 (*very important*), 6 (*extremely important*). Respondents who judged the practice not relevant are not included in the calculation of the average importance judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

Importance ratings by geographic region are disaggregated in Tables 9 and 10. For mathematics, average judgments across regions were above 5.0 except for HLP 6 (average = 4.88), whereas for ELA, all were above 5.0 except for HLP 5 (average = 4.87) and HLP 11 (average = 4.94). Using Northeast teachers as a reference group, in mathematics, importance ratings were higher than those for teachers in other regions for HLP 2 (ES range = 0.45–0.79), HLP 6 (ES range = 0.26–0.42), and HLP 9 (ES range = 0.27–0.60). For ELA, importance ratings were higher than those for teachers in other regions only for HLP 7 (ES range = 0.24–0.36).

Respondents only made ratings for importance if they indicated an HLP was relevant, which resulted in some missing cases. To compute ICC(2), we made the assumption that if an HLP was not considered relevant, it would also be considered not at all important. Therefore, we imputed a value of 1 for missing cases. The ICC(2) (Shrout & Fleiss, 1979) indexing agreement among educators regarding their importance ratings across the 20 HLPs is 0.94 (95% CI [0.90, 0.97]) for ELA and 0.89 (95% CI [0.81, 0.95]) for mathematics.

Table 9. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Geographic Region for Mathematics

HLP	Northeast	Midwest	South	West	Overall
1 ^a	5.44 (0.66)	5.11 (0.60)	5.29 (0.66)	5.31 (0.66)	5.27 (0.65)
2 ^a	5.47 (0.62)	4.91 (0.75)	5.18 (0.65)	5.12 (0.71)	5.14 (0.71)
3 ^a	5.39 (0.70)	5.13 (0.71)	5.27 (0.71)	5.07 (0.92)	5.22 (0.75)
4	5.44 (0.79)	5.26 (0.73)	5.29 (0.84)	5.39 (0.69)	5.33 (0.77)
5	5.12 (0.70)	4.90 (0.71)	5.08 (0.77)	4.89 (0.92)	5.01 (0.77)
6	5.09 (0.68)	4.81 (0.79)	4.90 (0.79)	4.75 (0.93)	4.88 (0.80)
7	5.47 (0.66)	5.24 (0.80)	5.38 (0.71)	5.41 (0.69)	5.36 (0.72)
8	5.55 (0.62)	5.51 (0.63)	5.68 (0.62)	5.57 (0.63)	5.58 (0.63)
9	5.39 (0.56)	4.94 (0.86)	5.20 (0.80)	5.11 (0.83)	5.15 (0.79)
10	5.24 (0.71)	5.19 (0.73)	5.18 (0.83)	5.32 (0.77)	5.22 (0.76)
11	5.09 (0.91)	5.04 (0.88)	4.98 (0.91)	4.93 (1.07)	5.01 (0.92)
12	5.39 (0.80)	5.02 (0.80)	5.11 (0.73)	5.25 (0.75)	5.16 (0.77)
13	5.09 (0.68)	5.27 (0.63)	5.11 (0.84)	5.29 (0.66)	5.18 (0.72)
14	5.36 (0.78)	5.24 (0.75)	5.26 (0.68)	5.14 (0.76)	5.26 (0.73)
15	5.38 (0.82)	5.26 (0.78)	5.33 (0.66)	5.46 (0.79)	5.34 (0.75)
16	5.27 (0.83)	5.22 (0.71)	5.25 (0.73)	5.34 (0.81)	5.26 (0.75)
17	5.18 (0.80)	5.02 (0.90)	5.18 (0.83)	5.43 (0.57)	5.17 (0.82)
18	5.50 (0.75)	5.41 (0.71)	5.52 (0.68)	5.33 (0.88)	5.45 (0.74)
19	5.13 (0.94)	5.11 (0.78)	5.37 (0.67)	5.21 (0.77)	5.22 (0.78)
20	5.03 (0.85)	5.08 (0.74)	5.16 (0.81)	5.17 (0.70)	5.11 (0.78)
Minimum	5.03	4.81	4.90	4.75	4.88
Maximum	5.55	5.51	5.68	5.57	5.58
Sample size	30–34	49–55	57–63	26–30	169–179

Note. Importance scale: 1 (*not at all important*), 2 (*of little importance*), 3 (*of some importance*), 4 (*moderately important*), 5 (*very important*), 6 (*extremely important*). Respondents who judged the practice not relevant are not included in the calculation of the average importance judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

Table 10. Summary of Importance Judgments for High-Leverage Practices (HLPs) by Geographic Region for English Language Arts (ELA)

HLP	Northeast	Midwest	South	West	Overall
1 ^a	5.32 (0.77)	5.28 (0.82)	5.40 (0.76)	5.41 (0.63)	5.36 (0.75)
2 ^a	5.00 (0.79)	4.98 (0.92)	5.18 (0.79)	5.23 (0.62)	5.10 (0.80)
3 ^a	5.25 (0.81)	5.32 (0.73)	5.38 (0.69)	5.46 (0.60)	5.36 (0.70)
4	5.28 (0.78)	5.30 (0.77)	5.33 (0.77)	5.49 (0.68)	5.35 (0.75)
5	4.94 (0.66)	4.57 (0.84)	4.97 (0.71)	5.03 (0.63)	4.87 (0.74)
6	5.17 (0.79)	4.92 (0.78)	5.00 (0.72)	5.11 (0.69)	5.03 (0.74)
7	5.61 (0.60)	5.43 (0.80)	5.41 (0.64)	5.36 (0.79)	5.44 (0.71)
8	5.62 (0.72)	5.51 (0.81)	5.61 (0.55)	5.71 (0.46)	5.61 (0.65)
9	5.19 (0.79)	4.96 (0.96)	5.26 (0.87)	5.12 (0.87)	5.13 (0.88)
10	5.35 (0.77)	5.28 (0.92)	5.32 (0.74)	5.29 (0.72)	5.31 (0.79)
11	5.00 (0.79)	4.78 (0.86)	5.03 (0.91)	4.95 (0.88)	4.94 (0.87)
12	5.08 (0.81)	5.02 (0.93)	5.08 (0.86)	4.95 (0.86)	5.04 (0.86)
13	5.06 (0.83)	5.08 (0.88)	5.21 (0.61)	5.02 (0.76)	5.11 (0.76)
14	5.42 (0.77)	5.06 (0.83)	5.38 (0.61)	5.24 (0.69)	5.27 (0.73)
15	5.27 (0.80)	5.34 (0.83)	5.53 (0.59)	5.21 (0.78)	5.36 (0.75)
16	5.19 (0.89)	5.04 (1.05)	5.38 (0.76)	4.98 (1.00)	5.16 (0.93)
17	5.19 (0.74)	5.23 (0.91)	5.25 (0.68)	5.12 (0.84)	5.20 (0.79)
18	5.32 (0.78)	5.27 (0.87)	5.45 (0.77)	5.50 (0.63)	5.39 (0.78)
19	5.09 (0.79)	4.88 (1.01)	5.19 (0.69)	5.22 (0.88)	5.10 (0.85)
20	5.27 (0.73)	5.06 (0.92)	5.20 (0.78)	5.24 (0.77)	5.18 (0.81)
Minimum	4.94	4.57	4.97	4.95	4.87
Maximum	5.62	5.51	5.61	5.71	5.61
Sample size	33–37	49–55	60–63	38–42	182–196

Note. Importance scale: 1 (*not at all important*), 2 (*of little importance*), 3 (*of some importance*), 4 (*moderately important*), 5 (*very important*), 6 (*extremely important*). Respondents who judged the practice not relevant are not included in the calculation of the average importance judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

Frequency

In addition to judgments regarding the relevance and importance of the HLPs for effective beginning practice, participants in the survey judged how frequently an HLP is applied by beginning elementary school teachers when teaching ELA or mathematics. Participants used a 6-point scale ranging from 1 (*never*) to 6 (*very frequently*). See Table 11.

In all cases except for one HLP in mathematics (HLP 12: Setting long- and short-term learning goals for students referenced to external benchmarks), more teachers judged an HLP as frequently applied or more frequently applied when compared to faculty. However, the absolute difference in judgments exceeded 0.50 (on a 6-point scale) only for two HLPs for ELA (HLP 4: Establishing norms and routines for classroom discourse and work that are central to the subject matter domain and HLP 18: Communicating about a student with a parent or guardian) and one HLP for mathematics (HLP 2: Leading a group discussion).

Table 11. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Subject Area for English Language Arts (ELA) and Mathematics

HLP	ELA			Mathematics		
	Teachers	Faculty	Overall	Teachers	Faculty	Overall
1 ^a	5.05 (0.83)	4.90 (0.73)	5.01 (0.81)	5.04 (0.77)	4.84 (0.77)	4.97 (0.77)
2 ^a	4.93 (0.82)	4.74 (0.90)	4.88 (0.85)	4.96 (0.81)	4.35 (0.81)	4.73 (0.86)
3 ^a	4.92 (0.92)	4.62 (0.96)	4.83 (0.94)	4.88 (0.78)	4.48 (0.87)	4.73 (0.84)
4	5.02 (0.85)	4.44 (0.97)	4.85 (0.92)	5.06 (0.73)	4.58 (0.86)	4.89 (0.81)
5	4.59 (0.81)	4.41 (1.02)	4.54 (0.88)	4.67 (0.85)	4.29 (0.93)	4.54 (0.89)
6	4.65 (0.93)	4.42 (0.97)	4.59 (0.95)	4.64 (0.89)	4.43 (0.85)	4.57 (0.88)
7	5.16 (0.84)	4.95 (0.87)	5.10 (0.85)	5.09 (0.74)	5.02 (0.72)	5.06 (0.74)
8	5.33 (0.81)	5.12 (0.81)	5.27 (0.82)	5.29 (0.80)	5.02 (0.71)	5.20 (0.78)
9	4.86 (0.82)	4.65 (0.96)	4.80 (0.87)	4.85 (0.88)	4.67 (0.89)	4.79 (0.89)
10	4.85 (0.91)	4.67 (0.96)	4.80 (0.93)	4.85 (0.90)	4.63 (0.85)	4.77 (0.89)
11	4.42 (0.87)	4.35 (0.94)	4.40 (0.89)	4.57 (0.92)	4.43 (0.94)	4.52 (0.93)
12	4.61 (1.06)	4.56 (0.90)	4.60 (1.01)	4.80 (0.89)	4.83 (0.85)	4.81 (0.87)
13	4.72 (0.91)	4.38 (1.04)	4.62 (0.97)	4.82 (0.86)	4.52 (0.89)	4.72 (0.88)
14	4.90 (0.93)	4.84 (0.87)	4.88 (0.91)	4.93 (0.83)	4.88 (0.80)	4.91 (0.81)
15	5.05 (0.94)	4.95 (0.84)	5.02 (0.91)	5.07 (0.85)	4.94 (0.86)	5.03 (0.85)
16	4.96 (0.91)	4.66 (1.01)	4.87 (0.95)	4.99 (0.82)	4.66 (1.02)	4.88 (0.90)
17	4.85 (0.87)	4.55 (1.00)	4.76 (0.92)	4.77 (0.93)	4.70 (0.93)	4.74 (0.93)
18	4.89 (0.84)	4.34 (0.96)	4.72 (0.91)	4.84 (0.93)	4.38 (0.86)	4.68 (0.93)
19	4.54 (1.09)	4.29 (1.13)	4.46 (1.11)	4.57 (0.98)	4.38 (1.07)	4.51 (1.01)
20	4.76 (1.03)	4.26 (0.94)	4.61 (1.02)	4.87 (0.96)	4.45 (1.01)	4.73 (0.99)
Minimum	4.42	4.26	4.40	4.57	4.29	4.51
Maximum	5.33	5.12	5.27	5.29	5.02	5.20
Sample size	182–196	76–82	261–278	170–179	85–102	256–281

Note. Frequency scale: 1 (*never*), 2 (*very rarely*), 3 (*rarely*), 4 (*occasionally*), 5 (*frequently*), 6 (*very frequently*). Respondents who judged the practice not relevant are not included in the calculation of the average frequency judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

Across the two subject areas, the average frequency judgments ranged from 4.40–5.27 for ELA and 4.51–5.20 for mathematics (where 4 = occasionally and 5 = frequently). The judgment of the participants surveyed supports the prevalence of the HLPs in the practice of beginning elementary school teachers. Tables 12 and 13 compare HLP frequency ratings between lower elementary teachers (K to Grade 3) and upper elementary teachers (Grades 4 to 6) in each subject area. In mathematics (Table 12), 11 of the 20 HLPs had effect sizes below 0.20, whereas for the other nine HLPs, all differences reflected higher means among lower elementary teachers (ES range = 0.21–0.43). In ELA (Table 13), all but four HLPs had effect sizes below 0.20. Two of these showed higher means among lower elementary teachers: HLP 9 (difference = 0.17; ES = 0.21) and HLP 18 (difference = 0.24; ES = 0.29). Two others showed higher means among upper elementary teachers: HLP 11 (difference = 0.23; ES = 0.27) and HLP 19 (difference = 0.25; ES = 0.24).

Table 12. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Current Grade Level Taught for Mathematics

HLP	Lower (K–3)	Upper (4–6)	Difference
1 ^a	5.08 (0.78)	4.91 (0.80)	0.16 (0.21)
2 ^a	5.02 (0.78)	4.83 (0.87)	0.19 (0.23)
3 ^a	4.88 (0.79)	4.81 (0.83)	0.07 (0.08)
4	5.10 (0.71)	5.08 (0.73)	0.02 (0.03)
5	4.68 (0.86)	4.70 (0.82)	0.02 (0.02)
6	4.63 (0.83)	4.66 (0.96)	0.03 (0.04)
7	5.04 (0.77)	5.14 (0.72)	0.10 (0.14)
8	5.33 (0.83)	5.25 (0.79)	0.08 (0.09)
9	4.95 (0.85)	4.77 (0.90)	0.18 (0.21)
10	4.87 (0.88)	4.87 (0.91)	0.00 (0.00)
11	4.64 (0.90)	4.47 (0.89)	0.17 (0.19)
12	4.92 (0.84)	4.55 (0.93)	0.37 (0.43)
13	4.91 (0.86)	4.64 (0.92)	0.27 (0.30)
14	5.02 (0.82)	4.73 (0.88)	0.29 (0.34)
15	5.14 (0.85)	4.89 (0.87)	0.25 (0.29)
16	5.03 (0.81)	4.89 (0.88)	0.14 (0.17)
17	4.73 (0.89)	4.82 (0.96)	0.09 (0.10)
18	4.90 (0.97)	4.69 (0.86)	0.21 (0.23)
19	4.59 (0.99)	4.48 (0.99)	0.11 (0.11)
20	4.95 (0.92)	4.66 (1.07)	0.29 (0.30)
Minimum	4.59	4.47	0.00
Maximum	5.33	5.25	0.37
Sample size	97–105	52–58	

Note. Frequency scale: 1 (*never*), 2 (*very rarely*), 3 (*rarely*), 4 (*occasionally*), 5 (*frequently*), 6 (*very frequently*). Respondents who judged the practice not relevant are not included in the calculation of the average frequency judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

Table 13. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Current Grade Level Taught for English Language Arts (ELA)

HLP	Lower (K–3)	Upper (4–6)	Difference
1 ^a	5.11 (0.81)	5.01 (0.92)	0.09 (0.11)
2 ^a	5.02 (0.80)	4.88 (0.83)	0.14 (0.18)
3 ^a	4.89 (0.97)	5.04 (0.85)	0.15 (0.16)
4	4.97 (0.83)	5.11 (0.87)	0.14 (0.16)
5	4.60 (0.82)	4.60 (0.76)	0.01 (0.01)
6	4.60 (0.96)	4.75 (0.85)	0.15 (0.16)
7	5.19 (0.77)	5.14 (0.88)	0.05 (0.06)
8	5.38 (0.74)	5.30 (0.87)	0.08 (0.10)
9	4.94 (0.84)	4.77 (0.79)	0.17 (0.21)
10	4.86 (0.90)	4.86 (0.95)	0.00 (0.00)
11	4.35 (0.82)	4.58 (0.92)	0.23 (0.27)
12	4.64 (0.98)	4.63 (1.07)	0.01 (0.01)
13	4.74 (0.89)	4.64 (0.94)	0.10 (0.11)
14	4.96 (0.88)	4.90 (0.95)	0.06 (0.06)
15	5.08 (0.95)	5.11 (0.84)	0.03 (0.03)
16	4.99 (0.84)	5.01 (0.89)	0.02 (0.03)
17	4.88 (0.74)	4.89 (0.93)	0.01 (0.01)
18	4.98 (0.85)	4.74 (0.84)	0.24 (0.29)
19	4.44 (1.05)	4.69 (1.06)	0.25 (0.24)
20	4.77 (0.94)	4.79 (1.12)	0.01 (0.01)
Minimum	4.35	4.58	0.00
Maximum	5.38	5.30	0.25
Sample size	95–103	67–74	

Note. Frequency scale: 1 (*never*), 2 (*very rarely*), 3 (*rarely*), 4 (*occasionally*), 5 (*frequently*), 6 (*very frequently*). Respondents who judged the practice not relevant are not included in the calculation of the average frequency judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

Frequency judgments by teacher race/ethnicity are disaggregated in Tables 14 and 15 for mathematics and ELA, respectively. In mathematics, by contrast, Black/African American teachers applied 15 of the 20 HLPs frequently or very frequently, while only five of the 20 HLPs were applied as often by Hispanic/Latino or White teachers. Relative to White teachers, Black/African American and Hispanic/Latino teachers more frequently applied three HLPs: HLP 4 (Black/African American and Hispanic/Latino, $ES = 0.43$), HLP 8 (Black/African American, $ES = 0.29$; Hispanic/Latino, $ES = 0.30$), and HLP 9 (Black/African American, $ES = 0.46$; Hispanic/Latino, $ES = 0.29$). There were 12 other HLPs more frequently applied by Black/African American teachers relative to White teachers (ES range = 0.21–0.71). Four other HLPs were more frequently applied by White teachers relative to Hispanic/Latino teachers: HLP 1 ($ES = 0.26$), HLP 3 ($ES = 0.21$), HLP 7 ($ES = 0.42$), and HLP 15 ($ES = 0.24$).

Table 14. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Race/Ethnicity for Mathematics

HLP	Black/African American	Hispanic/Latino	White	Overall
1 ^a	5.23 (0.61)	4.85 (0.88)	5.06 (0.79)	5.06 (0.78)
2 ^a	5.29 (0.72)	5.00 (0.84)	4.94 (0.83)	4.99 (0.82)
3 ^a	5.16 (0.60)	4.71 (0.85)	4.88 (0.81)	4.89 (0.80)
4	5.32 (0.65)	5.32 (0.65)	4.99 (0.77)	5.08 (0.75)
5	4.90 (0.70)	4.65 (0.67)	4.73 (0.88)	4.74 (0.83)
6	4.95 (0.59)	4.56 (0.92)	4.65 (0.95)	4.68 (0.91)
7	5.18 (0.59)	4.82 (0.73)	5.15 (0.78)	5.11 (0.76)
8	5.48 (0.60)	5.50 (0.61)	5.25 (0.88)	5.31 (0.82)
9	5.20 (0.70)	5.06 (0.80)	4.80 (0.89)	4.88 (0.87)
10	4.95 (0.80)	4.75 (1.12)	4.84 (0.89)	4.85 (0.91)
11	4.81 (1.03)	4.44 (1.04)	4.56 (0.91)	4.58 (0.94)
12	5.25 (0.64)	4.67 (0.84)	4.78 (0.87)	4.83 (0.85)
13	5.24 (0.62)	4.67 (0.77)	4.82 (0.85)	4.86 (0.82)
14	5.00 (0.79)	4.89 (0.88)	4.99 (0.81)	4.98 (0.81)
15	5.05 (0.67)	4.89 (0.94)	5.10 (0.87)	5.07 (0.85)
16	5.30 (0.66)	4.95 (0.71)	4.97 (0.79)	5.01 (0.77)
17	5.10 (0.94)	4.58 (0.84)	4.69 (0.94)	4.73 (0.94)
18	5.33 (0.58)	4.60 (1.05)	4.78 (0.91)	4.83 (0.91)
19	5.20 (0.52)	4.40 (1.10)	4.56 (0.95)	4.62 (0.95)
20	4.86 (0.79)	5.05 (1.00)	4.87 (0.94)	4.89 (0.93)
Minimum	4.81	4.40	4.56	4.58
Maximum	5.48	5.50	5.25	5.31
Sample size	19–22	18–22	112–119	151–161

Note. Frequency scale: 1 (*never*), 2 (*very rarely*), 3 (*rarely*), 4 (*occasionally*), 5 (*frequently*), 6 (*very frequently*). Respondents who judged the practice not relevant are not included in the calculation of the average frequency judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

Table 15. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Race/Ethnicity for English Language Arts (ELA)

HLP	Black/African American	Hispanic/Latino	White	Overall
1 ^a	5.10 (1.02)	4.87 (0.76)	5.09 (0.77)	5.06 (0.80)
2 ^a	5.05 (0.94)	4.87 (0.81)	4.93 (0.79)	4.94 (0.81)
3 ^a	4.81 (1.21)	4.78 (1.00)	4.98 (0.88)	4.93 (0.93)
4	5.05 (1.12)	5.00 (0.77)	5.01 (0.83)	5.01 (0.86)
5	4.37 (1.12)	4.83 (0.78)	4.58 (0.77)	4.59 (0.82)
6	4.11 (1.15)	4.83 (0.83)	4.72 (0.90)	4.66 (0.94)
7	5.10 (1.00)	5.13 (0.85)	5.20 (0.80)	5.18 (0.82)
8	5.62 (0.59)	5.25 (0.85)	5.34 (0.84)	5.36 (0.82)
9	5.05 (0.76)	4.74 (0.92)	4.82 (0.81)	4.84 (0.82)
10	4.68 (0.89)	4.75 (1.03)	4.94 (0.87)	4.89 (0.89)
11	4.05 (1.20)	4.33 (0.82)	4.50 (0.82)	4.43 (0.88)
12	4.60 (1.23)	4.75 (0.94)	4.60 (1.10)	4.62 (1.09)
13	4.95 (0.89)	4.91 (0.92)	4.66 (0.89)	4.73 (0.90)
14	5.00 (0.89)	4.96 (0.77)	4.93 (0.93)	4.94 (0.90)
15	4.90 (1.09)	5.00 (1.13)	5.08 (0.90)	5.05 (0.96)
16	4.95 (1.02)	5.17 (0.92)	4.94 (0.91)	4.97 (0.92)
17	4.95 (1.07)	4.71 (1.00)	4.90 (0.79)	4.88 (0.85)
18	5.14 (0.96)	5.17 (0.70)	4.83 (0.86)	4.91 (0.86)
19	4.86 (1.20)	4.58 (1.06)	4.49 (1.10)	4.55 (1.11)
20	4.85 (1.09)	4.92 (1.06)	4.69 (1.02)	4.74 (1.03)
Minimum	4.05	4.33	4.49	4.43
Maximum	5.62	5.25	5.34	5.36
Sample size	19–21	21–24	123–134	165–179

Note. Frequency scale: 1 (*never*), 2 (*very rarely*), 3 (*rarely*), 4 (*occasionally*), 5 (*frequently*), 6 (*very frequently*). Respondents who judged the practice not relevant are not included in the calculation of the average frequency judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

Black/African American teachers did not apply as many of the ELA HLPs frequently or very frequently (8) compared to mathematics, whereas corresponding rates for Hispanic/Latino and White teachers for ELA were similar to those for mathematics. White teachers more frequently applied HLP 10 and HLP 11 compared to Black/African American and Hispanic/Latino teachers (ES range = 0.21–0.52), whereas Black/African American and Hispanic/Latino teachers more frequently applied HLP 13 and HLP 18 compared to White teachers (ES range = 0.27–0.41). For HLP 5, the average frequency was higher for White teachers compared to Black/African American teachers (ES = 0.25) but lower compared to Hispanic/Latino teachers (ES = 0.32).

Tables 16 and 17 display disaggregated average frequency ratings by geographic region for mathematics and ELA, respectively. In mathematics, only five HLPs (1, 4, 7, 8, and 15) received average ratings of frequently or very frequently; however, by region, Northeast teachers

had such high ratings on far more HLPs (11) compared to other regions (Midwest = 3, South = 6, West = 4), with only two consistently showing average ratings above 5.0 for all regions (HLP 8 and HLP 15). For HLPs 4, 9, 12, and 15, relative to Northeast teachers, average ratings were higher compared to teachers from other regions (ES range = 0.20–0.57). Average ratings were also higher for Northeast teachers relative to Midwest teachers for HLPs 1, 11, 14, and 16 (ES range = 0.23–0.35); two of these were also relative to South teachers (HLP 11 and 16; ES range = 0.22–0.27), and the other two to West teachers (ES range = 0.23–0.34). Average ratings were higher for Northeast teachers relative to West teachers for HLPs 3, 5, 6, and 18 (ES range = 0.29–0.64) and one other relative to South teachers (HLP 18; ES = 0.30).

Table 16. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Geographic Region for Mathematics

HLP	Northeast	Midwest	South	West	Overall
1 ^a	5.15 (0.86)	4.96 (0.75)	5.08 (0.75)	4.97 (0.73)	5.04 (0.77)
2 ^a	5.06 (0.88)	4.85 (0.89)	5.00 (0.74)	4.96 (0.72)	4.96 (0.81)
3 ^a	4.97 (0.65)	4.89 (0.75)	4.88 (0.90)	4.76 (0.74)	4.88 (0.78)
4	5.24 (0.61)	4.98 (0.71)	5.02 (0.84)	5.07 (0.66)	5.06 (0.73)
5	4.76 (0.75)	4.67 (0.72)	4.71 (0.89)	4.50 (1.04)	4.67 (0.85)
6	4.79 (0.82)	4.67 (0.88)	4.64 (0.93)	4.43 (0.88)	4.64 (0.89)
7	5.15 (0.61)	5.11 (0.76)	5.10 (0.82)	4.96 (0.71)	5.09 (0.74)
8	5.24 (0.75)	5.35 (0.78)	5.26 (0.85)	5.29 (0.85)	5.29 (0.80)
9	5.16 (0.88)	4.76 (0.84)	4.82 (0.91)	4.75 (0.84)	4.85 (0.88)
10	4.85 (0.94)	4.94 (0.86)	4.72 (0.92)	4.93 (0.90)	4.85 (0.90)
11	4.76 (0.94)	4.50 (0.94)	4.50 (0.97)	4.62 (0.78)	4.57 (0.92)
12	5.13 (0.94)	4.66 (0.73)	4.79 (0.95)	4.71 (0.90)	4.80 (0.89)
13	4.88 (0.78)	4.86 (0.75)	4.76 (1.00)	4.79 (0.88)	4.82 (0.86)
14	5.12 (0.89)	4.83 (0.78)	4.97 (0.81)	4.82 (0.86)	4.93 (0.83)
15	5.24 (0.85)	5.00 (0.89)	5.07 (0.86)	5.00 (0.77)	5.07 (0.85)
16	5.13 (0.82)	4.95 (0.76)	4.95 (0.85)	5.03 (0.89)	4.99 (0.82)
17	4.82 (0.90)	4.72 (0.81)	4.75 (1.03)	4.82 (0.98)	4.77 (0.93)
18	5.06 (0.85)	4.96 (0.82)	4.77 (1.04)	4.50 (0.90)	4.84 (0.93)
19	4.66 (1.00)	4.53 (0.89)	4.60 (1.02)	4.52 (1.09)	4.57 (0.98)
20	4.79 (0.95)	4.96 (0.81)	4.87 (1.05)	4.80 (1.06)	4.87 (0.96)
Minimum	4.66	4.50	4.50	4.43	4.57
Maximum	5.24	5.35	5.26	5.29	5.29
Sample size	30–34	48–55	57–63	26–30	170–179

Note. Frequency scale: 1 (*never*), 2 (*very rarely*), 3 (*rarely*), 4 (*occasionally*), 5 (*frequently*), 6 (*very frequently*). Respondents who judged the practice not relevant are not included in the calculation of the average frequency judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

Table 17. Summary of Frequency Judgments for High-Leverage Practices (HLPs) by Geographic Region for English Language Arts (ELA)

HLP	Northeast	Midwest	South	West	Overall
1 ^a	5.00 (0.98)	5.09 (0.74)	5.21 (0.79)	4.80 (0.84)	5.05 (0.83)
2 ^a	4.86 (0.87)	4.89 (0.79)	5.08 (0.83)	4.83 (0.78)	4.93 (0.82)
3 ^a	5.03 (0.97)	4.77 (0.93)	5.10 (0.83)	4.73 (0.95)	4.92 (0.92)
4	5.06 (0.86)	4.91 (0.90)	5.13 (0.79)	4.95 (0.84)	5.02 (0.85)
5	4.58 (0.94)	4.49 (0.84)	4.77 (0.74)	4.46 (0.72)	4.59 (0.81)
6	4.71 (0.96)	4.54 (1.05)	4.82 (0.82)	4.47 (0.89)	4.65 (0.93)
7	5.25 (0.97)	5.08 (0.78)	5.27 (0.77)	5.02 (0.90)	5.16 (0.84)
8	5.41 (0.86)	5.33 (0.79)	5.44 (0.69)	5.12 (0.94)	5.33 (0.81)
9	5.00 (0.86)	4.67 (0.77)	5.03 (0.72)	4.73 (0.92)	4.86 (0.82)
10	4.79 (1.07)	4.74 (0.87)	4.90 (0.92)	4.95 (0.84)	4.85 (0.91)
11	4.39 (0.96)	4.31 (0.80)	4.61 (0.84)	4.33 (0.90)	4.42 (0.87)
12	4.67 (1.15)	4.55 (1.15)	4.81 (0.88)	4.34 (1.09)	4.61 (1.06)
13	4.78 (0.83)	4.69 (0.96)	4.95 (0.85)	4.37 (0.94)	4.72 (0.91)
14	4.94 (1.07)	4.77 (0.92)	5.13 (0.76)	4.69 (0.98)	4.90 (0.93)
15	4.95 (1.03)	5.15 (0.86)	5.26 (0.83)	4.69 (1.02)	5.05 (0.94)
16	5.00 (0.79)	4.87 (0.87)	5.20 (0.87)	4.71 (1.04)	4.96 (0.91)
17	4.65 (0.95)	5.02 (0.77)	4.95 (0.75)	4.68 (1.04)	4.85 (0.87)
18	4.86 (0.82)	4.80 (0.91)	5.08 (0.77)	4.74 (0.86)	4.89 (0.84)
19	4.44 (1.16)	4.37 (1.00)	4.87 (0.85)	4.32 (1.37)	4.54 (1.09)
20	4.68 (0.94)	4.56 (1.02)	5.12 (0.78)	4.56 (1.29)	4.76 (1.03)
Minimum	4.39	4.31	4.61	4.32	4.42
Maximum	5.41	5.33	5.44	5.12	5.33
Sample size	33–37	49–55	60–63	38–42	182–196

Note. Frequency scale: 1 (*never*), 2 (*very rarely*), 3 (*rarely*), 4 (*occasionally*), 5 (*frequently*), 6 (*very frequently*). Respondents who judged the practice not relevant are not included in the calculation of the average frequency judgment.

^aDenotes an HLP assessed by the NOTE assessment series.

In ELA, the same five HLPs were the only ones with average ratings of frequently or very frequently, but unlike in mathematics, South teachers had such high ratings on more HLPs (12) compared to other regions (Northeast = 7, Midwest = 5, West = 2). HLP 8 was consistently highly rated on average, as well as HLP 7. Regardless of statistical significance, teachers from the South generally had higher means on all HLPs compared to teachers from other regions, except for HLP 17 compared to Midwest teachers and HLP 10 compared to West teachers, but in both of these cases, effect sizes were below 0.10.

We employed the same strategy to compute ICC(2) for frequency judgments as importance judgments, given missing cases when respondents judged that an HLP was not relevant. Specifically, we imputed a value of 1 for missing cases. The ICC(2) (Shrout & Fleiss, 1979) indexing agreement among educators regarding their frequency ratings across the 20 HLPs are 0.96 (95% CI [0.93, 0.98]) for ELA and 0.94 (95% CI [0.89, 0.97]) for mathematics.

Another lens for interpreting the frequency ratings would be to examine the percentage of educators who judged the practice as frequently or very frequently applied as well as very or extremely important. Table 18 summarizes the percentage of educators (combining teachers and faculty) who judged an HLP as performed frequently or very frequently *and* as very important or extremely important. The following five HLPs were identified as high by approximately two thirds or more of the participants on both the importance and frequency scale:

- HLP 1: Making content and practices (e.g., specific texts, problems, ideas, theories, processes) explicit through explanation, modeling, representations, and examples
- HLP 7: Teaching a lesson or segment of instruction
- HLP 8: Implementing organizational routines, procedures, and strategies to support a learning environment
- HLP 14: Designing a sequence of lessons toward a specific learning goal
- HLP 15: Selecting and using particular methods to check understanding and monitor student learning during and across lessons

Table 18. Summary (Overall Sample) of High Importance and Frequency Ratings for High-Leverage Practices (HLPs) for English Language Arts (ELA) and Mathematics

HLP	ELA	Mathematics
1 ^a	72%	70%
2 ^a	65%	56%
3 ^a	64%	60%
4	63%	63%
5	48%	45%
6	52%	49%
7	76%	75%
8	79%	78%
9	61%	61%
10	57%	60%
11	39%	48%
12	53%	62%
13	54%	60%
14	67%	67%
15	71%	67%
16	59%	67%
17	59%	58%
18	56%	56%
19	49%	55%
20	51%	59%
Minimum	39%	45%
Maximum	79%	78%

^aDenotes an HLP assessed by the NOTE assessment series.

Also, HLP 16 (Composing, selecting, and interpreting, and using information from quizzes, tests, and other methods of summative assessment) was identified by approximately two thirds of the educators in reference to teaching elementary school mathematics.

Relative Importance

A method for examining relative importance (i.e., comparison within the set of 20 HLPs) is to ask educators to prioritize the practices. Given the number of HLPs, educators were not asked to rank all 20 practices. Rather, educators were asked to identify the five most and five least important HLPs. Tables 19 and 20 summarize the percentage of faculty and teachers judging a particular HLP as falling under either the five most or least important practices for ELA and mathematics, respectively. Comparing relative importance judgments for teachers and faculty, the percentages of each HLP categorized as among the most or least important differed by up to 23 percentage points. For least important, three HLPs for ELA and six for mathematics differed by more than 10 percentage points between the two groups. For most important, seven HLPs for ELA and nine for mathematics differed by more than 10 percentage points.

Table 19. Percentage of Teacher/Faculty Categorizations for High-Leverage Practices (HLPs) as Five Least and Most Important Practices for English Language Arts (ELA)

HLP	Least important		Most important	
	Teachers	Faculty	Teachers	Faculty
1 ^a	11.2%	11.0%	46.7%	51.2%
2 ^a	49.7%	42.7%	7.6%	13.4%
3 ^a	28.9%	24.4%	14.2%	25.6%
4	10.2%	24.4%	49.7%	26.8%
5	52.8%	45.1%	6.1%	4.9%
6	40.6%	42.7%	6.6%	12.2%
7	13.7%	8.5%	26.9%	34.1%
8	6.1%	11.0%	55.3%	52.4%
9	33.0%	29.3%	26.9%	12.2%
10	28.4%	20.7%	30.5%	31.7%
11	38.6%	22.0%	12.7%	30.5%
12	37.1%	43.9%	22.8%	7.3%
13	20.8%	26.8%	19.8%	23.2%
14	22.3%	18.3%	23.9%	39.0%
15	14.2%	12.2%	26.9%	30.5%
16	24.4%	31.7%	23.9%	18.3%
17	15.2%	11.0%	23.9%	25.6%
18	18.8%	23.2%	28.4%	19.5%
19	14.2%	12.2%	23.9%	30.5%
20	19.8%	39.0%	23.4%	11.0%
Minimum	6.1%	8.5%	6.1%	4.9%
Maximum	52.8%	45.1%	55.3%	52.4%

^aDenotes an HLP assessed by the NOTE assessment series.

Table 20. Percentage of Teacher/Faculty Categorizations for High-Leverage Practices (HLPs) as Five Least and Most Important Practices for Mathematics

HLP	Least important		Most important	
	Teachers	Faculty	Teachers	Faculty
1 ^a	12.8%	2.9%	60.6%	73.5%
2 ^a	47.3%	55.9%	11.7%	3.9%
3 ^a	22.3%	18.6%	21.8%	31.4%
4	16.0%	25.5%	38.3%	21.6%
5	36.2%	31.4%	11.7%	18.6%
6	37.8%	32.4%	10.1%	25.5%
7	17.0%	10.8%	23.9%	35.3%
8	8.5%	20.6%	46.8%	33.3%
9	37.8%	37.3%	21.8%	11.8%
10	29.3%	21.6%	27.1%	26.5%
11	48.4%	43.1%	13.3%	12.7%
12	28.2%	40.2%	18.1%	10.8%
13	23.4%	21.6%	20.2%	20.6%
14	28.7%	13.7%	22.9%	32.4%
15	13.3%	10.8%	29.8%	41.2%
16	18.6%	19.6%	22.9%	19.6%
17	17.6%	5.9%	29.3%	30.4%
18	20.2%	31.4%	23.9%	3.9%
19	10.6%	7.8%	28.7%	39.2%
20	26.1%	49.0%	17.0%	7.8%
Minimum	8.5%	2.9%	10.1%	3.9%
Maximum	48.4%	55.9%	60.6%	73.5%

^aDenotes an HLP assessed by the NOTE assessment series.

To more easily digest the relative importance judgments, we combined the *least* and *most* judgments to rank order the HLPs by perceived importance for beginning elementary school teachers. First, we ranked the HLPs by the percentage of teachers who identified the HLP as one of the five most important;⁹ the highest percentage received a rank of 1, and the lowest, a rank of 20. Second, we ranked the HLPs by the percentage of teachers who identified the HLP as one of the five least important; the highest percentage received a rank of 20, and the lowest, a rank of 1. Then the two rankings were summed with the lower value indicating the more relative importance. Calculations were completed for each subject area (ELA and mathematics) and for both teachers and faculty.

Focusing on teachers' relative importance rankings of the HLPs, while the order varied, nine HLPs were identified by teachers as being in the top 10 for both ELA and mathematics¹⁰ (see Table 21). The top three ranked HLPs for both subjects were the same: HLPs 1, 4, and 8. HLPs 3 and 16 were identified as top 10 by both teachers and faculty for mathematics but not

ELA; HLP 11 was identified as top 10 by faculty (but not teachers) for ELA but not mathematics.

While all 20 HLPs were judged to be important (average judgment 5.0 or higher on a 6-point scale), the relative importance points out some differences between the two groups of educators and between content areas.

Table 21. Relative Importance Rankings for English Language Arts (ELAs) and Mathematics High-Leverage Practices (HLPs)

HLP	ELA		Mathematics	
	Teacher	Faculty	Teacher	Faculty
1 ^a	3	2	2	1
2 ^a	18(T)	16	18(T)	20
3 ^a	15(T)	11	10(T)	7(T)
4	2	10	3(T)	12
5	20	20	17	14
6	18(T)	17	18(T)	13
7	4	3	7	4
8	1	1	1	7(T)
9	11(T)	15	16	15(T)
10	9	8	10(T)	9
11	17	9	18(T)	17
12	15(T)	19	14(T)	18
13	14	13	13	11
14	10	4	12	6
15	5	5(T)	3(T)	2(T)
16	11(T)	14	9	10
17	8	7	6	5
18	6	12	8	15(T)
19	7	5(T)	3(T)	2(T)
20	13	18	14(T)	19

Note. (T) indicates a tie in the ranking of the relative importance.

^aDenotes an HLP assessed by the NOTE assessment series.

Discussion

The purpose of the current investigation was twofold—(a) to further examine the relevance and importance of a core set of HLPs for teaching and (b) to explore the content-related validity evidence supporting the selection of several of these teaching practices for inclusion on the performance component in the NOTE assessment series. Content-validity evidence was gathered using an online survey of educators—practicing elementary school teachers and college faculty who prepare elementary teachers—who judged the relevance, importance, and frequency of each HLP for beginning teachers. The content-related validity questions were couched in teaching ELA or mathematics. Across both groups of educators and

for both subject areas, each of the HLPs was judged to be relevant and important for beginning elementary school teachers and to be more than occasionally applied by beginning teachers.

Based on the accumulated research evidence and experience preparing teacher candidates, Deborah Ball and her colleagues have identified and described 20 HLPs associated with quality teaching (Ball & Forzani, 2009; Ball & Hill, 2008). While research is ongoing for individual HLPs or sets of HLPs, the inclusion of one or more of these practices into a licensure process must be supported by the perceived relevance and importance of the practices by incumbents (i.e., elementary school teachers) and SMEs who prepare teachers (i.e., college faculty). The validity evidence collected through this study provides support for the initial inclusion of these practices in teacher licensure assessment.

The NOTE assessment series includes a performance component that addresses a teacher candidate's ability to model and explain content (HLP 1), lead group discussions (HLP 2), and elicit and interpret student thinking model (HLP 3). All three HLPs were judged—by teachers and faculty and across subject areas—as being very important (average judgment greater than 5.0) for effective beginning practice and at least occasionally (average judgment greater than 4.0) applied by beginning elementary school teachers. One of the HLPs, HLP 1, was among the top three for teachers and faculty across content areas and was judged as very or extremely important and frequently or very frequently applied by two thirds or more of the educators (teachers and faculty combined) for both ELA and mathematics. (HLPs 2 and 3 were judged as very or extremely important and frequently or very frequently applied by at least 60% of the educators for ELA; HLP 3 was similarly judged for mathematics.)

Results of this study showed that HLP 2 (Leading a class discussion) was not in the top 10 list. However, it is worth noting that mean importance ratings for this HLP still render this practice as very important. Among the set of practices considered very important by experts, leading a class discussion is ranked as less important than some of the other HLPs. The ranking does not discount the criterion-based judged importance of that HLP, but rather, it places that importance in the context of the other HLPs. This finding is in keeping with prior research that has shown that class discussion yields many positive effects for students. For instance, class discussion has been shown to improve student skills such as reasoning (Cazden & Beck, 2003), reading comprehension (Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009) and overall engagement (Henning, 2005).

The accumulation of content validity evidence is of critical importance in the construction and evaluation of licensure assessments (M. Kane, 2004; Sireci & Sukin, 2013). Overall, the content-related validity evidence collected supports the complete set of HLPs as important for a beginning elementary school teacher's ability to be an effective educator. Each HLP was judged to be relevant and important as well as frequently applied by beginning teachers. Regarding the performance component of the new licensure assessment under development, evidence from this study supports the inclusion of the three HLPs being measured. However, future research may explore the establishment of direct links between test specifications derived from this work and specific items from the assessment, as individual items should pertain to one or more test specifications (Tannenbaum & Rosenfeld, 1994).

Limitations

This study is not without limitations. First, although the structured survey methodology we used (cf. Raymond & Luecht, 2013) has many strengths that may outweigh its weaknesses, it produces simplified information that tends to be less rich than that obtained via focus groups or observations. Second, although our sample consisted of educators from around the country, our sample was neither nationally representative of beginning teachers, practicing teachers, or teacher education faculty. Generalizations from the sample provided to educators in general or to subgroups of interest need to be made with caution.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Ball, D. L., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, *60*, 497–511.
- Ball, D. L., & Hill, H. C. (2008). Measuring teacher quality in practice. In D. Gitomer (Ed.), *Measurement issues and assessment for teaching quality*. (p. 80–98). Thousand Oaks, CA: Sage.
- Boulet, J., & Zanten, M. (2014). Ensuring high-quality patient care: the role of accreditation, licensure, specialty certification and revalidation in medicine. *Medical Education*, *48*, 75–86.
- Cascio, W. F. (1982). *Applied psychology in personnel management*. Reston Publishing Company, VA: Reston.
- Cazden, C. B., & Beck, S. W. (2003). Classroom discourse. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 165–197). New York, NY: Routledge.
- Clauser, B. E., Margolis, M. J., & Case, S. M. (2006). Testing for licensure and certification in the professions. *Educational Measurement*, *4*, 701–731.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Danielson, C. (2006). *Teacher leadership that strengthens professional practice*. Retrieved from ERIC database. (ED509051)
- Ferguson, R. (1998). Teachers' perceptions and expectations and the Black White test score gap. In C. Jencks and M. Phillips (Eds.), *The Black-White test score gap* (pp. 273–317). Washington, DC: Brookings Institution Press.
- Gael, S. (1983). Job analysis. *A guide to assessing work activities*. San Francisco, CA: Jossey Bass.
- Goldhaber, D. (2002). The mystery of good teaching. *Education Next*, *2*, 50–55.
- Goldring, R., Gray, L., & Bitterman, A. (2013). *Characteristics of public and private elementary and secondary school teachers in the United States: Results from the 2011–12 Schools*

- and Staffing Survey* (NCES 2013-314). Retrieved from <http://nces.ed.gov/pubs2013/2013314.pdf>
- Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (1999). *Do higher salaries buy better teachers?* (No. w7082). Retrieved from <http://www.nber.org/papers/w7082.pdf>
- Harris, D., & Rutledge, S. (2010). Models and predictors of teacher effectiveness: A comparison of research about teaching and other occupations. *The Teachers College Record*, *112*, 914–960.
- Henning, J. E. (2005). Leading discussion: Opening up the conversation. *College Teaching*, *53*, 90–94.
- Kane, M. (2004). Certification testing as an illustration of argument-based validation. *Measurement*, *2*, 135–170.
- Kane, M. T. (1982). The validity of licensure examinations. *American Psychologist*, *37*, 911–918.
- Kane, M. T., Kingsbury, C., Colton, D., & Estes, C. (1989). Combining data on criticality and frequency in developing plans for licensure and certification examinations. *Journal of Educational Measurement*, *26*, 17–27.
- Kane, M. T., Miller, T., Trine, M., Becker, C., & Carson, K. (1995). The precision of practice analysis results in the professions. *Evaluation & the Health Professions*, *18*, 29–50.
- Mehrens, W. A. (1995). Methodological issues in standard setting for educational exams. In *Proceedings of joint conference on standard setting for large-scale assessments* (Vol. 2, pp. 221–263). Washington, DC: National Assessment Governing Board and National Center for Education Statistics.
- Mehrens, W. A., & Lehmann, I. J. (1984). *Measurement and evaluation in education and psychology*. Fort Worth, TX: Holt McDougal.
- Murphy, P. K., Wilkinson, I. A., Soter, A. O., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology*, *101*, 740–264.
- Raymond, M. R. (2001). Job analysis and the specification of content for licensure and certification examinations. *Applied Measurement in Education*, *14*, 369–415.

- Raymond, M. R. (2005). An NCME instructional module on developing and administering practice analysis questionnaires. *Educational Measurement: Issues and Practice*, *24*, 29–42.
- Raymond, M. R., & Luecht, R. M. (2013). Licensure and certification testing. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J. I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology, Vol. 3: Testing and assessment in school psychology and education* (pp. 391–414). Washington, DC: American Psychological Association.
- Raymond, M., & Neustel, S. (2006). Determining the content of credentialing examinations. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 181–224). Mahwah, NJ: Erlbaum.
- Rosenfeld, M., & Tannenbaum, R. J. (1991). *Identification of a core of important enabling skills for the NTE Successor Stage I Examination* (Research Report No. RR-91-37). Princeton, NJ: Educational Testing Service. <http://dx.doi.org/10.1002/j.2333-8504.1991.tb01404.x>
- Schmitt, K. (1995). What is licensure? In J. C. Impara (Ed.), *Licensure testing: Purposes, procedures, and practices* (pp. 3–32). Lincoln, NE: Buros Institute of Mental Measurements.
- Shimberg, B. (1981). Testing for licensure and certification. *American Psychologist*, *36*, 1138–1146.
- Shrout, P., & Fleiss, J. (1979). Intraclass correlations: Uses in assessing rater reliability [Abstract]. *Psychological Bulletin*, *86*, 420–428. doi:10.1037/0033-2909.86.2.420
- Sireci, S. G., & Sukin, T. (2013). Test validity. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology, Vol. 1: Test theory and testing and assessment in industrial and organizational psychology. APA handbooks in psychology* (pp. 61–84). Washington, DC: American Psychological Association,
- Smith, I. L., & Hambleton, R. K. (1990). Content validity studies of licensing examinations. *Educational Measurement: Issues & Practice*, *9*, 7–10.
- Tannenbaum, R. J., Robustelli, S. L., & Baron, P. A. (2008). Evidence-centered design: A lens through which the process of job analysis may be focused to guide the development of knowledge-based test content specifications. *CLEAR Exam Review*, *19*, 26–35.

- Tannenbaum, R. J., & Rosenfeld, M. (1994). Job analysis for teacher competency testing: Identification of basic skills important for all entry-level teachers. *Educational and Psychological Measurement, 54*, 199–211.
- Tannenbaum, R. J., & Wesley, S. (1993). Agreement between committee-based and field-based job analyses: A study in the context of licensure testing. *Journal of Applied Psychology, 78*, 975–980.
- Wilson, S. M. (2009). Measuring teacher quality at entry. In D. Gitomer (Ed.), *Measurement Issues and Assessment for Teaching Quality* (pp. 8–29). Thousand Oaks, CA: Sage.

Appendix. High-Leverage Practices (HLPs)

1. Making content and practices (e.g., specific texts, problems, ideas, theories, processes) explicit through explanation, modeling, representations, and examples

Making content and practices (e.g., processes, strategies) explicit is essential to providing all students with access to fundamental ideas and practices in a given subject or domain. Effective efforts to do this attend both to the integrity of the subject and to students' likely interpretations of it. They include strategically choosing and using representations and examples to build understanding and remediate misconceptions, using language carefully, highlighting core ideas while sidelining potentially distracting ones, and making one's own thinking visible while modeling and demonstrating.

2. Leading a group discussion

In a group discussion, the teacher and all the students work on specific content or practices together, using one another's ideas as resources. The purposes of a discussion are to build collective knowledge and capability in relation to specific instructional goals and to allow students to practice listening, speaking, and interpreting. In instructionally productive discussions, the teacher and a wide range of students contribute orally, listen actively, and respond to and learn from others' contributions.

3. Eliciting and interpreting individual students' thinking

Teachers pose questions or tasks that provoke or allow students to share their thinking about specific academic content and practices in order to evaluate student understanding, guide instructional decisions, and surface ideas that will benefit other students. To do this effectively, a teacher draws out a student's thinking through carefully chosen questions and tasks and considers and checks alternative interpretations of the student's ideas and methods.

4. Establishing norms and routines for classroom discourse and work that are central to the subject-matter domain

Each discipline or domain has norms and routines that reflect the ways in which people in the field construct and share knowledge. These norms and routines vary across subjects but often include establishing hypotheses, providing evidence for claims, and showing one's thinking in detail. Teaching students what they are, why they are important, and how to use them is crucial to building understanding and capability in a given subject. Teachers may use explicit explanation, modeling, and repeated practice to do this.

5. Recognizing particular common patterns of student thinking and development in a subject-matter domain

Although there are important individual and cultural differences among students, there are also common patterns in the ways in which students think about and develop understanding and skill in relation to particular topics and problems. Teachers who are familiar with common patterns of student thinking and development and who are fluent

in anticipating or identifying them are able to work more effectively and efficiently as they plan and implement instruction and evaluate student learning.

6. Identifying and implementing an instructional response or strategy in response to common patterns of student thinking

Specific instructional strategies are known to be effective in response to particular common patterns of student thinking. Teachers who are familiar with them can choose among them appropriately and use them to support, extend, or begin to change student thinking.

7. Teaching a lesson or segment of instruction

During a lesson or segment of instruction, the teacher sequences instructional opportunities toward specific learning goals and represents academic content and processes in ways that connect to students' prior knowledge and extend their learning. In a skillfully enacted lesson, the teacher fosters student engagement, provides access to new material and opportunities for student practice, adapts instruction in response to what students do or say, and assesses what students know and can do as a result of instruction.

8. Implementing organizational routines, procedures, and strategies to support a learning environment

Teachers implement routine ways of carrying out classroom tasks in order to maximize the time available for learning and minimize disruptions and distractions. They organize time, space, materials, and students strategically and deliberately teach students how to complete tasks such as lining up at the door, passing out papers, and asking to participate in class discussion. This can include demonstrating and rehearsing routines and maintaining them consistently.

9. Setting up and managing small-group work

Teachers use small-group work when instructional goals call for in-depth interaction among students and in order to teach students to work collaboratively. To use groups effectively, teachers choose tasks that require and foster collaborative work, issue clear directions that permit groups to work semi-independently, and implement mechanisms for holding students accountable for both collective and individual learning. They use their own time strategically, deliberately choosing which groups to work with, when, and on what.

10. Engaging in strategic relationship-building conversations with students

Teachers increase the likelihood that students will engage and persist in school when they establish positive, individual relationships with them. Brief, one-on-one conversations with students are a fundamental way of doing this, as they help teachers learn about students and communicate care and interest. They are most effective when teachers are strategic about when to have them and what to talk about and use what they learn to address academic and social needs.

11. Learning about students' cultural, family, intellectual, and personal experiences and resources

Teachers who know their students and know the resources that they bring to instruction are more likely to build effective relationships with them and to be able to connect learning to their students' experience. To do this, teachers make connections with and learn about the communities and families of their students, and they listen and observe their students.

12. Setting long- and short-term learning goals for students referenced to external benchmarks

Clear goals referenced to external standards help teachers ensure that all students learn expected content. Explicit goals help teachers to maintain coherent, purposeful, and equitable instruction over time. Setting effective goals involves analysis of student knowledge and skills in relation to established standards and careful efforts to establish and sequence interim benchmarks that will help ensure steady progress toward larger goals.

13. Appraising, choosing, and modifying tasks and texts for a specific learning goal

Teachers appraise and modify texts, tasks, problems or problem sets, and other curriculum materials to determine their appropriateness for helping particular students work toward specific learning goals. This involves considering students' needs and assessing what questions and ideas particular materials will raise and the ways in which they are likely to challenge students. Teachers choose and modify material accordingly, sometimes deciding to use parts of a text or activity and not others, for example, or to combine material from more than one source.

14. Designing a sequence of lessons toward a specific learning goal

Carefully sequenced lessons help students develop deep understanding of content and sophisticated skills and practices. Teachers design and sequence lessons with an eye toward providing opportunities for student inquiry and discovery and include opportunities for students to practice and master foundational concepts and skills before moving on to more advanced ones. Effectively sequenced lessons maintain a coherent focus while keeping students engaged; they also help students achieve an appreciation of what they have learned.

15. Selecting and using particular methods to check understanding and monitor student learning during and across lessons

Teachers use a variety of informal but deliberate methods to assess what students are learning during and between lessons. These frequent checks provide information about students' current level of competence and help the teacher adjust instruction during a single lesson or from one lesson to the next. They may include, for example, simple questioning, short performance tasks, or journal or notebook entries.

16. Composing, selecting, and interpreting, and using information from quizzes, tests, and other methods of summative assessment

Effective summative assessments provide teachers with rich information about what students have learned and where they are struggling in relation to specific learning goals. In composing and selecting assessments, teachers consider validity, fairness, and efficiency. Effective summative assessments provide both students and teachers with useful information and help teachers evaluate and design further instruction. Teachers analyze the results of assessments carefully, looking for patterns that will guide efforts to assist specific students and inform future instruction.

17. Providing oral and written feedback to students on their work

Effective feedback helps focus students' attention on specific qualities of their work, it highlights areas needing improvement, and delineates ways to improve. Good feedback is specific, not overwhelming in scope, and focused on the academic task, and supports students' perceptions of their own capability. Giving skillful feedback requires the teacher to make strategic choices about the frequency, method, and content of feedback and to communicate in ways that are understandable by students.

18. Communicating about a student with a parent or guardian

Regular communication between teachers and parents/guardians supports student learning. Teachers communicate with parents to provide information about students' academic progress, behavior, or development; to seek information and help; and to request parental involvement in school. These communications may take place in person, in writing, or over the phone. Productive communications are attentive to considerations of language and culture and designed to support parents and guardians in fostering their child's success in and out of school.

19. Analyzing instruction for the purpose of improving it

Learning to teach is an ongoing process that requires regular analysis of instruction and its effectiveness. Teachers study their own teaching and that of their colleagues in order to improve their understanding of the complex interactions between teachers, students, and content and of the impact of particular instructional approaches. Analyzing instruction may take place individually or collectively and involves identifying salient features of the instruction and making reasoned hypotheses for how to improve.

20. Communicating with other professionals

Teachers routinely communicate with fellow teachers, administrators, and other professionals in order to plan teaching, discuss student needs and secure special services for students, and manage school policies. They do this orally, in meetings and presentations, and in writing, in letters, e-mails, newsletters, and other documents. Skillful communication is succinct, respectful, and focused on specific professional topics. It uses clear, accessible language, generally in standard English, and is attentive to its specific audience.

Notes

- ¹ TeachingWorks is housed in the School of Education at the University of Michigan and focuses on the improvement of teacher preparation.
- ² One companion report is entitled *Investigating the Relevance and Importance of English Language Arts Content Knowledge Areas for Beginning Elementary School Teachers* (ETS Research Memorandum No. RM-16-08). The other is *Investigating the Relevance and Importance of Mathematical Content Knowledge Areas for Beginning Elementary School Teachers* (ETS Research Memorandum No. RM-16-10).
- ³ The *Standards for Educational and Psychological Testing* (AERA et al., 2014) generically refer to licensure and certification as *credentialing*.
- ⁴ See <http://www.teachingworks.org/work-of-teaching/high-leverage-practices> for more information
- ⁵ The list of HLPs established by TeachingWorks is an evolving document. The survey was conducted using the 20 HLPs described by TeachingWorks in Fall 2014. The list has since been refined and includes 19 HLPs.
- ⁶ Of the 387 respondents who indicated they were teachers, all but two indicated they held a current license to teach in their state. These two respondents were removed from the sample. Of the 202 responded who indicated they were faculty, 17 indicated they were not currently preparing elementary school teacher candidates. These respondents also were removed from the sample.
- ⁷ Response rate for the e-mailed surveys delivered to sampled educators.
- ⁸ Results based on the total sample are unweighted, given the similarity in judgments from teachers and faculty.
- ⁹ Percentages were calculated to one decimal place precision, and rankings allowed for ties.
- ¹⁰ When disaggregating data, eight of the top nine HLPs identified for the overall sample also were in the top eight for one of the two content areas or both for Black/African American and Hispanic/Latino teachers. When looking at the data by census region, six of the top nine HLPs for the overall sample also were in the top nine for one of the two content areas or both for at least three of the regions.