Chapter 6
Teacher Selection: History and Current Practices

Abstract In this chapter, we show that choosing the right people to teach has been an ongoing challenge in education, with researchers, practitioners, and policymakers pondering two key questions for as long as teachers have been appointed: What are the key personal characteristics related to teacher effectiveness? and How can these characteristics be assessed in a valid way when selecting prospective teachers? These two questions have received long-standing attention in education, but with little systematic research carried out to provide guidance to selectors. This chapter begins with an overview of historical issues in teacher selection, and then examines current practices for selecting teachers into ITE and into employment in a range of jurisdictions. The second half of the chapter reports a review of the research exploring the links between teacher selection practices and teacher effectiveness based on a recent meta-analysis that closely examined the research (Klassen & Kim, 2019).

In Chap. 5 we examined the selection practices in a range of fields outside of teaching and teacher education, including selection into medical education, legal education, and various kinds of jobs in organizational contexts. In this chapter we turn our attention towards the selection of teachers, first taking a historical perspective, and then critically reviewing current practices.

At the heart of teacher selection is the prediction of short-term and long-term teaching effectiveness. The question for selecting teachers for initial teacher education (ITE) or into employment is, at its essence, the same: *Will this person be, or develop into, an effective teacher*? Teachers become more effective, more reflective, and more knowledgeable about teaching as they gain experience (Antoniou et al., 2015; Atteberry et al., 2015), and the 'art of selection is to sample and evaluate personal attributes and behaviors that are believed to predict future effectiveness in the classroom. However, predicting teacher effectiveness is remarkably difficult, because teaching is a complex, multi-faceted job that is influenced by a host of interacting environmental and personal factors (e.g., Rimm-Kaufman & Hamre, 2010). Furthermore, predicting future behaviors from a brief sample of carefully curated behaviors during the selection process will always carry a measure of error. Choosing selection methods that are reliable, valid, and fair can improve the likelihood that we will make the best possible decisions about prospective teachers.

R. M. Klassen and L. E. Kim, *Teacher Selection: Evidence-Based Practices*,
https://doi.org/10.1007/978-3-030-76188-2_6

In this chapter, we show that choosing the right people to teach has been an ongoing challenge in education, with researchers, practitioners, and policymakers pondering two key questions for as long as teachers have been appointed: *What are the key personal characteristics related to teacher effectiveness?* and *How can these characteristics be assessed in a valid way when selecting prospective teachers?* These two questions have received long-standing attention in education, but with little systematic research carried out to provide guidance to selectors. This chapter begins with an overview of historical issues in teacher selection, and then examines current practices for selecting teachers into ITE and into employment in a range of jurisdictions. The second half of the chapter reports a review of the research exploring the links between teacher selection practices and teacher effectiveness based on a recent meta-analysis that closely examined the research (Klassen & Kim, 2019).

6.1 Historical Perspective on Teacher Selection

The question of how to select the most effective teachers has been asked for nearly a century. In 1922 F. B. Knight asked the questions that remain at the heart of teacher selection:

> *What facts concerning a candidate for a teaching position are of prognostic value? Of a hundred graduates of a normal college quite probably some will make excellent teachers, a larger number will do well, and a few will fail. By what system of interviewing can a superintendent increase his chances of picking more successful teachers and fewer failures than pure chance would account for? What qualities possessed by a candidate and ascertainable by a prospective employer are correlated highly enough with teaching success to be worth considering in a sound selective technique? (*Knight, 1922*, p. 207)*

Knight's work was an attempt to improve the likelihood of making good decisions about selection through establishing "statistically dependable facts to teacher selection" (p. 207). His study assessed a wide range of potential predictors: handwriting, age, experience, intelligence, ranking in teacher education program, amount of additional 'professional study' (defined as summer school and Saturday work in educational courses), and a 'trade test' assessing candidates' knowledge about teaching. No significant relationship with teacher effectiveness was found for handwriting, age, teaching experience, intelligence, amount of professional development, or standing in a teacher education program. The study found a statistically significant relation between teacher effectiveness and a tailor-made 'trade test', which was designed to measure knowledge about teaching practices. Knight concluded by questioning whether school district superintendents were able to reliably identify teacher effectiveness using their own intuition and called for "a genuinely scientific procedure of teacher selection" (p. 216).

Other early twentieth century educationalists recognized the importance of selection, with Tubbs (1928) leading the charge in promoting more systematic research on teacher selection: "Upon this one thing (i.e., teacher selection) more than any

other depends on success or failure in the education and training of future citizens" (p. 332). Tubbs' strong beliefs about the importance of teacher selection stemmed from teacher shortages of the time, partly due to the loss of teachers and potential teachers in WWI, and also from high rates of teacher attrition rates. Tubbs stated that attrition rates should be "regarded with alarm by everyone interested in American education" (p. 323), due to the one-third to one-half of the teaching population that were leaving the profession (or changing jobs within the profession) each year. He proposed five criteria for teacher selection: (a) educational background, (b) *experience*, which, according to Tubbs, builds adaptability, (c) *health* ("teachers are under a moral obligation to protect others from any possible contagion or infection" p. 328), a focus perhaps not surprising after the ravages of the influenza epidemic of the preceding decade), (d) *character* (described as the 'greatest' of the requisite qualities), and (e) *personality*, the lack of which "greatly handicaps the quality of service which a superior teacher should give" (p. 329).

Tubbs described the ability to "see below surface indications" in the selection of teachers as a gift without which "no (superintendent) can meet with more than a modicum of success" (p. 329). The 'problem' of teacher selection has historical roots—and current employers and ITE providers continue to focus on identifying the attributes associated with future effectiveness, and how to measure these attributes in a way that is reliable, valid, and fair.

6.2 Need for Teacher Selection

A process for teacher selection is needed when the number of applicants is greater than the number of available ITE places or jobs, when there is a need to identify unsuitable applicants ('selecting out') before beginning training or employment, and when there is a benefit in generating profiles of applicant strengths and weaknesses for future development. Systems for teacher selection are built on data gathered from existing records (e.g., evaluation of academic transcripts) and from new sources (e.g., face-to-face interviews, personality tests, teaching demonstrations) that are determined by employers or teacher education programs. Although selection methods have been the subject of in-depth research attention in some professions—especially medicine and business—the knowledge gained has nor often spilled over into education.

A selection process—for training or employment—is a predictive exercise that involves three steps: first, identification of the attributes needed for success in the endeavor, second, development of a method for assessing these attributes, and third, an assessment of the relationship between measured attributes and some kind of criterion or outcome measure. In order to make these predictions, selectors gather evidence that they believe can help them make valid selection decisions. Most selection methods will focus on the three personal characteristics discussed in Chap. 2: cognitive attributes, such as subject area knowledge; non-cognitive attributes, such as beliefs, motives, traits, and dispositions; and background experience, including previous relevant experiences. Higher-performing school systems tend to have more

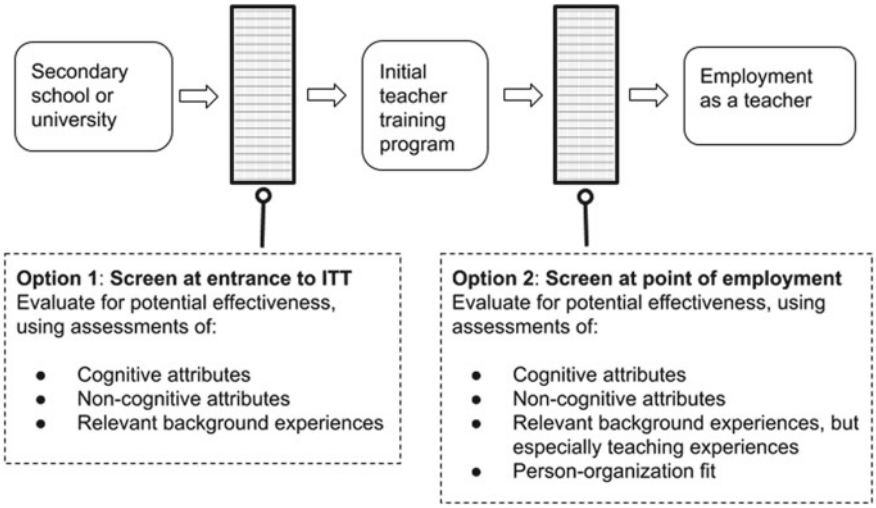

Fig. 6.1 Options for screening applicants

sophisticated selection systems (i.e., with multiple stages assessing multiple factors), with explicit recognition that poor selection methods influence the quality of the teaching workforce (Barber & Mourshed, 2007). Whatever selection process is implemented, the purpose of teacher selection is universal: using the best possible data to make the best possible decisions about prospective teachers.

Prospective teachers can be selected, or screened, at two key points: at entry into ITE or entry into employment. Figure 6.1 shows two potential screening points in the teacher selection process. Option 1 shows screening at the entrance to ITE, where cognitive attributes, non-cognitive attributes, and relevant background factors are evaluated. Option 2 shows that screening at the point of employment includes evaluation of the same factors, but with the addition of consideration of 'person-organization fit', where consideration is given to how well the applicant might fit into the school or school district based on additional, possibly non-evaluated factors. Higher performing education systems tend to have more effective processes to select candidates for ITE. Barber and Mourshed (2007) show that countries that perform well in international comparisons, such as Finland and Singapore, have selection procedures that are systematic, test a wide range of attributes, and filter applicants at the point of entry into ITE, rather than at the point of entry into the profession.

6.3 Selecting Candidates for ITE

Little research evidence is available supporting the predictive validity of selection methods in teacher education (e.g., Casey & Childs, 2011), and some of the methods

used, such as letters of reference and interviews have been shown to be biased against certain groups of candidates (Patterson et al., 2016). Not very much is known about the effectiveness of selection procedures into ITE, and what we do know suggests an arbitrariness in selection methods (e.g., Casey & Childs, 2017; Denner et al., 2001).

Selection into ITE also varies according to the structure of the ITE program. Two models of ITE are generally endorsed: a *consecutive* model in which students first complete an undergraduate degree in a particular subject, and then enroll in an ITE program. A *concurrent* model involves students studying a combination of a particular subject (or subjects) alongside courses involving pedagogical and theoretical instruction coupled with practical experience in teaching. In many countries, universities (and other initial teacher-education providers) offer different selection methods for undergraduate and postgraduate entrance for ITE programs. For example, in Finland, the two-phase selection process for direct entry 'class-teacher' education programs involves a nationwide literacy test (VAKAVA) which assesses the cognitive attributes of memorization, understanding, and the ability to apply knowledge from articles to practice (Malinen et al., 2012). The second phase of selection involves an 'aptitude test' developed by individual universities, and aimed at evaluating applicants' suitability, motivation, and commitment to teaching. The aptitude test varies across universities but may include an individual interview and a group discussion task (Malinen et al., 2012). Selection methods in Finland have recently been under review, with a consideration of alternative selection methods, including situational judgment tests (SJTs) and multiple mini-interviews (MMIs), currently underway (personal correspondence, R. Metsäpelto, May 2020).

In the UK, selection for ITE programs usually takes place at the postgraduate level. A survey of 74 university-based initial teacher education providers in England and Wales was conducted to understand how cognitive and non-cognitive attributes were assessed for selection (Klassen & Dolan, 2015). Cognitive attributes were assessed in multiple ways: through the use of a government mandated professional skills (literacy and numeracy) test, through evaluation of academic qualifications such as A levels, GCSE grades in English and Math, and through evaluation of university degree performance or 'class' (i.e., 1st, 2:1, 2:2, etc.). Non-cognitive attributes were assessed through individual and group interviews (97%), assessment of social behaviours through group activities (62%), and formal personality tests (3%); however, the survey revealed no evidence of the robustness of assessment practices, and published research on the topic is rare (Klassen & Kim, 2019).

The range of selection methods used for ITE programs varies across and within countries. In Table 6.1 we report how a sample of international ITE programs assess cognitive, non-cognitive, and background factors for selection into their programs. Most of the ITE programs included in the sample evaluate cognitive factors through an assessment of achievement level of the completed degree (e.g., minimum qualification standards such as degree class (in the UK) or GPA (in American settings). In some jurisdictions, cognitive attributes are further assessed at selection; for example, in Singapore, entrance proficiency tests are used to test subject knowledge in some subjects, and in the U.S., scores from a basic skills test in math, reading, and writing

Table 6.1 Selection criteria for initial teacher education programs (content accurate as of June 2021)

ITE program	Assessment of cognitive factors	Assessment of non-cognitive factors	Assessment of background factors/experience
Australia (University of New South Wales) • 1-year Master of Teaching (Primary) Postgraduate	• Completed UG degree with a major in one of six key learning areas	• Personal statement (150 words) outlining "reasons for wanting to be a teacher and why you are suited for that role" (scored on a 2-point scale)	Not listed
Canada (University of British Columbia) • 11-month BEd program (Elementary) Postgraduate	• Completed UG degree • 65% GPA on last 60 credits of post-secondary coursework, including English, laboratory science, mathematics, Canadian history or geography, Canadian studies	• Personal profile essay (maximum 900 words) responding to prompts: "Why do you want to become a teacher? Describe the kind of teacher you want to be." • Two reference letters "speaking to the applicant's experiences, interests, and abilities relevant of the teaching profession"	• A minimum of 100 hours of practical experience working with groups of 10 or more children or youth • Two reference letters "speaking to the applicant's experiences, interests, and abilities relevant to the teaching profession"
Finland (University of Helsinki) • 1-year STEP program) Postgraduate	• Completed Master's degree in teaching subject and Pedagogical Studies for Teachers	• Individual interview to evaluate suitability to work as a teacher (educability, motivation, and commitment). A 'multiple mini-interview is now being used in some Finnish contexts'	Not listed

(continued)

Table 6.1 (continued)

ITE program	Assessment of cognitive factors	Assessment of non-cognitive factors	Assessment of background factors/experience
Korea (Seoul National University) • 4-year Bachelor of Primary Education • 2.5-year Master of Primary Education	• Excellent high school grades (mark not specified but based on competitiveness) • Completed UG degree • Entrance proficiency test	• Individual interview with panel • (Letter of recommendation from school and letter of recommendation from a teacher) • Individual interview with panel	Not listed
Malawi (national selection for primary teacher education) Post-secondary	• Secondary school academic record • 'Aptitude test' at selection including tests of reasoning, mathematics, English	• Situational judgment test at selection assessing non-cognitive attributes (empathy, communication, organization, resilience, adaptability, integrity, community relationships, motivation, commitment, reflection, creativity, autonomy	Not listed
Singapore (National Institute of Education) • 1-year diploma in education Postgraduate	• Completed UG degree • Entrance proficiency tests (for some subjects)	• Interview (writing exercise; role play; individual interview with panel) • Evaluation of suitability during compulsory 4-month 'untrained teaching stint'	• Compulsory 4-month teaching stint (if selected for appointment) that must be passed before students continue in the program
United Kingdom (University of Newcastle) • 1-year certificate in education Postgraduate	• Completed UG degree at 2.2 level or above • Written and mathematics task on selection day	• Selection day activities include an individual interview assessing motivation, self-awareness, educational issues, classroom management	• 'Some experience of the English school system, usually through school observation visits'

(continued)

Table 6.1 (continued)

ITE program	Assessment of cognitive factors	Assessment of non-cognitive factors	Assessment of background factors/experience
United Kingdom (Teach First)	• A 2.1 degree or above • Grade C in GCSE Maths and English	• Two stage process. First, online application that assesses seven competencies (e.g., motivation, leadership, resilience). An SJT is used to provide 'a preview of life as a Teach First trainee'. Second, a 'development center' that includes a 1-1 competency-based interview, a group exercise with self-evaluation, and a teaching demonstration with self-evaluation	Not listed
USA (University of Washington) • 1-year Master's in Teaching program (Elementary) Postgraduate	• Completed UG degree with 3.0 GPA • WEST-B Basic Skills Test (Math, Reading, Writing) • ACT or SAT scores • NES (state-mandated content knowledge test) • 3.0 GPA from 60 credits	• Goal statement highlighting passion for teaching underserved populations and an interest in social justice + completion of 'character and fitness form' • Two letters of recommendation	• Minimum of 40 hours of documented education-relevant experience • Resumé detailing work and academic experience + prerequisite courses on education of an ethnic group, developmental psychology, and math for elementary teachers

Note Sources are publicly available ITE program websites except for Malawi, which is from Klassen et al. (2018)

provide further data beyond the information provided in undergraduate degree transcripts. The theme running through all selection processes is that decisions are made based on the evaluation of cognitive and non-cognitive attributes and sometimes including an evaluation of relevant background factors.

6.4 Selecting Teachers into Employment

The methods used to select prospective teachers into employment have been described as "ad hoc" (p. 24, Goldhaber et al., 2014) and "information poor" (p. 324, Liu & Johnson, 2006), with weak empirical and theoretical foundations supporting their use. The lack of research analyzing the effectiveness of selection methods is surprising in light of the importance of teachers in achieving societal goals of social equality and improving knowledge levels, and in light of the knowledge about selection we have accrued in other fields.

The kinds of methods chosen for selection depends on the volume of applicants and the degree of centralization of the hiring process. In more centralized systems with a large volume of applicants (e.g., Austria, Italy, Korea, Malaysia, Malta, Mexico, Singapore, Spain, and Turkey), selection methods tend to be standardized, with the specific methods set by central bodies. In less centralized systems (e.g., Belgium, Bulgaria, Hungary, Norway, Poland), schools, and especially school principals, have considerable autonomy in hiring. Other countries used a combined system where a central office may screen applicants at an initial stage, but individual schools make final hiring decisions (e.g., Australia, Canada, United States). Whether the system is centralized or less centralized, the methods chosen for selection typically target a combination of cognitive attributes, non-cognitive attributes, and background factors (e.g., teaching experience).

Based on data from PISA 2012, Han (2018) found that decentralized selection processes (i.e., school-based hiring) are associated with greater variance in the distribution of teacher quality across schools, and a greater gap in achievement between low- and high-SES students. Although Han's data did not speak to specific selection methods used, research from organizational psychology suggests that the methods used by smaller employing units (e.g., schools that might be hiring one or two teachers) tend to be more idiosyncratic and less reliable than more systematic and structured selection methods used by larger organizations.

In decentralized systems, individual school principals play a key role in deciding the elements assessed during the selection process. Engel and Finch (2015) interviewed 31 principals of urban schools in Chicago about their decision-making processes when making new teacher hiring decisions. Evaluation of cognitive attributes, especially subject area knowledge, was typically done through collaboration with school colleagues, but the attributes targeted during selection tended to be determined non-systematically, e.g., "We sort of sit down and talk a little bit about what we are looking for… You know, to decide, what kind of person do we want to have here?" (p. 32). Systematic differences were found in the strategies principals

used to recruit and hire new teachers: principals in lower achieving schools were more likely to hire substitute or student teachers than principals in higher achieving schools (who accessed larger social networks in hiring), and principals in primary schools tended to work more autonomously throughout the hiring process than principals in secondary schools.

Two case studies: hiring teachers in the U.S. (NYC) and Australia (NSW). The methods used for selecting teachers for employment are similar across contexts. The New York City Department of Education Hiring Guide (2018–2019) outlines the use of individual interviews to target cognitive attributes such as *content knowledge* (How would you make your content area relevant to daily life?) and *instructional practice* (What specific strategies do you use for classroom management?). The assessment of non-cognitive attributes includes individual interviews targeting *beliefs and strengths* (Why did you become a teacher? What are three words to describe yourself as a teacher?), *collaboration* (How do you feel about collaborative teaching?) and *student understanding* (Does a student's background influence his or her achievement?). Assessment of background factors, and especially teaching skills, is optionally assessed through a demonstration lesson where students are evaluated on their "poise and comfort in front of a group" and on how well applicants test for student understanding.

The New South Wales Department of Education is the biggest employer of teachers in Australia, with over 2,200 schools ranging from very remote to very urban settings. Current government policy requires teacher education providers to select teachers based on both cognitive and non-cognitive attributes to ensure suitability for teaching (Sheridan et al., 2021). The 'New standards for NSW's teachers' document (NSW government, 2018) highlights five main criteria required for graduates to be considered for teaching positions:

- A minimum credit grade point average
- Sound practical knowledge and ability
- Superior cognitive and emotional intelligence measured by psychometric assessment
- Commitment to the values of public education displayed in an interview
- Preference for face-to-face teaching degrees over online degrees

Both cognitive and non-cognitive attributes are assessed through Teacher Suitability Assessments (https://www.teach.nsw.edu.au/becomeateacher/approval-to-teach/faqs) that include measures of verbal reasoning (ACER Advanced Test), abstract reasoning (ACER APTS Abstract Reasoning Organisational), and emotional intelligence (Genos Emotional Intelligence Inventory). During the COVID-19 pandemic, online interviews were used to assess **knowledge** (including pedagogy and syllabus content), **critical experiences** (demonstration of actions that have contributed to student progress and wellbeing), and **skills and capabilities**. Targeted cognitive attributes are aligned with the Australian Professional Standard for Teachers, e.g., *Know students and how they learn; Know the content and how to teach it; Plan for and implement effective teaching and learning.* An assessment of

professional experience (and/or practicum reports) is used to identify readiness for success in the classroom.

6.5 How Valid Are Current Selection Methods?

In education, there has been little systematic research examining the efficacy of selection methods (Bowles et al., 2014; Liu & Johnson, 2006). Many of the existing selection methods are based on ad hoc decisions with little evidence supporting their use. A recent study conducted in the UK (Davies et al., 2016) explored how selection methods were developed for teacher education programs, with the finding that selectors emphasized their intuition when making selection decisions: "Really, you do only need one question and a bit of conversation to find out what you are looking for" (p. 298), with selectors tending to rely on a "gut feeling" to identify the "X factor" (p. 298). No evidence was gathered to support the selection methods used: "I wouldn't have any statistics… after they've left us," (p. 297).

Most people are confident that they can accurately judge personality and other personal characteristics through interviews (Dana et al., 2013), but research tells us otherwise. Research from organizational psychology suggests that interviewers, especially when conducting unstructured interviews, suffer from unreliable judgment and are influenced in the decisions they make by unconscious biases based on race, age, and appearance (Cook, 2009). In education, selection methods may pay lip service to well-developed teaching standards frameworks that reflect multiple competencies and values (e.g., Casey & Childs, 2017; Denner et al., 2001), but the methods chosen for selection may not reliably assess these competencies. In any field, selection methods require regular evaluation of their reliability (consistency over time, validity (evidence of predictive utility), and fairness for all applicants, regardless of age, gender, ethnicity, sexuality, and socio-economic status.

6.6 Reviews of Research on Teacher Selection Practices

Research on the efficacy of teacher selection methods is under-developed in comparison to research in other professional fields, such as medicine or business. In Chap. 5, we reviewed selection practices in other fields, and considered the evidence available for the selection methods used in these fields. We saw that systematic research on selection into employment and training is particularly well developed in medical education, where a systematic program of research has been conducted into the reliability, validity, and fairness of selection methods. However, much less research attention has been given to the methods used for teacher selection. Two review studies sum up the field: Metzger and Wu (2008) and, more recently, Klassen and Kim (2019).

Metzger and Wu's 2008 meta-analysis. In 2008, Metzger and Wu reviewed and meta-analyzed 24 studies that examined the predictive validity of one teacher

selection tool, the Gallup Teacher Perceiver Interview (TPI). The review proves a useful starting point in investigating selection practices in education because it examined the use of what was, in the 1980s and 1990s, one of the most widely used teacher selection tools in the United States. Metzger and Wu's meta-analysis used validity data from studies published from 1975 to 2003, largely from dissertations ($n = 16$), but also from reports from the Gallup Organization ($n = 7$) and from one journal article. Most of the studies included in their review (20/24) were published before 2000, with 6 studies from the 1970s, 10 studies from the 1980s, and 4 studies from the 1990s. Of the four post-2000 studies, one was a dissertation, (Buresh, 2003), one was a journal article (Young & Delli, 2002), and two were released by test companies (i.e., the Gallup organization). Overall, the authors found a range of −0.12 to 0.87 for the correlation between TPI scores and indicators of teaching effectiveness, with a weighted mean of $r = 0.28$, and a median r of 0.22, considered by the authors to be a moderate effect size. Although Metzger & Wu's meta-analysis provided a valuable snapshot of one selection tool at a particular point in time, more work is needed to provide a fuller, more accurate, and more up-to-date picture of the teacher selection landscape.

Klassen and Kim's 2019 meta-analysis. Klassen and Kim extended Metzger & Wu's, 2008 review in 2019 (2019) in order to broaden the coverage of all teacher selection methods in use, and to provide a more up-to-date look at selection (most of the studies Metzger and Wu included were published before 2000). The goal stated by Klassen and Kim was to examine the methods used for the selection of teachers for employment and prospective teachers entering ITE. Four key questions were posed in their review:

1. What is the predictive validity of the methods used to select teachers and teacher candidates?
2. Are there differences in the predictive validity of tests assessing cognitive and non-cognitive attributes?
3. Are there differences in the predictive validity of the methods used for selection into employment and for selection into ITE programs?
4. What is the relationship between cost and benefit (predictive validity of selection methods?

Method. The key indicator of effect size for the meta-analysis was Pearson's r, which is a measure of the size of relation between selection method and teacher effectiveness, and which can be interpreted as an indication of predictive validity. In educational research $r = 0.10$ describes a small effect, $r = 0.20$ describes a medium effect, and $r = 0.30$ describes a large effect. Coe (2002) proposed that an effect size of $d = 0.10$ (roughly $r = 0.05$) can result in important educational outcomes if the effect can be applied to all students (i.e., as in an effect involving teachers) and is cumulative over time. For context, predictive validity coefficients in other fields are as follows: 0.18 to 0.43 in dentistry (Patterson et al., 2012), 0.37 in medicine (Lievens & Patterson, 2011), between 0.06 to 0.50 in business (Christian et al., 2010), and 0.34 across multiple occupation groups (McDaniel et al., 2001).

The goal of the meta-analysis was to analyze studies that: (a) reported primary research in the form of journal articles, dissertations, and published research reports published between 2000 and 2017, (b) included participants who were job applicants or ITE candidates in the K-12 system, (c) included a selection measure (cognitive or non-cognitive) administered at the point of selection, and (d) included a measure of teacher effectiveness using an *external* source (i.e., not self-reported), either observation scores (from supervisor or principal) or classroom-level student achievement gains. The authors excluded (in contrast to Metzger and Wu) un-verified data provided by test companies in support of their commercial products. The search of relevant databases and key journals resulted in 1306 records which were then screened for relevance to the study. A series of further screens left a pool of 32 studies that met the criteria for inclusion, and which were included in further analyses.

Results. Table 6.2 presents a summary of the research questions, the results, and the implications for research and practice. An overall effect size of $r = 0.12$ was found for the relationship between selection method and teacher effectiveness. Out of the 32 studies, 28 showed positive effect sizes, and 4 showed negative effect sizes, but only 10 studies reported statistically significant findings, all positive. The moderator analyses, conducted in order to break down the relations between predictors and outcomes, showed that cognitive predictors ($r = 0.13$) were significantly more predictive of teacher effectiveness than non-cognitive predictors ($r = 0.10$). Methods to select candidates for ITE programs were nominally more predictive than selecting candidates into employment, but the difference was not significant. There was no indication that paying more for selection methods (in money and time) resulted in better outcomes.

Conclusions of the study. There are several key conclusions that can be drawn from the study. First of all, there has been much less research and development attention paid to selection methods in education than in other fields, with the result that the current methods are not as effective as in other professions. The existing methods are, in general, not very effective at predicting which candidates will be successful in ITE programs or as teachers in schools. Notwithstanding the fact that small validity coefficients can be usefully applied at the systems-level, there are several possible explanations for the lack of predictive validity of current selection methods in education. We know that in fields where selection methods are closely studied, there is growth and development in the methods that are used; for example, in medicine, the relatively extensive body of research has led to new selection methods—e.g., SJTs and MMIs—being developed, tested, and implemented. In education, most jurisdictions that were studied used commercial tools that have little published evidence of validity, or in-house methods that have been developed 'organically' but again, neither reflect best practices in current selection research, nor have a base of evidence supporting their use. Looking outside of teacher education and educational psychology to medical education and organizational psychology, where research on selection methods is extensive, is one way to refresh the current moribund state of teacher selection.

Table 6.2 Results from meta-analysis of research on teacher selection tools

Research question	Results	Implications
What is the overall predictive validity of the methods used to select teachers and teacher candidates?	Overall $r = 0.12$ (small effect)	Predictive validity of current teacher selection methods, on the whole, is modest, especially in comparison with other professions
Are there differences in the predictive validity of tests assessing cognitive and non-cognitive attributes?	Cognitive predictors: $r = 0.13$ Non-cognitive: $r = 0.10$ (Sig. difference)	Cognitive predictors are slightly better at predicting teacher effectiveness than non-cognitive predictors. Finding valid methods to evaluate prospective teachers' non-cognitive attributes remains a challenge
Are there differences in the predictive validity of the methods used for selection into employment and for selection into ITT programs?	Employment: $r = 0.11$ ITT programs: $r = 14$ (No sig. difference)	Although the validity of selection methods for ITT is nominally higher, the difference is not significant. The methods used to select teachers for employment and for training need further research and development
What is the relationship between cost and benefit (predictive validity) of selection methods?	Mean cost: US $104 per candidate Relationship between validity and cost: $r = -0.12$	Spending more money on teacher selection methods is no guarantee of success; the commercial methods currently available seem no better than 'in-house' methods
What are the overall conclusions?		Three main conclusions can be drawn: (a) there is a lack of research on teacher selection methods in comparison to other professions, and (b) the current methods are largely stagnant, and not very good, and (c) more attention should be paid to methods used in other fields

Note Data from Klassen and Kim (2019)

6.7 Chapter Summary

Questions about how to select the best possible teachers have been asked for at least a century, with researchers and practitioners in the early 1900s pondering the challenges of teacher selection. In this chapter we explored various models used to select prospective teachers into ITE and into employment and found that there was a general agreement that both cognitive and non-cognitive methods were necessary for

successful selection. Reviews of research on teacher selection methods found that there was less research in education than in other fields, and perhaps consequently, the methods used for selection were not very effective and were not reflective of leading-edge research and practice found in other professional and research fields. In the following chapter, we explore new approaches to teacher selection that have emerged in the last few years. In particular, we look at how situational judgment tests (SJTs) have been developed to identify prospective teachers who are most likely to experience success in ITE programs and in teaching jobs.

References

Antoniou, P., Kyriakides, L., & Creemers, B. P. M. (2015). The Dynamic Integrated Approach to teacher professional development: Rationale and main characteristics. *Teacher Development, 19*, 535–552.

Atteberry, A., Loeb, S., & Wyckoff, J. (2015). Do first impressions matter? Predicting early career teacher effectiveness. *AERA Open, 1*(4), 1–23.

Barber, M., & Mourshed, M. (2007). *How the world's best-performing schools systems come out on top*. McKinsey & Company.

Bowles, T., Hattie, J., Dinham, S., Scull, J., & Clinton, J. (2014). Proposing a comprehensive model for identifying teaching candidates. *The Australian Educational Researcher, 41*, 365–380.

Buresh, R. J. (2003). *The predictive validity of the Teacher Perceiver Interview in selecting effective elementary teachers in a mid-sized midwestern school district (Unpublished doctoral dissertation)*. University of North Dakota.

Casey, C., & Childs, R. (2011). Teacher education admission criteria as measure of preparedness for teaching. *Canadian Journal of Education, 34*, 3–20.

Casey, C., & Childs, R. A. (2017). Teacher education program admission criteria and what beginning teachers need to know to be successful teachers. *Canadian Journal of Educational Administration and Policy, 67*, 1–24.

Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology, 63*, 83–117. https://doi.org/10.1111/j.1744-6570.2009.01163.x

Dana, J., Dawes, R., & Peterson, N. (2013). Belief in the unstructured interview: The persistence of an illusion. *Judgment and Decision Making, 8*, 512–520.

Davies, P., Connolly, M., Nelson, J., Hulme, M., Kirkman, J., & Greenway, C. (2016). 'Letting the right one in': Provider contexts for recruitment to initial teacher education in the United Kingdom. *Teaching and Teacher Education, 60*, 291–302.

Denner, P. R., Salzman, S. A., & Newsome, J. D. (2001). Selecting the qualified: A standards-based teacher education admission process. *Journal of Personnel Evaluation in Education, 15*, 165–180.

Engel, M., & Finch, M. A. (2015). Staffing the classroom: How urban principals find teachers and make hiring decisions. *Leadership and Policy in Schools, 14*, 12–41.

Goldhaber, D., Grout, C., and Huntington-Klein, N. (2014). *It's selective but is it effective? Exploring the predictive validity of teacher selection tools*. CEDR Policy Brief 2014–9. University of Washington, Seattle, WA.

Han, S. W. (2018). School-based teacher hiring and achievement inequality: A comparative perspective. *International Journal of Educational Development, 61*, 82–91.

Klassen, R. M., & Dolan, R. (2015, September). *Selection for teacher education in the UK and the Republic of Ireland: A proposal for innovation*. Presented at the meeting of the European Conference on Educational Research, Budapest, Hungary.

Klassen, R. M., & Kim, L. E. (2019). Selecting teachers and prospective teachers: A meta-analysis. *Educational Research Review, 26*, 32–51.

Knight, F. B. (1922). Qualities related to success in elementary school teaching. *The Journal of Educational Research, 5*, 207–216.

Lievens, F., & Patterson, F. (2011). The validity and incremental validity of knowledge tests, low-fidelity simulations, and high-fidelity simulations for predicting job performance in advanced-level high-stakes selection. *Journal of Applied Psychology, 96*, 927–940.

Liu, E., & Johnson, S. M. (2006). New teachers' experiences of hiring: Late, rushed, and information poor. *Educational Administration Quarterly, 42*, 324–360.

Malinen, O.-P., Vaisanen, P., & Savolainen, H. (2012). Teacher education in Finland: A review of a national effort for preparing teachers for the future. *The Curriculum Journal, 23*, 567–584.

McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology, 86*, 730–740.

Metzger, S. A., & Wu, M.-J. (2008). Commercial teacher selection instruments: The validity of selecting teachers through beliefs, attitudes, and values. *Review of Educational Research, 78*, 921–940. https://doi.org/10.3102/0034654308323035

NSW government (2018). *New standards for NSW's teachers.* Retrieved from: https://www.nsw.gov.au/news/new-standards-for-nsws-teachers

Patterson, F., Ashworth, V., Mehra, S., & Falcon, H. (2012). Could situational judgement tests be used for selection into dental foundation training? *British Dental Journal, 213*, 23–26.

Patterson, F., Knight, A., Dowell, J., Nicholson, S., Cousans, F., & Cleland, J. (2016). How effective are selection methods in medical education? A systematic review. *Medical Education, 50*, 36–60.

Rimm-Kaufman, S. E., & Hamre, B. K. (2010). The role of psychological and developmental science in efforts to improve teacher quality. *Teachers College Record, 112*, 2988–3023. http://www.tcrecord.org

Sheridan, L., Coleman, B., & Durksen, T. L. (2021). Important non-academic attributes in Australian initial teacher education. *The Australian Educational Researcher*. https://doi.org/10.1007/s13384-021-00443-x

Tubbs, E. V. (1928). The selection of teachers. *Peabody Journal of Education, 5–6*, 323–332.

Young, I. P., & Delli, D. A. (2002). The validity of the Teacher Perceiver Interview for predicting performance of classroom teachers. *Education Administration Quarterly, 38*, 586–612. https://doi.org/10.1177/0013161X02239640