

# The development of a situational judgement test of personal attributes for quality teaching in rural and remote Australia

Tracy L. Durksen<sup>1</sup> · Robert M. Klassen<sup>2</sup>

Received: 18 April 2017 / Accepted: 19 November 2017 / Published online: 9 December 2017  
© The Australian Association for Research in Education, Inc. 2017

**Abstract** Education authorities in Australia are calling for valid tools to help assess prospective teachers' non-academic attributes, with a particular need for identifying those attributes necessary for effective teaching in specific contexts. With the New South Wales (NSW) Department of Education, we aimed to develop a scenario-based tool to help assess the non-academic attributes necessary for beginning teachers working in rural and remote settings. Using a situational judgement test (SJT) methodology, we worked with experienced teachers ( $n = 19$ ) to develop scenario-based items, which were then reviewed by school principals ( $n = 13$ ). The teachers also adapted items previously developed and piloted in the UK. Next, prospective NSW teachers ( $n = 99$ ) tried 32 new and adapted items targeting four clusters of attributes: empathy and communication, resilience and adaptability, organisation and planning, and culture and context. Item quality analyses revealed 22 acceptable items. We conclude by suggesting SJTs for promoting non-academic growth of prospective and beginning teachers.

**Keywords** Rural teaching · Personal characteristics · Non-academic attributes · Teacher recruitment · Teacher retention · Critical incidents

---

✉ Tracy L. Durksen  
tracy.durksen@unsw.edu.au

<sup>1</sup> School of Education, Faculty of Arts and Social Sciences, UNSW Sydney, 131 John Goodsell Bldg, Sydney, NSW 2052, Australia

<sup>2</sup> University of York, York, UK

## Introduction

Teachers frequently leave the profession after only a few years of teaching (Kutsyuruba et al. 2014). The situation in rural and remote Australian schools is even more urgent with a turnover rate up to six times higher than city schools and many teachers choosing to leave their school within 2 years (Lyons 2006). High turnover has been attributed to teachers recruited for ‘hard-to-staff’ schools that are ‘easy to leave’ with their experience viewed as a ticket to a permanent city position (Reid et al. 2010). Although motivations for a beginning teacher to leave can vary, isolation and adjusting to differing cultural or community practices have been identified as additional stresses that can complicate what Huberman (1989) has called the *survival* career phase. With this career phase often defined by *sink or swim* or *lost at sea* responses (Ingersoll and Strong 2011), critical personal characteristics or ‘non-academic’ attributes like resilience are paramount for quality teaching (AITSL 2015). Since the majority of teachers in rural and remote settings have less than 4 years experience, we sought to develop a tool that can not only help identify, but also help promote further development of key attributes deemed necessary for teaching, remaining, and thriving in rural or remote educational settings.

Several terms (e.g. country, bush, regional) are used when discussing Australian areas outside of metropolitan centres, but for the purposes of this study we use rural and/or remote to describe more isolated educational settings. Rural or remoteness in Australia is statistically defined in terms of road distance from major cities (Baxter et al. 2011). In the educational context, rural and remote may be characterised by population size or movement to city centres for key services with a recognition of specific needs being directly related to a family’s geographic location and access to a range of resources (MCEETYA 2001). Overall, a much higher proportion of the population living in remote areas of Australia are Indigenous (Baxter et al. 2011). The term Indigenous encompasses and identifies Aboriginal and Torres Strait Islander peoples as the First Peoples of Australia. The current study was conducted in the state of New South Wales (NSW), home to 7.5 million people (2.9% of whom identify as Indigenous) and the largest (33%) Indigenous population of Australia (Australian Bureau of Statistics 2017). With teaching demands varying across contexts (e.g. cultural context influences how teachers develop, how educational systems are framed, and how teaching practices are valued; King and McInerney 2016), we sought to develop a culturally and contextually responsive way of assessing and increasing the retention of quality teachers in rural and remote settings.

## Personal attributes for teaching

When considering a teaching position in a rural or remote setting, critical reflection is encouraged in order to determine whether a list of personal attributes includes “adaptable, resilient, independent, sensitive, practical, strong in heart and mind, tolerant, organised, respecting of others, resourceful and a lateral thinker” (Austin

2010, p. 30). According to a state strategy for rural and remote education, “graduates from NSW teacher education programs will have the skills and personal attributes for teaching in rural and remote schools” (NSWDEC 2013, p. 11). Not only are initial teacher education (ITE) providers expected to produce classroom-ready teachers (TEMAG 2014), but also to prepare context or ‘remote ready’ graduates. In response, the NSW Department of Education (hereafter, the Department) developed an interview protocol that explicitly asks questions to assess personal suitability for rural or remote teaching. Although the new interview protocol greatly enhanced the approval-to-teach process, the Department continued to seek best-practice approaches to developing related valid and reliable methods for measuring personal characteristics.

The challenge remains, however, with agreeing on a definition of ‘quality’ teacher or teacher effectiveness (Rice et al. 2017). Research on teacher quality has highlighted the importance of personal attributes (i.e. beliefs, attitudes, and dispositions) such as self-awareness, adaptability, motivation, and personality (Buehl and Fives 2009). For example, a recent review (Klassen and Tze 2014) found personal attributes as predictors of teacher effectiveness. These predictors included teacher self-efficacy (i.e. a teacher’s belief in his or her capabilities to influence student learning; Klassen et al. 2011) and teacher resilience (i.e. “what sustains teachers and enables them to thrive rather than just survive in the profession”, Beltman et al. 2011, p. 186). Research on teachers’ emotions (Frenzel et al. 2015), engagement (Klassen et al. 2013), and teacher–student interactions (e.g. Hamre et al. 2013) has also helped advance our understanding and build support for the need to identify and foster teachers’ non-academic attributes—particularly within the complex social environment of teaching (e.g. interpersonal experiences with students, parents, colleagues; Martin and Dowson 2009).

In Australia, some teacher educators claim “a successful selection process (for ITE programs) will identify candidates who are highly resilient and thus likely to manage any stress associated with teaching, without it impacting on their teaching performance” (Sautelle et al. 2015, p. 57). Yet others argue that the “focus should not be on who is allowed into teaching courses (in NSW) but who is admitted into the profession” (Spence as cited by McNeilage 2014, n.p). Initial teacher education providers are thus looking at adopting evidence-informed selection tools that can help measure prospective teachers’ non-academic attributes. While the Australian Institute for Teaching and School Leadership (AITSL 2015) has set new selection guidelines for ITE providers, there remains a concern over if, how, and which personal attributes should be assessed at the entrance stage of a developmental professional program (Gore et al. 2016).

Although professional programs frequently use selection practices such as interviews, personal statements, or reference letters when assessing non-academic attributes for admission, most methods lack evidence of validity and reliability (Goldhaber et al. 2014). University-wide admissions in Australia primarily set academic requirements based on standardised measures (e.g. Australian Tertiary Admission Ranking) despite mixed results as to whether academic scores accurately predict performance in ITE programs (Caldwell and Sutton 2010). With the inherent self-presentation bias through forms of self-report on non-academic attributes, ITE

providers are also wary of selection recommendations that would require tools like personality tests for entry decisions. In fact, the majority of ITE providers in NSW are embracing a post-selection focus on non-academic attributes by seeking valid and reliable tools that can help assess and confirm personal suitability early on and throughout their programs and when their graduates enter the profession (NSW Council of Deans of Education 2016).

While there is some evidence that prospective teachers' beliefs may change very little over the course of an ITE program (Stenberg et al. 2014), there is increasing evidence that related attributes can be actively developed. For example, Bahr and Mellor (2016) state that kindness and care can be taught and modelled during ITE programs and explored through scenario-based problems. In addition, researchers with a professional development lens found that teacher resilience can be developed based on context (Day and Gu 2014). Therefore, we aimed to develop a contextualised tool—a situational judgement test—that can not only help identify but broaden teachers' options when coping with everyday teaching challenges and thus, help strengthen the non-academic attributes required for effective and sustainable teaching.

### **Situational judgement test as a measure of attributes**

The emergence of situational judgement tests (SJTs)—scenario-based measurement methods designed to assess individuals' judgement in contextualised workplace settings (Ryan and Ployhart 2014)—provides a potentially valid, evidence-informed, and innovative approach to assessing key non-academic attributes deemed necessary for teaching. The SJT presents a range of scenarios and requires judgements on the appropriateness of possible responses. While critical incidents (i.e. typical but challenging situations; Tripp 1994) have long been used to help prospective teachers develop professional judgement, SJTs offer a rigorous procedure designed to gain insight into the psychological characteristics underpinning context-specific professional judgement (Patterson et al. 2015).

SJTs are considered a particularly effective methodology for competitive high-stakes situations such as selection into medical school (Patterson et al. 2015) or large-scale job application processes in other professions. Unlike personality measures, SJTs provide little opportunity for self-presentation or inter-group (gender, ethnicity) bias, and have a lower susceptibility to faking and coaching (e.g. Stemig et al. 2015). This is due in part to the indirect response styles (e.g. ranking, choose) of the SJT being different from typical response scales (i.e. multiple choice, strongly disagree to strongly agree). Literature on SJT scoring details a range of response formats (e.g. Whelpley 2014), two of which were piloted in the current study: (1) ranking five response options from most to least appropriate and, (2) choosing three as the most appropriate responses from a set of eight possible options. An example of an item is presented in the Appendix.

Previous work (e.g. Klassen et al. 2014) involved trialling SJTs with applicants to ITE programs in the UK since this methodology has been validated for selection in a range of professions. More recently, SJTs have been explored for training purposes (humanitarian disaster relief; Cox et al. 2017). The use of SJTs as large-scale

selection tools for entry into ITE has shown promise in the UK (Klassen et al. 2017), but to the best of our knowledge, the current study is the first to build the foundation required for using SJTs with beginning teachers and with the intent to identify and guide contextualised professional learning priorities during this critical ‘survival’ phase of teacher development.

## Theoretical framework

Since teaching is influenced by a combination of implicit and explicit beliefs, our study was framed by the theory of implicit trait policies (Motowidlo and Beier 2010; Schultheiss and Brunstein 2010). This framework holds that some aspects of beliefs and motivations operate ‘on the surface’ (explicitly) with other aspects operating implicitly and, thus, may not be readily accessible by self-report. For example, in a personal suitability interview, a new teacher may state that s/he is generally agreeable and easy-going but self-presentation bias may prevent a related conclusion that being agreeable is not the best response in a particular situation. Thus, the SJT aims to indirectly access both the implicit and explicit components of a prospective teacher’s reasoning process when faced with a challenging situation.

When prospective teachers make a judgement and respond to a challenging teaching scenario, they may rely on explicit and implicit attributes. With the SJT, beginning teachers read scenarios and consider what *should* be done to resolve the situation given the context, relying on explicit knowledge (*I know that this is the right action*) as well as implicit knowledge (*I sense that this might be the right action in this context*). This knowledge-based approach differs from the typical over-confident responses to critical incident activities designed to reveal self-reported behavioural tendencies (e.g. what *would* you do in this challenging situation?). A behavioural-based approach, such as personality testing, is more likely than the SJT to encourage over-confident responses—particularly in high-stakes situations (Whetzel and McDaniel 2016). Therefore, SJT-based scenarios may help reveal beginning teachers’ underlying beliefs and motivations when they grapple with any persisting implicit and explicit imbalances.

## The current study

Our goal was to develop and pilot a contextualised SJT representing key non-academic attributes for teaching in NSW. Previous iterations of teaching-specific SJTs have focused on three clusters of attributes: empathy and communication, organisation and planning, and resilience and adaptability (Klassen et al. 2014). In NSW, we also explored a fourth cluster specific to rural and remote teaching in NSW. As such, the current study lays the foundation for implementing a tool that could help identify quality teachers who are ‘remote ready’ at the end of their education program and that could also offer direction for a related professional learning plan. Overall, we aimed to (a) develop new scenario-based items specific for secondary teaching in NSW, (b) explore established target attributes and a potential NSW-specific focus on rural and remote teaching, (c) contextualise piloted

UK items for NSW, and (d) pilot new and contextualised items specific to NSW secondary and/or rural and remote teaching settings.

## Methodology

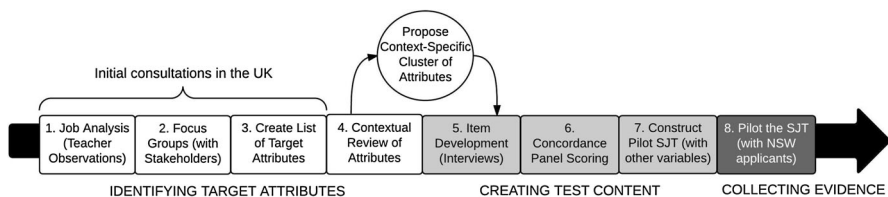
The study began with the development of SJT-based items (Phase One) and concluded with an initial SJT pilot (Phase Two). Building on previous proof-of-concept work for selection into UK teacher education programs (Klassen et al. 2014), we applied the rigorous and expert-informed procedure for developing teacher-specific SJT items at a different stage: entry into the teaching profession (see Fig. 1).

### Participant recruitment

During Phase One, 29 local educators recommended by the NSW Department of Education helped develop and review SJT items. For Phase Two, we recruited pilot study participants through the Department. We began with 51 applicants who were (a) seeking their approval to teach in NSW and (b) volunteered to participate immediately after their personal suitability interview. In addition, 16 attendees of a career fair (designed for teacher education students or recent graduates) consented to participate. Lastly, we received data from 32 attendees of a program that provides teacher education students with a short-term teaching experience in remote NSW.

### Phase one: development of scenario-based items

A number of steps are involved when developing and piloting a context-specific SJT-based questionnaire (see Fig. 1). Following the approach of Klassen et al.'s (2014) initial consultations in the UK (Steps 1–3), we enlisted voluntary NSW teachers to review the target non-academic attributes and indicate whether they were considered necessary for effective teaching in NSW (Step 4). We also invited feedback on a proposed NSW-specific cluster that aimed to highlight attributes thought to influence teacher quality and retention in rural or remote settings. In addition, each teacher was asked to describe at least one critical incident through a 30-min phone interview (Step 5). For each incident, we asked for possible responses and an indication of the best responses they would expect from a new teacher. Through the one-to-one phone conversations, teachers were also encouraged to



**Fig. 1** Process for developing a NSW-specific situational judgement test on teaching

identify which cluster(s) of attributes was the key target through the critical incident(s) provided.

The NSW teachers then attended a workshop to review the existing and new SJT items in small groups, with prompts to discuss whether contextual changes were needed for the scenarios and/or responses. Each participating teacher was also prompted to indicate what responses s/he believed were correct and whether consensus could be achieved within small groups. Lastly, we enlisted NSW principals to review and provide answers for the new scenario-based items. The principals formed our expert concordance panel (to review items and inform the scoring key) as each had experience conducting personal suitability interviews with those seeking their approval to teach in NSW government schools (Step 6).

## Phase 2: pilot study procedures and measures

The construction of the SJT questionnaire (Step 7) followed the format used in a pilot study with applicants to ITE in the UK (see Klassen et al. 2014). A selection of scales and demographic questions (e.g. school setting preference: rural/remote or city/metro) were included along with a feedback form. Three pools of potential participants were approached with an invitation to voluntarily complete our paper-pencil questionnaire within 1 h. During a two-month period, approximately 7% of the total number of approval-to-teach interviewees agreed to stay after their interview to independently complete the questionnaire.

Given the low response rate, we recruited additional participants from two opportunities: a career expo and a rural teaching experience. A small number of attendees (2.6%) volunteered to complete the questionnaire during the one-day career expo. The annual event—Australia's largest teaching careers expo—was held at a NSW university providing ITE and attracted students studying in education and students from other disciplines interested in a Masters of Teaching program. In addition, the majority (75%) of participants in a preservice rural teaching experience completed the questionnaire while on a week-long orientation and visit that showcased the lifestyle and career opportunities for teachers in rural and remote areas.

### *SJT*

The focus of the questionnaire was on presenting scenario-based items that targeted non-academic attributes (empathy and communication, organisation and planning, and resilience and adaptability) in addition to a NSW-specific target about rural and remote teaching: culture and context. Since the majority of new items created by NSW teachers required a choice of three responses, most of the items requiring ranked responses were selected from the pool of contextualised items (that resulted from the NSW teachers' review of piloted UK items). In the end, we compiled a pilot SJT that consisted of 32 scenario-based items: 22 items with *ranked* responses (eighteen contextualised from the UK to NSW and four new items) and ten new *choose three* items.

### *Additional measures*

We assessed three additional non-academic attributes using existing validated measures: personality, self-efficacy, and engagement. Personality was measured using the Ten-Item Personality Instrument (e.g. *I see myself as reserved, quiet*; Gosling et al. 2003) since previous SJT studies revealed correlations with personality measures. A reliable 6-item form of the Teachers' Sense of Efficacy Scale (Tschannen-Moran and Woolfolk Hoy 2001) measured teachers' self-efficacy in three domains: student engagement (e.g. *How much can you do to motivate students who show low interest in school work?*), classroom management (e.g. *How much can you do to calm a student who is disruptive or noisy?*), and instructional strategies (e.g. *How much can you do to implement a variety of assessment strategies?*). Teacher engagement was measured using the Engaged Teachers Scale (Klassen et al. 2013), which assesses cognitive engagement (e.g. *While teaching I pay a lot of attention to my work*), emotional engagement (e.g. *I feel happy while teaching*), social engagement with students (e.g. *In class, I show warmth to my students*), and social engagement with colleagues (e.g. *At school, I connect well with my colleagues*).

### **Analytical strategy**

We carried out descriptive and correlational analyses in order to summarise the pilot study responses. Using frequencies of SJT responses, we compared the results with the scoring suggestions provided through the UK pilot and by the NSW teachers and principals. Then we drafted a final scoring key. Next we used Patterson et al.'s (2013) formula<sup>1</sup> to determine the differences between each participant's score on each item and the scoring key. Each ranking item required participants to rank the appropriateness of options from 1 (most) to 5 (least) and each choice item required participants to choose three out of eight possible options as the most appropriate. Scores for each rank item were calculated out of 20 and each choice item out of 12. We then determined SJT item quality by examining the item partial—the degree of correlation between the item and the overall mean SJT score.

## **Results**

In this section, we present results from Phase One activities that highlight the developmental process of constructing a SJT for teaching (Steps 1–7) and results from Phase Two: the pilot study (Step 8).

---

<sup>1</sup> Please refer to Patterson et al.'s (2013) work for the formula as it is beyond the scope of this article to elaborate on the calculations.



**Table 1** Non-academic attributes and Australian Professional Standards for Teachers (APST)

| Non-academic attributes                                       | Definition   | APST <sup>a</sup> |
|---|--|-------------------|
| Empathy and communication                                     | Applicant demonstrates active listening, and engages in an open dialogue with both students and colleagues. Applicant is responsive to students' needs, and able to adapt the style of communication according to these needs  | 1, 2, 4, 5        |
| Organisation and planning                                     | Applicant has the ability to manage competing priorities and display time management skills effectively. Has an awareness of their level of competence, and willing to seek assistance when necessary  | 3, 5, 6, 7        |
| Resilience and adaptability                                   | Applicant demonstrates the capability to remain resilient under pressure. Demonstrates adaptability, and an ability to change lessons and the sequence of lessons accordingly where required. Has the confidence to make decisions independently, and is comfortable with challenges to own knowledge. Uses appropriate coping strategies  | 3, 5, 6, 7        |
| NSW-specific target for rural and remote teaching suitability | Initial draft definition: Applicant demonstrates the capability to adapt to remote settings. Demonstrates sensitivity to Indigenous cultural knowledge and practices. Applicant possesses the capacity to become culturally competent  |                   |
| ↓   |  |                   |
| <i>Culture and context</i>                                    | Refined definition: Applicant demonstrates the capability to adapt to remote settings and fit well into a small rural community. Recognises the importance of building relationships and maintaining professional behaviour in all aspects of community life. Demonstrates sensitivity to cultural knowledge and practices of Aboriginal and Torres Strait Islander peoples with the capacity to become culturally competent | 1, 2, 4, 6, 7     |

<sup>a</sup>The seven Australian Professional Standards for Teachers (APST; AITSL 2011) broadly align with the four clusters

1 = Teachers know their subject/content and how to teach that content to their students

2 = Teachers know their students and how students learn

3 = Teachers plan, assess, and report for effective learning

4 = Teachers communicate effectively with their students

5 = Teachers create and maintain safe and challenging learning environments through the use of classroom management skills

6 = Teachers continually improve their professional knowledge and practice

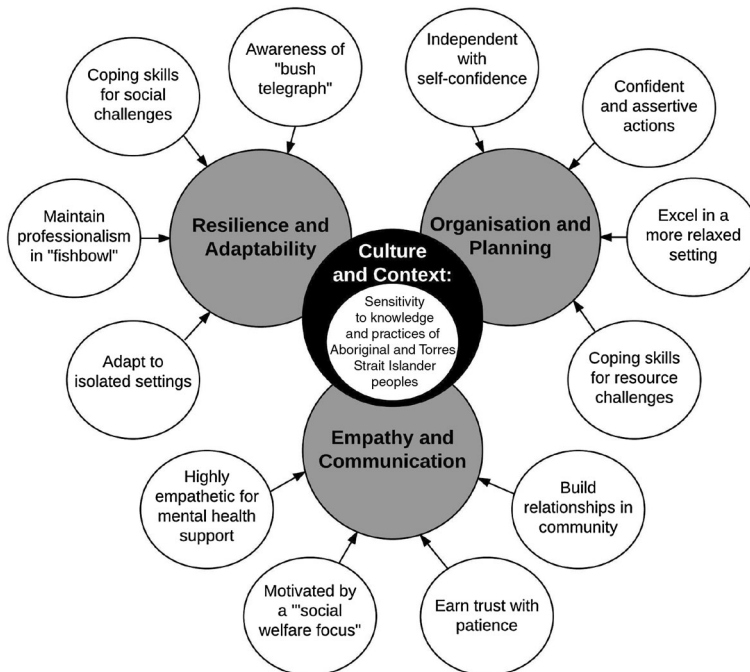
7 = Teachers are actively engaged members of their profession and the wider community

## Phase one

### *Rural and remote suitability*

Following previous job analyses, initial focus groups, and consultation with relevant stakeholders in the UK, three key clusters of non-academic attributes were defined (Klassen et al. 2014). We began the current study by contextualising the definitions of initial attribute clusters with NSW educators while proposing an additional target about rural and remote teaching in NSW. Together with the NSW educators, we confirmed the relevance of the initial three clusters and refined our proposed definition of a new cluster—culture and context—using indicators of rural and remote suitability. Table 1 presents the four key clusters of non-academic attributes as relevant for beginning teachers in NSW, with definitions broadly aligning with the Australian Professional Standards for Teachers (AITSL 2011).

NSW educators confirmed the initial three clusters of attributes as applicable to NSW-specific contexts and indicated that attributes specific to rural and remote suitability were worth further exploration. Some expressed the importance of assessing for rural and remote suitability since a teacher “[needs to fit] well into a small rural community—it is more than just turning up at school, they have to live



**Fig. 2** Possible indicators for rural and remote teaching suitability. The diagram highlights four clusters of non-academic attributes with outer circles displaying indicators provided by NSW experts

there too"! Others stated that, while important, rural and remote suitability need not be viewed or defined as a stand-alone target:

...all the [non-academic attributes] that make a good teacher in non-rural areas are the same ones that make you a good teacher in rural areas. A greater level of resilience and adaptability may be required, but that is already a stand-alone [cluster]. The only other real variable is desire. Does the teacher want to work and live in a rural area?

Thus, we present a visual summary through Fig. 2 of how each of the initial three target clusters overlap with a fourth cluster. The additional cluster of culture and context helped to provide a more holistic description of rural and remote suitability for teaching in NSW.

### *Item development*

One-to-one interviews were carried out with 11 teachers prior to the workshop, resulting in 37 draft scenario-based items. Most scenarios targeted Resilience and Adaptability (11) and Empathy and Communication (11). Organisation and Planning was the target for seven scenarios with the remaining eight scenarios targeting more than one cluster. For example, while a NSW-specific target on rural and remote teaching suitability (culture and context) was not independently identified, one scenario was a shared target with Empathy and Communication and Resilience and Adaptability.

Next, the teachers provided scoring suggestions and feedback on the 37 new NSW-specific items through a workshop. In addition, the teachers provided feedback and revised 32 existing UK items to better suit NSW teaching contexts. We collected written feedback and referred to overall comments when refining the items for the pilot study. For example, recommendations included references to head teacher, year advisor, or mentor teacher instead of head of department and senior or more experienced colleague. Other suggestions included changing the response content and format, confirming what the item was intended to measure (*what should you do?*) since some items appeared to focus on self-reported ability or behaviour (*what would you do?*). Contextual recommendations also included removing specific responses that may be obvious when expected or legal requirements in NSW school settings are taken into consideration (e.g. social media guidelines).

Since only one item that targeted rural and remote teaching resulted from the phone conversations, we guided small groups within the workshop to develop items specific to rural and remote settings. With only seven new items developed by small groups, we sought further rural and remote items by contacting thirteen additional NSW educators (from four rural or remote secondary school contexts) after the workshop. Of the thirteen, seven arranged a phone conversation that resulted in eleven new items specific to rural or remote contexts. In total, 55 new items were

**Table 2** Participant demographics ( $n = 99$ )

|                     |   |
|---------------------|---|
| Teaching level      | Primary: 47                                   |
|                     | Secondary: 37                                 |
|                     | Unspecified: 15                               |
| Teaching experience | Teacher education students: 56                |
|                     | Recent graduates: 34                          |
|                     | New NSW teachers: 1 (3 years experience)      |
|                     | Unspecified: 8                                |
| Teaching setting    | Rural/remote: 7 (current), 14 (preference)    |
|                     | City/metro: 69 (current), 46 (preference)     |
| Age                 | 20–55 years ( $M = 25.3$ years, $SD = 7.52$ ) |
| Gender              | Female: 59                                    |
|                     | Male: 27                                      |
|                     | Unspecified: 13                               |

developed during Phase One: eighteen new items targeting rural and remote suitability and 37 new items targeting a range of identified clusters.

### *Item review*

We chose 36 out of 55 items to be reviewed by principals (recommended by the Department). The items aligned with the interests of the Department by targeting Resilience and Adaptability (11), culture and context (9), Empathy and Communication (6), and Organisation and Planning (5). Five additional items overlapped with more than one cluster of attributes.

Eighteen principals (eight retired, ten employed) were invited to review and score the 36 new NSW-specific items. Eight experienced principals provided responses (on-site at the Department), each one scheduling a 30-min follow-up phone conversation. Five currently employed principals also provided written feedback. In total, we compiled responses from 13 who had up to 37 years of experience as a principal or in other school leadership roles in primary and/or secondary schools (the majority had been employed in city schools).

Principals reported that the scenarios developed by NSW teachers were appropriate, and responded with comments such as “impressive” and “realistic”. Overall, they viewed the SJT as presenting a balance of items across key non-academic attributes and considered most items as relevant for both primary and secondary teaching contexts. Of the 36 items, 17 achieved moderate to high scoring consensus among the principals. The inability for principals to reach consensus on more than half of the items was associated with the expressed perception that “not all responses were deemed appropriate” and therefore, “ranking most to least” felt “forced” or “inaccurate”. They suggested that some options would best suit a rating or dichotomous response format (i.e. appropriate/not appropriate).

**Table 3** Correlational results (examples)

|                            | 1            | 2           | 3            | 4            | 5           | 6            | 7          | 8              | 9            | 10             |
|----------------------------|--------------|-------------|--------------|--------------|-------------|--------------|------------|----------------|--------------|----------------|
| 1. Extraversion            | 4.90 (1.45)  |             |              |              |             |              |            |                |              |                |
| 2. Agreeableness           | -.03         | 5.32 (1.16) |              |              |             |              |            |                |              |                |
| 3. Conscientiousness       | .19          | -.06        | 6.08 (.95)   |              |             |              |            |                |              |                |
| 4. Emotional stability     | <b>.25*</b>  | <b>.26*</b> | 5.44 (1.16)  |              |             |              |            |                |              |                |
| 5. Open to new experiences | .14          | .11         | <b>.25*</b>  | .16          | 6.05 (.71)  |              |            |                |              |                |
| 6. Teacher self-efficacy   | <b>.22*</b>  | .09         | <b>.45**</b> | <b>.26*</b>  | <b>.25*</b> | 7.65 (.99)   |            |                |              |                |
| 7. Teacher engagement      | <b>.27**</b> | <b>.22*</b> | <b>.41**</b> | <b>.41**</b> | .15         | <b>.57**</b> | 6.47 (.51) |                |              |                |
| 8. SJT-part 1: ranking     | .04          | -.18        | .07          | -.08         | .05         | -.004        | -.09       | 102.80 (16.46) |              |                |
| 9. SJT-part 2: choose      | .08          | -.07        | -.02         | -.10         | .03         | .07          | .09        | -.23*          | 85.48 (9.45) |                |
| 10. SJT-total              | .14          | -.17        | .06          | -.03         | .09         | .10          | -.02       | <b>.79**</b>   | <b>.42**</b> | 187.35 (15.33) |

Variables 1–5 resulted from the personality measure. The mean (standard deviation) for each variable is displayed on the diagonal  
 Bold values indicate significant relationships (\*  $p < .05$  and \*\*  $p < .01$ )

**Table 4** Situational judgement test: item quality

| Item performance (partial range) | Rank | Choose | Total |
|----------------------------------|------|--------|-------|
| Good (0.25–0.46)                 | 8    | 2      | 10    |
| Satisfactory (0.17–0.24)         | 5    | 3      | 8     |
| Moderate (0.13–0.16)             | 3    | 1      | 4     |
| Limited (0.13 >)                 | 6    | 4      | 10    |

### Phase two: pilot study findings

Table 2 presents a demographic summary of pilot study participants ( $N = 99$ ). The average age was 25 years old (Range 20–55 years old), with more participants identified as female and primary trained. Questionnaire completion time averaged 46 min, with a range from 20 to 68 min. Two-thirds identified as Australian or Caucasian and 4% as Aboriginal or Torres Strait Islander people with a majority having experience teaching in the city and/or preferring a city position.

Correlational results revealed no significant relationships between the SJT and measures of self-efficacy, personality, and engagement (see Table 3), suggesting that the SJTs were measuring different constructs.<sup>2</sup> We also determined SJT item quality by looking at the item partial—the degree of correlation between the item and the overall mean SJT score. As summarised through Table 4, items were classified in terms of their quality, with good items exhibiting a partial above 0.25, satisfactory items between 0.24 and 0.17, moderate items between 0.16 and 0.13, and limited items with less than 0.13. In total, 22 out of 32 items were deemed moderate to good items. The average SJT score was low (187.35 out of 560) as was the reliability.<sup>3</sup>

Most participants (70) completed the feedback form to share their perceptions of the SJT. Participants indicated their level of agreement with several statements regarding item content. Overall, they agreed that the content was relevant ( $M = 4.31$  out of 5; 91% agreed/strongly agreed) and fair ( $M = 4.27$  out of 5; 94% agreed/strongly agreed). The level of difficulty was also considered appropriate ( $M = 4.31$  out of 5; 98.5% agreed/strongly agreed). Participants somewhat supported the idea ( $M = 3.78$  out of 5) that the tool would help differentiate teachers during the hiring process. While more participants agreed that the SJT could be used to measure important attributes, most were neutral as to whether the tool would be fair and appropriate as a selection method. Participants indicated that the content was “interesting” and “thought provoking”, but they also highlighted

<sup>2</sup> While inconclusive, a larger sample of participants may help determine the validity of the scenario-based items as a whole in relation to a range of related constructs. Since SJT research, overall, has struggled with conventional methods for assessing validity (Sorrel et al. 2016), additional methods in addition to use with a larger sample will be explored in future research.

<sup>3</sup> The low reliability may be due in part to the low sample size and higher number of primary trained participants (since the items were specific to secondary school settings). In addition, 10 items were revealed as being of limited quality (see Table 4). By excluding lower quality items, the reliability (as indicated by Cronbach’s alpha) increased from 0.16 to 0.42. Previous SJT studies have averaged an internal consistency of 0.46 (see Sorrel et al. 2016 for a discussion).

the need for different versions (i.e. primary and secondary settings) and different response formats (“ranking response format was confusing”).

## Discussion

The Department expressed interest in exploring value-added tools that could potentially complement their interview process. In response, we developed and trialled a scenario-based measure with the help of NSW educators. In this study, we sought to explore a tool that may be used to help measure beginning teachers’ non-academic attributes—one that can inform a related professional learning plan and potentially influence retention rates of newly placed quality teachers in rural or remote contexts. We received feedback from experts and participants in favour of developing a non-academic tool for developmental purposes—one that includes expert-informed and context-specific teaching scenarios set in primary and secondary settings within rural/remote and city/metro contexts. Results revealed SJT constructs that were separate from existing measures that assessed teachers’ self-efficacy, personality, and engagement—yet further research is needed to evaluate whether it is a fair and appropriate approach to measuring and promoting the development of beginning teachers’ non-academic attributes.

As seen in Table 1, non-academic attributes broadly align with the Australian Professional Standards for Teachers (AITSL 2011). The standards help assess quality teaching in Australia, but more research is needed on the relationship with what are deemed to be important teacher attributes—particularly in diverse settings. Expert teachers in NSW confirmed that the three clusters of attributes deemed necessary for teachers in the UK were relevant for local settings, although adaptation of content was necessary to reflect the Australian context. In addition, teachers provided support and content for a fourth target cluster—culture and context—that represented professional challenges in rural and remote settings. In seeking to define attributes specific to rural and remote suitability, culture and context highlighted the importance of teacher adaptability and cultural competence. Adaptability and cultural competence is also important as a teacher in culturally diverse city schools; however, our experts expressed the need for a specific focus on the cultural knowledge and practices of Indigenous Australians when considering suitability for teaching in rural or remote settings.

What began as an exploration of ‘rural and remote teaching suitability’ resulted with both a context (i.e. possible isolation-related challenges) and a cultural focus (i.e. cultural competence specific to communities with Indigenous Australians). Therefore, future research will benefit from inviting community members to contribute to the development of SJTs. This step in our development is crucial since perceptions of quality teacher attributes can not only vary across countries (Meng and Muñoz 2016), but between cultural populations and among rural communities within NSW. As such, our results provide an example of the need for constructing SJTs that acknowledge and celebrate the complex intersection of “rurality in both *geographic* and *cultural* terms” (Reid et al. 2010, p. 263).

With SJT-based tools, we can help develop and determine how well a new teacher is meeting Standards 1 and 2 (*Know students and how they learn; Know the content and how to teach it*). Specifically, graduates teaching in the city and the country are expected to use “strategies for teaching Aboriginal and Torres Strait Islander students” (Standard 1.4; AITSL 2011) and have the capacity to “understand and respect Aboriginal and Torres Strait Islander people to promote reconciliation between Indigenous and non-Indigenous Australians” (Standard 2.4; AITSL 2011). Yet, further work on SJT validation for NSW and identifying explicit links with related standards and teacher professional learning are still needed. Therefore, the current study serves as a foundational first step as we gather evidence that can support the use of SJTs as professional learning tools for new teachers (e.g. those working towards their *Proficient* level of accreditation). By recommending future SJT-based professional learning in NSW that targets culture and context as an overlapping cluster, outcomes may include identifying and supporting the development of beginning teachers who are demonstrating sensitivity to cultural knowledge and practices of Aboriginal and Torres Strait Islander peoples with the capacity to become culturally competent.

## Limitations

As noted by Whelpley (2014), “there is no objectively correct response to many SJT items and, in reality, the best response will likely vary based on the person and situation” (p. 21). Our biggest challenges to assessing reliability and validity were related to the complex and multidimensional format and scoring of the SJT, particularly since the lengthy administration procedure reduced the number of volunteer participants. While correlations with the additional measures of self-efficacy, personality, and engagement were non-significant, a larger sample may help determine the validity of the scenario-based items as a whole in relation to a range of related constructs. Previous SJT research has typically resulted in low indicators of reliability due to the multidimensional nature of the test (i.e. Cronbach’s alpha is unidimensional; Sorrel et al. 2016). Therefore, the low reliability of the SJT found in the current study does not necessarily equate with the development of a poor measure.

The focus and strength of this study was on the process of developing a contextualised SJT for teachers in NSW. For example, we involved two groups of experts (19 teachers and 13 principals) in the developmental and rigorous process recommended by SJT developers (e.g. Patterson et al. 2015). However, a clear limitation was the small sample size that piloted the SJT. While our multiple data collection sites did allow for a diverse sample, conclusions drawn from the SJT responses provided by 99 participants require further testing with a larger sample. In particular, future SJT research is needed that includes a greater number of participants who are interested in or have had experience in rural or remote contexts since only 21% who trialled the SJT expressed an interest in rural or remote teaching.

The overall mean SJT score was low, which may be due in part to the low sample size and higher number of primary trained participants (since the items were specific



to secondary school settings). Fatigue may have also influenced participants' scores given the length of time and cognitive load required to complete the questionnaire. A number of potential participants expressed interest, but only if they could complete the pilot questionnaire in a shorter period, at their convenience online, and/or if the scenario-based questionnaire somehow contributed to their formal application or resulted in evidence of their professional learning. Thus, further development is required with respect to items specific to primary and to secondary teaching in NSW and with different formats (and shorter) SJTs administered to larger samples.

## Future research

Our next steps will include categorising response types using theoretically and empirically tested models (Whelpley 2014). While SJTs are typically scored using experts' judgements, themes and categories can be applied when developing a scoring protocol that meets the needs of the Department and addresses the SJT-related issues with assessing validity and reliability. For example, applying Holland's (1997) RIASEC model (Realistic, Investigative, Artistic, Social, Enterprising, Conventional) is one option that may provide the developmental and environmental perspective needed for categorising and standardising the types of responses across scenarios targeting all attributes. Here, teacher profiles could be identified based on response patterns found across all SJT sections (e.g. "Social") and allude to the type of competency-developing support necessary for retention in rural and remote NSW settings.

The experts involved in the current study expressed an interest in seeing the further development of SJTs with a more fine-grained response format. This differs from the originally intended use of the SJT as a large-scale selection-focused assessment that uses a pre-determined scoring key. For example, principals involved in our Phase One activities warned of new teachers' potential over-reliance on options related to "seeking help from administrators". Occasionally, seeking help from principals or deputy principals is expected; however, when assessing teachers' non-academic attributes—especially in more isolating conditions—the SJT methodology may help with revealing patterns of dependency. For example, frequently selecting "seeking help" as the *most appropriate* response to SJT scenarios may indicate challenges relating to problem solving or level of required independence for some remote contexts. Such information may be useful for the Department and the new teacher when placing and planning for supportive resources.

Future research will involve trialling different SJT formats. For example, we have started to include a format that requires a rating (from most to least appropriate) for each of a scenario's response options. By using rating scales, we hope to gather additional insight through the responses judged as most and least appropriate. This will be a beneficial next step since Stemler et al. (2016) found that scoring based on the ability to identify the 'worst' or least appropriate response was more predictive of teacher effectiveness than identifying the best response. Moreover, data collected using the range of response formats will help contribute to our search for categorical patterns across items and attributes. As with SJTs used

for selection by medical education faculties (Patterson et al. 2015), our piloted format could only provide one overall non-academic score. Given the overlap across key attributes, clear factor scores for each cluster are also not emerging in the literature (Sorrel et al. 2016). Therefore, future research exploring alternate scoring options will help construct a tool that can be better suited for the needs of the Department and, consequently, lead to targeted support for newly placed teachers in state schools. For example, Cox et al. (2017) have set the stage as the first to study SJT use for training with other professionals, and we are confident that SJTs will also be successful when training and supporting teachers.

Future activities also include piloting the remaining items developed through this project (that achieved scoring consensus among NSW principals) and engaging in an iterative process that involves discussion among other subject matter experts (e.g. teacher educators) for the purpose of refining low-quality items that did not pass initial review in this study. We are also planning to train item writers and enlist expert item reviewers as we develop new, high-quality reliable items (for primary and secondary school contexts) specific to measuring and supporting beginning teachers in a range of NSW settings. Since fewer items specific to rural and remote teaching were developed, further research on the scenarios targeting a range of attributes is needed in order to confirm whether the four clusters adequately and parsimoniously represent the non-academic make-up of quality NSW teachers.

Given the feedback received during this study from our Department collaborators, study participants, and fellow researchers, next steps also include various SJTs administered multiple times during the first few years of teaching. By trialling scenario-based items over time with new teachers, we can investigate how and to what extent non-academic attributes are being fostered and developed. For example, we can start with teachers who provided consent to link their questionnaire results with other related data collected by the Department and run trials of the SJT for professional learning purposes. New graduates could also try scenario-based items prior to the personal suitability interview with the Department and then interviewers would use some of the responses for discussion points within the interview.

## Conclusions

Unfortunately, rural and remote educational settings have been “judged in terms of a deficit discourse... rather than a diversity discourse” (Reid et al. 2010, p. 267). So we wondered: how can we influence the recruitment and retention of quality and diverse beginning teachers in state schools—particularly those in rural or remote contexts? Through our research, we found that the SJT methodology has the potential for identifying, measuring, and developing non-academic attributes in beginning teachers in NSW. Based on feedback from experienced NSW educators and findings from our initial pilot study, there is potential in using the SJT when making evidence-based recommendations for rural or remote teaching contexts in NSW. Contextualised scenario-based items can be used by the Department to provide one more perspective on an applicant’s personal characteristics and suitability for a rural or remote school setting. Experts involved in Phase One

expressed interest in the SJT methodology as one tool that can help identify a potential teacher's personal 'fit' to a particular school or setting while at the same time identify areas where the teacher's professional learning may be best supported. Moreover, this tool may be a beneficial addition to supporting preservice development of non-academic attributes and confirming suitability for teaching as teacher education students' progress through their professional program (NSWCDE 2017).

As a result of government recommendations (e.g. TEMAG 2014) for quality teaching in Australia, the pursuit of a measurement of personal attributes has focused on selecting entrants into ITE programs. We argue that since ITE programs in NSW are focused on developing academic and non-academic attributes deemed necessary for teaching, the use of a context-specific SJT may be better positioned when graduates are entering the profession. Until recently, SJT development has been focused on large-scale selection into training programs, but with the foundation of the current study, our next steps will see SJTs in NSW being developed for professional learning and development purposes—within and beyond teacher education programs. Alongside educational and organisational efforts in rural and remote Australia aimed at improving the attraction, recruitment, and retention of effective teachers, SJT research can help define quality *teachers*, contribute to frameworks aimed at identifying quality *teaching*, and promote professional development to ensure beginning teachers are equipped with the attributes to thrive in culturally and contextually diverse settings.

**Acknowledgements** We gratefully acknowledge the assistance provided by Lisa Kim (University of York, UK) and the consistent dedication and commitment shown by our team of collaborators at the New South Wales (NSW) Department of Education. In particular, we are grateful for the help provided by Mark Anderson, Ian Quintos, John Wilson, and Veronica Willmott of the Human Resources Directorate. We also express our utmost thanks to the NSW teachers and principals who shared their insights and expertise and to the pilot study participants for their time and feedback.

## Appendix: Example of a situational judgement test item for teachers

You are teaching a lesson and have asked the students to individually complete an exercise that requires them to write down their responses. You have explained the exercise to the students and answered all of the questions that they have asked. As the students begin writing, one student, Ruby, starts to throw paper around and is clearly distracting the students sitting nearby. You know from previous incidents that Ruby often becomes frustrated when she does not understand how to complete activities, and that she often displays her frustration by being disruptive.

Choose the three most appropriate actions to take in this situation (alternatively, Rank the items in the most appropriate order)

- Send Ruby out the class if she continues to be disruptive
- Ask Ruby if she understands what the activity requires her to do
- Check in 5 min to see if Ruby has made progress with the exercise

- Tell Ruby that you are disappointed in her behaviour
- Ask Ruby's classmate to discreetly provide help
- Stop the exercise and discuss the classroom behaviour plan with the whole class
- etc. (eight total response options)

*Note* This is an example only, and is adapted from an item in a SJT for primary teachers (Klassen et al. 2017, p. 905).

## References

- Austin, J. (2010). *Remote ready*. Remote Health Recruitment, Northern Territory Government. Charles Darwin University, Australia. Retrieved from [http://www.remoterecruitment.nt.gov.au/docs/remote\\_ready.pdf](http://www.remoterecruitment.nt.gov.au/docs/remote_ready.pdf).
- Australian Bureau of Statistics. (2017). *Census 2016 QuickStats: Aboriginal and/or Torres Strait Islander people*. Belconnen, ACT: Commonwealth of Australia. Retrieved from <http://www.abs.gov.au/>.
- Australian Institute for Teaching and School Leadership (AITSL). (2011). *Accreditation of initial teacher education programs in Australia: Standards and procedures*. Melbourne, VIC: Ministerial Council for Education, Early Childhood Development and Youth Affairs.
- Australian Institute for Teaching and School Leadership (AITSL). (2015). *Action now: Selection of entrants into initial teacher education—guidelines*. Melbourne, VIC: Australian Institute for Teaching and School Leadership.
- Bahr, N., & Mellor, S. (2016). *Australian education review: Building quality in teaching and teacher education*. Camberwell, VIC: Australian Council for Educational Research.
- Baxter, J., Gray, M., & Hayes, A. (2011). Families in regional, rural and remote Australia (Fact Sheet). Australia Institute of Families Studies, Melbourne, VIC: Commonwealth of Australia. Retrieved from <https://aifs.gov.au/publications/families-regional-rural-and-remote-australia>.
- Beltman, S., Mansfield, C., & Price, A. (2011). Thriving not just surviving: A review of research on teacher resilience. *Educational Research Review*, 6(3), 185–207. <https://doi.org/10.1016/j.edurev.2011.09.001>.
- Buehl, M. M., & Fives, H. (2009). Exploring teachers' beliefs about teaching knowledge: Where does it come from? Does it change? *The Journal of Experimental Education*, 77(4), 367–407.
- Caldwell, B., & Sutton, D. (2010). *Review of teacher education and induction. First report—full report*. Education Queensland <http://education.qld.gov.au/students/higher-education/resources/review-teachereducation-school-induction-first-full-report.pdf>.
- Cox, C. B., Barron, L. G., Davis, W., & de la Garza, B. (2017). Using situational judgment tests (SJTs) in training. *Personnel Review*, 46(1), 36–45. <https://doi.org/10.1108/PR-05-2015-0137>.
- Day, C., & Gu, Q. (2014). *Resilient teachers, resilient schools*. New York, NY: Routledge Taylor Francis.
- Frenzel, A. C., Becker-Kurz, B., Pekrun, R., & Goetz, T. (2015). Teaching this class drives me nuts!—Examining the person and context specificity of teacher emotions. *PLoS ONE*, 10(6), e0129630. <https://doi.org/10.1371/journal.pone.0129630>.
- Goldhaber, D., Grout, C., & Huntington-Klein, N. (2014). *Screen twice, cut once: Assessing the predictive validity of teacher selection tools. (CEDR Working Paper No. 9)*. Seattle, WA: Center for Education Data and Research. <http://www.cedr.us/papers/working/CEDR%20WP%202014-9.1.pdf>.
- Gore, J., Barron, R. J., Holmes, K., & Smith, M. (2016). Who says we are not attracting the best and brightest? Teacher selection and the aspirations of Australian school students. *The Australian Educational Researcher*, 43(5), 527–549. <https://doi.org/10.1007/s13384-016-0221-8>.
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the big-five personality domains. *Journal of Research in Personality*, 37, 504–528. [https://doi.org/10.1016/S0092-6566\(03\)00046-1](https://doi.org/10.1016/S0092-6566(03)00046-1).
- Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., et al. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4000 classrooms. *The Elementary School Journal*, 113(4), 461–487.

- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Odessa, FL: Psychological Assessment Resources.
- Huberman, M. (1989). The professional life cycle of teachers. *Teachers College Record*, *91*, 31–57.
- Ingersoll, R. M., & Strong, M. (2011). The impact of induction and mentoring programs for beginning teachers: A critical review of the research. *Review of Educational Research*, *81*, 201–233. <https://doi.org/10.3102/0034654311403323>.
- King, R. B., & McInerney, D. M. (2016). Culture and motivation. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of motivation at school* (pp. 275–310). New York: Routledge.
- Klassen, R. M., Durksen, T. L., Kim, L., Patterson, F., Rowett, E., Warwick, J., et al. (2017a). Developing a proof-of-concept selection test for entry into primary teacher education programs. *International Journal of Assessment Tools in Education*, *4*(2), 96–114. <https://doi.org/10.21449/ijate.275772>.
- Klassen, R. M., Durksen, T. L., Patterson, F., & Rowett, E. (2017b). Filtering functions of assessment for selection into initial teacher education programs. In J. Clandinin & J. Husu (Eds.), *The Sage handbook of research on teacher education* (Vol. 1, pp. 893–909). New York: Sage Publications Ltd.
- Klassen, R. M., Durksen, T. L., Rowett, E., & Patterson, F. (2014). Applicant reactions to a situational judgment test used for selection into initial teacher training. *International Journal of Educational Psychology*, *3*, 104–125. Retrieved from <http://www.hipatiapress.com/hpjournals/index.php/ijep>.
- Klassen, R. M., & Tze, V. M. C. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review*, *12*, 59–76.
- Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998–2009: Signs of progress or unfulfilled promise? *Educational Psychology Review*, *23*, 21–43. <https://doi.org/10.1007/s10648-010-9141-8>.
- Klassen, R. M., Yerdelen, S., & Durksen, T. L. (2013). Measuring teacher engagement: The development of the engaged teacher scale (ETS). *Frontline Learning Research*, *1*, 33–52. Retrieved from <http://journals.sfu.ca/flr/index.php/journal/article/view/44/37>.
- Kutsyruba, B., Godden, L., & Tregunna, L. (2014). Curbing early-career teacher attrition: A pan-Canadian document analysis of teacher induction and mentorship programs. *Canadian Journal of Educational Administration and Policy*, *161*, 1–42.
- Lyon, T. (2006). Science, ICT and mathematics education in rural and regional Australia, SiMERR National Research Centre [Report]. <http://apo.org.au/node/2732>.
- Martin, A. J., & Dowson, M. (2009). Interpersonal relationships, motivation, engagement, and achievement: Yields for theory, current issues, and practice. *Review of Educational Research*, *79*, 327–365.
- Mayer, D., Allard, A., Bates, R., Dixon, M., Doeke, B., & Kline, J., et al. (2015). *Studying the effectiveness of teacher education (Final Report)*. Geelong, VIC: Deakin University. <http://www.setearec.com.au/publication-reports/>.
- MCEETYA. (2001). National framework for rural and remote education in Australia. *Journal of Research in Rural Education*, *17*(2), 112–119.
- McNeillage, A. (2014, January 21). Adrian Piccoli says teacher selection criteria should be more strict. *The Sydney Morning Herald*. <http://www.smh.com.au/national/education/adrian-piccoli-says-teacher-selection-criteria-should-be-more-strict-20140120-314vn.html>.
- Meng, L., & Muñoz, M. (2016). Teachers' perceptions of effective teaching: A comparative study of elementary school teachers from China and the USA. *Educational Assessment, Evaluation and Accountability*. <https://doi.org/10.1007/s11092-015-9230-9>.
- Motowidlo, S. J., & Beier, M. E. (2010). Differentiating specific job knowledge from implicit trait policies in procedural knowledge measured by a situational judgment test. *Journal of Applied Psychology*, *95*(2), 321.
- NSW Council of Deans of Education. (2016, July). Position paper: Selection methods into initial teacher education programs [Draft v2]. NSWCEDE and BOSTES Joint Forum: Assuring Graduate Quality through the New Accreditation Standards. Received at University of Technology, Sydney.
- NSW Council of Deans of Education. (2017, January). Assessing and confirming suitability for teaching: A selection framework for NSW teacher education providers [Brief Report]. Sydney, NSW.
- NSW Department of Education and Communities (NSWDEC). (2013, November). *Rural and remote education: A blueprint for action*. NSW Department of Education and Communities. [www.dec.nsw.gov.au/](http://www.dec.nsw.gov.au/).
- Patterson, F., Ashworth, V., & Good, D. (2013). *Situational judgement tests: A guide for applicants to the UK foundation programme*. London: Medical School Council.

- Patterson, F., Zibarras, L., & Ashworth, V. (2015). Situational judgment tests in medical education and training: Research, theory and practice: AMEE Guide No.100. *Medical Teacher*, *00*(00), 1–15. <https://doi.org/10.3109/0142159X.2015.1072619>.
- Reid, J.-A., Green, B., Cooper, M., Hastings, W., Lock, G., & White, S. (2010). Regenerating rural social space? Teacher education for rural—Regional sustainability. *Australian Journal of Education*, *54*(3), 262–276.
- Rice, S., Richardson, P. W., & Watt, H. M. G. (2017). Hard-to-staff Australian schools: How can we ensure that all students have access to quality teachers? In T. Bentley and G. C. Savage (Eds.), *Educating Australia: Challenges for the decade ahead* (Chapter 17). Melbourne, VIC: Melbourne University Publishing (Kindle Edition).
- Ryan, A. M., & Ployhart, R. E. (2014). A century of selection. *Annual Review of Psychology*, *65*, 20.1–20.25.
- Sautelle, E., Bowles, T., Hattie, J., & Arifin, D. (2015). Personality, resilience, self-regulation and cognitive ability relevant to teacher selection. *Australian Journal of Teacher Education*, *40*(40), 1–19. <http://doi.org/10.14221/ajte.2015v40n4.4>.
- Schultheiss, O., & Brunstein, J. (2010). *Implicit motives*. UK: Oxford University Press.
- Sorrel, M. A., Olea, J., Abad, F. J., de la Torre, J., Aguado, D., & Lievens, F. (2016). Validity and reliability of situational judgement test scores. *Organizational Research Methods*, *19*(3), 506–532. <https://doi.org/10.1177/1094428116630065>.
- Stemig, M. S., Sackett, P. R., & Lievens, F. (2015). Effects of organizationally endorsed coaching on performance and validity of situational judgment tests. *International Journal of Selection and Assessment*, *23*(2), 174–181.
- Stemler, S. E., Aggarwal, V., & Nithyanand, S. (2016). Knowing what NOT to do is a critical job skill: Evidence from 10 different scoring methods. *International Journal of Selection and Assessment*, *24*(3), 229–245.
- Stenberg, K., Karlsson, L., Pitkaniemi, H., & Maaranen, K. (2014). Beginning student teachers' teacher identities based on their practical theories. *European Journal of Teacher Education*, *37*(2), 204–219.
- TEMAG: Teacher Education Ministerial Advisory Group. (2014). *Action now: Classroom ready teachers*. Canberra: Australian Government Department of Education. <http://www.studentsfirst.gov.au>.
- Tripp, D. (1994). Teachers' lives, critical incidents, and professional practice. *International Journal of Qualitative Studies in Education*, *7*(1), 65–76. <https://doi.org/10.1080/0951839940070105>.
- Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education*, *17*, 783–805.
- Whelpley, C. E. (2014). How to score situational judgment tests: A theoretical approach and empirical test [Unpublished Doctoral Dissertation]. Virginia Commonwealth University, USA.
- Whetzel, D. L., & McDaniel, M. A. (2016). Are Situational Judgment Tests better assessments of personality than traditional personality tests in high-stakes testing? In U. Kumar (Ed.), *The Wiley Handbook of Personality Assessment* (pp. 205–214). West Sussex: Wiley.

**Tracy L. Durksen** is a Vice Chancellor's Postdoctoral Research Fellow at UNSW Sydney in Australia. Her research interests centre on preservice and practicing teachers' professional learning, interpersonal skills, motivation, and engagement. Tracy's program of research includes a three-year funded project with the NSW Department of Education. Through this partnership project she is studying methods that can help assess and promote the development of teachers' academic and non-academic capabilities. Tracy previously worked as a primary school teacher in Canada and currently teaches postgraduate students in the School of Education at UNSW Sydney.

**Robert M. Klassen** is Professor and Chair of the Psychology in the Education Research Centre at the University of York in the UK, and is an adjunct Professor in Educational Psychology at the University of Alberta. His research background is in motivation, engagement and emotions in school settings. He currently focuses on applying educational psychology research to the problem of teacher selection. Rob has numerous international collaborations, and his work is funded by a European Research Council Consolidator Grant. Before entering academia, Rob worked as a teacher and school psychologist in Canada. He is a Chartered Psychologist in the UK.