

2021

## Does the ATAR Predict Pre-service Teacher Capacity for Inclusive Classroom Practice?

Greg Auhl  
*Charles Sturt University, Bathurst*

Alan Bain  
*Charles Sturt University, Bathurst*

Follow this and additional works at: <https://ro.ecu.edu.au/ajte>



Part of the [Other Teacher Education and Professional Development Commons](#), and the [Special Education and Teaching Commons](#)

---

### Recommended Citation

Auhl, G., & Bain, A. (2021). Does the ATAR Predict Pre-service Teacher Capacity for Inclusive Classroom Practice?. *Australian Journal of Teacher Education*, 46(4). <https://doi.org/10.14221/ajte.2021v46n4.5>

This Journal Article is posted at Research Online.  
<https://ro.ecu.edu.au/ajte/vol46/iss4/5>

## **Does the ATAR Predict Pre-service Teacher Capacity for Inclusive Classroom Practice?**

Greg Auhl  
Alan Bain  
Charles Sturt University

*Abstract: In this study, we examined whether the Australian Tertiary Admissions Rank (ATAR) predicted pre-service teachers' schema development for inclusive classroom teaching. Where previous studies have employed grade point average (GPA) as a criterion measure, this study employed a validated measure linked to standards-based classroom practice. The study was conducted in the final semester of the teacher education programs at three Australian universities. The results showed the explanatory power of the ATAR was close to zero for the students studied. The implications of the findings for teacher education and for using the ATAR as an indicator of teacher quality are discussed.*

### **Introduction**

There has been extensive recent popular debate in Australia about raising the Australian Tertiary Admissions Rank (ATAR) entry requirements for teacher preparation programs to improve graduate teacher quality (e.g., Conifer 2019; Urban, 2018). The interest has been stimulated by reports in the popular media and literature that universities are admitting students with some of the lowest levels of ATAR performance (Wilson, 2020; Robinson, 2018). Low ATAR has been linked to university student drop out rates, difficulty in tertiary study and lower levels of intrinsic interest in the course of study (Norton & Cherastidtham, 2018; Baik, Naylor, Arkoudis & Dabrowski, 2019; Wilson, 2020). Much of the interest in raising requirements for entry to teacher education programs is predicated on the belief that teaching candidates with higher prior levels of academic achievement exert a greater positive influence on the achievement of the students they will teach. Most studies that focus specifically on the achievement history of teaching candidates and the achievement of their students show only modest support for this contention (Clotfelter, Ladd & Vigor, 2010; Goldhaber, 2007; Hanushek, Piopiunik, & Wiederhold, 2019; Rockoff, Jacob, Kane, & Staiger, 2011). It is also important to note that only a small percentage (one in four applicants) are chosen singularly on the basis of their ATAR (Aspland, 2019; McGraw & Fish, 2018; TEMAG, 2014).

Wurf & Croft-Piggin (2015) studied the relationship between ATAR and grade point average as well as measures of engagement, motivation and the emotional intelligence of first year pre-service teachers. The authors found a statistically significant correlation of .36 between ATAR and grade point average while ATAR predicted 6% of the year-end variance in grade point average. However, it was not the strongest predictor, which was the students' self-report about their motivation and engagement (21%). Emotional intelligence did not contribute significantly to the predictive model (Wurf & Croft-Piggin, 2015).

Wright (2015) sought to establish the predictive validity of the ATAR for both grade point average and teaching performance in a professional experience placement. The study examined the performance of students in early childhood and primary degrees in years 2, 3 and 4 of their programs. The author also employed ratings of students by their supervising teachers in a professional experience broadening the approach to include a measure of teaching performance (Wright 2015). ATAR was found to be moderately correlated year-over-year with grade point average for the students in the primary program (.442-.562) but only weakly correlated for students in the early childhood program (.086-.239). No clear relationship was found between ATAR and the ratings of student performance in their teaching placements (Wright, 2015). While two of the six correlations between ATAR and teaching placement were statistically significant, they were generally low ranging from -0.030 to .232. This finding may (in part) reflect the unknown validity and reliability of the criterion placement measure.

McGraw & Fish, (2018) conducted a qualitative study of pre-service teachers with low ATARs who may not have been admitted to university on that basis. The authors found that students with lower performance in the last year of high school were not prevented from succeeding in their teacher preparation programs and that the more difficult to define quality of student character should be a key focus when determining admission to programs of teacher preparation (McGraw & Fish, 2018).

The debate about the predictive value of ATAR for entry to programs of teacher preparation also pertains in large measure to conjecture about the range of skills and capacities required to be a successful teacher and whether they are adequately represented by an ATAR (e.g., Donnelly, 2019). Critics make the claim that being a successful teacher requires a broader range of skills and capacities than those measured by the ATAR. This argument is strengthened by the use of grade point average as a criterion measure in studies of the predictive qualities of the ATAR. Much of the academic behaviour that underpins the ATAR (e.g., taking exams and tests, writing reports and essays etc.) is similar to the kind of activity employed to build a grade point average at university irrespective of the field of study. As such, it is not surprising that ATAR would predict university performance to some extent especially for those students with a strong record of performance at secondary school (Norton & Cherastidham, 2018; TEMAG, 2014). The degree to which the knowledge and skills required to build a grade point average adequately represents or captures the full range of skills/qualities for successful classroom teaching is a more open question. The rationale for the study described here is to establish the predictive value of the ATAR when using a valid and reliable criterion measure that specifically relates to what teachers need to know and be able to do in planning for routine classroom practice.

In this study, students were asked to complete the Schema for Teaching Scale (Auhl, 2019; Auhl & Bain, 2020) at the end of their teacher preparation programs. The Schema for Teaching Scale measures the extent to which preservice teachers possess a schema for the kind of practice circumstances commonly encountered by teachers in inclusive classrooms. In the scale, participants are asked to “Use the knowledge acquired in your teacher education course to prepare a sample lesson” (Auhl, 2019). For this lesson, participants are asked to frame their response in terms of aspects of lesson development such as providing a description of “the intended learning goals/outcomes of the lesson, including the way they would be differentiated”; and “a brief narrative describing how you would differentiate the lesson in terms of:

- what you teach (the content)
- how you teach it (the process/pedagogies used)

- how learning will be assessed (what product(s) would show that learning had taken place) to cater for a diversity of learners, in particular students with disabilities.” (Auhl, 2019).

A schema is defined as “a vehicle of memory” that allows individuals to organise their experiences from similar areas of life... to “draw inferences, make estimates, create goals and develop plans using the framework”; and it allows the utilisation of skills, procedures or rules when faced with a problem that is relevant for a particular framework (Marshall, 1995, p. 40). Research both within teaching (e.g., Borg, 2015) and within other fields (e.g., Seel, Darabi & Nelson, 2006 in engineering) has established a relationship between schema development and behaviour where the capacity to cognitively plan a course of action relates to successfully executing that plan in practice. Schemas for practice are necessary to assist teachers to problem-solve and reduce cognitive load as they manage the complex and diverse activity that occurs in all classrooms (Feldon, 2007; Sweller, 1988).

For preservice teachers, schema development can be expected as a result of participation in a teacher education program and specifically through involvement in activities and input related to the acquisition of knowledge and skills required to be a successful teacher in an inclusive classroom (Auhl & Bain, 2020). From an inclusive education perspective this schema development relates to the skills required to address the needs of students in an inclusive classroom setting. These are the inputs required in order to graduate what the Australian Institute for Teaching and School Leadership (AITSL) describes as classroom ready teachers that meet AITSL standards. The term 'inclusive' when applied here refers to a regular classroom where all students are included irrespective of aptitudes and achievement histories and taught by a regular (non- specialist) classroom teacher. This study sought to establish whether the ATAR predicted schema development for inclusive classroom teaching of preservice teachers by establishing the predictive value of the ATAR using a criterion measure that related more specifically to the normal work of inclusive classroom teaching practice.

## Method

### Research Design

*The study employed a post-test-only control and comparison group design (Cook & Campbell, 1979), with more than one experimental group, to determine the extent to which the ATAR scores predicted performance of students from undergraduate teacher education programs. The study addressed the following research question:*

*To what extent does the ATAR predict the performance of graduate teachers on a measure of their capacity to develop instruction for an inclusive classroom?*

### Participants

The participants included 136 final year undergraduates drawn from three university teacher-education degree programs. The sample included female and male candidates (97 female and 39 male) enrolled in primary and secondary teaching degrees. While participants represented a variety of ages, as described in Table 1, the majority were in the 21– 25-year age bracket. The participants also represented a range of ATAR scores, as described in Table 2. While all institutions accepted students through a variety of pathways, only those admitted

on the basis of ATAR were included in the study. For all institutions, this included ‘bonus’ points added to the ATAR for various school and community activities.

	Age					Total
	17-20	21-25	26-30	31-40	>40	
Total	3	122	7	3	1	136

**Table 2. Number of participants by ATAR categories**

ATAR ranges						
<50	51 – 60	61 – 70	71 – 80	81 – 90	91 – 100	Total
0	14	23	50	36	13	136

**Table 1. Age breakdown of participants**

## Settings

The participating universities included regional inland, large city urban, and regional coastal urban providers. All of the universities offered nationally accredited programs (AITSL, 2011) in primary and secondary education. Key personnel at the sites were contacted and provided with information about the study. These details, and the researcher’s institutional ethics protocols, were approved by the deans of each faculty at the participating universities.

## Variables

The variable employed to predict pre-service teacher schema development in the study was the Australian Tertiary Admission Rank (ATAR). "The ATAR is a number between 0.00 and 99.95 that indicates a student’s position relative to all the students in their age group (i.e., all 16 to 20 year olds in NSW)"(UAC. 2019). The ATAR is employed by universities as part of the process for admission to higher education courses. The criterion variable for the study was schema development for inclusive practice – to what degree have participants developed a coherent schema for professional practice as measured by the Schema for Teaching Scale (Auhl, 2019; Auhl & Bain, 2020).

### Criterion Measure

The Schema for Teaching Scale (STS) requires preservice teachers to complete an open-ended written response to a classroom teaching scenario describing the way they would build a differentiated lesson for students at high, low and medium risk for learning difficulties. This includes the way those preservice teachers would develop learning goals and how the lesson would be taught, differentiated and assessed based upon AITSL graduate teacher standards 1.5, 2, 3.1, 3.6, 5.1 and 5.2 (AITSL, 2011). Those standards pertain to inclusive lesson goal development, pedagogy, differentiation, assessment, and feedback.

The content validity of the scenario employed in the Schema for Teaching Scale was established using the Behaviour Analytic Method (Goldfried & D’Zurilla, 1969) which involves determining the representativeness of a situation or scenario for the purposes of eliciting behavioural responses. The Schema for Teaching Scale scenario was validated by three participant groups – expert judges (eight school leaders and academics), teachers in their first year of practice (14 participants) and three pilot groups of preservice teachers (Goldfried & D’ Zurilla, 1969) as a realistic representation of a normal work situation faced by graduating teachers in an inclusive classroom.

The internal consistency of the Schema for Teaching Scale was established using Cronbach's Alpha (Cronbach, 1951). One hundred and twelve responses were evaluated producing a correlation of .94 among the rubric items. Inter-rater reliability was calculated on a sample of 20 double-blind marked responses that produced an inter-rater agreement of 94%. Cohen's Kappa (Cohen, 1988) was also employed to determine inter-rater reliability on 112 responses and yielded a reliability rating of .704, designated as good reliability in the interpretation of the Kappa statistic.

Concurrent validity refers to the extent to which a test can distinguish among groups that should be distinguished by their test performance (Trochin, 2021). The concurrent validity of the Schema for Teaching Scale was determined by comparing responses from 17 practicing inclusive education teachers with masters' level qualifications in inclusive education and a group of 18 graduating majors in an undergraduate exercise physiology degree. The mean scores for the masters qualified practicing teachers was 30.15 (extended abstract/execution level) with a standard deviation of 3.7, while the exercise physiology students scored 5.28 (uni-structural/identification level) with a standard deviation of 2.39.

The Schema for Teaching Scale is scored from 0-40, one absolute score (0) and four score bands using a metric that combines Marshall's (1995) knowledge levels for schema development and the Solo Taxonomy (Biggs & Collis, 1982). The score bands were: Pre-structural-pre-schematic responses characterized by no understanding of a professional response or valid strategies- scored 0; Uni-structural response that indicates limited understanding of an appropriate professional response, few valid strategies, and limited connections among them, scored 1-9; Multi-structural/elaboration response that indicates some understanding (albeit incomplete) of a professional response and an ability to elaborate on valid strategies identified along with some basic evidence of how to apply them, scored 10-19; Relational/ planning level response that indicates an effective understanding of a functional professional response to the scenario, including a number of interconnected valid strategies and how to implement them in a workable plan for classroom practice, scored 20-29; Extended abstract/ advanced execution level response indicates a complete professional response including significant interrelationships and connections among different valid strategies and a sophisticated interconnected plan for execution, scored 30-40. Scoring within the score bands is undertaken using a rubric that provides descriptors of responses at each level.

#### Procedures

After organising permission from participating institutions, data were collected at each of the three sites by the first author. On attending scheduled sessions, the purpose of the research was explained to prospective participants in an introductory manner, both verbally and in a written statement. It was also made clear all responses would remain confidential; have no bearing on performance in the course, and there was no penalty for declining involvement. Informed consent was obtained for those who expressed a willingness to participate. Five potential participants declined to participate at Site 1, one at Site 2 and six at Site 3. Participants who had completed informed consent forms were then provided with a copy of the Schema for Teaching Scale (Auhl, 2019; Auhl & Bain, 2020) and told they would have an indicative time of 30 minutes for completion of the task, although all participants were allowed as much time as they deemed necessary to complete their responses. Participants provided written responses for later analysis, with all students completed the task within a 30-45-minute time frame.

## Results

The overall mean score for all participants was 8.33 (from a total possible score of 40) with a standard deviation of 3.95. This compared to an overall mean of 30.15 and a standard deviation of 3.7 for the Master's level comparison group. The scores indicate that mean preservice teacher responses for the cohorts fell within the score range for the uni-structural level of schema development, which is characterised by low levels of professional understanding and implementation. None of the participants from any of the universities had developed a functional schema for inclusive practice based upon their responses to the Schema for Teaching Scale scenario. Examples of responses for the differentiation section of the scenario at level 0/1 and 3/4 are included below:

### Level 0/1

*Vary outcomes, and expectations (Participant 111)*

*Teacher differentiates learning by providing different activities of different levels that are visual, kinaesthetic and interactive learning. (Participant 3)*

### Level 3/4

*The lesson will involve mainstream students presenting a detailed assignment, with definitions, explanations and diagrams of the water cycle and its importance to life on earth. They may use a variety of presentation formats. (Tomlinson, 2004) and I will encourage the use of technologies, especially new technologies (Bain, 2007). The differentiated lesson will include "must know" and some "should know" for the 15% and as much "should know" as possible, for the 5%. Both the 15% and the 5% will have necessary aide support and will demonstrate their learning without or with aide support as necessary.*

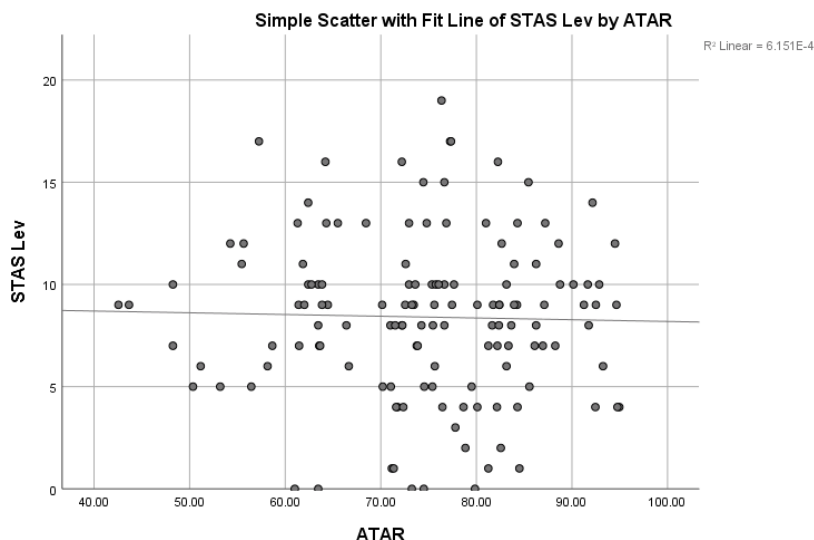
*What I teach will be differentiated on content: must know, should know, could know. Within the 80%, learning and response will vary depending on student interest and ability, and each student's ability with technology and literacy. The use of meta-language will vary according to literacy levels so as to reach all students. (Participant 307)*

*The task of differentiating the assessment tasks will involve considering factors from the specific role in my place of work, using the readings as a reference guide and referring to the three tier model by Walker, Hill and, Horner (1996), where tertiary prevention strategies apply in terms dealing with delinquent and antisocial behavior. With reference to the students at the school I operate in, the candidates require the most intensive intervention, as outline in level 3 of the three tier model. In the following examples I will give some attention to, as is referred to in Tomlinson (2004), as 'struggling learners', p 90.*

*In the decision making for the appropriate assessment tasks I also used detail whereby classroom, curriculum and materials adjustments need to be made. In this case the use of modified materials due, to safety concerns; use of computer assisted tasks and manual activities to 'by-pass difficulty of task'; and 'use of visual supports'. This thinking is supported by the ideas described by Tomlinson (2004), p 80, where the students will work 'with varied amounts of teacher or peer support' (scaffolding), while the differentiation of content and process is based on 'student readiness, 'student interest' and 'student learning profile'. (Participant 252)*

A regression analysis was employed to determine the extent to which the ATAR predicted performance on the Schema for Teaching Scale. The scatterplot in Figure 1

provides a descriptive visual depiction of the relationship between the ATAR and Schema for Teaching Scale scores.



**Figure 1. The relationship between the ATAR and Schema for Teaching Scale scores**

The scatterplot shows a very weak negative relationship between the ATAR and the preservice teachers' scores on the Schema for Teaching Scale indicating that overall the scores on the ATAR and Schema for Teaching Scale were poorly correlated. Table 3 describes the standardized regression coefficients for the analysis indicating that when ATAR was at zero the model estimated Schema for Teaching Scale level is 9.038; the Schema for Teaching Scale would decrease by 0.008 units as every one unit of ATAR increases (i.e., negatively correlated). The ATAR effect was not statistically significant ( $p$ -value = 0.774).

Model	<i>B</i>	<i>SE B</i>	$\beta$	<i>t</i>	<i>p</i>
Constant	9.04	2.23		4.04	.000
ATAR	-.008	0.03	-.02	0.29	.774

**Table 3. Standardized regression coefficients**

The model summary (Table 4) below indicates the explanatory power of ATAR scores for performance on the Schema for teaching Scale. The R square for the model was .001 a level that should be considered as none/nil (adjusted R square = 0) indicating that the ATAR had virtually no predictive relationship with the preservice teachers' scores on the Schema for Teaching Scale.

Model	<i>R</i>	<i>R Square</i>	Adjusted R Square	<i>Std Error</i>
1	.025	.001	-.007	3.981

**Table 4. Model summary indicating the explanatory power of ATAR scores for performance on the Schema for Teaching Scale**



## Discussion

The discussion that follows is predicated on the reasonable assumption, supported by the literature (e.g., Borg, 2015), that the criterion of schema development as measured by the Schema for Teaching Scale is both a precursor to and predictor of teaching behaviour, and that participants who possess a functional schema for their teaching intentions are more strongly positioned to enact those intents. If graduates cannot adequately describe specific professional knowledge and intentions with respect to the standards for inclusive practice (as was the case here), it is unlikely they will be able to effectively deploy such professional knowledge to address the challenges of the inclusive classroom. This claim is supported by a recent study that found students more or less capable of articulating their intentions for practice were more or less effective when observed during actual classroom delivery (Lancaster & Bain, 2019).

The results of this study indicate that when the ATAR is employed to predict performance in an AITSL standards-based task that required actual planning for inclusive classroom teaching it has little predictive power for those graduating pre-service teachers from the universities included in this study. While the results of earlier work (e.g., Wurf & Croft-Piggin, 2015; Wright, 2015), shows some relationship between ATAR and academic performance at university this was not the case here when the criterion measure was more specifically related to the normal work of teaching.

The findings also point to a question/concern that relates to the poor performance of preservice teachers on the Schema for Teaching Scale task overall. The scenario to which participants responded was referenced directly to the standards and practices their preparation programs were accredited to deliver. We also know given the construct validation of the Schema for Teaching Scale that the results are not explained by a floor effect (i.e., where scores on a criterion measure predominantly fall at the low end of possible scores). The Schema for Teaching Scale was shown to be capable of detecting high levels of schema development when completed by a sample of experienced teachers. As such, it is reasonable to ask whether the programs included in the study are capable of delivering an experience that can assure a criterion level of performance to which predictive measures like the ATAR can be applied? While as noted by Aspland (2019) making a great teacher requires more than an ATAR, whether the programs studied have a clear handle on what "more" means and how it is assured in the competency of graduate teachers remains a question.

While critics of admissions standards seek to ensure appropriate threshold levels of academic capability for entrants to teacher preparation programs by employing measures like the ATAR, it is clear for the participants in this study that when a measure of actual professional capacity was used as a criterion, the ATAR had virtually no predictive power. As such, extreme caution should be exercised when linking ATAR to the construct of teaching quality and the overall professional capability (for inclusive practice) of graduate teachers from the institutions participating in this study.

Further, a case has been made by advocates of the ATAR suggesting that the diminution in the use of the ATAR in admission to teacher education programs constitutes a lack of transparency or implied obfuscation in the admissions process (Wilson, 2020). The findings described here suggest that given the lack of predictive value provided by the ATAR, seeking alternative or additional approaches constitutes a sensible course action for the institutions studied.

## Limitations and Conclusion

There are a number of limitations of this research that need to be acknowledged. Post-test only designs are observational and do not seek to provide a causal explanation for why

those conditions exist or provide an explanation for why any differences occur. The design was selected for this study to establish how a recognized predictor of admission to teacher preparation programs functioned when employed to predict professional outcomes for students graduating from teacher preparation programs. There was no intent to make comparisons or claims of specific causal attribution about the content of the different programs studied beyond the claim that the stated expectations and standards were not being met.

A further limitation is the restricted geographic area from which participants were obtained. All of the institutions represented within the study were drawn from NSW, Australia. Follow-up research in other jurisdictions would prove worthwhile in examining whether the predictive power of the ATAR changes under those circumstances. This study only examined one aspect of the teaching standards established by AITSL (2011) for graduate-level teachers. The discussion has been careful to limit conclusions specifically to the institutions and preservice teachers studied. An obvious target for future research is establishing the extent to which the need/problem identified here is common to other jurisdictions, standards and content areas.

An additional limitation of the research is that it did not include observational data derived from watching what preservice teachers could do in classrooms. This means that for this study, it is important not to confuse the cognitive structure for inclusive practice with actual inclusive teaching behaviour; although as noted earlier this relationship has been established (Borg, 2015; Lancaster & Bain, 2019).

We also recognise that while the Schema for Teaching Scale was subject to extensive validation, including reference to literature, standards, experts and pilot groups, the demand associated with completing a written response in a time constrained situation could influence scores. Importantly, we found no examples where preservice teachers' failure to complete the task resulted in a lower score. There was also little within-response variability where students responded more effectively in responses at the beginning of the task when compared to those completed nearer the end. Irrespective, the demand associated with the assessment context is acknowledged.

In concluding, it is important to emphasise the risk associated with conflating the use of ranking measures for tertiary admission with expectations about what programs of tertiary education are expected to produce. For graduate teachers, the outcomes of teacher education represent an interaction of core academic skills and content expertise with professional knowledge, a range of affective characteristics and dispositions, as well as the capacity to represent professional knowledge, skills, and affective considerations in actual practice. Efforts to predict teacher performance require a more expansive approach that address the aforementioned capacities. While ATAR scores may predict higher education program completion rates and is a modest predictor of academic skills, this should not be confused with the knowledge, and schema required for successful teaching in an inclusive classroom.

Our results suggest that expecting the ATAR to predict those outcomes is an over-reach for the participants in this study. While the ATAR may have some predictive value when linked to the kind of academic performance from which it is derived, caution should be applied to efforts that suggest a relationship between ATAR selectivity; the professional schema of graduates, and the performance of students, given the findings for the participants described here and when viewed within the context of prior research.

Finally, it is also clear from the findings that caution should also be exerted when using the ATAR to bundle issues of program completion and competence or capacity for teaching. While ATAR may successfully predict low completion rates, for the preservice teachers in this study it did not predict their schema development for inclusive classroom

practice. Low completion rates are associated with a range of issues in addition to ATAR score (e.g., Naylor, Baik, & Arkoudis, 2018).

## References

- Aspland, T. (2019, January 9). It takes more than a great ATAR to make a great teacher. *EduResearch Matters*. Retrieved from <https://www.aare.edu.au/blog/?p=3608>
- Auhl, G. K. (2019). *The development of pre-service teachers' schema for teaching students with diverse learning needs*. Unpublished doctoral dissertation, Charles Sturt University, Bathurst, Australia.
- Auhl, G., & Bain, A. (2020). Do pre-service teachers develop a schema for inclusive classroom practice? In press *Asia Pacific Journal of Teacher Education*.  
<https://doi.org/10.1080/1359866X.2020.1777530>
- Australian Institute for Teaching and School Leadership. (2011). *Accreditation of initial teacher education programs in Australia: Frequently asked questions, standards and procedures*. Carlton South, Victoria: Education Services Australia.
- Baik, C., Naylor, R., Arkoudis, S., & Dabrowski, A. (2019). Examining the experiences of first-year students with low tertiary admission scores in Australian universities, *Studies in Higher Education*, 44(3), 526-538. <https://doi.org/10.1080/03075079.2017.1383376>
- Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy (Structure of the Observed Learning Outcome)*. New York, NY: Academic Press.
- Borg, S. (2015). *Teacher cognition and language education: Research and practice*. London: Bloomsbury.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in High School: A cross-subject analysis with student fixed effects. *Journal of Human Resources*, 45(3), 655-681. <https://doi.org/10.1353/jhr.2010.0023>
- Conifer D. (2019). Low scoring students to be barred from becoming teachers under Labor government. *ABC News*. Retrieved from <https://www.abc.net.au/news/2019-01-06/low-scoring-atar-students-to-be-barred-from->
- Donnelly, K. (2019). It's not academic: Raising ATAR for teachers won't raise standards. *Sydney Morning Herald*. Retrieved from <https://www.smh.com.au/national/it-s-not-academic-raising-atar-for-teachers-won-t-raise-standards-20190108-p50q6e.html>
- Feldon, D. F. (2007). Cognitive load and classroom teaching: The double-edged sword of automaticity. *Educational Psychologist*, 42(3), 123-137.  
<https://doi.org/10.1080/00461520701416173>
- Goldfried, M. R., & D'Zurilla, T. J. (1969). A behavioral-analytic model for assessing competence. In C. D. Spielberger (Ed.), *Current topics in clinical and community psychology* (Vol. 1, pp. 151-196). London: Academic Press.  
<https://doi.org/10.1016/B978-1-4831-9972-6.50009-3>
- Goldhaber, D. (2007). Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources*, 42(4), 765-794.  
<https://doi.org/10.3368/jhr.XLII.4.765>
- Hanushek, E. A., Piopiunik, M., & Wiederhold, S. (2019). The Value of Smarter Teachers: International Evidence on Teacher Cognitive Skills and Student Performance. *Journal of Human Resources*, 54(4), 857-899. <https://doi.org/10.3368/jhr.54.4.0317.8619R1>
- Lancaster, J., & Bain, A. (2019). Teacher preparation and the inclusive practice of pre-service teachers: A comparative follow-up study *International Journal of Inclusive Education*, [doi.org/10.1080/13603116.2018.1523954](https://doi.org/10.1080/13603116.2018.1523954).

- Marshall, S. P. (1995). *Schemas in problem solving*. Cambridge: Cambridge University Press.  
<https://doi.org/10.1017/CBO9780511527890>
- McGraw, A., & Fish, T. (2018). Selection and rejection in teacher education: qualities of character crucial in selecting and developing teacher education students. *Asia-Pacific Journal of Teacher Education*, 46(2), 120132. <https://doi.org/10.1080/1359866X.2017.1355048>
- Naylor, R., Baik, C., & Arkoudis, S. (2018). Identifying attrition risk based on the first-year experience, *Higher Education Research & Development*, (37)2, 328-342. <https://doi.org/10.1080/07294360.2017.1370438>
- Norton, A. & Cherastiditham, I. (2018). *Dropping out: The benefits of trying university*. Retrieved from Grattan Institute Website: <https://grattan.edu.au/report/dropping-out/>
- Palmer, N., Bexley, E., & James, R. (2011). *Selection and Participation in Higher Education: University selection in support of student success and diversity of participation. Prepared for the Group of Eight*. Centre for the Study of Higher Education: The University of Melbourne.
- Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2011). Can you recognize an effective teacher when you recruit one? *Education finance and Policy*, 6(1), 43-74. [https://doi.org/10.1162/EDFP\\_a\\_00022](https://doi.org/10.1162/EDFP_a_00022)
- Robinson, N. (2018). Students with lowest ATAR scores being offered places in teaching degrees: Secret report. *ABC News*. Retrieved from <https://www.abc.net.au/news/2018-09-18/students-lowest-atar-scores-teaching-degree-offers-secret-report/10200666>
- Seel, N. M., Darabi, A. A., & Nelson, D. W. (2006). A dynamic mental model approach to examine schema development in performing a complex troubleshooting task: Retention of mental models. *Technology, Instruction, Cognition & Learning*, 4(3/4), 303-329.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285. [https://doi.org/10.1207/s15516709cog1202\\_4](https://doi.org/10.1207/s15516709cog1202_4)
- TEMAG (2014). *Action now: Classroom ready teachers*. Retrieved from [https://docs.education.gov.au/system/files/doc/other/action\\_now\\_classroom\\_ready\\_teachers\\_print.pdf](https://docs.education.gov.au/system/files/doc/other/action_now_classroom_ready_teachers_print.pdf)
- Trochin, W. (2021). *Measurement validity types*. <https://conjointly.com/kb/measurement-validity-types/>
- UAC. (2019). Australian Tertiary Admission Rank: What is the ATAR? Retrieved from <https://www.uac.edu.au/future-applicants/atar>
- Urban R. (2018, April 23). Pressure on universities to raise entry level for teaching courses. *The Australian*. Retrieved from <https://www.theaustralian.com.au/nation/nation/pressure-on-universities-to-raise-the-entry-level-for-teaching-courses/news-story/290dced1b1ced0e564b8fa79703050a1>
- Wilson, R. (2020) *The Profession at Risk: Trends in standards for Admission to Teaching*. Commissioned report. NSW Teachers Federation
- Wright, V. (2015). Is ATAR useful in predicting the success of Australian students in initial teacher education? *Australian Journal of Teacher Education*, 40(9), 1-15. <https://doi.org/10.14221/ajte.2015v40n9.1>
- Wurf, G., & Croft-Piggin, L. (2015). Predicting the academic achievement of first-year, pre-service teachers: the role of engagement, motivation, ATAR, and emotional intelligence. *Asia-Pacific Journal of Teacher Education*, 43(1), 75-91. <https://doi.org/10.1080/1359866X.2014.932328>