Chapter 6
The Emotional Work of Being an Assessor: A Reflective Writing Analytics Inquiry into Digital Self-assessment

Jill Willis and Andrew Gibson

Abstract Digital tools are reshaping how we understand assessment and evaluation in educational contexts as they create new forms of digital assessment data. Such data has been critiqued by educational scholars as it is increasingly associated with high stakes accountability, with the computational interpretation and abstraction occurring at a distance from the authors and their learning experiences. This chapter explores an alternative, educative vision for digital tools and the assessment data they collect. It is a case study of how a digital self-assessment tool captured evidence of early career teachers developing their assessment capability. It also positions an innovative digital research methodology within global sociological concerns about digital assessment tools, to consider how they might inform locally meaningful data stories.

1 Introduction

Currently there are few insights about how early career teachers navigate their assessment capabilities within a rapidly changing data driven educational landscape. 20 Early career teachers (ECTs) in three Australian contexts reflected and self-assessed their experiences in their first year of teaching using an online digital reflective writing application called GoingOK.[1] The ECTs self-reflections were recorded in a personal digital journal, alongside a plotline of point-in-time emotional responses to the prompt 'How are you going?' The digital data was analysed in a transdisciplinary process that combined interpretive and computational insights, that together enables pragmatic identification of meaningful data stories. One of the unexpected data stories was about the emotional work of becoming an assessor. This chapter illustrates

[1] http://goingok.org.

J. Willis (✉) · A. Gibson
Queensland University of Technology, Brisbane, Australia
e-mail: jill.willis@qut.edu.au

A. Gibson
e-mail: andrew.gibson@qut.edu.au

J. Fox et al. (eds.), *Teacher Education in Globalised Times*,
https://doi.org/10.1007/978-981-15-4124-7_6

the relationships between reflective self-assessment and learning to develop assessment capabilities as an early career teacher, and the methodological processes of foregrounding a hidden data story about the emotional work of early career teachers when they are learning to be assessors. The chapter concludes with implications for teacher education in the development of teacher assessment capability, and for local innovations in digital tools and research to address global assessment concerns.

Assessment in education has become strongly associated with accountability and performativity in contexts around the world. Assessment data are used to regulate the practices of teachers, schools and teacher education with concerning consequences such as narrowed curriculum, increased stress, greater time pressures and limited freedom being well documented (Ball, 2003; Klenowski, 2011; Thompson & Mockler, 2016). Instead of a local experience of performance to inform a student and teacher of the next steps in learning, assessment information can quickly become a data package that can travel beyond the local context in national and global systems. The affordances of digital technology enable large data sets to be created, leading to comparisons between educational systems which in turn contribute to educational governance systems that are increasingly networked (Sellar, Rutkowski, & Thompson, 2017). These systems help to create dominant teacher quality discourses of effectiveness, performance and readiness that reflect a rational, scientific approach to education which can lead to a technical form of professionalism (Eacott, 2017; Churchward & Willis, 2019; Bourke, Ryan, & Lloyd, 2016). In this policy context, assessment outcomes play a significant role, becoming part of algorithmic accountability calculations that are often opaque and used for decisions far removed from the control of the original assessment actors. Yet digital data innovations that create performativity discourses, also point to possibilities for disrupting assessment practices to enable small everyday data to generate local, collaborative assessment responses (Cope & Kalantzis, 2016).

Importantly online digital tools also have the potential to enable new local forms of assessment. When pedagogic and computational systems are designed together, there is the potential for digital tools to support teachers and students to collaboratively inquire into learning and reclaim some ownership over assessment decisions. One example is through online digital self-assessment tools that can assist learners to record their own reflective self-assessment over time, and through seeing patterns, engage in self-regulation. The teacher can quickly gather this information, see immediate patterns from a group of students, and make adjustments to their pedagogical practices in response. There are a number of digital tools like online polling software, learning management systems and apps such as Class Dojo that enable such self-assessment activities.

While such digital formats of online formative assessment tools are convenient and easy to use, some with international popularity and reach, they have also raised new questions. Ownership of data and on-selling of data is one concern (Slade & Prinsloo, 2013). When digital tools become ubiquitous, they can also be creating an always-measured sense of self that is fostered throughout childhood through schooling practices as well as wearable devices and digital games (Smith, 2017; Lupton, 2016). It is not yet clear what the extended temporal consequences might be

for authors when aggregated smaller scale assessment performances like digital self-assessments are adopted, as they have an existence long after the events, and might be read by unknown audiences (Williamson, 2016). Use of digital self-assessment tools in research in learning situations raises complex ethical questions (Gibson & Willis, in press). Researchers need to consider ethical use of digital tools and data alongside their potential to identify new patterns and possibilities for learning. We also need to know more about the experience of teachers as they engage in their everyday experiences of being assessors.

Teachers are front line actors in the changing digital assessment story. Teachers make decisions to use digital assessment tools, and their assessment decisions, aggregated by digital means, are also scrutinised. To understand how to prepare preservice teachers to be adaptive, ethical assessors in an increasingly digitally mediated environment, we need to know more about the day-to-day assessment experiences of early career teachers. This chapter reports on insights about being an assessor; insights that became evident in self-assessment reflections of three groups of ECTs. Reflections were gathered through an online reflective writing application and analysed using Reflective Writing Analytics (RWA). In analysing the reflections we noticed how learning to be an assessor is related to intensive emotional labour. This chapter highlights the findings about the emotional work of being an assessor and points towards possibilities for local and global actions for teacher educators and for researchers drawing on digital self-assessment data.

2 Reflective Writing Analytics: A Socio-Technical Interactive Ecosystem

To realise the potential of digital tools informing local evidence stories and empowering authors, the computer, human users and the social context need to be considered as part of a socio-technical ecosystem. Computational analysis promises quick analysis of large amounts of data, but it is often not meaningful. Human socio-cultural analysis can find patterns in human experience, but it is labour intensive and the generalisability of interpretations conclusions can challenged when applied to data at scale. This is particularly the case when dealing with reflective writing, where data is personal writing as part of a process of self-reflection. Finding sensitive ways to uncover the meaning of reflective writing is a key focus of Reflective Writing Analytics (RWA). RWA uses natural language processing (NLP) technologies to computationally analyse reflective text, interacting with human insights for the purposes of scalable meaning-making (Gibson, 2017). Our approach has been to bring social and technical philosophies and approaches together through a continuously interactive dialogue, pragmatically focused on achieving good outcomes for participants. We draw on Peirce's (1905) pragmatic maxim that the full understanding of a concept is found in its practical effects, and apply this idea in a socio-technical process of inquiry (Gibson & Lang, 2018). This is an application of Transepistemic Abduction (TeA)

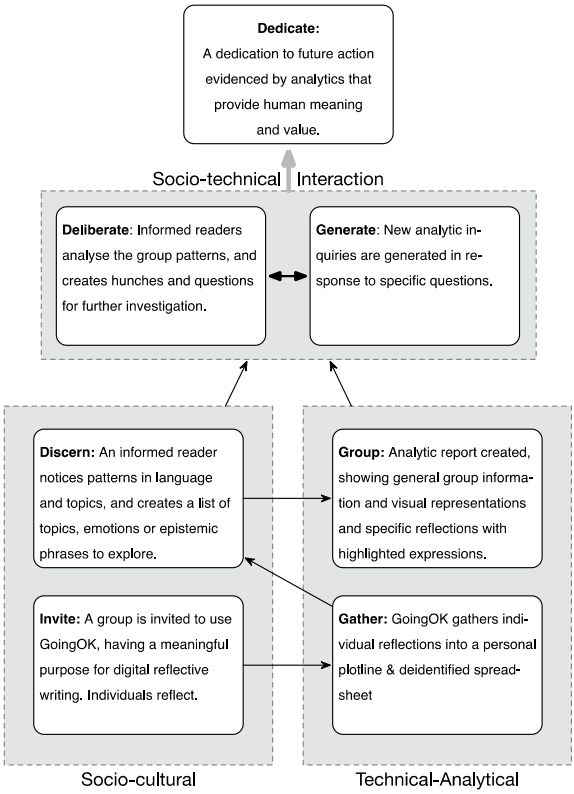

Fig. 1 A pragmatic socio-technical approach to Reflective Writing Analytics (RWA) with GoingOK

where abductive reasoning occurs across distinct epistemic domains towards a productive end (Gibson, 2017). Abductive reasoning is sometimes thought of as the guess which best explains a surprising phenomenon. The way that TeA informs our socio-technical approach is illustrated in Fig. 1. Interactions between the socio-cultural and the technical-analytical domains lead to an abductive interaction (a kind of interactive 'guessing') that is inherently socio-technical. It is a dialogue between human deliberations and generated computational analytics. This dialogue is focused towards a dedication to a practical outcome—the application of the analytics to support further action. We detail this process in the remainder of this chapter showing the potential for the theory of TeA and application of RWA in Assessment for Learning (AfL).

3 Invitation: Early Career Teachers Reflect

GoingOk was created to support a group of early career teachers in their first year of teaching through facilitating digital personal reflective writing (Gibson, 2017). The digital tool and the pedagogic processes have since been refined across seven

projects from 2013, and as of January 2020, GoingOK has been used by over 2600 people who have written more than 14,600 reflections. This growing database allows for the creation of computational models of reflective writing characteristics, which can help inform the socio-technical approach to RWA. It also allows for analysis that is difficult to do manually, such as feature comparisons between subsets (e.g. a single cohort of authors) and all reflections in the database. This can help reveal characteristics of reflection that are unique to a specific group of authors, or which might be found more generally in the reflective writing of authors from different contexts.

Three groups of early career teachers (ECTs) were invited to reflect in general terms about their experiences in teaching using GoingOK, authoring their own teaching story. The ECTs were located in remote Queensland (BCQ), South Australia (BCSA) and metropolitan and regional Queensland (BW), Australia (see Table 1). Ethical approval from the host universities was sought and participants consented for the data to be compared in future analyses such as this.

When an author logs into their personal GoingOK page and create a reflection, they create three different data types (see Fig. 2). Firstly the author uses a visual slider to indicate how they are going, which creates the **reflection point**. The range from *distressed* to *soaring* with *going ok* is captured as a number between 0 and

Table 1 GoingOK projects with ECTs

Year	Project name	Code	Number of authors	Number of reflections
2013	Becoming colleagues (Queensland)	BCQ	7	118
2014–15	Becoming colleagues (South Australia)	BCSA	8	129
2014–15	Beginning well	BW	5	58
			Total 20	Total 305

Fig. 2 The GoingOk user interface

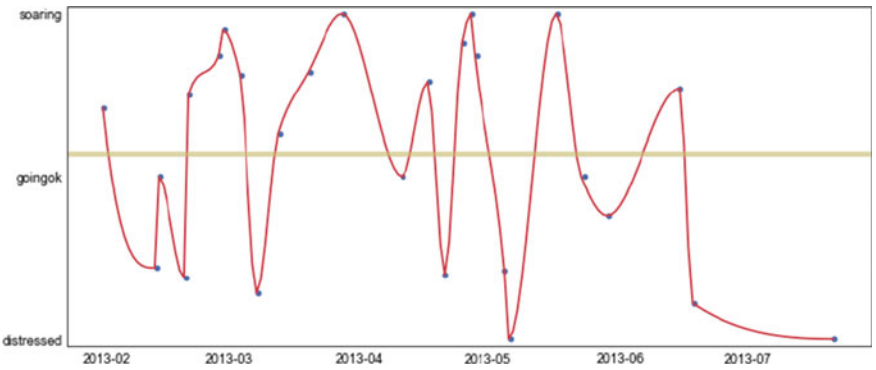

Fig. 3 An example of a digital plotline over time (kolsag, BCQ)

100. Secondly the **reflection text** in response to the prompt *Describe how you are going…* is captured as text. Thirdly when the author clicks the 'save' button, the server adds a **reflection timestamp** to the reflection entry.

Each participant had their own account, and for each reflective entry would slide the slider to indicate how they were going, a decision that was recorded by the digital tool with 0 representing distressed, 100 representing soaring, and 50 representing going ok. As the entries accumulated over time, they were visible to the beginning teacher as a plotline, with the ups and downs recorded (Fig. 3).

Authors also wrote reflective text in the text box each time, that created a longitudinal narrative. The data set is a collection of rich, point in time reflections about the day-to-day experiences of beginning teachers (Crosswell, Morrison, Gibson, & Ryan, 2017; Crosswell, Willis, Morrison, Gibson, & Ryan, 2018).

The first group of ECTs (BCQ) worked with the research team[2] while they were in their final year of their undergraduate program and during their first year of teaching engaging in collaborative interpretation of the data. The sliding scale was modelled on the idea of a common Assessment for Learning 'traffic light' self-assessment activity. However the ECTs and researchers were keen to avoid a deficit framing of ECT capabilities that might be implied by medical metaphors of pulse checking, or external judgment metaphors of performance. Instead the decision was made to recognise the personal agency of the ECTs who were *authors* of their professional identity, creating *plotlines* of their experiences, and through the creation of *reflective narratives* that could empower them to be resilient colleagues (Morrison, Willis, Crosswell, & Gibson, 2014). Reflective writing was seen by the ECTs and researchers as a strategy for professional and personal development.

[2]The research team included the authors, along with Leanne Crosswell and Chad Morrison.

4 Gathering Reflective Writing—ECT Professional Self-assessment and Personal Reflexivity

Reflection is a well-established practice in teacher education, with deep roots in the work of Dewey (1933) and Schön (1983). Reflection is an active and intentional inquiry that enables the author to look back on experience, to question and make connections, and in doing so create new knowledge to inform current and future practice. Through reflection teachers can frame and reframe problems of practice (Loughran, 2002), and craft a responsive professional identity (Beijaard, Meijer, & Verloop, 2004). Expressively writing a story about self, as is done in reflective writing also provides personal well-being benefits (Pennebaker & Seagal, 1999). These benefits were also experienced by the early career teachers who indicated that GoingOK felt like 'a big listening ear'.

As the digital tool also required the author to make a judgement about their emotional state using a sliding scale and then explain it, the reflections could also be seen as a form of self-assessment. Reflection and self-assessment are closely related. Self-assessment is most often associated with cognitive views of assessment for learning or formative assessment pedagogies where the learner/author evaluates their performance against formal criteria and/or specific curriculum goals to develop self-regulation in formal learning settings (Panadero, Andrade, & Brookhart, 2018). However other assessment scholars write about self-assessment as a more informal and internalised process of reflection that legitimises the experiences of the author as they explore new learning experiences (Swaffield, 2008; Bourke & Mentis, 2013). Self-assessment when viewed from this socio-cultural theoretical lens is a process of social negotiation. Explicit and tacit, group and individual knowledge is always being worked out and resolved by the individual as they engage in the practices of the group (Willis & Cowie, 2014). At the heart of self-assessment is a process of questioning and evaluating, of self judgement that can lead to decisions about what to do next (Boud, 1995; Sadler, 1989). Margaret Archer's (2003, 2007) theories inform this analysis, especially her proposal that in an increasingly complex world, individuals draw on their inner evaluative conversations to make their way through the world, reflexively *discerning* and *deliberating* about their situations in light of their concerns, to *dedicate* towards what to do next.

When the first sets of data were read by the researchers it became evident that GoingOK reflective entries typically were highly self-evaluative. The early career teachers were evaluating their own experiences at that moment of the entry by sliding the slider, and in their text often began by discerning what they were feeling and thinking about, deliberating about what was important or not, and often by the end of a short reflective entry dedicating towards a course of action. For example,

> Having a fabulous week in terms of teaching and being organised but am feeling a little bit stressed about reporting for term 1 [Discerning feelings]. I'm mostly feeling this way because I just don't know how it's going to turn out. We have requested parent interview for the beginning of term 2 and am feeling a little bit anxious about this. [Deliberating about why reporting is causing anxiety]. I probably don't have anything to worry about, it's just

> that it will be my first time. [Dedicating to press on with some reassuring self-talk] (kolsag BCQ, plotpoint 82).
>
> I have just finished report writing with interviews next week. This week has probably been the most full on and stressful week of my teaching career so far [Discerning context and evaluating emotion]. I did have my first teacher meltdown this week with many different layers responsible for this, [Deliberating about why] but I have been proactive in dealing with stressors and I have been communicating strongly with my class. I have great external support. I am really looking forward to next week. I feel like I'm starting fresh and have a really clear mind [Dedicating towards positive next steps] (dobpoc BCSA, plotpoint 63).

These internal conversations were ECTs working through their concerns to recognise what they were caring about and how they might adapt their practices to have a better chance of realising what matters most to them (Archer, 2003). The GoingOK entries of ECTs recorded deliberations about a wide range of topics like moving to a new town, figuring out expectations of principals and the parents of students in their classes, recording highs and lows of classroom teaching, managing a sense of professional identity, and trying to eat healthy and exercise to be resilient (Crosswell et al., 2018). This analysis focused on their experiences as assessors, using a socio-technical analytic approach.

5 Discerning Assessment Capability Through Socio-Cultural Qualitive Analysis

The first qualitative analysis of the reflective entries for assessment capability was informed by a socio-cultural view of assessment that acknowledges that learning is a social process, embedded in contexts and relationships. This view is reflected in much of Australia's research about teacher assessment capability and informed the inductive qualitative coding. Teachers have multiple social roles in assessment practices including those of assessors, pedagogy experts, student partners, motivators, teacher learners, and stakeholder partners (Alonzo, 2016). Teachers also have to be literate in multiple assessment contexts from high stakes standardised tasks, to classroom interactions and culture fair assessment (Willis, Adie, & Klenowski, 2013). Assessing involves critical reflexivity and collaboration (Willis & Klenowski, 2018; Wyatt-Smith & Gunn, 2009), and considerations about beliefs and feelings about assessment (Looney, Cumming, van der Kleij, & Harris, 2017). Charteris and Dargusch (2018) also emphasise that assessment capability transcends a checklist approach and is an ongoing development that involves bodies and minds in assessment decision making in situ. The initial coding was done by one author (Willis) who noted the phrases that referred to assessment practices, relationships, beliefs, emotions and concerns. A second reading using Archer's scheme of 3Ds of reflexivity (*deliberation, discernment, and dedication*) highlighted how the ECTs were engaging in reflexive self-evaluation.

Assessment seemed to be a recurring topic for reflection, often interwoven with other topics. The early career teachers confidently reflected on how they used data

to evaluate how well their students were learning, and how to integrate assessment into their daily planning, assessment and reporting. They discerned their excitement and pride when students showed progress in pre and post test data, or where they successfully exhibited portfolio work.

> Feeling good about where things are heading. We had a successful exhibition of learning last night that the students came in with parents to show their work. It was a long process to ensure that everything was completed in time, but it was very rewarding (gixbob, BCSA, plotpoint 83).
>
> I'm going ok. Assessment time and things have shifted up a gear. Late night marking and phone calls home is what I love. Assessment time has caught me by surprise, but I am really enjoying seeing mine and the students hard work paying off when they are receiving great marks. Obviously it is disappointing to see those students not meeting their potential, but fortunately I haven't had too many of those yet. (doclip, BCQ, plotpoint 85).

However even such positive statements were located near reflections about being tired, and managing their time and workload. These concerns were even more evident in a dominant assessment topic of concern—marking, reporting and sharing reports with parents in interviews:

> Am going ok. Teaching seems to be going well & behaviour management is a lot easier now. Am struggling with the workload though **- trying to fit everything in… marking…assessment…etc etc…so many things to do.** So many things I WANT to do but just don't have the time. Hard to find a work-life balance! (catpub, BCQ, Mid April plotpoint 42).

Two months later the same ECT reflected that they now expected report time to be distressing, with the emotions of feeling judged by others heightened.

> Prior to and during report cards, **I was 'distressed'**… as can be expected during this time! It felt like I was writing multiple uni assignments all over again. Plus with the added stress of my Principal and parents reading it. (catpub, BCQ, plotpoint 41).

Other ECTs were similarly emotionally challenged by preparations to share assessment results with parents.

> **Very exhausted this week.** Busy preparing for parent teacher interviews. Am **feeling very lonely this week**. (domsiv, BCSA, plotpoint 32).
>
> Report writing is **really stressful!** I'm currently wading through my data which includes work samples, anecdotal observations and photos. I know the children really well and the comments section is fine but assigning a grade is **awful**…I'm so concerned he will be deflated and stop loving school again. I never knew there was so much to report writing. Also those darn parents who question everything will want hard evidence of their child's grade. **Already a bit sweaty thinking** about it. (juzvun, BCSA, plotpoint 50).
>
> I am currently finishing off my first lot of report cards. I haven't been stressing too much but I didn't finish them when they were meant to be originally due (last Wednesday), but neither did any of the other teachers. Our principal was not impressed with us. I found having them all due in on Wednesday during week 8 meant that every piece of assessment had to be completed by week 7, which for some of my units that I only have once a week is very **ridiculous and difficult.** I don't know whether experienced teachers would find it hard to teach everything in five weeks and then start to assess…**I am extremely tired all of the time but still feel like I'm not doing an awesome job even though I'm putting in my all**. (rulguz, BCQ, plotpoint 50).

Assessment capability included being able to manage the emotions associated with being judged, while also managing their time and energy. In these entries ECTs discerned their emotional states by directly labelling them—*distressed, exhausted, lonely, stressed, concerned, extremely tired*—or by implied in evaluations of their situations as they deliberated about why there were experiencing these emotions, with cultural constraints that were—*awful, ridiculous, difficult.*

A less obvious, but still emotion laden assessment capability was using diagnostic, formative assessment or Assessment for Learning (AfL) to be responsive to student needs. It was less obvious as it was often entangled with stories of managing student behaviour:

> This entire week has been an absolute drag. [Discerns current emotional state and reasons for it] I've had misbehaving students, too excited students who can't calm down…etc…it's been a tough week. To top that off, I also realised at the end of the week that I haven't been formatively assessing anything as yet. I've jumped into preparation of formative assessment for the next week [Decides to take action], but I feel like the stress of how my students were behaving caused me to forget about the other parts of teaching [Deliberates about why action didn't happen earlier] (kolsag, BCQ, plotpoint 21).

AfL capabilities were implied when an ECT was struggling to teach a student with disabilities. The ECT reflected that they knew they should identify what his learning needs were, but still felt out of their depth:

> Am currently feeling distressed [Discerns current emotional state and reasons for it] as I'm finding is extremely difficult with one of my new students who has high functioning autism and ADHD. He has been in detention every day for the past 4 days, has hit and punched kids and has swore repetitively towards other students and teachers. It is hard to get him to do any work as he just says 'no', runs off and swears. I know it's still early days (this is his 3rd week), but it doesn't make it any easier. I'm getting sick from lack of sleep and stress. [Deliberates and weighs up all of the factors to consider] Am finding it hard to teach the other kids when he'll talk over me, interrupt every lesson etc. I want the best for him. **I want him to love school and learn at his pace and from where he's at. I know it will take time to get to know him and get to know strategies which work best with him,** but these early days are still hard. There are no support staff trained/specialised with working with students with disabilities/high learning needs. The school is having to train up a teacher aide, so at the moment it feels like I'm on my own. No university course can prepare you for this! I only had 1 subject about teaching students with disabilities. So I am feeling out of my depth. [Decides to reframe the problem into a positive light] I know this is a great experience for 1st year out, and will only make me stronger. (catpub, BCQ, plotpoint 0).

University preparation in Australia emphasises an expectation that ECT will differentiate the curriculum for the individual using AfL strategies to find out what the student knows, in order to plan differentiated practices. In the reflection above, the ECT seems to refer to these assessment and differentiated planning practices in the phrase 'I want him to love school and learn at his pace and from where he's at'. The ECT's self-assessment commentary is that they feel their work is 'hard' and they were losing sleep and getting sick. They reflect that their university learning has not fully prepared them to manage differentiated practices in ways that were simultaneously urgent and extended over months. AfL would most likely have been taught as a separate assessment capability at university. In practice, AfL was not separated from

classroom management, or from managing the resources available in the context, and the emotions of feeling uncertain.

This entry highlights that assessment capability is a holistic experience (Hill, Ell, & Eyers, 2017). What was apparent throughout this entry and many others was that in teaching 'emotion and cognition, self and context, ethical judgement and purposeful action are all entwined', with the emotions showing what is at stake as teachers make sense of the complexities of teaching (Kelchtermans, 2005. p. 1000). In their self-assessing online entries the ECTs were evaluating their performances and their contexts, seeking to turn challenges into learning experiences. During the inductive, qualitative analysis, the relationship between assessment and emotions seemed to be quite clear, and new territory to explore. This led to the next phase of data analysis using Reflective Writing Analytics to identify patterns.

6 Grouping Language Features for Emotions and Assessment Using Reflective Writing Analytics (RWA)

Up until this point, the digital tool had just collected the qualitative data that could be analysed in traditional qualitative ways. The inductive hunch that emotions were strongly related to assessment capability could have remained as a discussion point and hypothesis for further explanation. Instead, the affordances of digital quantitative analysis enabled us to explore these inductive hunches further. Assessment key words that were evident in reflections analysed qualitatively were used to select relevant reflective entries for further analysis (Table 2).

Those assessment words were displayed as red in a RWA report which annotated the reflective text with features of interest (see Fig. 4). In this report, almost all of the authors (19 of 20) had written reflections that included assessment words. Reflections with assessment words represented nearly a quarter of all reflections (28% of BW,

Table 2 Assessment words used to select relevant reflective entries

	Assessment words	
Assess	Grade	Interview
Assessment	Test	Parent
Exam	Quiz	Report
Examination	Formative	Report cards
Assignment	Summative	Reports
Mark	Workload	Reporting
Marking	Time	Feedback
	Performance	
	Data	

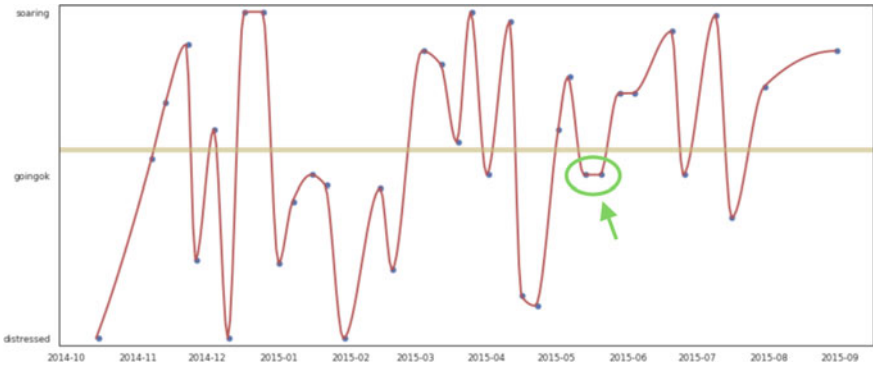

Fig. 4 An example of an assessment reflections with RWA analytic highlights for movbuc, BCSA

Table 3 Frequency of assessment reflections

Year	Code	Number of authors	Number of authors who wrote about assessment	Number of reflections	Number of reflections with assessment words
2013	BCQ	7	7	118	33 (27%)
2014–15	BCSA	8	7	129	20 (15%)
2014–15	BW	5	5	58	17 (29%)

27% of BCQ), and 15% of BCSA). The number reflections for each cohort is shown in Table 3.

The entries where AfL might be implied in pedagogic discussions were not identified by the lexical approach. Simple word matching where there is a lack of ambiguity in word meanings provides a reasonable filtering approach, but for a more detailed analysis, an NLP technique known as topic modelling may need to be employed in future work. Nevertheless, even considering that not all assessment related reflections were identified in the word matching approach, the significant number of reflections identified suggests that assessment is an important topic of reflection for these teachers.

In the RWA report, affective words with a high level of valency and arousal (Warriner, Kuperman, & Brysbaert, 2013) were identified and highlighted in pink (see Fig. 4 for an example). With this lexical analysis, words were identified in the reflections that had previously been quantified (by a crowd-sourcing process) in terms of valence (the positivity or negativity), arousal (the affective strength), and dominance (controlling or being controlled). What was interesting to note was that phrases that might convey emotions to a human reader such as ‘really stressful’, ‘is awful’, ‘I’m so concerned’ and ‘feeling a bit sweaty’, or the implied pleasure in phrases like ‘for the first time EVER’ and ‘win-win’ were not highlighted. This was due to both the

lexical nature of the analysis, and that the affective character of ordinary words is different to words (and expressions) that are intended to convey specific emotions. This means that rarely is a single affective feature sufficient for analysis of the complexity of emotion, and multiple approaches need to be taken. In this analysis, the approach was to extend the analysis to include non-lexical textual features that are proxies for emotional expression. These were annotated in orange and included the use of all capitals (e.g. WOW!, first time EVER), the use of multiple explanation and question marks (e.g. !!!, ?!?!) and the use of repeated vowels (e.g. I am soooooo stressed right now). The RWA also included the extraction of epistemic expressions (e.g. I think, I wonder) annotated with blue highlighting. These tend to be self-evaluative statements that were associated with deliberation in the qualitative analysis, such as 'I'm trying', 'I know these are', 'I realize that this is probably' and 'I am still not really happy with'. These deliberative phrases are highlighted in green and show a temporal relationship between previous and current or future understanding. Finally, the RWA also included 'my' expressions (annotated in yellow). An abductive hunch was that these expressions might be helpful in distinguishing between a personal perspective (e.g. my classroom, my students) and a more objective perspective (e.g. the classroom, the students). This might be helpful for analysis as personal language tends to be associated with a deeper process of reflection indicative of the 'deliberating' and 'dedicating' in the qualitative analysis.

Our RWA enabled the selection of emotion rich entries, foregrounding them for the researcher's close attention, and highlighting features that were not necessarily noticeable from other ECT data. For example the plotpoints related to the reflective entries from Fig. 4 are circled in green in the ECT's plotline in Fig. 5. These reflections are within the middle, or 'GoingOK' range, and would not have triggered a researcher's attention to a particularly high or low emotional 'turning point', if just the plotline was used as a data source.

This nonlinear relationship between emotions in the reflective entries and the measures like the plotline pointed to the need for a range of different reflective writing analytics, to avoid over-interpreting what look like clear data pictures about emotions. Similarly, careful qualitative analysis was required to avoid drawing over-simplistic interpretations based on the purely computational analysis.

7 Socio-technical Inquiry: Investigating Specific Relationships Between Emotions and Assessment Using Reflective Writing Analytics (RWA)

We found in this work that extensive socio-technical interaction was needed to deliberate about possible meanings and generate specific inquiry questions. Socio-technical interaction brings together the strengths of both the computational and socio-cultural analyses, to support the end practical goal of the research. In this case

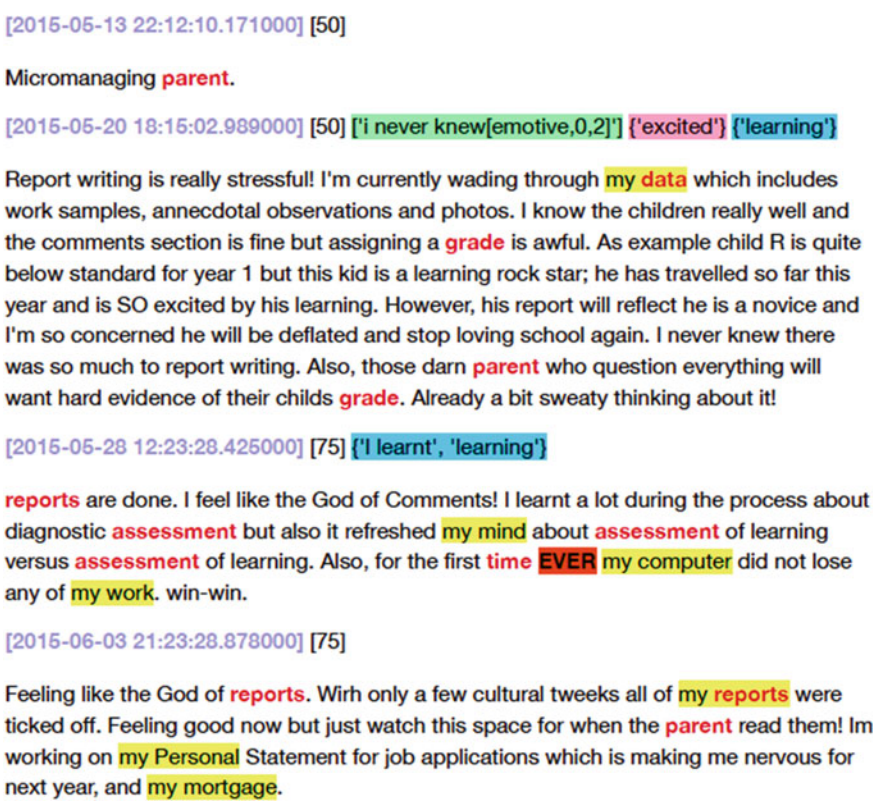

Fig. 5 Corresponding entries in the overall plotline for movbuc, BCSA

that goal was to obtain a better understanding of the relationship between assessment and emotion for early career teachers.

The socio-cultural hunch that assessment reflections were highly associated with emotional work was the founding 'guess' of the abductive interaction process referred to previously. This hunch spawned a series of informed guesses about how aspects of assessment, emotion, and learning might be captured in the teachers' reflective writing, and what kind of analytical features help make them visible in the RWA report. The features described previously are the results of this abductive interaction process.

However, identifying key features in hundreds of separate reflections does not make visible any relationships between those features. Hence, our socio-technical interactive approach raised another question: Do the features identified by RWA support the hunch of a relationship between assessment reflections and emotional work? And if so, to what extent can this be made visible?

This question led us into a further abductive interactive process centred on using RWA to visualise the relationship between assessment words and affective features (both lexical and non-lexical). The result of this inquiry process is shown in the

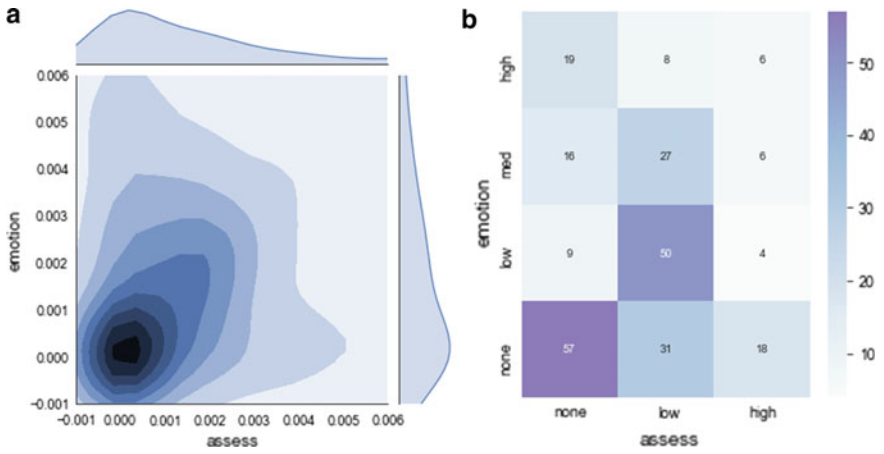

Fig. 6 The relationship between quantities of assessment words and affective features as visualised in **a** a heatmap showing categorical relationships, and **b** a contour plot which relates the distributions

assessment-emotion heatmap (Fig. 6a) and the assessment-emotion contour plot (Fig. 6b).

Prior to creating these visualisations, both the assessment words and the affect features were scaled with respect to the length of each individual reflection. This was to remove the impact of the length of each reflection, which was highly variable. Further, all reflections (including those with no assessment words) were used in the analysis. This ensures that we have a full picture of the reflections, not just those which are identified as relevant to assessment.

The heatmap (Fig. 6a) involved categorising the quantity of assessment words as 'none', 'low', or 'high', and categorising the quantity of emotion words as 'none', 'low', 'medium', and 'high'. This process was done by setting the values that separate the categories (known as bins) to ensure (a) that none accurately reflected 0 features, and that (b) the other categories had roughly similar quantities across the dataset as a whole. The resultant heatmap displays the number of reflections which fall into the two intersecting categories, and higher quantities are visualised with a darker colour. A strong correlation between assessment and emotion in the 'none' and 'low' categories is obvious in the heatmap. Emotional language is still used without assessment words, but it is more common for there to be no emotion when there is no reference to assessment. There are no obvious correlations in the other categories which is not surprising as more assessment words (in any one reflection) do not necessarily mean that the text is more about assessment. Similarly, writing with more affective features (in any one reflection) does not necessarily mean that the author was more emotional at that point in time. This analysis did not include the analysis of authors assessment-emotion use over time, nor did it include an analysis of authors. Those kinds of analyses may prove fruitful in future work.

The contour plot (Fig. 6b) takes the assessment word and emotion/affect feature values for each individual reflection and displays the density of the resultant points.

The darker part of the contour represents and corresponds to correlations between the number of reflections that had 'none' or 'low' assessment words or emotion features that we saw in the heatmap. The lighter parts of the contour represent less reflections. The distributions on each axis show that there were greater numbers of reflections associated with the lower end of the spectrum, particularly for assessment words. This was to be expected, as 235 reflections (77%) did not contain assessment words at all. This contour plot, which is not constructed from categorical data, displays more obviously that when assessment words appear in the reflective writing, there is an increase in emotional expression. That is, there is a clear correlation between assessment words and affective features. Assessment is emotional work.

These two visualisations (Fig. 6) drawn from all 305 reflections provide computational support for our socio-cultural hunch that assessment reflections were highly associated with emotional work. This in turn supports further socio-technical inquiry and the investigation of questions like

(a) What topics feature frequently alongside assessment topics?
(b) Do report cards remain a topic of concern over time?

While the RWA has been able to add value to a socio-technical inquiry on a relatively small dataset, it has also shown the potential for scaling the same analysis to larger quantities of reflective writing. As the dataset of reflections grows, some of the observed patterns, together with the additional questions raised, may be able to provide some answers at scale about the assessment experiences of early career teachers, which in turn might inform teacher preparation courses and mentor training for ECTs.

8 Discussion

8.1 Assessment as a Topic of ECT Concern

In the reflective entries, assessment was a topic of 'concern' that was evident for early career teachers. Archer (2003) argues that the reflective process of *discerning*, *deliberating* and *dedicating* ourselves in our daily concerns is a pursuit of our ultimate concern, to live a *modus vivendi*, or life worth living. Assessment topics were situated, and entangled with concerns about student success, collaborations with colleagues, parent expectations, and balancing personal and professional life in their performance of their new teaching roles. Positive emotions were evident when ECTs confidently used data to monitor student progress, saw student success, got positive feedback from parents, or successfully completed new tasks like writing report cards. Emotional words like 'fun', 'excited', 'fantastic' and 'happy' associated with reflective assessment entries. Importantly, there were other emotions highlighted like 'down', 'anxious', 'scared', 'angry' and quite frequently 'worried'.

Managing emotions, tiredness and the assessment relationships with parents, support staff, principals were associated with assessment capability, however these do not feature in Australian ITE policy statements about assessment. For example,

Australian Professional Standard for Teachers 5.5 that focuses on reporting to parents, highlights knowledge of reporting processes. Graduates need to able to '*Demonstrate understanding of a range of strategies for reporting to students and parents/carers and the purpose of keeping accurate and reliable records of student achievement*' (AITSL, 2011). The cognitive and skills focus in the policy document is an important but insufficient descriptor of assessment capability. What is clear from these plotlines, and the reflective entries is the common occurrence of multiple highs and lows, with the turbulence of emotions a stark contrast to the linear representations of assessment capability in policy documents. This information can support and inform teacher educators who already advocate and teach preservice teachers about professional identity and resilience.

8.2 Emotions and Assessment

While it might be easy to dismiss the emotions of ECTs who are engaging with assessment activities like writing report cards or managing their tiredness as just a reality that the ECTS have to deal with, or a right of passage, the RWA results invite us to look again. Emotions reflect an embodied knowledge where the reflective author is engaging in self-assessment in anticipation of social performances, to shape how the author will find meaning in the event (Archer, 2003). Anticipation of events generates a lot of emotional and intellectual work as 'our emotions go before us to meet the future. This means that emotions do not just happen as internal events (which may be true of moods) …they entail cognition about the intentional object' (Archer, 2003, p. 202). Through emotional commentary, authors negotiate practical, social and embodied realities in their lives, trying to strike a balance between their concerns. For teacher educators, these insights can inform preparation programs for preservice teachers, and, in particular, how self-assessment is integral to pedagogical practices in the classroom. Paying attention to the concerns and representations of ECTs can also inform supporters like mentors, principals and system designers, who provide social support for the emotional work of assessment.

8.3 Potential for Collaborative Socio-technical Analysis Between Educators and Learning Analytics Specialists

Assessment data reported at scale is socially convincing and has a washback effect on the work of teachers, often in unhelpful ways (Spina, 2017). As evident in the flurry of commentary in the wake of 2019 international PISA assessment results, large-scale data can lead to normative performance stories that in turn drive reactive policy decisions. An alternative future for assessment data may be possible through learning analytics that enable more local and contextualised educational digital assessment data story telling. Personal self-assessing reflections such as those gathered by GoingOK

have the potential to provide authentic experiential data, especially for small scale assessment data stories as inductive qualitative social analysis is time-consuming work. For larger scale assessment data stories, Reflective Writing Analytics have the potential to honour the experiences of the reflectors while enabling large-scale data analysis that can have meaning for educators and computational analysts alike. Digital tools by themselves cannot achieve this work. Nuanced computational analysis through Reflective Writing Analytics relies on the co-construction of meaning with users. The field of Learning Analytics is still new and there are calls for it to mature beyond exploratory small studies towards larger scale systems and validating studies (Dawson, Joksimovic, Poquet, & Siemens, 2019). More opportunities for collaboration, author facing analytic tools and reports are likely to emerge as the field grows. We propose further investigation through a collaborative socio-technical approach involving teacher educators, classroom teachers and school mentors and leaders.

9 Conclusion

'Teacher assessment capability' is powerful phrase that acknowledges the situated experience of learning to be an assessor over time, yet it is not a concept that has been well researched or understood in Australia to date. This analysis of the emotional work of being an early career assessor points to a productive avenue for research into how to prepare preservice teachers to be assessment capable in their early career experiences. Reflective self-assessment data gathered using GoingOK has enabled day-to-day assessment experiences of early career teachers to be acknowledged and explored. From the analysis in this chapter, reporting and managing the workload of multiple assessment deadlines, and integrating AfL, differentiated pedagogy and behaviour management are highlighted as possible areas for further development in initial teacher education courses. It is acknowledged that university cannot prepare new teachers for every circumstance and that valued learning happens during embedded experience. Knowing more about where the challenges might be for groups of ECTs does provide important data to inform ongoing support.

There are implications for policy as the analytic computational representations of teacher experiences have potential to provide new forms that can enable the local voices of the authors in assessment contexts to be heard. Data from preservice teachers who are learning to assess is currently being gathered in four countries using GoingOK and will enable the research team to understand whether the experiences of Australian preservice teachers are similar or different to those in Canada, England and New Zealand partner universities. It may be that learning to manage the emotional work of assessing is an international assessment capacity that needs to be taught in practice and in initial teacher education.

Glossary

Abductive reasoning Abductive reasoning is sometimes known as inference to the best explanation, and can be thought of as the guess which provides the most likely explanation for a surprising phenomenon. It is contrasted to inductive reasoning where generalisations are made from specific observed phenomena, and deductive reasoning which involves reaching a logical conclusion from given premises by following rules of deductive logic.

Reflective Writing Analytics (RWA) RWA uses natural language processing (NLP) technologies to computationally analyse reflective text, interacting with human insights for the purposes of scalable meaning-making (Gibson, 2017).

Socio-technical analysis Socio-technical analysis is a dialogue between human deliberations and generated computational analytics. This dialogue is focused towards a dedication to a practical outcome—the application of the analytics to support further action.

Transepistemic Abduction (TeA) TeA is the application of abductive reasoning that occurs across two or more distinct epistemic domains towards a productive end that cannot be reached from a single domain (Gibson, 2017).

References

AITSL (Australian Institute for Teaching and School Leadership). (2011). National professional standards for teachers. https://www.aitsl.edu.au/teach/standards.

Alonzo, D. (2016). *Development and application of a teacher assessment for learning (AfL) literacy tool (Unpublished doctoral dissertation)*. Australia: University of New South Wales.

Archer, M. (2003). *Structure, agency and the internal conversation*. Cambridge: Cambridge University Press.

Ball, S. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy, 18,* 215–228. https://doi.org/10.1080/0268093022000043065.

Beijaard, D., Meijer, P. C., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education, 20*(2), 107–128.

Boud, D. (1995). Assessment and learning: contradictory or complementary. *Assessment for Learning in Higher Education,* 35–48.

Bourke, T., Ryan, M., & Lloyd, M. (2016). The discursive positioning of graduating teachers in accreditation of teacher education programs. *Teaching and Teacher Education, 53,* 1–9.

Bourke, R., & Mentis, M. (2013). Self-assessment as a process for inclusion. *International Journal of Inclusive Education, 17*(8), 854–867.

Charteris, J., & Dargusch, J. (2018). The tensions of preparing pre-service teachers to be assessment capable and profession-ready. *Asia-Pacific Journal of Teacher Education, 46*(4), 354–368.

Churchward, P., & Willis, J. (2019). The pursuit of teacher quality: Identifying some of the multiple discourses of quality that impact the work of teacher educators. *Asia-Pacific Journal of Teacher Education, 47,* 1–14. https://doi.org/10.1080/1359866X.2018.1555792.

Cope, B., & Kalantzis, M. (2016). Big Data Comes to School: Implications for Learning, Assessment, and Research. *AERA Open, 2*(2), 2332858416641907.

Crosswell, L., Willis, J., Morrison, C., Gibson, A., & Ryan, M. (2018). Early career teachers in rural schools: Plotlines of resilience. In *Resilience in education* (pp. 131–146). Cham: Springer.

Dawson, S., Joksimovic, S., Poquet, O., & Siemens, G. (2019, March). Increasing the impact of learning analytics. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 446–455). ACM.

Dewey, J. (1933). Philosophy and civilization. *Philosophy, 8*(31), 360–361.

Eacott, S. (2017). School leadership and the cult of the guru: the neo-Taylorism of Hattie. *School Leadership & Management,* 1–14.

Gibson, A. P. (2017). *Reflective writing analytics and transepistemic abduction* (Doctoral dissertation, Queensland University of Technology).

Gibson, A. P., & Lang. C. (2018). The pragmatic maxim as learning analytics research method. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (LAK '18) (pp. 461–465). New York, NY, USA: ACM. DOI: https://doi.org/10.1145/3170358.3170384.

Gibson, A. P., & Willis, J. (in press). Ethical challenges and guiding principles in facilitating personal digital reflections. In C. Burr & L. Flioridi (Eds.), *Ethics of digital well-being: A multi-disciplinary approach.* Cambridge: Cambridge University.

Hill, M. F., Ell, F. R., & Eyers, G. (2017). Assessment capability and student self-regulation: The challenge of preparing teachers. In *Frontiers in education; assessment, testing and applied measurement* (Vol. 2, p. 21). https://doi.org/10.3389/feduc.2017.00021.

Kelchtermans, G. (2005). Teachers' emotions in educational reforms: Self-understanding, vulnerable commitment and micropolitical literacy. *Teaching and Teacher Education, 21*(8), 995–1006.

Klenowski, V. (2011). Assessment for learning in the accountability era: Queensland. *Australia. Studies in Educational Evaluation, 37*(1), 78–83.

Loughran, J. J. (2002). Effective reflective practice: In search of meaning in learning about teaching. *Journal of Teacher Education, 53*(1), 33–43.

Looney, A., Cumming, J., van der Kleij, F., & Harris, K. (2017). Reconceptualising the role of teachers as assessors: teacher assessment identity. *Assessment in Education: Principles, Policy & Practice.* https://doi.org/10.1080/0969594x.2016.1268090.

Lupton, D. (2016). *The quantified self*. Wiley.

Morrison, C., Willis, J., Crosswell, L., & Gibson, A. (2014). Turning points in narratives of research design: Research innovation stimulating unique responses to existing challenges for beginning rural teachers. *The Journal of Educational Enquiry, 13*(1), 3–17.

Panadero, E., Andrade, H., & Brookhart, S. (2018). Fusing self-regulated learning and formative assessment: A roadmap of where we are, how we got here, and where we are going. *The Australian Educational Researcher, 45*(1), 13–31.

Peirce, C. S. (1905). What pragmatism is. *The Monist, 15*(2), 161–181.

Pennebaker, J. W., & Seagal, J. D. (1999). Forming a story: The health benefits of narrative. *Journal of Clinical Psychology, 55*(10), 1243–1254.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144.

Schon, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.

Sellar, S., Rutkowski, D., & Thompson, G. (2017). *The global education race: Taking the measure of PISA and international testing*. Edmonton, Alberta: Brush Education.

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist, 57*(10), 1510–1529.

Smith, R. (2017). The emergence of the quantified child. *Discourse: Studies in the Cultural Politics of Education*, *38*(5), 701–712.

Spina, N. (2017). The global testing culture: Shaping education policy, perceptions and practice. *Journal of Education for Teaching, 43*(2), 267–269.

Swaffield, S. (Ed.). (2008). *Unlocking assessment: Understanding for reflection and application*. Routledge.

Thompson, G., & Mockler, N. (2016). Principals of audit: Testing, data and 'implicated advocacy'. *Journal of Educational Administration and History, 48*(1), 1–18.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods, 45*(4), 1191–1207.
Williamson, B. (2016). Coding the biodigital child: The biopolitics and pedagogic strategies of educational data science. *Pedagogy, Culture & Society, 24*(3), 401–416. https://doi.org/10.1080/14681366.2016.1175499.
Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. Australian Educational Researcher, 40, 241–256. https://doi.org/10.1007/s13384-013-0089-9
Willis, J., & Cowie, B. (2014). Assessment as a generative dance. In Claire Wyatt-Smith, Val Klenwoski, & Peta Colbert (Eds.), *Designing assessment for quality learning* (pp. 23–37). Dordrecht: Springer.
Willis, J., Crosswell, L., Morrison, C., Gibson, A., & Ryan, M. (2017). Looking for leadership: The potential of dialogic reflexivity with rural early-career teachers. *Teachers and Teaching,* 1–16.
Willis, J., & Klenowski, V. (2018). Classroom assessment practices and teacher learning: An Australian perspective. In *Teacher learning with classroom assessment* (pp. 19–37). Singapore: Springer.
Wyatt-Smith, C., & Gunn, S. (2009). Towards theorising assessment as critical inquiry. In *Educational assessment in the 21st century* (pp. 83–102). Springer, Dordrecht.

Jill Willis is an Associate Professor in Education at Queensland University of Technology. She is interested in how learners of all ages can be supported to experience greater learning success and agency, especially in complex assessment systems. Jill researches in transdisciplinary partnerships to better understand how learners navigate performance expectations, and identify socio-cultural conditions that generate inclusive, mutual learning experiences.

Andrew Gibson is a Lecturer in the Information Systems School at Queensland University of Technology (QUT), Brisbane, Australia. Andrew's research centres on the interaction between people and technology, particularly in Learning Analytics. This work includes Transepistemic Abduction, Reflective Writing Analytics, and Peircian Pragmatism. Andrew created and maintains the open-source software GoingOK and TAP (Text Analytics Pipeline). Originally a secondary school music teacher, specialising in music technology, Andrew has worked in IT management, tech start-up management, and creative production management. Andrew's qualifications include a PhD from QUT, a Bachelor degree in Educational Studies, Postgraduate Diploma IT, and Diploma of Teaching.