7-1-2019

The development of a student engagement instrument for the responding strand in visual arts

Julia Morris
Edith Cowan University

10.1007/s13384-018-0296-5
This is a post-peer-review, pre-copyedit version of an article published as: Morris, J. E. (2019). The development of a student engagement instrument for the responding strand in visual arts. *The Australian Educational Researcher*, *46*(3), 449-468.
https://doi.org/10.1007/s13384-018-0296-5

The development of a student engagement instrument for the 'responding' strand in visual arts

Julia E. Morris
School of Education, Edith Cowan University

Abstract

The 'responding' strand of the Australian visual arts curriculum promotes 21st century learning skills through students' analytical engagement with artworks and artists. Assessing students' experiences and engagement is one strategy to improve teaching and learning in responding. However, there are no validated, subject-specific student engagement instruments for teachers to use. This study sought to develop a student self-report diagnostic instrument that provides information on past experiences with visual arts and factors affecting both cognitive and psychological engagement, with implications for improving teaching and learning. The instrument was piloted with 266 Year 10 to 12 students, as responding has an approximate 50% assessment weighting in the Australian visual arts course for senior school students. This paper reports on the exploratory and confirmatory factor analyses conducted in the development of the instrument. It adds to the body of knowledge on developing engagement instruments, recognising that student engagement in secondary education is context-dependent.

Keywords

Student engagement, diagnostic assessment, instrument development, responding, visual arts

Introduction

Contemporary education research and policy calls for students to develop as 21st century learners, as they begin to understand and apply skills in collaboration, critical thinking and creativity (Greiff et al. 2014; Saavedra and Opfer, 2012). The aim of promoting 21st century learning is that it will better place students to engage as active citizens in a global community (Commonwealth of Australia, 2017). Such 21st century learning skills are taught and assessed in visual arts education. While both the making (practical) and responding (analytical) strands of the Australian Curriculum: Visual Arts encompass 21st century learning skills, the responding strand particularly focuses on students' development of critical and creative thinking (School Curriculum and Standards Authority, 2014). In responding, students are taught how to decode artworks using principles of art, how to make meaning from an artwork based on their analysis and understanding of context, and how to encode this information in their own arts practice and interpretation (School Curriculum and Standards Authority, 2015). While not all states and territories have embraced the Australian Curriculum terminology of making and responding, and there has been contention about the curriculum's implementation, there are similar analytical tasks in art criticism and history across all jurisdictions. These broader analytical skills are not unique to the Australian curriculum, they have been common in visual arts curricula internationally since the 1970s (Caldwell and Vaughan 2011; Eisner 1987; Kim and Geahigan 2004; Macdonald 2005).

While the responding strand of the visual arts curriculum is congruent with 21st century learning skills literature, students cannot develop these skills if they are not engaged in visual arts education. While student engagement is not the only factor affecting their acquisition of these skills, it is a significant factor that impacts learning (Christenson, Reschly and Wylie 2012; Davis and McPartland 2012; Fredricks, Blumenfeld and Paris 2004; Pekrun and Linnenbrink-Garcia 2012; Sinatra, Heddy and Lombardi 2015). Consequently, it is important to measure student engagement so that teachers can monitor levels of engagement, and act when students are disengaged from learning within their own context (Appleton, Christenson and Furlong 2008; Appleton, Christenson, Kim and Reschly 2006; Fredricks et al. 2004; Sinatra et al. 2015). Students often disengage from responding to art as it can be isolated from their art making, and students report that it seems more like an analytical English lesson than visual arts (Author 2015). This research sought to address the issue of measuring student engagement in visual arts, specifically in the responding strand. It is uncommon to use quantitative methods in visual arts, and subsequently, the research also sought to explore validity of quantitative instruments within an arts-based subject. Lastly, the study aimed to create a diagnostic assessment instrument that could be used by teachers to collect meaningful data about students' prior learning and engagement, so that these data could be used for reporting and to inform teachers' practice. This paper reports on the validation of the diagnostic assessment itself; building on reported findings about the impact of non-school based arts experiences on student engagement in the *Australian Educational Researcher* (Author 2018).

Diagnostic assessment and data reporting

This research sought to explore the potential for quantitative measures to be introduced as a means of diagnostic assessment in visual arts. Creating a diagnostic assessment instrument was a key aim due to teachers' accountability to collect and report on student data (AITSL 2011; A. D. Gilbert 2016; Hardy 2015). Measuring student achievement can be linked to government funding, such as in the United States (A. D. Gilbert 2016; McDonnell 2012) or it can be used for public scrutiny, as seen on Australia's *MySchool* website (Cumming and Mawdesley 2013). Whatever the purpose may be, data collection and reporting are affecting the education community (Cumming and Mawdesley 2013; Hardy 2015) and consequently, the development and accessibility of valid assessment tools is of increasing importance for teachers.

Diagnostic assessment refers to assessment conducted to determine students' prior knowledge and/or attitudes about a particular topic or construct (Scaife and Wellington 2010; Van der Kleij, Vermeulen, Schildkamp and Eggen 2015). Diagnostic assessment often takes place in classrooms as an informal process, described by Kemp and Scaife (2012) as a 'moment-by-moment' process. However, this type of assessment has received criticism for not being used effectively in schools, in that the feedback from diagnostic assessment should be timely so that it may be used to effect appropriate changes in teaching instruction (Hattie and Brown 2011; Van der Kleij et al. 2015). In Australia the use of diagnostic assessment is inherent in the *National Professional Standards for Teachers*, which specify that teachers should be using a range of assessment strategies to adapt learning experiences to meet students' needs (AITSL 2011). Consequently, collecting evidence about diagnostic assessment processes becomes an important aspect of teaching practice in the need to formalise or supplement the 'moment-by-moment' process (Kemp and Scaife 2012).

Defining student engagement

While it is important to conduct diagnostic assessment about what students *know*, it is also important to measure how students *feel*, as affective responses to education impact on students' achievement and retention in schooling (Appleton et al. 2008; Gray and Hackling 2009; Mansour et al. 2016). Furthermore, engagement influences students' perceptions about the importance of learning itself (Christenson et al. 2012). Two overarching types of engagement, cognitive and psychological engagement, were explored in the development of the diagnostic assessment instrument. These types of engagement imply advanced knowledge acquisition, defined as the ability to apply knowledge to a range of problems or contexts, to make meaning about the world and to identify gaps in learning that need to be addressed (Efland 1990). This definition links closely to 21st century learning skills through developing global understanding and critical thinking, as well as reflexivity (Greiff et al. 2014).

Cognitive engagement is about students' engagement with a subject on a deeper level, because they are inherently interested in learning about the subject (Appleton et al. 2006; Author 2017, 2018). It was measured through three core indicators of autonomy, intrinsic motivation and metacognition. Autonomy is related to cognitive engagement as students internalise information and processes when they are engaged, as these processes become part of the students' identity (I. Gilbert 2013; Moller, Deci and Ryan 2006). Autonomous students make decisions about their learning because it intersects with their beliefs and interests (Dodge and Kaufman 2009; Winchmann 2011). Intrinsic motivation is students' desire to learn for knowledge, accomplishment and to stimulate the senses (Carbonneau, Vallerand and Lafrenière 2012) without the need for external reward (Ryan and Deci 2000). Motivation is often conceptualised as a continuum, whereby students can be extrinsically motivated or exhibit varying degrees of intrinsic motivation or learning (Ryan and Deci 2000; Vansteenkiste, Simons, Lens, Soenens and Matos 2005). An example of a student in the middle of the continuum is one who does not enjoy the task set by the teacher, but is motivated to complete the task as they desire high grades (Author 2015). A true intrinsically motivated student is one who is both interested in the task itself as well as the benefit of learning (Carbonneau et al. 2012; Ryan and Deci 2000). The final indicator of cognitive engagement explored is metacognition. Metacognition refers to the act of self-regulating cognitive behaviour (Wiley and Jee 2011). Three types of metacognition were considered: knowledge of self, task knowledge, and strategic knowledge (Tarricone 2011). Knowledge of self refers to a students' awareness of their goals, strengths and weaknesses (Proust 2010; Wiley and Jee 2011). Task knowledge refers to the students' awareness of task objectives and the procedure of completing the task to meet the objectives (Tarricone 2011). Strategic knowledge is an awareness of the methods to complete the task and adapting known methods based on the specific task (Tarricone 2011).

Psychological engagement is about the affective aspect of engagement, or how students feel about the subject (Christenson et al. 2012). It had two key indicators: self-efficacy and positive relationships. Self-efficacy is students' belief that they are achieving to the best of their ability, with a positive attitude that they can reach their goals despite potential challenges (Bandura 2012; Bandura and Locke 2003; Martin 2007). Self-efficacy can be diminished when students feel isolation or a sense of failure (Ainley 2012), and as such, self-efficacy is affected by prior achievements (Hattie 2009; Author 2014). Positive engagement was defined as supportive relationships with peers and the visual arts specialist teacher, specifically as relationships that supported the student to feel a sense of belonging in the class and to feel relatedness (Appleton et al. 2008; Gray and Hackling 2009). Relationships were particularly important as students who report higher interest in learning often report increased mutual respect between teacher and student (Dodge and Kaufman 2009; Gray and Hackling 2009).

These factors of engagement are frequently reported by the literature and integrated in assessment measures on student engagement in the school context more broadly (Appleton et al. 2008; Dulfer, Rice and Clarke 2017; Mazer 2012; Moreira and Dias 2018; Yonezawa, Jones and Joselowsky 2009). However, secondary students' engagement may vary from subject to subject, and holistic measures of engagement do not provide adequate information for teachers wanting to effect change (Author 2015). In measuring engagement both the individual and their context need to be considered (Sinatra et al. 2015). Consequently, subject-specific instruments are necessary to provide teachers with valid information that can be used to support teaching and learning.

Responding in the curriculum: Developing 21st century learning skills

The contemporary literature on 21st century learning skills provided the impetus for this research project. Students cannot be expected to develop these skills at school if they are not engaged in learning experiences that promote them. In visual arts and the arts learning area more broadly, critical thinking and communication are not new skills. A. D. Gilbert (2016) reflects this in her framework of 21st century learning skills, whereby she argues that the 4C's of creativity, critical thinking, communication and collaboration are inherent in music curricula. In visual arts specifically, the United States introduced art criticism into curricula in the 1970s, and particularly in the 1980s with the advent of Discipline Based Art Education (DBAE) (Eisner 1990; Greer 1987). DBAE advocated for students to be involved in studio practice as artists, but to also develop knowledge in art history and skills to critique art through critical discourse (Eisner 1990; Greer 1987). DBAE shaped education in many countries, including the United Kingdom and Australia's decision to introduce art criticism into curricula (Boughton 1989; McKeon 2002). Contemporary responding curricula have grown from DBAE and are similar across the three countries as well; with all curricula including a visual culture and global issues focus (Cherry 2004; Department of Education 2013; Smith-Shank 2008). In the Western Australian context, where this research was conducted, there are three main sections in the responding strand of visual arts. These sections, known as threads, are: analysis, social cultural and historical contexts, and interpretation/response (School Curriculum and Standards Authority 2015). Together they account for 50% of a student's final grade in visual arts. In analysis, students learn to decode artworks using the principles of art in order to make meaning from the artwork (School Curriculum and Standards Authority 2015). In social, cultural and historical contexts, students learn about art through history and the value or purpose of creating artworks within society (School Curriculum and Standards Authority 2015). In interpretation/response, students learn to make judgements about artworks and to give evidence to justify their interpretation (School Curriculum and Standards Authority 2015). Embedded in each of these sections is critical thinking to synthesise knowledge, to encode personal meaning and articulate their interpretations of visual artworks, all of which are essential aspects of 21st century learning (Greiff et al. 2014; Saavedra and Opfer 2012). The high assessment weighting for this strand of the curriculum highlights its value and the necessity for students to engage with this content in order to develop these essential skills.

Method

Sample

A power calculation indicated a sample of 263 students was required for a 95% confidence level in this study. The sample comprised 266 secondary students from metropolitan schools in Perth, Western Australia. These students were enrolled in Year 10 visual arts and Year 11 or Year 12 ATAR (Australian Tertiary Admissions Rank – university-pathway) visual arts. Students enrolled in these courses were purposively sampled as responding tasks have higher assessment weighting in their courses, with most Year 10 classes weighting responding at 40% of a students' grade, and Year 11 and 12 courses mandated to have 50% of a student's grade determined by their performance on responding tasks. Of the total 18 schools in the sample, six schools were government, seven were independent and five were Catholic to represent the spread of school sectors. Fifteen of the schools were co-educational and three schools were single sex. All schools had an Index of Community Socio-Educational Advantage (ICSEA) between 900 and 1200. ICSEA was used in developing the sample because:

> Research shows that there is a strong relationship between the educational advantage a student has, as measured by the parents' occupation and level of education completed, and their educational achievement … developed to enable fair and meaningful comparisons … of students in a given school with that of similar schools serving students with statistically similar backgrounds … (ACARA 2012, p. 2)

The inclusion of ICSEA alongside school sector was to ensure the sample represented a range of schools and students; however, it is important to acknowledge the final sample still tended towards the middle-upper ranges of ICSEA. Within the Year 10 sample, 24.7% identified as being male and 75.3% identified as female. These percentages were comparable to the Year 11 sample, in which 23.3% identified as male and 76.8% as female. Most of the students were 16 years old (56.3%), 29.1% were 15 years old, 10.2% were 17 years old, and 4.3% were 14 years old.

Measure

The instrument completed by the students was largely developed from the Student Engagement Instrument (Appleton et al. 2008; Appleton et al. 2006; Moreira and Dias 2018), which is a generalised measure of student engagement across four types of engagement: academic, behavioural, cognitive and psychological. The full version of the Student Engagement Instrument is comprised of 30 items measuring cognitive engagement (academic and cognitive scales) and 26 items related to psychological engagement (behavioural and psychological scales). Statistical analysis of this instrument found six underlying factors, three related to cognitive engagement and three to psychological engagement. The three cognitive factors were control and relevance of school work (α = .80), future aspirations and goals (α = .78) and extrinsic motivation (α = .72). The three factors affecting psychological engagement were teacher-student relationships (α = .88), peer support for learning (α = .82) and family support for learning (α = .76) (Appleton et al. 2006). These factors showed little to moderate inter-correlation, suggesting that each factor was assessing unique variance related to the students' engagement (Appleton et al. 2006).

The Student Engagement Instrument was explored as the basis for this study as its scales relate to higher order thinking and were consistent with the rationale for the responding strand in the Western Australian Curriculum: Visual Arts (School Curriculum and Standards Authority 2015). Furthermore, the Student Engagement Instrument has been re-examined to test for validity over large samples within the secondary school context (n = 35,900) (Betts, Appleton, Reschly, Christenson and Huebner 2010; Lovelace, Reschly, Appleton and Lutz 2014).

As the items in the Student Engagement Instrument relate to general engagement (e.g., 'At my school, teachers care about students'), these items were amended using key words and ideas that made them applicable to responding to visual arts. For example, 'School is important for achieving my future goals' (Appleton et al. 2006) became 'Studying visual arts theory will help me in the future' in the motivation subscale of cognitive engagement. In this example the concept of the future and the importance of learning to future aspirations is maintained in both items, but the context of the item has changed to be applicable to visual arts responding tasks. The new instrument had three cognitive engagement subscales:

autonomy, intrinsic motivation and metacognition. It had two psychological engagement subscales: positive relationships and self-efficacy. Each scale had at least four items with the exception of positive relationships, which had three. Students' responses to each item were measured on a five-point Likert scale (1 = strongly disagree; 5 = strongly agree).

In addition to the items about students' engagement, the instrument also had sections that elicited students' prior engagement with visual arts responding in primary and middle school (Years 7-9). Questions in these sections included an overview of how time was allocated in visual arts (balance of making and responding, mostly making or mostly responding) and art movements discussed by the teacher. A separate section asked students about their current personal engagement with visual arts outside of the school context, for example, how often students attend art exhibitions, if their families were interested in art and if they also practice visual arts after school. These questions were included to give the teacher information about the students' involvement with art beyond the context of their current schooling.

Procedure

The development of the instrument occurred in three phases, two pilot phases and the final study. This paper reports on the phase three findings as these represent the most recent analyses of the instrument; however, a description of phases one and two are provided as background information about the overarching study. In phase one the instrument was piloted with a class of Year 11 students for face validity, to determine if the students could understand the questions clearly and if they felt that the items were appropriate to measure their engagement in responding tasks. A group of three teachers also gave feedback on the instrument to determine its validity based on their responding tasks (i.e., suitability of terminology) and if the instrument would provide feedback useful to their reporting purposes.

Furthermore, two academic researchers gave feedback on the fidelity of the instrument in relation to the engagement construct and visual arts.

In phase two, the instrument was piloted with 137 Year 11 students. The instrument was delivered through an online questionnaire that was accessed on school issued devices (iPad or laptop) or on a university set of iPads where the school did not have a device per child. In this phase exploratory factor analyses were conducted to reduce the subscales and items down to the three cognitive and two psychological scales that appear in the instrument (Author 2015).

In phase three, the sample was extended to Year 10 and Year 12 students to validate the instrument across a larger sample. The aim of this phase was to conduct confirmatory factor analyses to ensure the appropriateness of the instrument in measuring students' engagement in visual arts responding. Also, the initial instrument used a six-point scale (1 = strongly disagree, 2 = disagree, 3 = slightly disagree … 6 = strongly agree), and students reported finding the six-point scale difficult to respond to. Consequently, the phase three study changed the measurement to a five-point scale and confirmatory factor analyses were used to determine if this impacted the measurement of the instrument.

Ethical clearance was sought from the university's Human Research Ethics Committee for each phase of the research, and research approval was sought from both the Department of Education and Catholic Education WA for each of the three phases. In phase three, school principals were emailed an invitation to participate and were recruited by return email. The school principals were provided with an information letter and consent document, which was returned to the researcher. After school approval was given, visual arts teachers received information and consent documents to read and return to the researcher. The teachers were then given information letters and consent documents to circulate to parents/guardians and students. In addition to parental consent, the students opted-in to the research by agreeing to participation on the first page of the online questionnaire. The primary researcher administered the instrument in person at each of the schools so that the students had an opportunity to ask questions about the research or the instrument while they were completing it. The data from the instrument were downloaded by the researcher onto a secure university server, and were aggregated to form an overall sample. Unique school data sets were maintained so that the researcher could send reports to the visual arts teachers based on the aggregate data from their school or class.

Statistical analysis

The phase three data set were opened in a SPSS 24 data file and examined for missing data. Firstly, data were examined for missing response patterns and none were found. When a case had more than 5% missing data in the engagement scales it was removed from the data set. Where the case had fewer than 5% missing, the median scores were used to replace the missing values. Of the possible 5586 missing values (21 items multiplied by 266 participants), 26 were replaced by the median value. The replaced values represent less than 1% of the data set.

Even though exploratory factor analyses had been conducted on the phase two data, they were repeated with the phase three data to ensure that changing from a six-point to five-point response scale had not affected the discrete subscales of engagement. Performing both exploratory and confirmatory factor analyses on the same data was important as it meant only methodological explanations would account for differences between the analyses, and excluded the potential for sampling factors to effect conclusions (van Prooijen and van der Kloot 2001). In the exploratory phase, a cutoff of 0.6 was used for the Kaiser-Meyer Olkin measure of sampling (Cohen, Manion and Morrison 2011). Principal axis factoring with oblique rotation (direct oblimin) was used as the subscales were likely to correlate (Field

2013). Scree plots and eigenvalues were used to determine the number of factors, and items that loaded less than .40 were removed (Field 2013). After the exploratory factor analysis, the phase three data were further explored with confirmatory factor analyses using SPSS AMOS version 24. It was decided that if the measurement models were a bad fit, the data could be re-examined using exploratory factor analyses. As the sample were quite small, both the Chi-square test and the Chi-square degrees of freedom ratio (χ^2/df) were used to determine good model fit. The researchers sought a χ^2/df of approximately 2 to suggest good model fit (Sun 2005). Other measures of good model fit included the Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) as these are robust even in smaller samples (Sun 2005). A cut-off of 0.90 was used for both CFI and TLI, with 0.95 indicating a better fit (Hu and Bentler 1998). In addition, the Root Mean Square Error of Approximation (RMSEA) and the Root Mean Square Residual (RMR) were used to assess model fit, with a cut-off of 0.80 and a better fit indicated by a value less than 0.50 (Hu and Bentler 1998, 1999). The cut-offs were used as a guideline only, as there are no 'golden rules' when determining model fit (Marsh, Hau and Wen 2004).

After determining model fit, the items in the measurement models were examined for reliability and validity. Reliability was measured through internal consistency using Cronbach's alpha coefficient (Punch 2009). An alpha over 0.70 was considered reliable due to the content of the items (Muijs 2011). Stability could not be determined as the instrument was only administered once to the students, and phase three findings could not be compared to phase two as the measurement scale changed between these iterations of the instrument development. Content and construct validity were considered in this study through ANOVAs with post hoc multiple comparisons. However, these validations have focused on instrument alignment with construct and further investigation is needed validity through falsification testing (Goldstein 2015), and including concurrent and predictive validity.

Results

Exploratory factor analysis

Both the cognitive and psychological engagement exploratory factor analyses confirmed the constructs identified in the phase two research. Three factors were extracted for cognitive engagement and two for psychological engagement. The Kaiser-Meyer-Olkin measure verified sampling adequacy, KMO = .838. The three cognitive engagement factors had eigenvalues over 1.0 (4.26, 1.31, 1.22 respectively), and the scree plot confirmed the extraction of these factors. Factor loadings for the retained items can be found in Table 1.

Table 1. Factor loadings for cognitive engagement factors.

Item number	Factor 1	Factor 2	Factor 3	Item text
Q13			.544	I view others' artworks to influence my own visual arts practice
Q14			-*.330*	I cannot make decisions about what visual artworks I view
Q15			.429	My teacher lets me view artworks that I am interested in
Q16			.439	I think it is important to study visual arts/artists
Q17			*.375*	I am responsible for my own learning in visual arts
Q18	.483			I like being challenged to make meaning from visual artworks
Q19	.504			I enjoy experiencing new artworks
Q20	.582			I like learning about history by studying visual arts/artists
Q21	.703			Studying visual arts will help me in the future
Q22	.630			I do not want to learn about visual artists
Q23		.490		When I see an artwork, I know what to do to understand its meaning
Q24		.682		When I see an artwork, I know what knowledge I will need in order to analyse it

Item number	Factor 1	2	3	Item text
Q25		.529		I know where to get the information I need to help me analyse artworks
Q26		.495		I can explain how different techniques influence the meaning we make from artworks

Factor 1 = intrinsic motivation, Factor 2 = metacognition, Factor 3 = autonomy.

The exploratory factor analysis for the cognitive engagement scale had two key changes to the pilot data, in that Question 14 (*I cannot make decisions about what visual artworks I view*) and Question 17 (*I am responsible for my own learning in visual arts*) had loadings less than the .40. Both of these items related to the autonomy subscale. As both items had content validity, it was decided to leave these items in for the initial confirmatory factor analysis and to remove them if they compromised model fit.

In measuring psychological engagement, two factors had eigenvalues over 1.0 (3.39 and 1.44 respectively), and the scree plot confirmed the extraction of these factors. Factor loadings for the retained psychological engagement items can be found in Table 2.

Table 2. Factor loadings for the psychological engagement scale.

Item number	Factor 1	2	Item text
Q27	-.634*		I give up when visual arts responding tasks become challenging (* reverse coded due to negative wording)
Q28	.506	.422	My friends encourage me to achieve to the best of my ability in responding tasks
Q30	.588		The skills I learn from studying visual arts responding help me in everyday life
Q31	.416		I believe I am achieving to the best of my ability in visual arts responding
Q35	.680		I feel like I belong in my visual arts class
Q36		.739	I enjoy visual arts because I like my teacher
Q37		.749	My teacher encourages me to achieve to the best of my ability in responding tasks

Factor 1 = self-efficacy, Factor 2 = positive relationships.

All of these items were retained for the confirmatory factor analyses, despite crossloading on Question 28. While there were limited number of items in the second factor, it was decided to test the psychological engagement model with the confirmatory factor analysis by testing a unidimensional model (self-efficacy factor only) and a multivariate model that included both factors.

Confirmatory factor analyses

The measurement models of the instrument were examined using confirmatory factor analyses on a random sample of half of the data, and then compared to the second half to ensure their integrity. The results reported are for the confirmatory factor analyses of the full data set (n = 266). These analyses were used to determine model fit, and to reduce the number of items until the most parsimonious model was found for both cognitive and psychological engagement. Consideration was given to ceiling and floor effects in determining the most parsimonious model, and subsequently the number of items in each scale were kept similar to ensure a more balanced metric (Dell-Kuster et al. 2014). It was noted that some subscales, such as the positive relationships subscale, would need further investigation due to its low number of items. Three measurement models were computed for cognitive engagement and two for psychological engagement.

The measurement models for cognitive engagement included:

- Model 1: Q13, Q14, Q15, Q16, Q17 (autonomy); Q18, Q19, Q20, Q21, Q22 (intrinsic motivation); Q23, Q24, Q25, Q26 (metacognition)
- Model 2: Q13, Q14, Q15, Q16, Q17 (autonomy); Q18, Q19, Q20, Q21 (intrinsic motivation); Q23, Q24, Q25, Q26 (metacognition)
- Model 3: Q13, Q15, Q16, Q17 (autonomy); Q18, Q19, Q20, Q21 (intrinsic motivation); Q23, Q24, Q25, Q26 (metacognition)

The measurement models for psychological engagement included:

- Model 1: Q27, Q28, Q30, Q31, Q35 (self-efficacy); Q36, Q37 (positive relationships)

- Model 2: Q27, Q30, Q31, Q35 (self-efficacy); Q28, Q36, Q37 (positive relationships)
- Model 3: Q27, Q28, Q30, Q31, Q35 (self-efficacy)

Table 3 shows the model fit statistics for both cognitive and psychological engagement scales.

Table 3. Model fit statistics for each measurement model.

Engagement construct	Model	χ^2	*df*	χ^2/df	CFI	TLI	RMR	RMSEA	$\Delta\chi^2$	Δdf	*p*
Cognitive	1	135.1	74	1.86	.915	.896	.044	.057			
	2	106.4	62	1.72	.936	.919	.037	.053	28.7	12	<.01
	3	97.8	51	1.92	.930	.910	.038	.060	8.6	11	
Psychological	1	21.7	13	1.67	.976	.961	.040	.052			
	2	51.9	13	3.99	.892	.825	.075	.140	*		
	3	2.90	2	1.45	.993	.980	.022	.042	24.6	11	<.05

*Psychological model 2 could not be compared as it was not nested.

Model 1 for cognitive engagement did not have adequate goodness-of-fit measurement (TLI >.90), and was rejected. Model 2 had good fit statistics (CFI and TLI >.91) and very good values for the RMR and RMSEA (<.055). As Model 2 included Questions 14 and 17, which had poor factor loadings on the EFA, a third model was tested with these items removed. Model 3 had goodness of fit, and although the RMSEA value was slightly higher than model 2, it was still within the acceptable .80 cutoff (Hu and Bentler 1999). Differences in the χ^2 tests were consulted to determine the best model. While Model 2 returned a significant result ($\Delta\chi^2 = 28.7$, $\Delta df = 12$, $p < .01$) there was no significant difference between Models 3 and 2. Therefore, Model 2 was deemed to be the most parsimonious fit for the cognitive engagement scale at this stage in the analysis.

Psychological engagement was initially tested as a multidimensional scale. As question 28 crossloaded on both subscales, each model was tested with that item on a different factor. The fit statistics were good on Model 1, which had Item 28 loaded on self-efficacy. The second model moved the item to positive relationships; however, this model did not have adequate goodness-of-fit (CFI and TLI <.90) and the RMSEA was also higher than the .80 cutoff. Model 3 was a unidimensional model using only the self-efficacy items. It had a good fit (CFI and TLI >.98, RMR and RMSEA <.05) but its relative chi-square was quite low at 1.45, when it is preferable to be closer to 2 (Carmines and McIver 1981). As Models 1 and 3 were nested, a difference in χ^2 test was consulted to determine the best model. The unidimensional model returned a significant result at the .05 level ($\Delta\chi^2 = 24.6$, $\Delta df = 11$, $p < .05$).

Reliability and validity

After the measurement models were established, the scales were examined for reliability and validity. Internal consistency was computed using Cronbach's alpha coefficient. As the three scales of cognitive engagement were related, an alpha was computed for each subscale as well as the overarching construct. The alpha was lowest for the autonomy subscale ($\alpha = .61$), while metacognition ($\alpha = .68$) and intrinsic motivation ($\alpha = .72$) were slightly higher. As Items 14 and 17 had lower factor loadings, the alpha was retested with these factors removed. If Item 14 was removed the alpha dropped to .59 and with Item 17 removed $\alpha = .58$; therefore, these items were retained in the final subscale. The overall alpha for cognitive engagement was more acceptable, $\alpha = .82$. As the unidimensional model for psychological engagement was the best fit only one alpha was computed, returning a reasonable result $\alpha = .71$. The data could not be tested for stability as only there was only one occasion of testing.

Content validity relates to the instrument containing sufficient items to measure the latent trait (Cohen et al. 2011). While engagement is a very complex construct, Punch (2009) suggests that content validity can be improved by having a clear definition of the construct and then developing indicators that relate to each part of the definition. The wording for the initial items was developed using the Student Engagement Instrument (Appleton et al. 2006) and its validation over time (Appleton et al. 2008; Betts et al. 2010; Lovelace et al. 2014). A systematic literature review also examined indicators of engagement from a variety of sources (see Author 2015) to create robust definitions from which the items stemmed. In addition, the instrument was piloted with senior school visual arts students (the target audience), as well as visual arts teachers and academic researchers prior to data collection.

Construct validity is the comparison of the instrument's measurement against findings related to the construct (Creswell 2014; Punch 2009). As validated instruments measuring engagement in visual arts were not found, one-way ANOVA tests were computed for each of the scales. These tests were conducted to explore if the scales were adequately measuring the nature of the construct they were intended to measure. The independent variables used in these analyses were computed using the aggregate of the overall scale, which was then scaled into five groups (i.e., so the score for cognitive engagement overall was scaled to match Likert responses on individual items or subscales). It was anticipated that participants with lower scores on each subscale should have a lower score for the overall scale as well. In conducting these analyses, the null hypothesis was that there would be no difference in the mean score of each subscale in comparison to the overall score. For the cognitive engagement scale, it was probable that Levene's test would return a significant result as there was not true independence between groups; therefore, Welch's F is reported for autonomy and intrinsic motivation in Table 4. As psychological engagement was a unidimensional scale, it was compared to an item measuring self-efficacy that was removed in the exploratory factor analysis.

Table 4. One-way ANOVA results showing significance (n = 266).

Dependent variable	Independent variable	df_M	df_R	*F*	*p*
Autonomy	Cognitive engagement (scaled)	2	40	39.75	<.001
Intrinsic motivation	Cognitive engagement (scaled)	2	86	223.70	<.001
Metacognition	Cognitive engagement (scaled)	2	244	56.32	<.001
Self-efficacy	Item 33: I enjoy talking about visual art I have created	4	*242*	*4.84*	<001

As the ANOVA results suggested significance, post hoc multiple comparisons were computed to investigate where the significant differences in the group means lie. Within this study the Tukey HSD multiple comparison approach was used due to its moderately conservative nature, although Type I error cannot be ruled out due to the unequal samples sizes. Results from the Tukey homogenous subset comparisons for the cognitive engagement subscales are listed in Tables 5, 6 and 7.

Table 5. Tukey HSD multiple comparison results for autonomy and the independent variable (n = 266).

		Subset for alpha = .05		
Cognitive Engagement	*N*	*1*	*2*	*3*
Neither agree nor disagree	36	2.69		
Agree	170		3.42	
Strongly agree	41			4.00
Significance		*1.00*	*1.00*	*1.00*

Table 6. Tukey HSD multiple comparison results for intrinsic motivation and the independent variable (n = 266).

		Subset for alpha = .05		
Cognitive Engagement	*N*	*1*	*2*	*3*
Neither agree nor disagree	36	2.36		
Agree	170		3.25	
Strongly agree	41			3.98
Significance		*1.00*	*1.00*	*1.00*

Table 7. Tukey HSD multiple comparison results for metacognition and the independent variable (n = 266).

		Subset for alpha = .05		
Cognitive Engagement	*N*	*1*	*2*	*3*
Neither agree nor disagree	36	2.56		
Agree	170		3.08	
Strongly agree	41			3.76
Significance		*1.00*	*1.00*	*1.00*

The cognitive engagement subscales showed that mean scores for each response set on the independent variables steadily decline from the strongly agree to the strongly disagree categories, with each response set mean being significantly different from the others. However, the post hoc comparison for self-efficacy was not highly significant as there were not clearly homogenised groups, as shown in Table 8.

Table 8. Tukey HSD multiple comparison results for self-efficacy and the independent variable (n = 266).

Item 33: I enjoy talking about visual art I have created	*N*	*Subset for alpha = .05* 1	2	3
Strongly disagree	16	3.44		
Disagree	31	3.45		
Neither agree nor disagree	83	3.66	3.66	
Agree	87		3.96	3.96
Strongly agree	30			4.13
Significance		*.68*	*.39*	*.86*

In addition, correlations between the cognitive and psychological engagement scales were computed. There was moderate correlation between the two scales, $r = 5.22$, $p < .001$, which was anticipated as both scales measure latent traits of the broader engagement construct. Table 9 shows the correlations between the subscales.

Table 9. Correlations between survey subscales (n = 266)

	Autonomy	Intrinsic motivation	Metacognition	Self-efficacy
Autonomy	1			
Intrinsic motivation	.56*	1		
Metacognition	.48*	.41*	1	
Self-efficacy	.43*	.29*	.54*	1

* $p < .001$

All correlations were small-moderate, with the exception metacognition and self-efficacy which was lower ($r = .29$, $p < .001$). This was anticipated as self-efficacy theory is related to an individual's perceived autonomy and motivation (Bandura 2012; Bandura and Locke 2003; Ryan and Deci 2006) but is not linked to metacognitive processes (Tarricone 2011).

Discussion

Measurement of student engagement is an ongoing area of interest, as researchers work to develop instruments that can help support students' learning and increase teacher impact (Appleton et al. 2006; Lovelace et al. 2014; Moreira and Dias 2018). This research sought to explore the development of a visual arts responding engagement instrument, which could measure students' engagement in a context-specific way that accounted for the type of learning that occurs within the visual arts environment. The visual arts instrument was initially modelled on the Student Engagement Instrument (Appleton et al. 2006), but adapted to meet visual arts responding content. Consequently, definitions for the constructs included were developed from a review of the literature prior to two phases of piloting. This paper reported on the third phase of the research, to perform confirmatory factor analyses on the data set of WA senior school student responses (n = 266).

The confirmatory factor analyses showed good model fit for cognitive engagement, with all three subscales correlating. However, the autonomy subscale requires further investigation due to the low factor loadings of two items (14 and 17) on the exploratory factor analysis and the low Cronbach alpha coefficient ($\alpha = .61$). Although it could be preferable to remove these items from the subscale, it would increase the chance of ceiling effects as the subscale would only retain three items (Dell-Kuster et al. 2014). Furthermore, removal of Item 14 on Model 3 did not return a statistically significant result in the difference in χ^2 test, and therefore, did not produce the most parsimonious model for cognitive engagement. While the construct validity testing supported each subscale, the use of the aggregate cognitive engagement scale was not truly independent and this could inflate significance. However, the aggregate scale was deemed appropriate as it was anticipated that students' scores would vary across each subscale.

The psychological model was more complex in that there was crossloading of item 28 on both self-efficacy and positive relationships, as well as only three items loading on the positive relationships subscale. The confirmatory factor analyses suggested a unidimensional model including only self-efficacy items was the best fit; however, this is not consistent with the literature. Research suggests that a range of factors, including teacher and peer interaction, and prior achievements all impact on students' psychological engagement (Appleton et al. 2008; Christenson et al. 2012; Gray and Hackling 2009; Hattie 2009; Author 2014). In addition, the ANOVA with post hoc Tukey HSD multiple comparisons suggested that the self-efficacy scale does not produce homogenous groups. This could be anticipated due to the choice of independent variable, 'I enjoy talking about visual art I have created'. It was hypothesised that students who had higher self-efficacy would be more confident in their ability to talk about their work, even if they found it challenging. However, this assumption does not account for alternate reasons for a students' response to this item; for example, a student may be shy and dislike speaking in front of others, but still have high self-efficacy to complete responding tasks in written or alternative forms.

Engagement is a complex construct to measure and research is suggesting that any measurement of engagement needs to consider the individual as well as the context (Sinatra et al. 2015). The visual arts responding engagement instrument included additional background questions about students' current personal experiences and prior school experiences in relation to responding. As such, the instrument was designed to consider both person and context in providing practical information for teachers, which is important in any educational assessment due to the level of accountability teachers face around collecting, acting and reporting on student data (AITSL 2011; A. D. Gilbert 2016; Hardy 2015). However, these background data are only useful if they can give context to students' levels of engagement, and consequently, developing a valid and reliable subject-engagement assessment instrument is critical to informing practice. Specifically, this instrument aimed to provide information that would inform teachers' 'moment-by-moment' decisions (Kemp and Scaife 2012), by highlighting factors that could impede students' engagement in responding tasks. However, the impact of this type of instrument in teachers' practice is yet to be measured and its replicability across jurisdictions could vary depending on localised art contexts, including how curriculum is implemented by teachers. Engaging students in responding is not only important to facilitate their development of 21st century learning skills, but also because responding is weighted as 50% of the students' final visual arts grade (School Curriculum and Standards Authority 2015).

Conclusion

This research sought to explore the issue of measuring student engagement in visual arts, specifically in the responding strand. With the promotion of 21st century learning skills teachers require mechanisms to monitor their students' engagement in learning, as engagement (or disengagement) can affect students' acquisition of these essential skills (A. D. Gilbert 2016; Saavedra and Opfer 2012). This research aimed to develop an instrument that could be used to diagnostically assess student engagement within a subject-specific context. However, Stobart (2004) notes that developing a valid and reliable multi-dimensional instrument that has implications for practice is a complex task. While this instrument is in its beginning stages of development, it is evident that more investigation is required to ensure it is a robust measurement of students' cognitive and psychological engagement. In particular, sourcing and testing the instrument against new independent variables would improve construct validity and provide greater rigour to validity where it has largely depended on subjective measures (Goldstein 2015). The length of the scales also needs investigation to counteract floor or ceiling effects due to instrument length (Dell-Kuster et al. 2014). Furthermore, the scales could be repeated with a larger sample to test the instrument's stability over time. While the instrument has potential for replication beyond the Western Australian context due to similarities in international arts curricula (Cherry 2004; Department of Education 2013; Smith-Shank 2008), its items were created from the state-based curricula. Consequently, replication would need to consider if the items have content validity for other jurisdictions.

References

ACARA. (2012). *A guide to understanding ICSEA*. Retrieved from Sydney, Australia: http://www.acara.edu.au/verve/_resources/Guide_to_understanding_ICSEA.pdf

Ainley, M. (2012). Students' interest and engagement in classroom activities. In S. L. Christenson, A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement*. New York, NY: Springer.

AITSL. (2011). The National Professional Standards for Teachers. Melbourne, Australia: Ministerial Council for Education, Early Childhood Development and Youth Affairs (MCEECDYA).

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the schools, 45*(5), 369-386.

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology, 44*(5), 427-445.

Author. (2015). Removed for peer review.

Author. (2014). Removed for peer review.

Author. (2017). Removed for peer review.

Author. (2018). Removed for peer review.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*(9), 9-44.

Bandura, A., & Locke, E. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*(1), 87-99.

Betts, J. E., Appleton, J. J., Reschly, A. L., Christenson, S. L., & Huebner, E. S. (2010). A study of the factorial invariance of the student engagement instrument (SEI): Results from middle and high school students. *School Psychology Quarterly, 25*(2), 84-93.

Boughton, D. (1989). The changing face of Australian art education: New horizons or sub-colonial politics? *Studies in Art Education, 30*(4), 197-211.
Caldwell, B., & Vaughan, T. (2011). *Transforming education through the arts*. Oxford, England: Routledge.
Carbonneau, N., Vallerand, R. J., & Lafrenière, M. K. (2012). Toward a tripartite model of intrinsic motivation. *Journal of Personality, 80*(5), 1147-1178.
Carmines, E. G., & McIver, J. P. (1981). Analyzing models with unobserved variables. In G. W. Bohrnstedt & E. F. Brogatta (Eds.), *Social measurement: Current issues*. Beverly Hills, L.A.: Sage Publications.
Cherry, D. (2004). Art history visual culture. *Art History, 27*(4), 479-493.
Christenson, S. L., Reschly, A., & Wylie, C. (Eds.). (2012). *Handbook of research on student engagment*. New York, NY: Springer.
Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). New York, NY: Routledge.
Commonwealth of Australia. (2017). National Innovation and Science Agenda. Retrieved from https://www.innovation.gov.au/page/national-innovation-and-science-agenda-report
Creswell, J. (2014). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (4th ed.). Essex, England: Pearson.
Cumming, J., & Mawdesley, R. (2013). Australia, quality education and the 'best interests of the child'. *Australian Journal of Education, 57*(3), 292-309.
Davis, M. H., & McPartland, J. M. (2012). High school reform and student engagement. In S. L. Christenson, A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement*. New York, NY: Springer.
Dell-Kuster, S., Sanjuan, E., Todorov, A., Weber, H., Heberer, M., & Rosenthal, R. (2014). Designing questionnaires: Healthcare survey to compare two different response scales. *BMC Medical Research Methodology, 14*, 96-109.
Department of Education. (2013). Statutory guidance: National curriculum in England: Art and design programmes of study. Retrieved from https://www.gov.uk/government/publications/national-curriculum-in-england-art-and-design-programmes-of-study/national-curriculum-in-england-art-and-design-programmes-of-study
Dodge, T., & Kaufman, A. (2009). Student perceptions and motivation in the classroom: exploring relatedness and value. *Social Psychology of Education, 12*(1), 101-112.
Dulfer, N., Rice, S., & Clarke, K. (2017). Student engagement, non-completion and pedagogy: Development of a measurement tool. *Australian Journal of Education, 61*(1), 40-53. doi:10.1177/0004944116685621
Efland, A. (1990). *A history of art education: Intellectual and social currents in teaching the visual arts*. New York, NY: Teachers College Press.
Eisner, E. (1987). The role of disciplined-based art education in America's schools. *Art Education, 40*(5), 6-26 & 43-45.
Eisner, E. (1990). Discipline-based art education: Conceptions and misconceptions. *Educational Theory, 40*(4), 423-430.
Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). London, England: Sage Publications.
Fredricks, J., Blumenfeld, P., & Paris, A. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.
Gilbert, A. D. (2016). The Framework for 21st Century Learning: A First-Rate Foundation for Music Education Assessment and Teacher Evaluation. *Arts Education Policy Review, 117*(1), 13-18.

Gilbert, I. (2013). *Essential motivation in the classroom*. New York, NY: Routledge.
Goldstein, H. (2015). Validity, science and educational measurement. *Assessment in Education: Principles, Policy & Practice, 22*(2), 193-201. doi:10.1080/0969594X.2015.1015402
Gray, J., & Hackling, M. (2009). Wellbeing and retention: A senior secondary student perspective. *The Australian Educational Researcher*, *36*(2), 119-145.
Greer, D. W. (1987). A structure of discipline concepts for DBAE. *Studies in Art Education, 28*(4), 227-233.
Greiff, S., Wüstenberg, S., Csapó, B., Demetriou, A., Hautamäki, J., Graesser, A. C., & Martin, R. (2014). Domain-general problem solving skills and education in the 21st century. *Educational Research Review, 13*, 74-83.
Hardy, I. (2015). Data, numbers and accountability: The complexity, nature and effects of data use in schools. *British Journal of Educational Studies, 63*(4), 467-486. doi:10.1080/00071005.2015.1066489
Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Oxon, England: Routledge.
Hattie, J. A., & Brown, G. T. L. (2011). Assessment and evaluation. In C. M. Rubie-Davies (Ed.), *Educational psychology: Concepts, research and challenges*. New York, NY: Routledge.
Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model specification. *Psychological Methods, 3*(4), 424-453.
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. doi:10.1080/ 10705519909540118
Kemp, S., & Scaife, J. (2012). Misunderstood and neglected? Diagnostic and formative assessment practices of lecturers. *Journal of Education for Teaching: International Research and Pedagogy, 38*(2), 181-192.
Kim, C., & Geahigan, G. (2004). South Korean art education during the third and fourth republics, 1960-1979: Economic development and nationalistic aspirations. *Journal of Cultural Research in Art Education, 22*, 69-85.
Lovelace, M. D., Reschly, A. L., Appleton, J. J., & Lutz, M. E. (2014). Concurrent and predictive validity of the student engagement instrument. *Journal of Psychoeducational Assessment, 32*(6), 509-520. doi:10.1177/0734282914527548
Macdonald, S. (2005). *A Century of Art and Design Education: From arts and crafts to conceptual art*. Cambridge, England: Lutterworth Press.
Mansour, M., Martin, A. J., Anderson, M., Gibson, R., Liem, G. A. D., & Sudmalis, D. (2016). Student, home, and school socio-demographic factors: Links to school, home, and community arts participation. *The Australian Educational Researcher, 43*, 221-244. doi:10.1007/s13384-015-0199-7
Marsh, H. W., Hau, K., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal, 11*(3), 320-341. doi:10.1207/s153280 07sem1103_2
Martin, A. (2007). The relationship between teachers' perceptions of student motivation and engagement and teachers' enjoyment of and confidence in teaching. *Asia-Pacific Journal of Teacher Education, 34*(1), 73-93.
Mazer, J. P. (2012). Development and validation of the student interest and engagement scales. *Communication Methods and Measures, 6*(2), 99-125.
McDonnell, L. M. (2012). Educational accountability and policy feedback. *Educational Policy, 27*(2), 170-189.

McKeon, P. (2002). The sense of art history in art education. *Journal of Aesthetic Education, 36*(2), 98-109. doi:10.2307/3333760

Moller, A. C., Deci, E. L., & Ryan, R. M. (2006). Choice and ego-depletion: The moderating role of autonomy. *Personality and Social Psychology Bulletin, 32*(8), 1024-1036.

Moreira, P. A. S., & Dias, M. A. (2018). Tests of factorial structure and measurement invariance for the Student Engagement Instrument: Evidence from middle and high school students. *International Journal of School & Educational Psychology*, 1-13. doi:10.1080/21683603.2017.1414004

Muijs, D. (2011). *Doing quantitative research in education with SPSS* (2nd ed.). London, England: Sage Publications.

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In S. L. Christenson, A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement*. New York, NY: Springer.

Proust, J. (2010). Metacognition. *Philosophy Compass, 5*(11), 989-988. doi:10.1111/j.1747-9991.2010.00340.x

Punch, K. (2009). *Introduction to research methods in education*. London, England: Sage Publications.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54-67. doi:10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2006). Self-regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will? *Journal of Personality, 74*(6), 1557-1586. doi:10.1111/j.1467-6494.2006.00420.x

Saavedra, A. R., & Opfer, V. D. (2012). Learning 21st-century skills requires 21st-century teaching. *Phi Delta Kappan, 94*(2), 8.

Scaife, J., & Wellington, J. (2010). Varying perspectives and practices in formative and diagnostic assessment: A case study. *Journal of Education for Teaching: International Research and Pedagogy, 36*(2), 137-151.

School Curriculum and Standards Authority. (2014). *Visual arts ATAR course: Year 12 syllabus*. Perth, Australia: School Curriculum and Standards Authority.

School Curriculum and Standards Authority. (2015). *Visual arts ATAR course: Year 12 syllabus*. Perth, Australia: School Curriculum and Standards Authority.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The Challenges of Defining and Measuring Student Engagement in Science. *Educational Psychologist, 50*(1), 1-13.

Smith-Shank, D. (2008). Visual culture and issues-based curricula. *Australian Art Education, 31*(1), 6-15.

Stobart, G. (2004). Developing and improving assessment instruments. *Assessment in Education: Principles, Policy & Practice, 11*(3), 243-245. doi:10.1080/0969594042000304573

Sun, J. (2005). Assessing goodness of fit in confirmatory factor analysis. *Measurement and Evaluation in Counseling and Development, 37*(4), 240-256.

Tarricone, P. (2011). *The taxonomy of metacognition*. New York, NY: Psychology Press.

Van der Kleij, F. M., Vermeulen, J. A., Schildkamp, K., & Eggen, T. J. H. M. (2015). Integrating data-based decision making, Assessment for Learning and diagnostic testing in formative assessment. *Assessment in Education: Principles, Policy & Practice, 22*(3), 324-343. doi:10.1080/0969594X.2014.999024

van Prooijen, J. W., & van der Kloot, W. A. (2001). Confirmatory Analysis of Exploratively Obtained Factor Structures. *Educational and Psychological Measurement, 61*(5), 777-792. doi:10.1177/00131640121971518

Vansteenkiste, M., Simons, J., Lens, W., Soenens, B., & Matos, L. (2005). Examining the motivational impact of intrinsic versus extrinsic goal framing and autonomy-supportive versus internally controlling communication style on early adolescents' academic achievement. *Child Development,, 76*(2), 493-501.

Wiley, J., & Jee, B. D. (2011). Cognition: Overview and recent trends. In V. G. Aukrust (Ed.), *Learning and cognition in education* (pp. 3-8). Oxford, England: Academic Press.

Winchmann, S. S. (2011). Self-determination theory: The importance of autonomy to well-being across cultures. *Journal of Humanistic Counseling, 50*(1), 16-26.

Yonezawa, S., Jones, M., & Joselowsky, F. (2009). Youth engagement in high schools: Developing a mulitdimensional, critical apporach to improving engagement for all students. *Journal of Educational Change, 10*(2), 191-120.