



European Journal of Educational Research

Volume 9, Issue 3, 1177 - 1187.

ISSN: 2165-8714

<http://www.eu-jer.com/>

Ethnic Differences in Students' Attitudes to the Arts: Providing Validity Evidence to Make Comparisons

Mustafa Asil*

University of Aberdeen, UNITED
KINGDOM

Alison Gilmore

University of Otago, NEW ZEALAND

Haci Bayram Yilmaz

Alanya Alaadin Keykubat University,
TURKEY

Received: April 7, 2020 ▪ Revised: June 1, 2020 ▪ Accepted: July 2, 2020

Abstract: Previous research suggests that non-cognitive factors play an important role in promoting success at school and beyond, aligning with the multifaceted goals of education. Enhancing students' attitudes to learning in school is expected to have positive impacts on various schooling outcomes. To date, very few studies have focused on measuring and understanding students' attitude to the arts. This study aims to address a gap in current research in this area by introducing instruments designed to measure attitude to dance, drama, music and visual arts. Confirmatory factor analysis and measurement invariance analyses are employed to examine the factorial validity and measurement equivalence of the scales of attitude to the arts disciplines for different ethnic groups in New Zealand. Findings support the utility of the scales as valid measures of attitude to dance, drama, music and visual arts. Noticeable differences are reported among New Zealand European, Maori, Pasifika and Asian students regarding their attitudes to dance, drama, music and visual arts.

Keywords: *Attitude to arts, measurement invariance, ethnicity.*

To cite this article: Asil, M., Gilmore, A., & Yilmaz, H. B. (2020). Ethnic differences in students' attitudes to the arts: Providing validity evidence to make comparisons. *European Journal of Educational Research*, 9(3), 1177-1187. <https://doi.org/10.12973/eu-jer.9.3.1177>

Introduction

Mounting evidence from research suggests that non-cognitive factors play an important role in promoting success at school and beyond, aligning with the multifaceted goals of education (Bertling et al., 2016). Attitude to a learning area is one such factor (Lipnevich et al., 2013) that has received growing recognition by educators and researchers in its ability to influence life outcomes and predict students' academic success (Popham, 2005).

Enhancing students' attitudes to learning in school is expected to have positive impacts on student achievement. However, investigations of relationships between attitudes and achievement present complex and sometimes conflicting patterns. The relationships are mostly positive, statistically significant but weak. The magnitude of the relationship may vary for specific groups, such as gender and ethnicity, for different learning areas or cultural contexts, or for how attitudes are measured and their psychometric quality and validity for different groups (Aiken, 1970, 1976; Hansford & Hattie, 1982; Logan & Johnston, 2009; Ma & Kishor, 1997; Silverman & Subramaniam, 2000; Steinkamp & Maehr, 1983; Weinburg, 1995). Berry et al. (2002) argue that cultural factors may have strong impacts on affective measures, such as attitudes, threatening the generalizability of psychological constructs across groups.

The National Monitoring Study of Student Achievement (NMSSA) is designed to assess and understand student achievement across the New Zealand Curriculum (NZC). As well as gathering student achievement data, the study collects contextual information about attitudes, learning experiences and learning environments, from students, teachers and principals. As indicated in their reports, NMSSA is designed to recognize and take account of the identity, language and culture of a multi-cultural population.

Although student success is generally measured by narrow standardized tests and the curricula are focused on the disciplines that are measured through those tests in most countries, research suggests positive effects of artistic engagement in cognitive development, and social and emotional well-being (Caldwell & Vaughan, 2011). For example,

* **Corresponding author:**

Mustafa Asil, School of Medicine, Medical Sciences and Nutrition, Polwarth Building, Foresterhill, Aberdeen AB25 2ZD, UK.

✉ mustafa.asil@abdn.ac.uk

Werner (2001) showed that using dance to teach math resulted in a significant difference in students' math scores and motivation to learn. Skinner and Pitzer (2012) state that arts-integrated lessons provide high levels of academic engagement. Furthermore, artistic behaviors are considered tools for communicating and expressing ideas and feelings, and for creating and maintaining social bonds among young students (Blatt-Gross, 2013), thus helping them to improve team-work skills and become active members of society.

The number of studies on students' attitudes towards the arts is not only limited in the literature but the approaches to measuring attitudes to the arts also vary. For example, Smithrim and Uptis (2005) used two item stems (i.e., "I like..." and "I would like to do more...") for music, drama and dance; these comprised one factor. The National Assessment Education Progress (NAEP) uses various numbers of items to survey eighth grade students' attitudes towards music and visual arts; however, results are examined in terms of individual items rather than scales. Xu et al. (2018) use partial least squares structural equation modelling to predict student achievement performance. The formative measurement models used in their study showed that three of the seven items about students' attitudes and behaviors (i.e., personal preferences) were not significant predictors of achievement. Pavlou and Kambouri's (2007) research measured elementary school students' attitudes to art education through a scale including 34 items in four dimensions: enjoyment, confidence, usefulness and support needed. They found high intercorrelations (0.81 to 0.90) among enjoyment, confidence and usefulness, and the comparative fit index (CFI) and goodness of fit index (GFI) were below the threshold of 0.90, which indicates a possible lack of fit. Considering the lack of available literature, NMSSA sought to develop items to create new scales to measure students' attitudes to the arts as part of its assessment program.

The use of the same measurement instrument does not ensure comparability of the results across groups, as the items or statements may have different meanings for students from different ethnic, language or cultural subgroups. Therefore, it is important that NMSSA provides evidence that the measurement of a non-cognitive construct like attitude is equally valid for all the subgroups that will be compared. NMSSA uses item response theory (IRT) to investigate the quality and validity of cognitive and non-cognitive scales, including differential item functioning (DIF, item-level measurement invariance (MI)) to examine whether the items are equivalent across subgroups. However, it is also important to examine MI at the scale level to see whether the accumulation of item-level DIFs yields non-invariance at the scale level. Measurement non-invariance at the scale level would suggest that the meaning of the construct being measured may not be the same for subgroups that are compared.

Thus, the goal of this study was to investigate the extent to which NMSSA scales of attitude to the arts were valid measurements for use across different ethnic groups.

Importance of Testing Measurement Invariance

To make meaningful and valid inferences, measurement instruments need to be interpreted in a similar way and responded to using the same reference framework across subgroups. Therefore, establishing MI across subgroups is regarded as a prerequisite before making comparisons between those subgroups (Bauer, 2017; Ercikan & Lyons-Thomas, 2013). Quite recently, the International Test Commission (ITC) published guidelines for the large-scale assessment of linguistically diverse populations (ITC, 2017). The commission underlined the necessity and importance of testing MI across subgroups by introducing Guideline 2.12, which requires evidence to:

Evaluate the invariance of the internal factor structure of the assessment across the L1 (refers to the home/first language (s) of a test taker) and L2 (refers to the second language or foreign language of a test taker) populations. (p. 14)

Similarly, the second edition of the ITC guidelines for translating and adapting tests (ITC, 2016) advises test developers to "only compare scores across populations when the level of invariance has been established on the scale on which scores are reported" (Guideline: SSI-2 (16), p. 17). MI across subgroups should not simply be assumed or taken for granted but should be conceptually and statistically supported before making comparisons between subgroup means (Asil & Brown, 2016; Van de Vijver & He, 2016).

MI is defined as "the equivalence of measured constructs in two or more independent groups to assure that the same constructs are being assessed in each group" (Chen et al., 2005, p. 472). In other words, it evaluates whether there is any differential scale functioning for different groups. Do all groups interpret and understand the construct in the same way?

MI is generally tested with a multi-group confirmatory factor analytic approach (MG-CFA) with nested model comparisons. MI is accepted if the differences between nested models are small. There are four levels of invariance that are sequentially tested by putting additional constraints on the item parameters; configural invariance, weak (metric) invariance, strong (scalar) invariance, and strict (residual) invariance (Meredith, 1993).

Configural invariance is the least stringent and requires similar item-factor patterns in each group (Horn & McArdle, 1992) while allowing factor loadings to differ. Metric invariance implies that the unit of measurement (factor loadings) is the same across groups. Even though metric invariance is achieved, comparing factor-latent mean differences is still not possible as the origin of the scales may vary (Chen et al., 2005). Scalar invariance refers to the equivalence of

intercepts (scale origins) across groups. Achieving scalar invariance indicates that differences on latent factor means can be attributed to group characteristics rather than to measurement artefacts (Brown et al., 2017). There seems to be a consensus in the literature that scalar invariance is sufficient to compare latent means meaningfully across subgroups (Schmitt & Kuljanin, 2008).

To establish the invariance of the construct of attitudes to the arts across different ethnic groups, this study applied confirmatory factor analysis (CFA). This procedure has been adopted by large-scale international surveys, such as the Programme for International Student Assessment (PISA), the Teaching and Learning International Survey (TALIS), and the Programme for the International Assessment of Adult Competences (PIAAC) (Van de Vijver et al., 2019).

National Monitoring Study of Student Achievement

NMSSA assesses students' achievement and attitudes across all learning areas of the NZC in a five-year cycle (NMSSA, 2017a). NMSSA is a collaboration between the Educational Assessment Research Unit (EARU) of the University of Otago, and the New Zealand Council for Educational Research (NZCER).

The purpose of NMSSA is to get a broad picture of student achievement in New Zealand (NZ) state and state-integrated primary schools and to collect contextual information to understand factors related to student achievement. This includes, but is not limited to attitude, engagement and opportunities to learn in the learning areas of the curriculum. NMSSA collects data from a linguistically, culturally and socio-economically heterogeneous NZ sample, and seeks to investigate which contextual factors may influence achievement.

While NZ participates in international assessment programs, NMSSA is the only data source for the NZ Ministry of Education that covers all learning areas of the NZC. This unique feature of NMSSA provides an opportunity for the project to produce system-level information that informs national policy and practice. Consequently, NMSSA results are highly valued by educators and used to improve student learning and accelerate progress.

Given their importance, many large-scale national and international studies (e.g. PISA, TIMSS, TALIS and PIRLS) collect information from students about their non-cognitive skills in their efforts to explain achievement differences and trends. However, most of these large-scale studies and researchers have focused mainly in the learning areas of reading (Yurdakal, 2019), mathematics (Aktas & Tabak, 2018) and science (Kayacan & Sonmez, 2019) rather than other learning areas, such as the arts.

There is a vast body of research (for example, Iwai, 2002; Kairavuori & Sintonen, 2012) that shows how arts education improves students' aesthetic, socio-emotional, socio-cultural, and cognitive developments and their academic achievement. Thus, many countries nowadays have implemented arts education in their curricula. As de Eça et al. (2017) state, "art education is preparing students with not only knowledge about art and the contemporary world they live in but also to think critically, use problem-solving skills and develop skills for living together in harmony" (p. 104). These are all considered important 21st century citizen skills. Therefore, it is important for today's schools to offer equal opportunities and experiences in arts education, which will encourage students to develop positive attitudes. As we know, attitudes are formed and shaped largely by experiences, and these attitudes, in turn, influence motivation for behavior (Morris & Stuckhardt, 1977). Students come to schools from various cultural and socioeconomic backgrounds with different childhood experiences, which influence their attitudes to certain arts disciplines. Moreover, population demographics in NZ are changing dramatically. In light of these factors, it was important in this study to seek information about students' attitudes to arts within the context of multicultural NZ setting. Encouraging positive attitudes and improving learning outcomes for all students are not unique to the NZ context, but they are global challenges. Therefore, the findings of this study are expected to be of interest to educators and researchers in multicultural societies.

Research Goal

The goal of this study was twofold: (i) use the CFA to examine the factorial validity of the attitude to arts scales in dance, drama, music and visual arts for different ethnic groups in NZ: NZ European, Maori, Pasifika, and Asian, and (ii) if MI is established, i.e. the construct is similar for all ethnic groups, investigate any differences in attitudes between NZ European, Maori, Pasifika, and Asian students.

The research questions were:

- To what extent is there evidence to support the factorial validity of the attitude to arts scales (dance, drama, music and visual arts) for use with different ethnic groups in the NZ context?
- If appropriate, are there any differences in attitudes to dance, drama, music and visual arts by ethnicity?

Methodology

Sample

NMSSA reports on achievement in different learning areas for year 4 and year 8 student populations in NZ. Nationally representative samples are drawn so that all year 4 and year 8 students in schools (with at least eight students at the respective year level) have an approximately equal chance of being selected. To achieve this, NMSSA carries out a two-step sampling procedure. The first step involves stratifying all eligible schools that have year 4 students and all eligible schools that have year 8 students using region, decile and school size as stratification variables. A stratified random sample of 100 schools at year 4 and 100 schools at year 8 are selected. Within each school selected, a random sample of up to 25 students from the year 4 or year 8 cohort is selected.

Approximately 5000 students are sampled each year from 200 schools. More detailed information about the NMSSA sampling algorithm and the characteristics of the 2015 sample can be found in 2015 NMSSA Technical Information report (NMSSA, 2017f). The 2015 sample was deemed to be nationally representative across gender, ethnicity, school decile (socio-economic status), type of school and region.

The assessment program was developed by the NMSSA team and used a variety of approaches to collect the data (see (NMSSA, 2017a) for details of the assessment framework for the study). Twelve experienced teachers were employed and specifically trained in 2015 to conduct the program of assessments and other data collection over two and half days within each school. Data about students' attitudes to dance, drama, music and visual arts were collected via questionnaires presented on laptops.

Because of the specific needs of the program, NMSSA used a matrix booklet design in 2015 for the student questionnaire, which meant that students in a random half of the schools responded to questions related to drama and music, and students in the other half of the schools responded to questions related to dance and visual arts. The number of students who responded to each arts discipline by ethnicity and included in this study is given in Table 1. For the study reported in this paper, the year 4 and year 8 students were combined as they were asked the same questions regarding their attitudes to the art disciplines.

Table 1. Number of Students for Each Arts Discipline by Ethnicity

Discipline	Ethnicity	N	Percentage
Dance Visual Arts	NZ European	1196	57.4
	Maori	450	21.6
	Pasifika	215	10.3
	Asian	223	10.7
Drama Music	NZ European	1202	54.3
	Maori	499	22.6
	Pasifika	251	11.3
	Asian	261	11.8

The total number of students responding to attitude items in dance and visual arts was 2084 (47.6% female), and in drama and music 2213 (49.8% female).

Measures

Attitude Statements and Scales

The NZC categorizes 'the arts' as four distinct but related disciplines; dance, drama, music and visual arts. The attitude statements and scales were developed by the NMSSA team in 2014-2015. The statement generation process began by reviewing the relevant literature and other large-scale studies that discussed indicators of attitudes to art disciplines. The NMSSA team worked with teachers, principals, members of the reference groups (Technical, Maori, Pasifika and Special education), members of the curriculum advisory panel in the arts, Ministry of Education officials and academic researchers to discuss and re-shape the statements within each discipline to ensure that they validly represented the attitude construct. The questionnaires were extensively peer-reviewed and piloted to ensure that the crafted statements would be understood by students. Following this, the revised versions were trialed with a total of 941 year 4 and year 8 students. At the end of this process, five items were retained for the final versions to be used in the 2015 study.

Each scale comprised five statements (e.g. "I like doing dance at school"). Students were asked to indicate how much they agreed with each statement on a 4-point Likert scale ("do not agree at all", "agree a little", "mostly agree", "totally agree"). Each scale was a measure of the extent to which students enjoyed and felt confident about their learning in the respective arts.

Data Analysis

Data analysis was conducted in three stages using Mplus 7.1 software (Muthen & Muthen, 2015) to answer each research question. Firstly, the factorial structure of each attitude scale was examined to determine whether the scales were unidimensional and represented the construct of attitude to each discipline of the arts. To determine whether the scales were unidimensional, we used CFA with maximum likelihood estimation with robust standard errors (MLR), which is a robust estimator to non-normality (Sass et al., 2014). Goodness-of-fit was evaluated using multiple criteria (Cheung & Rensvold, 2002; Fan & Sivo, 2005, 2007; Hu & Bentler, 1999; Vandenberg & Lance, 2000; Wu et al., 2007). Because of the sensitivity of the χ^2 statistic to sample size (Brown, 2006; Kline, 2005), acceptable and good model fit was determined by these measures: the root mean square error of approximation (RMSEA) and the standardized root mean square residuals (SRMR) with values < 0.08 (acceptable) or < 0.05 (good); and the comparative fit index (CFI) and the Tucker-Lewis index (TLI) with values > 0.90 (acceptable) or 0.95 (good).

Secondly, to see if the same construct was being measured in the same way for different ethnic groups, the MI for each attitude scale was examined by ethnicity using MG-CFA with means and covariance structure (MACS) (Sorbom, 1974). To determine if valid and meaningful comparisons between the ethnic groups could be made, we tested three progressively stringent estimates of MI: configural, metric and scalar. We applied the decision rule of $\Delta CFI \leq .01$ (Cheung & Rensvold, 2002; Vandenberg & Lance, 2000; Wu et al., 2007) because of the over-sensitivity of the chi-square difference test to large sample sizes.

Thirdly, (assuming scalar invariance was established), we compared the latent mean differences between ethnic groups to investigate if different ethnic groups had similar attitudes to dance, drama, music and visual arts.

Results

The results are presented in three sections. The first section reports the descriptive statistics for the attitude scales. The second section displays the results of the CFA analyses conducted to determine the unidimensionality of each scale and to determine if it was the same for each ethnic subgroup. The third section presents the comparisons of attitudes of NZ European, Maori, Pasifika, and Asian students across the arts disciplines.

1. Statements and Scales of Attitude to Dance, Drama, Music and Visual Arts

Table 2 presents the descriptive statistics for the five attitude statements, and Cronbach's Alpha values for each scale of attitude to dance, drama, music and visual arts. Overall, students had positive attitudes to the art disciplines, particularly for visual arts. They indicated the lowest attitudes, on average, to dance. The skewness and kurtosis values for the statements were within the acceptable range, not violating the univariate normality. Cronbach's alpha internal consistency estimates for the respective scales of attitude to the arts were above the recommended level of 0.70.

Table 2. Descriptive Statistics for Each Attitude Statement and Cronbach's Alpha for the Scales of Attitudes to Dance, Drama, Music and Visual Arts

Attitude scale and statements	M	SD	Cronbach's α
Dance			.93
1. I like learning about dance at school	2.43	1.08	
2. I like doing dance at school	2.40	1.14	
3. I would like to do more dance at school	2.37	1.18	
4. I would like to keep learning about dance as I grow up	2.28	1.15	
5. Dance is my favorite subject at school	1.89	1.05	
Drama			.92
1. I like learning about music at school	2.51	1.06	
2. I like doing music at school	2.55	1.08	
3. I would like to do more music at school	2.52	1.15	
4. I would like to keep doing music as I grow up	2.09	1.07	
5. Music is my favorite subject at school	1.84	.98	
Music			.85
1. I like learning about drama at school	2.90	.92	
2. I like doing drama at school	2.94	.96	
3. I would like to do more drama at school	2.81	1.06	
4. I would like to keep doing drama as I grow up	2.68	1.09	
5. Drama is my favorite subject at school	2.16	1.05	
Visual Arts			.85
1. I like learning art at school	3.09	.85	
2. I like doing art at school	3.29	.85	
3. I would like to do more art at school	3.07	1.01	
4. I would like to keep learning art as I grow up	2.82	1.02	
5. Art is my favorite subject at school	2.52	1.12	

2. Confirmatory Factor Analysis

Separate unidimensional CFAs for each attitude scale were performed using the MLR estimation method. The model-data-fit statistics are summarized in Table 3.

Table 3. Confirmatory Factor Analysis Results for the Scales of Attitude to Dance, Drama, Music and Visual Arts

Discipline	MLR- χ^2	df	RMSEA (90% CI)	CFI	TLI	SRMR
Dance	37.94	4	.06 (.05, .08)	.99	.98	.01
Drama	63.36	4	.08 (.07, .10)	.99	.97	.02
Music	18.99	4	.04 (.02, .06)	.99	.99	.01
Visual Arts	22.61	4	.05 (.03, .07)	.99	.99	.01

Note. RMSEA=Root Mean Square Error of Approximation; CFI=Comparative Fit Index; TLI=Tucker-Lewis Index; SRMR=Standardized Root Mean Square Residual.

As can be seen from Table 3, unidimensional models of each attitude scale provided acceptable (e.g., RMSEA < .08) to good (e.g., RMSEA < .05) fit indicating that the one-factor model of these measures was supported. Therefore, we concluded that CFA results were in line with the IRT estimates provided in the NMSSA report (NMSSA, 2017a).

The factor loadings of the attitude to dance, drama, music and visual arts scales are presented in Table 4. All standardized factor loadings ranged from .60 to .93, and were statistically significant, indicating that these statements were good indicators of each arts discipline-specific construct. The CFA results provided support for the factorial validity of each scale.

Table 4. Standardized Factor Loadings for the Scales of Attitudes to Dance, Drama, Music and Visual Arts

Statements	Dance	Drama	Music	Visual Arts
1	.85	.88	.76	.73
2	.90	.93	.81	.83
3	.89	.86	.80	.79
4	.78	.71	.60	.60
5	.76	.71	.66	.70

Measurement Invariance Analyses

After establishing good model fit, we used MG-CFA to test the MI of each scale. MI analyses by ethnicity results for each attitude scale are summarized in Table 5. The configural models fitted the data reasonably well (range of RMSEA: .05-.08, range of CFI/TLI: .97-.99, and range of SRMR: .01-.02) for all ethnic groups, indicating that students from different ethnic subgroups used the same frame of reference to answer the attitude items. Metric and scalar invariance results supported making meaningful comparisons of the latent mean scores between ethnic subgroups. Overall, the analyses supported the MI of the unidimensional attitude scales for each ethnic group in the sample. That is, NZ European, Maori, Pasifika, and Asian students all interpreted the attitude statements in the same way.

Table 5. Measurement Invariance Results

	MLR- χ^2	df	RMSEA	CFI	TLI	SRMR	Δ CFI
Dance							
Configural	46.25	16	.06	.99	.98	.01	-
Metric	71.35	28	.06	.99	.99	.03	-.003
Scalar	100.27	40	.05	.99	.99	.03	-.004
Drama							
Configural	76.62	16	.08	.99	.97	.02	-
Metric	100.93	28	.07	.98	.98	.03	-.003
Scalar	129.20	40	.06	.98	.98	.03	-.003
Music							
Configural	50.29	16	.06	.99	.97	.02	-
Metric	66.52	28	.05	.99	.98	.03	-.001
Scalar	109.82	40	.06	.98	.98	.04	-.010
Visual Arts							
Configural	37.09	16	.05	.99	.98	.02	-
Metric	55.80	28	.04	.99	.99	.04	-.002
Scalar	89.05	40	.05	.98	.98	.05	-.007

3. Comparison of Attitudes to the Arts by Ethnicity

Table 6 presents the comparisons of the latent mean scores of attitudes to dance, drama, music and visual arts for different ethnic groups; specifically, to determine whether Maori, Pasifika, and Asian students differed from NZ European students (which was treated as the reference groups). Effect sizes (ES) were used to judge the meaningfulness of the differences. Cohen's (1988) *d* estimates of ES are judged to be small when $d = 0.20$; medium when $d = 0.50$; and large when $d = 0.80$.

Pasifika students had higher mean attitude scale scores than NZ European students across all arts disciplines. However, the ES for music was medium and for dance, drama and visual arts, it was small. Maori students had higher mean attitude scale scores than NZ European students for dance, music and visual arts, and a lower mean scale score for drama. The ES for music was small and negligible for the others. Asian students had higher mean scale scores for drama, music and visual arts, although the mean differences were negligible or small.

A small effect size does not mean that it is trivial. According to some scholars in education (Slavin, 1990) effect sizes above .25 are considered large enough to be educationally meaningful.

Table 6. Latent Mean Scores and Effect Sizes for the Scales of Attitudes to Dance, Drama, Music and Visual Arts by Ethnic Subgroups Relative to NZ European Students

Ethnicity	Dance		Drama		Music		Visual Arts	
	LMS	ES	LMS	ES	LMS	ES	LMS	ES
NZ European	.00	-	.00	-	.00	-	.00	-
Maori	.11	.12	-.14	.15	.18	.26	.05	.07
Pasifika	.33	.36	.21	.23	.40	.57	.14	.23
Asian	-.03	.03	.03	.03	.11	.16	.13	.21

Note. LMS = Latent Mean Score; ES = Effect Size

Discussion

MI is an important prerequisite for making meaningful and valid comparisons between subgroups on cognitive and non-cognitive measures. Comparisons may be made on the basis of numerous variables, such as gender, ethnicity, socio-economic status, region etc. Given the increasing cultural diversity of NZ and interest in cross-cultural research, a variable of particular interest was ethnicity. This study described the MI of scales to measure students' attitudes to four disciplines of art for different ethnic groups in NZ and investigated the differences in attitudes of those subgroups. This study provided validity and reliability evidence for measuring students' attitudes to dance, drama, music, and visual arts whereas previous research either had a narrow focus on the areas of the arts, such as drama and visual arts (Smithrim & Upitis, 2005), considered art as one subject – art making (Pavlou & Kambouri, 2007), or had not investigated the arts at all. The present study adds to the attitude to arts research literature, by providing empirical evidence supporting the MI of the unidimensional attitude to dance, drama, music and visual arts scales across four ethnic subgroups in NZ primary schools.

The findings of the study supported the scales of attitude to dance, drama, music and visual arts as unidimensional scales that were equally valid for each ethnic group. These findings corresponded with the IRT (Rasch) approach used by NMSSA to construct cognitive and non-cognitive scales (NMSSA, 2017a) and provided support for an endorsement of the procedures used by NMSSA for constructing scales in the learning areas of the NZC.

The attitudes to dance, drama, music and visual arts by ethnic subgroups showed findings similar to those reported in the NMSSA reports (NMSSA 2017b, 2017c, 2017d, 2017e). Mean scores on the attitude scales indicated generally positive attitudes towards the arts disciplines. However, there were attitudinal differences with respect to ethnicity. Similar to the other learning areas assessed by NMSSA, Pasifika students tended to have higher mean attitude scores than the other ethnic groups in the art disciplines, but especially for dance and music. Mean attitude scale score differences between Asian and NZ students were generally negligible, while the mean attitude to music scale score for Maori students was higher than their NZ European friends.

The psychometric investigation of this study, and the similarity of findings with the published NMSSA findings gives us confidence that reported comparisons between subgroups on cognitive and non-cognitive measures are actual differences, not influenced by extraneous psychometric variables.

Performing arts, especially music and dance are important elements of Pasifika and Maori cultures. Thompson et al., (2009) and Whitinui (2008) have argued that Pasifika and Maori students engage in and learn more through the art of performing by enabling them to make stronger links with their own cultural identity and heritage. These results may partly explain our findings about Pasifika and Maori students' more positive attitudes to these disciplines and provides further construct validity evidence.

According to the NZC, the arts and culture are inextricably interconnected. The importance of culturally inclusive programs in the arts is well highlighted in the NZC document: "The Arts in the New Zealand Curriculum places emphasis on all students having opportunities to learn about the indigenous heritage of Maori and the diverse traditions of the European, Pacific, and other cultures that make up our nation" (Ministry of Education, 2007, p. 20).

Conclusion

This study has demonstrated that the IRT procedures adopted by NMSSA to construct cognitive and non-cognitive scales meets the psychometric expectations required to make comparisons between subgroups in large-scale assessment programs. Having established that the scales of attitudes to dance, drama, music and visual arts were unidimensional, multiple criteria were used to evaluate and establish the measurement invariance of the scales. Scale level MI results reported in this study were found to be similar to those item-level MI findings reported by NMSSA. Therefore, the readers and users of NMSSA findings can be reassured of the veracity of the psychometric procedures used, and the findings reported.

Scales measuring student attitudes to the arts have seldom been constructed. Understanding attitudes to different learning areas and possible ethnic differences on those constructs may help educators and policy makers reflect on their practices and provide strategies and intervention programs aimed at encouraging positive attitudes, and thereby improving engagement and schooling outcomes for *all* students.

Suggestions

Further studies are needed to understand the relationship between achievement in school learning areas and attitude to that learning area. NMSSA has sought to report this for all learning areas in the NZC, but other measures of attitudes also merit investigation. The literature suggests that engagement in the arts may help increase achievement as well as critical thinking and problem-solving skills. Studies examining the relationship between attitudes to the arts and achievement in both the learning area of arts and other curriculum areas are encouraged. The roles of external variables, such as socio-economic background, gender, teacher and school characteristics on students' attitudes to the arts could also be investigated through further research. MI studies of these attitude scales over time also merit investigation given the rapid technological advances in the practice of the arts.

Limitations

This study is not immune to limitations. The data used in this research came from self-report instruments. New studies are encouraged to evaluate the criterion validity of the scales using various measures. For example, the relationship between student-reported attitudes and in-class observations during art classes or the achievement scores awarded by art teachers could be investigated.

Acknowledgements

The National Monitoring Study of Student Achievement (NMSSA) project is funded by the Ministry of Education, New Zealand.

References

- Aiken, L. R. (1970). Attitudes toward mathematics. *Review of Educational Research*, 40(4), 551–596.
- Aiken, L. R. (1976). Update on attitudes and other affective variables in learning mathematics. *Review of Educational Research*, 46(2), 293–311.
- Aktas, M. C., & Tabak, S. (2018). Turkish adaptation of math and me survey: A Validity and reliability study. *European Journal of Educational Research*, 7(3), 707–714. <https://doi.org/10.12973/eu-jer.7.3.707>
- Asil, M., & Brown, G. T. (2016). Comparing OECD PISA reading in English to other languages: Identifying potential sources of non-invariance. *International Journal of Testing*, 16(1), 71–93. <https://doi.org/10.1080/15305058.2015.1064431>
- Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological Methods*, 22(3), 507–526. <https://doi.org/10.1037/met0000077>
- Bertling, J. P., Borgonovi, F., & Almonte, D. E. (2016). Psychosocial skills in large-scale assessments: Trends, challenges, and policy implications. In A. A. Lipnevich, F. Preckel & R. D. Roberts (Eds.), *Psychosocial skills and school systems in the 21st century* (pp. 347–372). Springer International Publishing.
- Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2002). *Cross-cultural psychology: Research and applications* (2nd ed.). Cambridge University Press.
- Blatt-Gross, C. (2013). Toward meaningful education: Investigating artful behavior as a human proclivity in the classroom. *International Journal of Education and the Arts*, 14(7), 1–22.

- Brown, A. T. (2006). *Confirmatory factor analysis for applied research*. Guilford Press.
- Brown, G. T., Harris, L. R., O'Quin, C., & Lane, K. E. (2017). Using multi-group confirmatory factor analysis to evaluate cross-cultural research: Identifying and understanding non-invariance. *International Journal of Research & Method in Education*, 40(1) 66–90.
- Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling*, 12(3), 471–492.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233–255.
- Cohen, J. (1988) *Statistical power analysis for the behavioral sciences* (2nd ed.) Hillsdale.
- de Eça, T. T., Milbrandt, M. K., Shin, R., & Hsieh, K. (2017). Visual arts education and the challenges of the millennium goals. In G. Barton & M. Baguley (Eds.), *The palgrave handbook of global arts education* (pp. 91–105). Palgrave Macmillan.
- Ercikan, K., & Lyons-Thomas, J. (2013). Adapting tests for use in other languages and cultures. In K. Geisinger (Ed.), *APA handbook testing and assessment in psychology* (Vol. 3, pp. 545–569). American Psychological Association.
- Fan, X., & Sivo, S. A. (2005). Sensitivity of fit indexes to misspecified structural or measurement model components: Rationale of two-index strategy revisited. *Structural Equation Modeling*, 12(3), 343–367. https://doi.org/10.1207/s15328007sem1203_1
- Fan, X., & Sivo, S. A. (2007). Sensitivity of fit indices to model misspecification and model types. *Multivariate Behavioral Research*, 42(3), 509–529.
- Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/ performance measures. *Review of Educational Research*, 52(1), 123–142.
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18(3), 117–144.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.
- International Test Commission (ITC). (2016). *The ITC guidelines for translating and adapting tests (Second edition)*. <https://www.intestcom.org/page/16#:~:text=About%20the%20ITC%20Guidelines%20for,in%20testing%20technology%20and%20practices>.
- International Test Commission (ITC). (2017). *The ITC guidelines for large-scale assessment of linguistically diverse populations*. <https://www.intestcom.org/page/31>
- Iwai, K. (2002). The contribution of arts education to children's lives. *Prospects*, 32(4), 407–420. <https://doi.org/10.1023/A:1022102017139>
- Kairavuori, S., & Sintonen, S. (2012). Arts education: Instruments of expression and communication. In H. Niemi, A. Toom & A. Kallioniemi, (Eds.), *The Miracle of education: The principles and practices of teaching and learning in Finnish schools* (pp. 211–225). Sense Publishers.
- Kayacan, K., & Sonmez Ektem, I. (2019). The effects of biology laboratory practices supported with self-regulated learning strategies on students' self-directed learning readiness and their attitudes towards science experiments. *European Journal of Educational Research*, 8(1), 313–299. <https://doi.org/10.12973/eu-jer.8.1.313>
- Kline, R. (2005). *Principles and practice of structural equation modeling* (2nd ed.). Guilford Press.
- Lipnevich, A. A., MacCann, C., & Roberts, R. D. (2013). Assessing non-cognitive constructs in education: A review of traditional and innovative approaches. In D. H. Saklofske, V. L. Schwann, & C. R. Reynolds (Eds.), *The Oxford handbook of child psychological assessment* (pp. 750–772). Oxford University Press.
- Logan, S., & Johnston, R. (2009). Gender differences in reading ability and attitudes: Examining where these differences lie. *Journal of Research in Reading*, 32(2), 199–214.
- Ma, X., & Kishor, N. (1997). Attitude toward self, social factors, and achievement in mathematics: A meta-analytic review. *Educational Psychology Review*, 9(2), 89–120.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543.
- Ministry of Education. (2007). *The New Zealand Curriculum*. Learning Media.
- Morris, J. W., & Stuckhardt, M. H. (1977). Art attitude: Conceptualization and implication, *Studies in Art Education*, 19(1), 21–28.

- Muthen, L. K., & Muthen, B. O. (2015). *Mplus user's guide* (7th ed.). Muthen & Muthen.
- National Monitoring Study of Student Achievement. (2017a). *The arts 2015: Key findings* (No. 10.1). Educational Assessment Research Unit & New Zealand Council for Educational Research. <https://www.educationcounts.govt.nz/publications/series/nmssa/all-nmssa-publications>
- National Monitoring Study of Student Achievement. (2017b). *Visual arts 2015: Key findings* (No. 10.5). Educational Assessment Research Unit & New Zealand Council for Educational Research. <https://www.educationcounts.govt.nz/publications/series/nmssa/all-nmssa-publications>
- National Monitoring Study of Student Achievement. (2017c). *Music- Sound arts 2015: Key findings* (No. 10.4). Educational Assessment Research Unit & New Zealand Council for Educational Research. <https://www.educationcounts.govt.nz/publications/series/nmssa/all-nmssa-publications>
- National Monitoring Study of Student Achievement. (2017d). *Dance 2015: Key findings* (No. 10.2). Educational Assessment Research Unit & New Zealand Council for Educational Research. <https://www.educationcounts.govt.nz/publications/series/nmssa/all-nmssa-publications>
- National Monitoring Study of Student Achievement. (2017e). *Drama 2015: Key findings* (No. 10.3). Educational Assessment Research Unit & New Zealand Council for Educational Research. <https://www.educationcounts.govt.nz/publications/series/nmssa/all-nmssa-publications>
- National Monitoring Study of Student Achievement. (2017f). *Technical information 2015* (No. 11). Educational Assessment Research Unit & New Zealand Council for Educational Research. <https://www.educationcounts.govt.nz/publications/series/nmssa/all-nmssa-publications>
- Pavlou, V. & Kambouri, M. (2007). Pupils' attitudes towards art teaching in primary school: An evaluation tool. *Studies in Educational Evaluation*, 33(3-4), 282–301.
- Popham, W. J. (2005). Students' attitudes count. *Educational Leadership*, 62(5), 84-85.
- Sass, D. A., Schmitt, T. A., & Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(2), 167–180.
- Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review*, 18(4), 210–222. <https://doi.org/10.1016/j.hrmr.2008.03.003>
- Silverman, S., & Subramaniam, P. R. (2000). Student attitude toward physical education and physical activity: A review of measurement issues and outcomes. *Journal of Teaching in Physical Education*, 19(1), 97–125.
- Slavin R. E. (1990). On making a difference. *Educational Researcher*, 19 (3), 30–44.
- Smithrim, K. & Uptis, R. (2005). Learning through the arts: Lessons of engagement. *Canadian Journal of Education*, 28(1 & 2), 109–127.
- Sorbom, D. (1974). A general method for studying differences in factor means and factor structures between groups. *British Journal of Mathematical and Statistical Psychology*, 27(2), 229–239.
- Steinkamp, M. W., & Maehr, M. L. (1983). Affect, ability, and science achievement: A quantitative synthesis of correlational research. *Review of Educational Research*, 53(3), 369–396.
- Thompson, S., McDonald, F., Talakai, M., Taumoepeau, V., & Te Ava, A. (2009). *Training manual for teachers working with Pacific students: Engaging Pacific learners*. Ako Aotearoa. <https://gifted.tki.org.nz/assets/Uploads/files/Training-manual-for-teaching-working-with-pacific-students-engaging-pacific-learners.pdf>
- Van de Vijver, F. J., Avvisati, F., Davidov, E., Eid, M., Fox, J. P., Le Donne, N., Lek, K., Meuleman, B., Paccagnella, M. & van de Schoot, R. (2019). *Invariance analyses in large-scale studies*, (No. 201). OECD Publishing. <https://doi.org/10.1787/254738dd-en>.
- Van de Vijver, F. J., & He, J. (2016). Bias assessment and prevention in noncognitive outcome measures in context assessments. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds), *Assessing contexts of learning: An international perspective* (pp. 229–253). Springer International Publishing.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4–70.
- Whitinui, P. (2008). Kapa Haka counts: Improving participation levels of Maori students in mainstream secondary schools. *Mai Review*, 3(8), 1–14.

-
- Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research & Evaluation, 12*(3). <https://doi.org/10.7275/mhqa-cd89>
- Xu, L., Diket, R., & Brewer, T. (2018). Predicting student performance via NAEP secondary art analysis using partial least squares SEM. *Arts Education Policy Review, 119*(4), 231–242.
- Yurdakal, I. H. (2019). Examination of correlation between attitude towards reading and perception of creative reading. *European Journal of Educational Research, 8*(2), 443–452. <https://doi.org/10.12973/eu-jer.8.2.443>