

A Rasch analysis of the Teachers Music Confidence Scale

Hoi Yin Bonnie Yim

University of South Australia, School of Education *hoi.yim@unisa.edu.au*

Sabry Abd-El-Fattah

University of South Australia, School of Education *sabry.abd-el-fattah@unisa.edu.au*

Lai Wan Maria Lee

University of South Australia, School of Education *mlee@vtc.edu.hk*

This article presents a new measure of teachers' confidence to conduct musical activities with young children; Teachers Music Confidence Scale (TMCS). The TMCS was developed using a sample of 284 in-service and pre-service early childhood teachers in Hong Kong Special Administrative Region (HKSAR). The TMCS consisted of 10 musical activities. Teachers rated their confidence levels to conduct each activity on a scale from 1 (Not confident at all) to 5 (Very confident). An exploratory factor analysis retained a 10-item single factor that was replicated using confirmatory factor analysis procedures. All items of the TMCS fitted the Rasch model adequately. In-service teachers showed higher confidence levels to conduct several musical activities with young children than pre-service teachers. Implications of these findings for measuring teachers' confidence to conduct musical activities with young children were discussed.

Music education, early childhood education, confidence,
in-service and pre-service teachers, Rasch analysis

INTRODUCTION

Music in early childhood education encompasses different areas of teaching, including singing, moving, dancing, playing percussive instruments, and listening. Several research studies have highlighted that involvement in musical activities is thought to develop one's reading and neuroanatomical abilities, verbal learning and retention (Butzlaff, 2000; Ho, Cheung, & Chan, 2003) while also promoting understanding of language, improving the ability to recall information, fostering creativity, and creating an environment more conducive to learning in other areas (Neelly, 2001; Rauscher, 2002; Rauscher & LeMieux, 2003; Vaughn, 2000). The merits associated with involvement in musical activities have encouraged many countries to incorporate music into their national curriculum from pre-school to postsecondary education (Snyder, 1997).

Furthermore, there have been growing research efforts to investigate factors that may contribute towards improving music teaching within a school context (Hamann, Baker, McAllister, & Bauer, 2000; Hennessy, Rolfe, & Chedzoy, 2001; Russell-Bowie & Dowson, 2005). One possible important factor is teachers' confidence levels to conduct musical activities. Overall, confidence is meant to refer to one's faith in one's ability. Several researchers have established a linkage between teachers' confidence levels to conduct musical activities and several desirable educational outcomes. For example, Mills (1989) reported that music taught by a confident teacher helped children appreciate music as part of the whole curriculum, and enabled greater opportunities to be provided for music. Rainbow (1996) argued that a confident music teacher was meant to help new learners master musical skills more quickly. Rainbow explained that music teachers' mastery of various musical activities such as singing and aural perception was

essential before introducing such activities to children. Similarly, Tillman (1988) and Glover and Ward (1993) highlighted that teachers' own musical skills and their levels of confidence in these skills, as well as their general teaching abilities, could be sufficient to help children learn music.

However, music teachers seem to be presented with different levels of confidence both in their own musical abilities and their abilities to teach music in a school context. For example, Mills (1989, 1995-6) and Russell-Bowie (1993) indicated that approximately 60 to 70 per cent of primary teacher education students entered their primary teachers training having minimal, if any, formal music education experiences and consequently lower levels of confidence to conduct musical activities. Similarly, Lawson, Plummeridge, and Swanwick (1994) expressed concern that there might be insufficient teachers in primary schools with the necessary confidence and expertise to implement fully the music program. Moreover, Hennessy (2000) highlighted that "many teachers believe that music requires gifts that are only attainable by, or given to, a chosen few" (pp. 183-184). Beauchamp and Harvey (2006) argued that music could be one of the problem areas for managerial and administrative staff in the school.

Furthermore, Holden and Button (2006) asked a sample of 141 British teachers to indicate their levels of confidence to teach 10 national curriculum subjects, including music, on a scale from 1 (highest levels of confidence) to 10 (lowest levels of confidence). Participants were also requested to attend a semi-structured interview. Results of the study showed that music was given the lowest ranking of confidence to teach. In addition, the interviewees showed high levels of uncertainty about music and described it as a specialist area. The results also revealed non-significant differences between Key Stage 1 (ages 4-7) and Key Stage 2 (ages 7-11) teachers in their confidence levels to teach music. However, there was a positive and significant relationship between teachers' confidence levels to teach music and teachers' musical qualifications, musical experience and training, and attitudes toward music. The semi-structured interview further revealed that singing was the most difficult aspect of music to practise confidently although it was the activity taught most frequently.

AIM OF THE STUDY

Despite the above concerns about music teacher's confidence levels, there seem to be little research that investigates teachers' confidence levels to conduct musical activities with young children. The present study attempts to build on the work of Holden and Button (2006) through developing a scale that aims at measuring teachers' confidence levels to conduct musical activities with young children; *Teachers Music Confidence Scale* (TMCS). One goal of the present study is to test the factorial structure of the TMCS using both exploratory and confirmatory factor analysis techniques. A second goal is to investigate whether the items of TMCS fit the Rasch model. A third goal is to test whether there are any differences between in-service and pre-service teachers' confidence levels to conduct musical activities with young children.

METHODS

Participants

The present study included 284 early childhood teachers (165 pre-service and 119 in-service) in Hong Kong Special Administrative Region (HKSAR). Of the whole sample, 66 per cent were aged 25 years or below. Pre-service teachers were from a local tertiary institute, and in-service teachers were from 18 local preschools. Although a cluster sample design was employed sample random simple statistics have been employed and reported in this article. Consequently, in the use of the tests some allowances must be made for the cluster sample design.

Measurements

The Teachers Music Confidence Scale (TMCS) is designed according to the Guide to the Pre-primary Curriculum (Hong Kong Curriculum Development Council, 2006; Hong Kong Curriculum Development Institute, 1996); South Australian Curriculum, Standards and Accountability Framework (Department of Education and Children's Services, 2004) and the National Standards for Music Education (1994). The TMCS is a 10-item scale that intended to measure teachers' confidence levels to conduct musical activities with young children. The question of the TMCS stated "On a scale of 1-5, how confident are you in undertaking the following musical activities with young children?" This question is followed by a list of 10 music-related activities. Teachers express their confidence level to conduct each musical activity on a scale from 1(Not confident at all) to 5 (Very confident). Scores on all items of the TMCS can be summed up to obtain a total score which represents teachers' overall confidence levels to conduct musical activities with young children.

Procedures

The TMCS was originally prepared in English. The first author translated the English version to Chinese. Two early childhood bilingual professionals compared the English and the Chinese versions of the TMCS and found the translation to be satisfactory. For pre-service teachers, the TMCS was administered and collected in-person in the same session. For in-service teachers, the TMCS was sent out by mail and returned within a period of a week.

RESULTS

Exploratory Factor Analysis

An exploratory factor analysis of the TMCS yielded a 10-item single factor (Cronbach $\alpha = 0.89$) which explained 50.5 per cent of the total variance extracted. The factor loadings of all items of the TMCS are presented in Table 1.

Table 1: Exploratory factor analysis of the TMCS (N= 284)

Factor/Statement	Factor loadings
1. Singing.	0.81
2. Dancing/Moving/Dramatising with music.	0.74
3. Playing percussive instrument(s).	0.73
4. Listening to music.	0.72
5. Composing / improvising music.	0.72
6. Integrating music into curriculum.	0.70
7. Providing various types of music materials.	0.70
8. Using multimedia tools to facilitate teaching.	0.69
9. Identifying children's musical potentials.	0.68
10. Knowing about children's musical interests.	0.60
Eigenvalue	7.1

Unidimensionality

In order to test whether the items of the TMCS fitted the Rasch model, it was necessary to examine whether or not the items of the TMCS were unidimensional since the unidimensionality of items was one of the requirements for the use of the Rasch model (Anderson, 1994; Hambleton & Cook, 1977).

Consequently, confirmatory factor analysis procedure was used to test the unidimensionality of TMCS items. Confirmatory factor analysis is a statistical procedure used for investigating relations between a set of observed variables and the underlying latent variables (Byrne, 2001; Kim & Mueller, 1978). Thus, confirmatory factor analysis assumes that the observed variables are derived from some underlying source variables (Kim & Mueller, 1978). Factor analysis may also be used as an appropriate method for identifying the minimum number of hypothetical variables

that account for the observed covariation, and thus as a means of exploring the data for possible data reduction (Kim & Mueller, 1978). However, one of the main purposes of confirmatory factor analysis is to examine the common underlying dimensions associated with a number of observed variables.

The AMOS 6.0 program (Arbuckle, 2005) was used to run a confirmatory factor analysis of the TMCS using the full information maximum likelihood estimation procedure (Bollen, 1989). The analysis showed that the TMCS could be described as a one factor model, presented in Figure 1, $\chi^2 (35, N = 284) = 45.5, p = 0.11$, Root-Mean-Square Error of Approximation (RMSEA) = 0.02, Standardized Root-Mean-Square Residual (SRMR) = 0.01, Adjusted Goodness of Fit Index (AGFI) = 0.98, Parsimonious Goodness of Fit Index (PGFI) = 0.32, Tucker-Lewis Index (TLI) = 0.99, Parsimony Ratio (PRATIO) = 0.84, and Parsimony Normed Fit Index (PNFI) = 0.85. All the hypothesized regression path coefficients of the TMCS model, presented in Table 2, were statistically significant because the critical ratio (CR) for a specific regression path coefficient was $> \pm 1.96$ (Byrne, 2001).

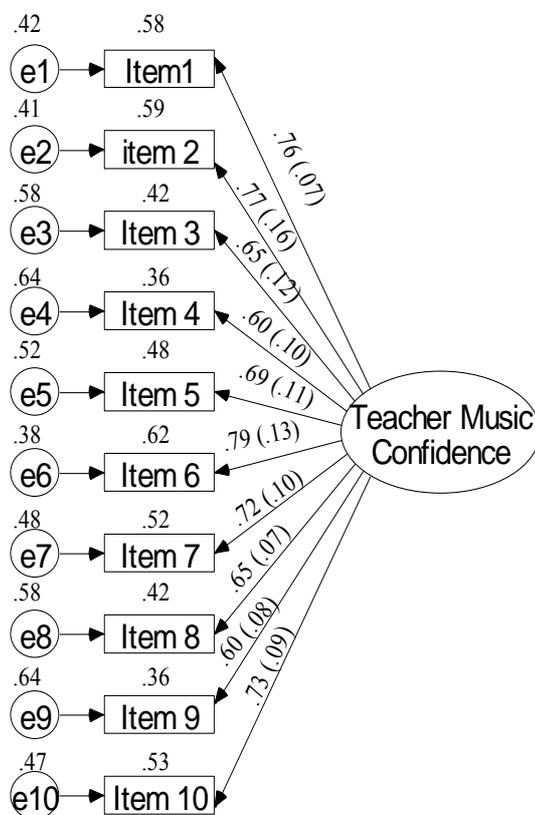


Figure 1: Confirmatory factor analysis of the TMCS

Rasch Analysis

It is common within classical test theory to sum individual item response values to obtain a total score. However, this approach has been criticised and reviews have been made by Andrich (1978), Masters (1988), and Wright and Masters (1982). For example, Bond and Fox (2001) highlighted that the summing of individual item response values had two underlying assumptions. First, each item was measured on an equal interval scale. Thus, each item was contributing equally to the underlying trait. Second, the distances or the steps among the response categories were equal for an item and through all items of a scale, that is, the level of the underlying trait required to move from one response category to another was the same for an item and was equal across all items of a scale. Bond and Fox concluded that those two assumptions were counterintuitive and mathematically inappropriate.

Table 2: Standardized loadings, standard error, critical ratio, error variance, and R2 of the second-order confirmatory factor analysis of the TMCS (N = 284)

Paths	Standardized loadings	Standard error	Critical ratio	Error variance	R ²
Teachers' Music Confidence					
1	0.76	0.07	10.9	0.42	0.58
2	0.77	0.16	4.8	0.41	0.59
3	0.65	0.12	5.4	0.58	0.42
4	0.60	0.10	6.0	0.64	0.36
5	0.69	0.11	6.3	0.52	0.48
6	0.79	0.13	6.1	0.38	0.62
7	0.72	0.10	7.2	0.48	0.52
8	0.65	0.07	9.3	0.58	0.42
9	0.60	0.08	7.5	0.64	0.36
10	0.73	0.09	8.1	0.47	0.53

The basic Rasch model is a dichotomous response model (Rasch, 1960; Wright & Store, 1979) that represents the conditional probability of a binary outcome as a function of a person's trait level (*B*) and an item's difficulty (*D*). The Rasch dichotomous response model is given by:

$$P_{ni} = \frac{\exp(\beta_n - \delta_i)}{1 + \exp(\beta_n - \delta_i)}$$

where P_{ni} is the probability of an endorsed response (a yes response to an item), β_n is the trait (or ability) parameter of person n , and δ_i is the difficulty of endorsing item i . When $\beta_n > \delta_i$, $\beta_n = \delta_i$, and $\beta_n < \delta_i$, the chances of a 'yes' response is greater than 50 per cent, equal to 50 per cent, and less than 50 per cent, respectively.

Andrich (1978; 1988) is credited for extending Rasch dichotomous response model to the rating scale. The rating scale model is an additive linear model that describes the probability that a specific person (n) will respond to a specific Likert-type item (i) with a specific rating scale step (x). It is important to note that the Likert scale can be modelled with either the rating scale or the partial credit model (Masters, 1988; Wright & Masters, 1982). The partial credit model allows the item format and the number of categories to vary from item to item (e.g., some items are scored with a 5-point scale and others with a 6-point scale). When the item format is inconsistent from item to item, the partial credit model is useful in providing estimates of the psychological distance between each set of the ordinal categories (Masters, 1988). However, the rating scale model restricts the step structure to be the same for all items (Wright & Masters, 1982). In essence, the rating scale models are a subset of the partial credit models (Andrich, 1978).

The simple dichotomous response model can be extended to provide an appropriate model for use with polytomous response categories by the addition of an additional difficulty parameter; either a second δ parameter or a τ parameter. The Rasch rating scale model is given by:

$$p_{nij} = fn \left(\frac{\exp(\beta_n - \delta_i - \tau_j)}{1 + \exp(\beta_n - \delta_i - \tau_j)} \right)$$

or

$$p_{nij} = \frac{\sum_{k=1}^m \exp(\beta_n - \delta_{ij})}{1 + \sum_{k=1}^m \exp(\beta_n - \delta_{ij})}$$

where n = subscript for persons, i = subscript for items, and j = response categories (0, 1, 2).

In the present analysis, the QUEST program (Adams & Khoo, 1993) was used to run the Rasch analysis for the TMCS. All the reported results were obtained from the QUEST program. The

RUMM program (Andrich, Sheridan, & Luo, 2000), however, was used to plot the Item Characteristic Curve and Category Probability Curve with thresholds for an example item of the TMCS.

Item Fit Statistics

One important item fit statistics was the infit mean square (INFIT MNSQ). The infit mean square measured the consistency of fit of the cases to the Item Characteristic Curve (ICC) for each item with weighted consideration given to those cases close to the 0.5 probability level. The acceptable range of the infit mean square statistic for each item of the TMCS was taken to be from 0.77 to 1.30 (Adams & Khoo, 1993). Items that had infit mean square above 1.30 indicated that the relevant items did not discriminate well, and below 0.77 indicated that the relevant items provided redundant information. Items that had INFIT MNSQ outside the acceptable range were deleted from the analysis (Wright & Store, 1979). Figure 2 shows that, in the present analysis, no items of the TMCS had been deleted because all items had INFIT MNSQ values within the acceptable range of 0.77 to 1.30. Specifically, the range of the INFIT MNSQ for all items of the TMCS were 0.90 to 1.20.

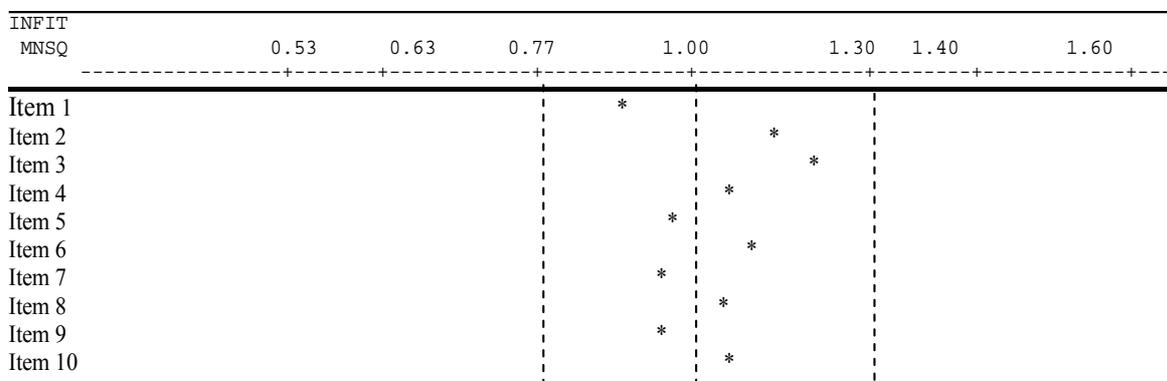


Figure 2. Plot of all Infit Mean Squares for all items of the TMCS

The RUMM program could divide the examined sample into a specified number of groups or Class Intervals (CIs) for each item. The average ability of individuals within each CI was calculated and represented by a dot on the ICC for each item. If an item fit the Rasch model, the dots should fall on or as close as possible to the ICC. Any deviations of any of these dots from the ICC represented a difference between the observed mean ability of the CI that these dots represented and the expected mean ability of the CI as predicted by the Rasch model. In the present analysis, the RUMM program divided the sample of the study (N = 284) into eight CIs that were plotted along the ICC for each item. Figure 3 shows the ICC for Item 1 of the TMCS.

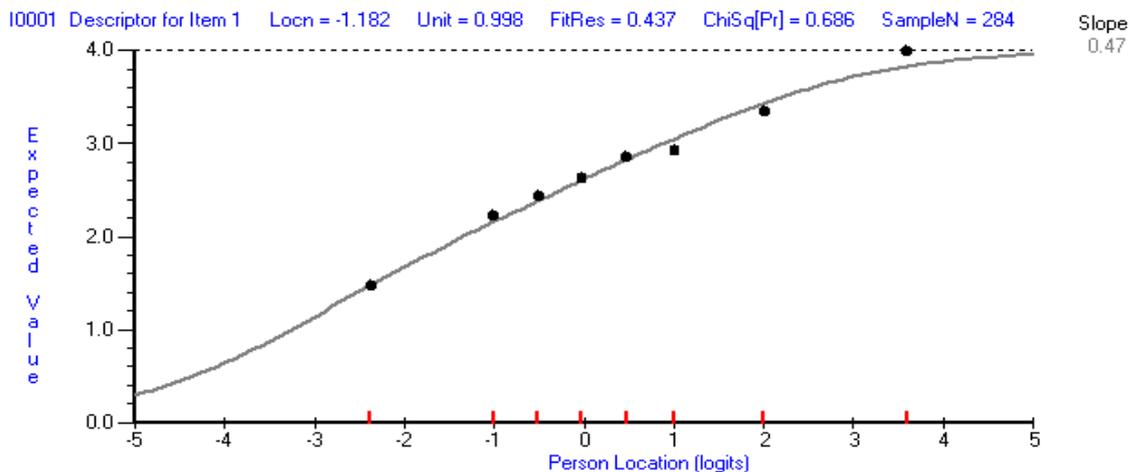


Figure 3. Item Characterise Curve for Item 1 of the TMCS

Figure 4 shows the threshold values for item 1 of the TMCS. The threshold values reflect the item difficulty for each item. According to Bond and Fox (2001) a threshold is “the level at which the likelihood of failure to endorse a given response category (below the threshold) turns to the likelihood of endorsing the category (above the threshold)” (p. 234). For example, in the case of four response categories, there are three thresholds that mark the boundaries between the four response categories: SD (*Strongly Disagree*)-D (*Disagree*)-A (*Agree*)-SA (*Strongly Agree*) and all are ordered. That is, the data are regarded as ordinal and the Rasch model transforms the counts of the endorsement of these ordered Likert categories into interval scales (Bond & Fox, 2001).

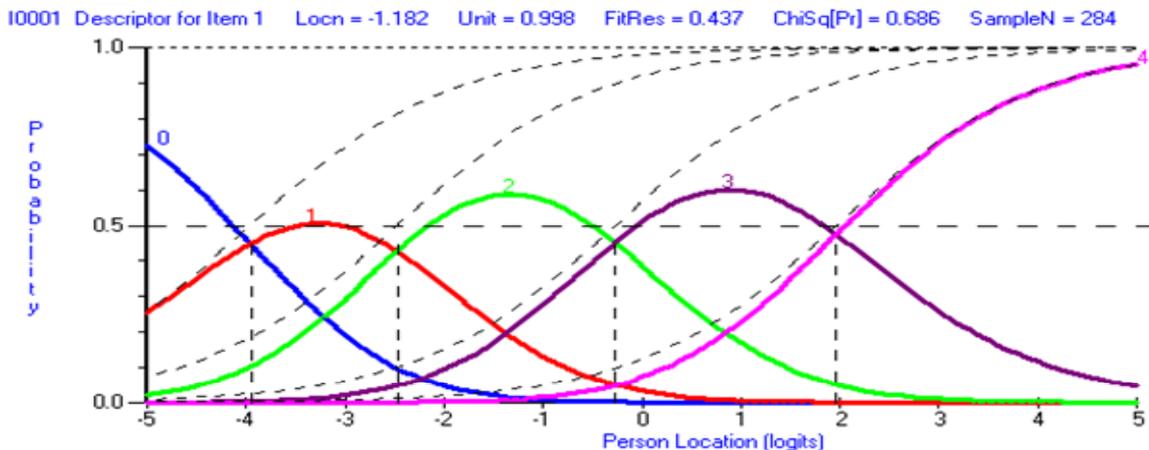


Figure 4. Category Probability Curve and thresholds for item 1 of the TMCS

Case Estimates

It is also important when investigating the fit of the Rasch scale to data to examine the estimates for each case. The case estimates give the performance level of each student on the total scale. In order to identify whether the cases fit the Rasch scale or not, it is important to examine the case OUTFIT mean square statistic (OUTFIT MNSQ) which measures the consistency of the fit of the persons to the student characteristic curve for each student, with special consideration given to extreme items. In the present analysis, the general guideline used for interpreting t as a sign of misfit was if $t > \pm 5$ (Wright & Stone, 1979). Thus, if the OUTFIT MNSQ value for a person had a t -value greater than $t > \pm 5$, that person did not fit the scale and was deleted from the analysis. In the present analysis, no person was deleted because the t -value for all cases fell within the acceptable range of ± 5 . Specifically, the OUTFIT MNSQ for all cases had t -values between - 3.4 to + 2.7, and since the normal t -value tests were not employed because of a cluster sample design, no cases were deleted.

Mean Testing

A series of independent-sample t tests is presented in Table 3 and shows that in-service teachers have higher confidence levels to conduct musical activities with young children than the pre-service teachers, including, (a) singing, (b) dancing/moving/dramatising with music, (c) playing percussive instruments, (d) composing / improvising music, (e) integrating music into curriculum, (f) identifying children’s musical potentials, and (g) knowing about children’s musical interests. In addition, in-service teachers show higher overall levels of confidence to conduct musical activities with young children than pre-service teachers. It should be noted that the t -tests associated with the differences between the mean values did not make allowance for a cluster sample design.

DISCUSSION

Building on the work of Holden and Button (2006), the main goal of the present study was to develop a quick and accessible measure of teachers’ confidence to conduct musical activities with

young children; *Teacher Music Confidence Scale (TMCS)*. The main aim was to establish the psychometric proprieties of the TMCS using appropriate statistical and measurement procedures such as exploratory factor analysis, confirmatory factor analysis, and Rasch analysis. Mean testing procedures were also used to examine differences between in-service and pre-service teachers' confidence levels to conduct musical activities with young children.

Table 3: Differences between in-service and pre-service teachers confidence levels to conduct musical activities with young children (N= 284)

Variable	Background	M	SD	df	t
Singing	Pre-service	3.4	0.95	281	- 4.5 *
	In-service	3.9	0.71		
Dancing/Moving/Dramatising with music	Pre-service	3.0	0.88	281	- 6.1 *
	In-service	3.6	0.71		
Playing percussive instrument(s)	Pre-service	2.9	0.90	281	- 5.8 *
	In-service	3.4	0.71		
Listening to music	Pre-service	3.4	0.91	281	- 0.94
	In-service	3.5	0.76		
Composing / improvising music	Pre-service	2.3	1.00	279	- 3.6 *
	In-service	2.8	0.90		
Integrating music into curriculum	Pre-service	2.8	0.88	282	- 4.7 *
	In-service	3.2	0.76		
Providing various types of music materials	Pre-service	2.8	0.84	280	- 1.8
	In-service	2.9	0.76		
Using multimedia tools to facilitate teaching	Pre-service	2.9	0.89	280	0.43
	In-service	2.8	0.93		
Identifying children's musical potentials	Pre-service	2.7	0.81	282	- 4.4 *
	In-service	3.1	0.79		
Knowing about children's musical interests	Pre-service	3.0	0.83	281	- 5.7 *
	In-service	3.5	.65		
Overall confidence	Pre-service	29.0	6.2	276	- 5.2 *
	In-service	32.6	5.2		

Note. * $p < 0.05$

Findings from the exploratory factor analysis showed that the TMCS could be represented by a 10-item single factor that had a satisfactory internal consistency reliability. A confirmatory factor analysis successfully replicated these findings with all 10 items showing acceptable loadings on the latent trait of 'teachers' confidence'. In addition, all items of the TMCS fitted the Rasch model satisfactorily, indicating that the 10 items of the TMCS measured teachers' levels of confidence to conduct musical activities with young children.

Mean testing showed that in-service teachers had higher confidence levels to conduct musical activities with young children than the pre-service teachers, including (a) singing, (b) dancing/moving/dramatising with music, (c) playing percussive instruments, (d) composing / improvising music, (e) integrating music into curriculum, (f) identifying children's musical potentials, and (g) knowing about children's musical interests. In addition, in-service teachers showed higher overall levels of confidence to conduct musical activities with young children than pre-service teachers.

One possible explanation that can account for these differences between in-service and pre-service teachers is practical work experience and on-job training. In-service teachers may regularly interact with young children and consequently gain broader insights about children's musical needs, and the appropriate musical activities to get children involved and foster their musical growth. This familiarity with children's musical preferences and interests may have fortified in-service teachers' levels of confidence to conduct musical activities with young children. On the contrary, pre-service teachers are more likely to be presented with lower levels of experience to handle a music class, unfamiliarity with the appropriate musical activities to attract young children, and consequently a possibility of failure to meet children's expectations

about a music class. This lack of experience may add difficulty to pre-service teachers and challenge their confidence levels to conduct musical activities with young children.

Another possible interpretation for these differences between in-service and pre-service teachers may be due in part to the nature of music itself. Music can be regarded as a unique discipline or mode of discourse that entails a unique set of practices, procedures and skills (Finney, 2000, p. 208). In the present study, for example, singing, dancing/moving/dramatising with music, playing percussive instruments, and composing/improvising music are comparatively more skill-based activities relative to other activities represented in the TMCS. In-service teachers may have higher levels of musical skill and knowledge due to their possible regular practices with young children. Consequently, in-service teachers may have higher confidence levels to conduct these musical activities than pre-service teachers. This notion seems to be consistent with findings by Holden and Button (2006) that singing and composition were more difficult to conduct than other musical activities by non-music specialist teachers in the United Kingdom.

In summary, the TMCS represents a promising measure of teachers' confidence to conduct musical activities with young children. Unlike the single question that measures teachers' overall confidence levels to teach music (Holden & Button, 2006), the TMCS measures teachers' confidence levels to conduct 10 musical activities. The scores on all the 10 musical activities can be summed to provide a total score which represents a teacher's overall confidence level to conduct musical activities with young children.

REFERENCES

- Adams, R. J., & Khoo, S. T. (1993). *Quest- The Interactive Test Analysis System*. Hawthorn, Victoria: ACER.
- Anderson, L. W. (1994). Attitude measures. In T. Husen (Ed.), *The International Encyclopaedia of Education* (2nd ed., Vol. 1, pp. 380-390). Oxford: Pergamon.
- Andrich, D. (1978). Rating formulation for ordered response categories. *Psychometrika*, 43, 561-573.
- Andrich, D. (1988). *Rasch models for measurement*. Newbury Park, CA: Sage publications.
- Andrich, D., Sheridan, B., & Luo, G. (2000). *RUMM2010: A Windows interactive program for analyzing data with Rasch unidimensional Models for Measurement [Computer Program]*. Perth, Western Australia: RUMM Laboratory.
- Arbuckle, J. L. (2005). *Amos 6.0 [Statistical Program]*. Chicago, IL: SPSS.
- Beauchamp, G., & Harvey, J. (2006). 'It's one of those scary areas': Leadership and management of music in primary schools. *British Journal of Music Education*, 23(1), 5-22.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley.
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Butzlaff, R. (2000). Can music be used to teach reading? *Journal of Aesthetic Education*, 34(3-4), 167-178.
- Byrne, B. (2001). *Structural equation modelling with AMOS: Basic concepts, applications, and programming*. New Jersey: Lawrence Erlbaum Associates.
- Department of Education and Children's Services. (2004). *South Australian curriculum, standards and accountability framework*. Adelaide, South Australia: Department of Education and Children's Services.
- Finney, J. (2000). Curriculum stagnation: the case of singing in the English national curriculum. *Music Education Research*, 2(2), 203-211.
- Glover, J., & Ward, S. (1993). *Teaching music in the primary school*. London: Cassell.
- Hamann, D. L., Baker, D. S., McAllister, P. A., & Bauer, W. I. (2000). Factors affecting university music students' perceptions of lesson quality and teaching effectiveness. *Journal of Research in Music Education*, 48(2), 102-113.
- Hambleton, R. K., & Cook, L. L. (1977). Latent trait models and their use in the analysis of educational test data. *Journal of educational measurement*, 14, 75-96.
- Hennessy, S. (2000). Overcoming the red-feeling: the development of confidence to teach music in primary school amongst student teachers. *British Journal of Music Education*, 17(2), 183-196.

- Hennessy, S., Rolfe, L., & Chedzoy, S. (2001). The Factors which influence student teachers' confidence to teach the arts in the primary classroom. *Research in Dance Education*, 2(1), 53-71.
- Ho, Y. C., Cheung, M. C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology*, 17(3), 439-450.
- Holden, H., & Button, S. (2006). The teaching of music in the primary school by the non-music specialist. *British Journal of Music Education*, 23(1), 23-38.
- Hong Kong Curriculum Development Institute. (1996). *Guide to the pre-primary curriculum*. Hong Kong: Educational Department.
- Hong Kong Curriculum Development Council. (2006). *Guide to the pre-primary curriculum*. Hong Kong: Educational and Manpower Bureau.
- Kim, J., & Mueller, C. W. (1978). *Factor analysis statistical methods and practical issues*. London: Sage.
- Lawson, D., Plummeridge, C., & Swanwick, K. (1994). Music and the national curriculum in primary schools. *British Journal of Music Education*, 11(1), 3-14.
- Masters, G. N. (1988). The analysis of partial credit scoring. *Applied Measurement in Education*, 1(4), 279-297.
- Mills, J. (1989). The generalist primary teacher of music: a problem of confidence. *British Journal of Music Education*, 6(23), 125-138.
- Mills, J. (1995/6). Primary student teachers as musicians. *Bulletin of the Council for Research in Music Education*, 127, 122-126.
- Neelly, L. P. (2001). Developmentally appropriate music practice: Children learn what they live. *Young Children*, 56(3), 32-37.
- Rainbow, B. (1996). Onward from Butler: Schools music 1945–1985. In G. Spruce (Ed.), *Teaching Music* (pp. 9-20). Milton Keynes: The Open University.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment test*. Copenhagen: The Danish Institute of Educational Research.
- Rauscher, F. H. (2002). Mozart and the mind: Factual and fictional effects of musical enrichment. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 267-278). San Diego, CA: Academic Press.
- Rauscher, F. H., & LeMieux, M. T. (2003, April). *Piano, rhythm, and singing instruction improve different aspects of spatial-temporal reasoning in Head Start children*. Paper presented at the annual meeting of the Cognitive Neuroscience Society, New York.
- Russell-Bowie, D. (1993). Where is music education in our primary schools? *Research Studies in Music Education*, 1, 40-51.
- Russell-Bowie, D., & Dowson, M. (2005). *Effects of background and sex on confidence in teaching the creative arts: Tests of specific hypotheses*. Paper presented at the Australian Association for Research in Education Conference, Sydney, Australia.
- Snyder, S. (1997). Development musical intelligence: Why and how. *Early Childhood Education Journal*, 24(3), 165-171.
- The National Association for Music Education. (1994). *National Standards for Music Education*. Retrieved 06 April, 2006, from <http://www.menc.org/publication/books/standards.htm>
- Tillman, J. (1988). Music in the primary school and the national curriculum. In W. Salaman & J. Mills (Eds.), *Challenging assumptions: New perspectives in the education of music teachers*. Exeter: Association for the Advancement of Teacher Education in Music.
- Vaughn, K. (2000). Music and mathematics: Modest support for the oft-claimed relationship. *Journal of Aesthetic Education*, 34(3-4), 149-166.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago: MESA Press.
- Wright, B. D., & Store, M. (1979). *Best test design*. Chicago: MESA Press.