THE DEVELOPMENT OF A PSYCHOMETRICALLY-SOUND INSTRUMENT TO MEASURE TEACHERS' MULTIDIMENSIONAL ATTITUDES TOWARD INCLUSIVE EDUCATION

Marian Mahat
Monash University

The Multidimensional Attitudes toward Inclusive Education Scale (MATIES) *was developed to effectively measure affective, cognitive and behavioural aspects of attitudes, within the realm of inclusive education that includes physical, social and curricular inclusion. Models within Item Response Theory and Classical Test Theory were used for calibrating the subscales. Using a sample of primary and secondary school regular teachers in Victoria, pilot study analyses indicate that the final three subscales of eighteen items successfully met standards for internal reliability, content validity, construct validity, criterion validity and convergent validity; and provide preliminary evidence to warrant further use of this instrument for the purpose of measuring teachers' attitudes toward inclusive education.*

Inclusive education is based on the notion that schools should, without question, provide for the needs of all students, whatever their level of ability or disability. This means that students with disabilities are educated in the company of their regular age peers in a regular school and classroom and provided with instructions that effectively and efficiently meet their educational needs. The ideal of inclusive education is that schools not only accept every child that walks through their doors but also ensure that students with disabilities are *considered as full members of the classroom learning community, with their special needs met there* (Friend & Bursuck, 1996, p4).

Hence inclusive education suggests that no child should be excluded from the neighbourhood school because of perceived learning differences. However, inclusive education is more than just being there. While the physical presence of a child in regular classrooms in their local neighbourhood school is agreed by most as a prerequisite to the actual act of including a child, inclusive education also encompasses the inclusion of a child within the social and curricular milieu of the educational environment (Wills & Jackson, 2001).

Profound changes in the provision of educational services to people with disabilities have resulted from international, national and state policies. These changes have seen a move from increasing the integration of students with disabilities in regular schools to providing an educational environment that includes all. Researchers over several decades have concluded that teachers' attitudes are one of the most crucial variables in the success of inclusive education (Chow & Winzer, 1992; Hayes & Gunn, 1988; Williams & Algozine, 1977). These attitudes can create positive (or negative) expectations and behaviours which increase (or limit) the successful inclusion of students with a disability in educational environments. It is important to obtain an accurate picture of teachers' attitudes toward inclusive education as these attitudes are predictors of the success of inclusion efforts for both students with and without disabilities (Schumm & Vaughn, 1995; Van Reusen, Shoho, & Barker, 2000; Villa, Thousand, Meyers, & Nevin, 1996). An understanding of these attitudes is essential for curriculum planning and in-service and pre-service training programs; and could have a significant impact on current and future educational policy, program planning and funding decisions.

Research concerning attitudinal barriers to inclusive education requires psychometrically sound instruments that will allow researchers, practitioners and policy makers to respond to factors that may facilitate or impede the formation and modification of attitudes toward inclusive education. Over the years, as the movement toward inclusive education accelerated, scales to measure attitudes toward various aspects of inclusive education have been proposed and constructed.

While most attitudinal instruments were designed for specific research situations and used only once (e.g. Barton, 1992; Daane, Beirne-Smith, & Latham, 2000), a number of attitudinal instruments in inclusive education have some psychometric characteristics that other researchers have deemed sufficient enough to justify further use (e.g. Berryman, Neal, & Robinson, 1980; Larrivee & Cook, 1979; Reynolds & Greco, 1980; Schmelkin, 1981; Wilczenski, 1992). However, in some cases, either the psychometric properties of the attitudes instruments have not been reported fully or the instruments have psychometric properties that are somewhat unclear. For instance, Reynolds and Greco (1980) failed to report on the characteristics of the items and scales when calibrating the Educational Attitude Survey. In the case of Attitudes toward Mainstreaming Scale (Berryman et al., 1980), the psychometric properties are unclear as different factorial structures were found on a number of occasions (Berryman & Neal, 1980; Berryman et al., 1980; Green & Harvey, 1983). This lack of evidence of psychometric adequacy raises concerns for the validity and reliability of some of the instruments.

A majority of attitudinal instruments (e.g. Berryman et al., 1980; Larrivee & Cook, 1979; Moberg, Zumberg, & Reinmaa, 1997; Reynolds & Greco, 1980; Schmelkin, 1981; Semmel, Abernathy, Butera, & Lesar, 1991; Sideridis & Chandler, 1995; Villa et al., 1996; Wilczenski, 1992) measured a single dimension of attitudes, particularly the cognitive aspects of attitudes. In a review of attitude scales (Mahat, 2007), only two studies were found to have employed the affective, cognitive and behavioural aspects of attitudes to measure attitudes toward inclusive education (Avramidis, Bayliss, & Burden, 2000; Cochran, 1997). In the same review, Mahat (2007) also found only one instrument that had items measuring attitudes toward the physical, academic, behavioural and social aspects of inclusion (Wilczenski, 1992). No study has attempted to incorporate both the different dimensions of inclusive education and attitudes while measuring teachers' attitudes toward inclusive education.

The purpose of the present study was to develop a multidimensional instrument that could effectively measure affective, cognitive and behavioural aspects of attitudes, within the realm of inclusive education that includes physical, social and curricular inclusion. While a number of studies have attempted to include one or the other (Avramidis et al., 2000; Cochran, 1997; Wilczenski, 1992), this study extends previous research and contributes to further understanding of the theoretical nature and structure of attitudes and the knowledge base for the provision of inclusive education particularly when inclusive education, warranted or not, is becoming a global phenomena that cannot be ignored.

Theoretical framework

Over the last few years, a number of attitudinal studies in inclusive education has claimed a theoretical base with regard to explaining and predicting behaviour (e.g. Conatser, Block, & Gansneder, 2002; Hodge & Jansma, 2000; Roberts & Lindsell, 1997; Roberts & Smith, 1999; Subban & Sharma, 2005). The current study is aligned with the Theory of Planned Behaviour which is an extension of the original Theory of Reasoned Action; and incorporates perceived control over behaviour achievement as an additional aggregate of intention (Ajzen, 1985).

The Theory of Planned Behaviour provides a framework for understanding the effects of factors such as relationships between attitudes toward behaviours, normative beliefs, perceived behavioural control, intention and behaviour. According to the theory, the most important determinant of a person's behaviour is behaviour intent and it specifically puts forward three conceptually independent determinants of intentions (Ajzen, 1987), i.e. attitudes toward the behaviour, subjective norms and the degree of perceived behavioural control. In line with the Theory of Planned Behaviour, this study postulates that the formation of intentions (within inclusive education that includes physical, social and curricular) is influenced by (multidimensional) attitudes toward the behaviour, perceived social pressure to perform or not perform the behaviour (subjective norms) and perceived ease or difficulty of performing the behaviour (perceived behavioural control) reflected by previous experience and knowledge; and newly acquired knowledge

(Ajzen, 1991). The more favourable the attitudes and subjective norm with respect to behaviour, and the greater the perceived control behaviour, the stronger should be the individual's intention to perform the behaviour. This will provide useful information in understanding these behaviours and for the implementation of intervention strategies to effectively change these behaviours.

Measurement framework

In developing the questionnaire, this study merged the measurement frameworks described by Wilson (2003; 2005) and DeVellis (2003). The measurement framework provided a systematic way of developing the questionnaire and is based on a combination of both theoretical and psychometric approaches to scale development. The development of the attitudes instrument, within this measurement framework, is also guided by the following criteria:

- Brevity - administration would not be a deterrent for its use;
- Ease of administration - requiring no extensive instructions or trained examiners;
- Flexibility - for use with different groups of educators;
- Valid – fulfil sufficient evidence of validity; and
- Reliable – fulfil sufficient evidence of reliability.

The development of the *attitudes toward inclusive education* construct involved stating the research aims and questions, review of literature and consultation of other instrumentation used to measure teachers' attitudes toward inclusive education. In this study, attitudes were seen as multidimensional consisting of affective, cognitive and behavioural, and inclusive education was seen to encompass physical, social and curricular inclusion. Specifically, inclusive education was defined *as the education of all students in age appropriate regular classrooms, regardless of the degree or severity of a disability. It involves students accessing the regular curriculum; with the necessary support; and within a welcoming social atmosphere.*

A construct map (Masters, Adams, & Wilson, 1990) was used to develop the theoretical framework for the questionnaire based on the scope and definition of the construct. The construct map illustrates the ordering of respondents concurrently to the difficulty levels of the items (see Figure 1). Respondents who are more positive about inclusive education are placed above respondents who are less positive. Likewise, items that are harder to agree with are located above items that are easier to agree. Such a map depicts an idea rather than an exact technical representation (Wilson, 2005). Because a respondent may lie at different levels on the inclusion continuum depending on their cognitive, affective and behavioural attitudes, a respondent would tend to respond differently depending on where he or she lies on the different continua. Thus a separate construct map is applied to each in turn.

In transforming the theoretical framework into a number of statements, more than a hundred items were constructed based on a synthesis of previous research focusing on teachers' attitudes toward inclusive education. Many of these were eliminated because of their ambiguity and their similarity with other items. A pool of 41 items was initially chosen to fit across the three dimensions of attitudes, i.e. items were formulated and chosen to correspond with affective, cognitive and behavioural attitudes. These items were then sorted according to the relative strength needed to agree with them.

In this study, a Likert-type scale was used as the outcome space or format for measurement. The Likert-type scale is regarded as a softer form of data collection that clearly acknowledges that the questions require merely expressed opinions (Bond & Fox, 2001). It is also not difficult to create, can include a large number of items that can be answered quickly, provides precise information about a respondent's degree of attitudes, and provides high reliability (Oppenheim, 1992). In this study, participants were asked to rate each item on a six point rating scale of *Strongly Agree*, *Somewhat Agree*, *Agree*, *Disagree, Somewhat Disagree* and *Strongly Disagree.* A *Neutral* or *Undecided* or *Uncertain* category was not included as it tended to attract responses from respondents who do not understand the statement (Bond & Fox, 2001); who do not wish to participate (Wright & Masters, 1982); and responses that may imply equal attraction to both agreement and disagreement irrespective of its precise wording (DeVellis, 2003). The number of categories that maximises reliability were tested empirically with the aid of the measurement model but were not reported here.

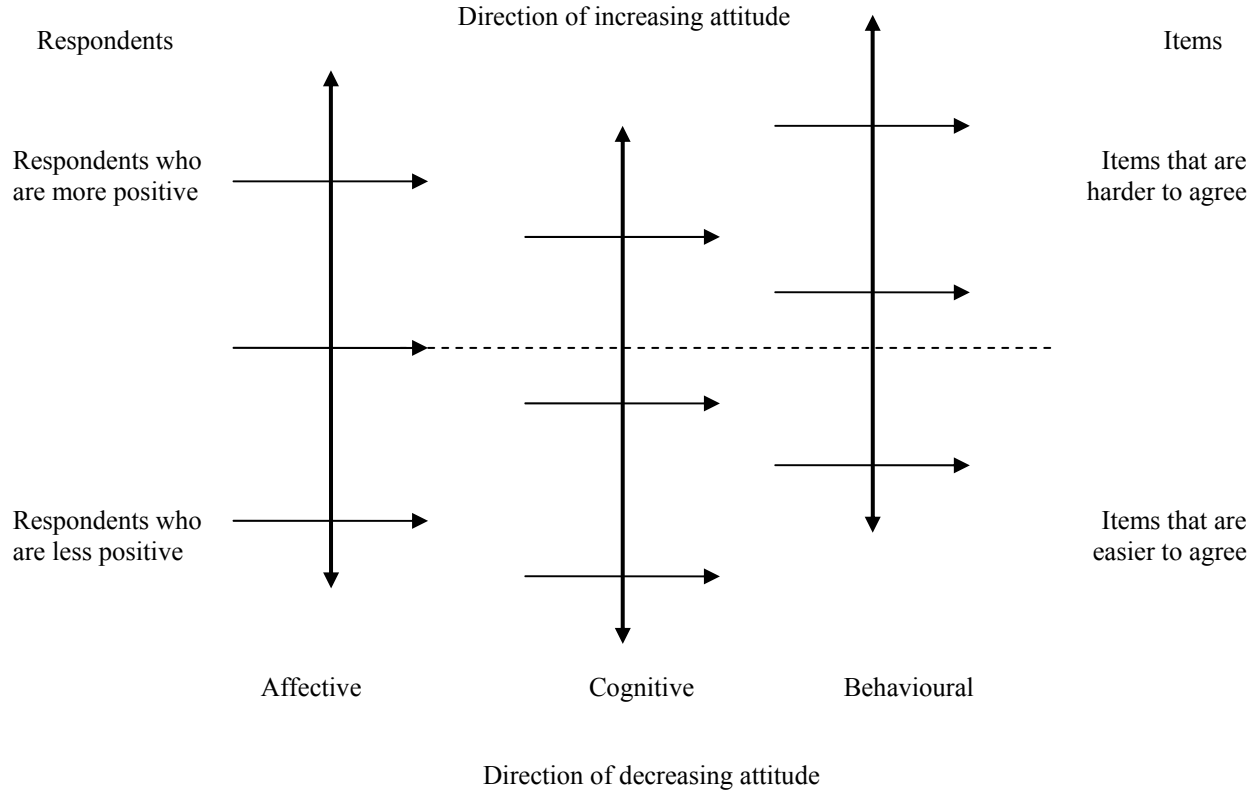

Figure 1
Construct map for three dimensions of attitudes

The next step involved asking a group of experts knowledgeable in the content area to review the item pool. Seven experts, in the area of special education, inclusive education and measurement, were approached to review the content and structure of items and questionnaire. The panel of experts in special education and inclusive education (and hence familiar with the construct) were also asked to determine whether a wide range of stringency was represented among the items. Following feedback, five items were deleted while a number of items were rephrased to ensure clarity. The amended questionnaire was then reviewed by a group of fourteen teachers, consisting of both special education and regular teachers. Teachers were asked to complete the questionnaire, providing comments on the clarity of the items and how the questionnaire might be improved. The questionnaire was generally *user friendly* and unambiguous in its instructions as few errors were made. There was no specific pattern to the errors identified indicating there was no significant fault in the design of any single item that may cause respondent errors. In addition, there was no discerning pattern to missing responses indicating that there was not a problem with any single item with regards to non-response.

The final pilot instrument, *Multidimensional Attitudes toward Inclusive Education Scale* (*MATIES*), consisted of 36 items inquiring about teachers' attitudes toward inclusive education. Within each dimension of attitudes, there were items that described physical, social and curricular inclusion. Items within the affective dimension of attitudes represent teachers' feelings and emotions associated with inclusive education and include items such as *I am pleased that students with a disability are able to attend the local neighbourhood school*. Items within the cognitive dimension of attitudes reflect teachers' perceptions and beliefs about inclusive education and these include *I believe that all students should be able to study side by side in the regular classroom regardless of ability*. Finally items such as *I am willing to modify the physical environment to include students with a disability in the regular classroom* are statements of behavioural intent and imply teacher's intention to act in a certain manner toward inclusive education. The response

continuum was reversed for half of the items in affective and cognitive dimensions. Due to the nature of the items, the response continuum of all items in the behavioural dimension was also reversed.
In the past, most instruments measuring attitudes in inclusive education (or mainstreaming and integration) have been calibrated using classical test models and procedures. Hambleton and Swaminathan (2004; 1985) identified a number of shortcomings of classical test theory (CTT) and related models and practices that make them not well-suited for some of the demands being placed on measurement models today. Item response theory (IRT) evolved in response to some of the shortcomings of CTT. It is a family of models that describes the interaction between examinees and items using probabilistic models. Notwithstanding the many developments in IRT, it has been argued that CTT continues to be an important framework for test construction (Bechger, Maris, Verstralen, & Beguin, 2003). In calibrating the attitudes instrument for this study, the measurement models used were located within IRT and CTT. These provided different dimensions of information and hence add valuable information about the validity and reliability of the instrument.

The analysis of the pilot data consisted of three stages. In order to locate respondents and items responses on the construct map, a model within IRT known as the Multidimensional Random Coefficients Multinomial Logit model (MRCML) (Adams, Wilson, & Wang, 1997) was employed. Items were discarded if the analyses of fit do not conform to the model. In the second stage, confirmatory factor analysis within CTT was utilised to identify the relationship to the underlying variable. At this stage, items may be discarded if they demonstrate a weak relationship with the underlying variable, i.e. items with high factor loadings were retained. Finally analyses of reliability and validity within items and subscales were reported using a combination of models within both theories.

Conducting a factor analysis first using the original observations can lead to misleading results and generate what Wright (1996) called *illusory factors*. Further Linacre (1998) argued that exploratory factor analysis can report items clustering at different performance levels as different factors; and that factor analysis alone cannot tell whether each factor is a dimension or a slice of a shared dimension. Item response theory facilitates the development of a construct theory and interpretation of levels of attitudes by providing item and respondent locations on the variable. Once the respondents and item responses have been defined, confirmatory factor analysis can then support the model of discrete factors that are consistent with the data.

The questionnaire was administered to a development (or calibration) sample. In this study, a stratified random sampling method was employed to select a sample in one of the school regions in Victoria. The target group was expected to closely approximate the actual school population breakdown in that region.

In evaluating the items, analyses were conducted using the Statistical Package for the Social Sciences (SPSS) and ConQuest computer program (Wu, Adams, & Wilson, 1998). Items were evaluated using fit analysis via the INFIT and OUTFIT mean square statistics (Wright & Masters, 1982), confirmatory factor analysis, construct and criterion validity and reliability via the item and person separation indices (Wright & Masters, 1982) and coefficient alpha internal consistency index (Cronbach, 1951). Those items that contributed least to the overall internal consistency were considered for exclusion or amendment. In arriving at the final scale that is optimal (given all circumstances), the evaluation of each item was both theory driven and data driven.

Sample
The target group was primary and secondary school regular teachers in Victoria. Using a random selection method, twenty schools within the Eastern Metropolitan region of Victoria were invited to participate in the study, of which seven principals (2 primary and 5 secondary schools) agreed to participate. Questionnaire equal to the number of teachers (216 primary and secondary) were distributed, of which 115 questionnaires were collected for a return rate of 53.2%. Three questionnaires were discarded because they were not completed. The sample characteristics are provided in Table 1.

Table 1
Number of years teaching, gender and age of participants

	All		Primary		Secondary	
	N	%	n	%	n	%
Gender						
Female	82	73.87	31	88.57	50	67.57
Male	29	26.13	4	11.43	24	32.43
	111	100.00	35	100.00	74	100.00
Age						
Less than 30 years	30	29.13	9	26.47	21	31.34
30 – 39 years	19	18.45	6	17.65	13	19.40
40 – 49 years	19	18.45	9	26.47	10	14.93
50 - 59 years	30	29.13	8	23.53	20	29.85
60 years and above	5	4.85	2	5.88	3	4.48
	103	100.00	34	100.00	67	100.00
No. of years teaching						
Below 2 years	14	12.96	3	8.57	11	15.49
2 – 5 years	28	25.93	8	22.86	20	28.17
6 – 10 years	9	8.33	5	14.29	4	5.63
11 – 20 years	16	14.81	7	20.00	8	11.27
21 – 30 years	29	26.85	9	25.71	20	28.17
More than 30 years	12	11.11	3	8.57	8	11.27
	108	100.00	35	100.00	71	100.00

Results

An initial exploratory factor analysis was used only to detect the dimensionality of the scale. The analysis found that all 36 items identified eight factors meeting Kaiser's eigenvalue criterion of greater than 1, accounting for 70.6% of the variance in the data. Although the scree plot suggested the presence of one single dominant factor, it also provided evidence of other factors in the data. Further, an IRT calibration of all items revealed that at least nine items showed a lack of relationship with a unidimensional variable. Together these analyses indicate that the individual subscales measure distinct dimensions of teachers' attitudes and suggest that total score on the whole instrument would be inappropriate. Both item response modelling and confirmatory factor analysis were used to follow up this initial exploration of the data.

Since all items used the same six response categories (*Strongly Agree* to *Strongly Disagree*), both the rating scale (Andrich, 1978) and partial credit (Masters, 1982) models were available within MRCML. The rating scale model assumes that respondents will use the six categories consistently with all items while the partial credit model is used if respondents' interpretations and uses of the categories were not consistent across all items. To decide between the models used in this study, the deviances and parameters of both models were examined. A comparison of the two models indicated deviances were 10580.27 (46 parameters) for the rating scale model and 10236.04 (184 parameters) for the partial credit model. The reduction in deviance was 344.24 for 138 additional parameters, hence the partial credit model was chosen for subsequent analyses. In item response modelling, the fit of the model to the data is used to predict response patterns to all items by all teachers. Where the data are shown to misfit, it is a signal that the item data need to be examined and in most cases, the item is excluded from the scale. The infit mean square unit provides a measure of item fit to the IRT model of the variable against which it is calibrated and a range of 0.77 and 1.30 (Wright & Masters, 1982) was chosen to indicate item fit. Items with an infit of 1.00 show acceptable fit to the model. Items with a fit below 0.77 show patterns of deterministic behaviour in the variable context, that there is dependence on another item or that there is redundancy in the scale. Items with fit greater than 1.30 show patterns of randomness or items that all or perhaps none of the teacher agrees with. Fit indices indicated 11 items lie outside the range. An examination of the items resulted in a decision to delete eight of the items. Three items (with an infit greater than 1.3 but smaller than 1.4) were retained, which is still an acceptable assessment of items in a rating scale (Wright & Linacre, 1994).

The revised scale of 28 items was subjected to confirmatory factor analysis. Consistent with the prevailing conceptualisation of teachers' attitudes, three factors were retained and rotated orthogonally as well as

obliquely. The factors were found to be correlated, hence the oblique solution using a promax rotation (kappa = 4) was used for further analyses. An examination of both the pattern and structure matrices

Table 2
Mean item difficulty, item fit and factor loadings for item

Item	IRT		CTT
	Logits	Infit MS	Factor loadings
Cognitive			
I believe that an inclusive school is one that permits academic progression of all students regardless of their ability.	-0.29	1.05	0.76
I believe that students with a disability should be taught in special education schools.	0.71	1.07	0.79
I believe that inclusion facilitates socially appropriate behaviour amongst all students.	-0.09	1.05	0.64
I believe that any student can learn in the regular curriculum of the school if the curriculum is adapted to meet their individual needs.	0.15	1.04	0.64
I believe that students with a disability should be segregated because it is too expensive to modify the physical environment of the school.	-0.69	0.91	0.76
I believe that students with a disability should be in special education schools so that they do not experience rejection in the regular school.	0.21	1.04	0.72
Affective			
I get frustrated when I have difficulty communicating with students with a disability.	0.26	0.94	0.72
I get upset when students with a disability cannot keep up with the day-to-day curriculum in my classroom.	-0.30	0.84	0.78
I get irritated when I am unable to understand students with a disability.	-0.35	0.89	0.79
I am uncomfortable including students with a disability in a regular classroom with other students without a disability.	0.12	1.21	0.62
I am disconcerted that students with a disability are included in the regular classroom, regardless of the severity of the disability.	0.34	1.14	0.52
I get frustrated when I have to adapt the curriculum to meet the individual needs of all students.	-0.07	0.94	0.73
Behavioural			
I am willing to encourage students with a disability to participate in all social activities in the regular classroom.	-0.53	0.82	0.90
I am willing to adapt the curriculum to meet the individual needs of all students regardless of their ability.	0.01	1.08	0.88
I am willing to physically include students with a severe disability in the regular classroom with the necessary support.	0.50	1.06	0.86
I am willing to modify the physical environment to include students with a disability in the regular classroom.	0.32	1.01	0.87
I am willing to adapt my communication techniques to ensure that all students with an emotional and behavioural disorder can be successfully included in the regular classroom.	0.01	0.95	0.86
I am willing to adapt the assessment of individual students in order for inclusive education to take place.	-0.30	0.81	0.88

resulted in a decision to retain items that loaded on at least 0.50 on one factor. This resulted in a removal of a further ten items. The remainder 18 items had a number of items that loaded on more than one factor. However the pattern coefficients for these items reflect a direct path from the factor to the variable when the influence of the other variables is partialled out. As the components are correlated, sums of squared loadings cannot be added to obtain a total variance. The factor loadings for each item are provided in Table 2.

The final instrument had six items within each subscale. Each of the subscales was re-calibrated in turn. All items fit within the range of 0.77 and 1.30 and had factor loadings greater than 0.5. Table 2 also provides the logit and fit values for each item. In IRT, the items and teachers are placed along the same latent variable using the location measure in units called logits. Negative logit values represent items that are easier to agree with and teachers' estimates that represent teachers who are less positive about inclusive education. Similarly, positive logit values represent items that are harder to agree with and teachers who are more positive about inclusive education. For instance, the item *I believe that students with a disability should be taught in special education schools* has the highest positive logit in the cognitive subscale, suggesting that this item was hardest to agree.

Reliability estimates in IRT, given in Table 3, indicate the extent to which location measures can be separated along the variable, given their error estimates. The possible range for separation indices varies from 0.0 to 1.0. In the case of item separation indices, a value of 0.0 indicates all items are located at the same position on the variable and there is complete redundancy in the items' capacities to measure agreement, while a value of 1.0 indicates that the items are completely separated along the variable and each contributes a unique amount to the interpretation of the variable. Wright and Masters (1982) referred to this as an index of construct validity. Similarly, the person (teacher) separation index is interpreted in the same way. A value closest to 1.0 shows the extent to which the subscale can discriminate between teacher agreement levels on the variable. This can be interpreted as an index of criterion validity. The Item Separation Index had values close to 1.0 for each of the subscale, indicating that the items are well separated along the variable being measured. Similarly, the Teacher Separation Index provided sufficient evidence of the capacity of the subscales to discriminate between differing levels of teachers' attitudes. The Cronbach reliability for each subscale was substantial, returning alpha coefficients between 0.77 and 0.91.

Table 3
Indices for item and teacher estimates on each subscale

Subscale	Item separation	Teacher separation	Cronbach □
Cognitive	0.98	0.81	0.77
Affective	0.95	0.94	0.78
Behavioural	0.91	0.75	0.91

Covariation between subscales was examined to explore relations between the subscales. The analysis, presented in Table 4, showed that the subscales independently form constructs. As the response continuum was reversed for a number of items, the covariation between the subscales indicated positive correlation. There was a medium positive correlation between the affective and cognitive variables; and high positive correlation between affective and behaviour and between cognitive and behaviour variables (Cohen, 1988). This means that high levels of one dimension of attitudes are associated with high levels of another dimension of attitudes; particularly positive affective and cognitive attitudes are associated with positive behavioural intentions in inclusive education.

Table 4
Inter-subscale correlations

	Cognitive	Affective	Behaviour
Cognitive	1.00		
Affective	0.48**	1.00	
Behavioural	0.61**	0.62**	1.00

**p<0.01

Discussion

Theories of attitudes measurement and a review of previous research on teachers' attitudes toward inclusive education suggested the need for the development of an instrument that measure teachers' attitudes toward inclusive education, one that is reliable, valid and multidimensional and within the realm of inclusive education that includes physical, social and curricular inclusion (Mahat, 2007). *MATIES* was developed in response to such a need.

The pilot sample was small, dictated by principals self-nominating their schools. Although a total pilot sample of 112 is an acceptable minimum for survey research (Borg & Gall, 1989) and for factor analysis (Oppenheim, 1992), it is acknowledged that a larger group is likely to give more reliable results. Further work is needed to include larger groups of teacher, particularly those teaching in lower grades.

Using models within IRT and CTT, eighteen items were retained across the three subscales. The subscales were demonstrated to have content validity by experts during the development stage. Items had highly weighted items that fit within the acceptable range, demonstrating that items had sufficient measurement qualities for the particular subscale. Construct validity were further established as items revealed high factor loadings. Items with high loadings on the first factor deal with behavioural dimension of attitudes, conceived as the teacher's intention to act in a certain manner toward inclusive education. Items with high loadings on the second factor reflect teachers' ideas, thoughts, perceptions, beliefs or opinions about inclusive education, i.e. cognitive aspects of attitudes. Finally, items which loaded on the third factor represent teachers' feelings and emotions associated with the attitudes object. These are teachers' affective responses.

The covariation between the subscales indicated positive correlation, indicating that teachers who hold positive affective and cognitive attitudes would engage in behaviours that support or enhance inclusive education whilst teachers who hold negative attitudes would engage in behaviours that avoid or hinder inclusive education. There are times, however, when people's attitudes are not consistent with their behaviour. For instance, several studies have shown that behavioural intention is closely related to norms of behaviour, i.e. what society thinks one should do (e.g. de Vries, Dijkstra, & Kuhlman, 1988; Nash, Edwards, & Nebauer, 1993; e.g. Triandis, Vassiliou, & Nassiakou, 1968). Hence a teacher, who has favourable attitudes toward inclusive education, may behave in ways that appear negative because the school in which he or she teaches do not support an inclusive philosophy. Further work with larger groups of teachers using *MATIES* is in progress to examine the effects of attitudes, subjective norms and perceived control behaviour toward behavioural intention in inclusive education.

The item and person separation indices and internal consistency reliability of the subscales have also provided substantial evidence of construct and criterion validity and reliability. Hence, the measurement framework, with its extensive steps including early theoretical conceptualisation, development of item pool and response format through to pilot study, data analysis and validation, have provided an effective tool for the development of psychometric subscales which are reliable, valid and stable. A well-validated instrument of teachers' attitudes toward inclusive education is the impetus for fundamental improvement in inclusive education.

The success of any inclusive policy depends upon the cooperation and commitment of those most directly involved, that is teachers. *MATIES* can provide the means to present evidence of such cooperation and commitment. Results of the pilot study provide preliminary evidence to warrant the use of this instrument for the purpose of measuring teachers' attitudes. Further work examining factors that affect teachers' attitudes toward inclusive education may confirm the subscales as an instrument which might record differences and shifts in attitudes. Such evidence may encourage a further acceptance of the changes needed within the structure and organisation of school systems for successful implementation of an education for all.

References

Adams, R. J., Wilson, M., & Wang, W. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement, 21*(1), 1-23.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckman (Eds.), *Action-control: From cognition to behavior* (pp. 11-39). Heidelberg: Springer.

Ajzen, I. (1987). Attitudes, traits and actions: Dispositional predictions of behavior in personality and social psychology. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 1-63). San Diego: Academic Press, Inc.

Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes, 50*, 179-211.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*(4), 561-573.
Avramidis, E., Bayliss, P., & Burden, R. (2000). Student teachers' attitudes towards the inclusion of children with special educational needs in the ordinary school. *Teaching and Teacher Education, 16*(3), 277-293.
Barton, M. L. (1992). *Teachers' opinions on the implementation and effects of mainstreaming*: ERIC Document No ED 350 802.
Bechger, T. M., Maris, G., Verstralen, H. H. F. M., & Beguin, A. A. (2003). Using classical test theory in combination with item response theory. *Applied Psychological Measurement, 27*(5), 319-334.
Berryman, J. D., & Neal, W. R. (1980). The cross validation of the Attitudes Toward Mainstreaming Scale (ATMS). *Education and Psychological Measurement, 40*, 467-474.
Berryman, J. D., Neal, W. R., & Robinson, J. E. (1980). The validation of a scale to measure attitudes toward the classroom integration of disabled students. *Journal of Educational Research, 73*, 199-203.
Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
Borg, W. R., & Gall, M. D. (1989). *Educational Research*. White Plains, New York, U.S.A.: Longman.
Chow, P., & Winzer, M. M. (1992). Reliability and validity of a scale measuring attitudes toward mainstreaming. *Educational and Psychological Measurement, 52*, 223-227.
Cochran, H. K. (1997, November). *The development and psychometric analysis of the Scale of Teachers' Attitudes Toward Inclusion (STATIC).* Paper presented at the Annual meeting of the Mid-South Educational Research Association, Memphis, Tennessee.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
Conatser, P., Block, M., & Gansneder, B. (2002). Aquatic instructors' beliefs toward inclusion: The theory of planned behavior. *Adapted Physical Activity Quarterly, 19*, 172-187.
Cronbach, N. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.
Daane, C. J., Beirne-Smith, M., & Latham, D. (2000). Administrators' and teachers' perceptions of the collaborative efforts of inclusion in the elementary grades. *Education, 121*(2), 331-338.
de Vries, H., Dijkstra, M., & Kuhlman, P. (1988). Self-efficacy: the third factor besides attitude and subjective norm as a predictor of behavioural intentions *Health Education Research, 3*(3), 273-282.
DeVellis, R. F. (2003). *Scale Development* (2nd ed.). Thousand Oaks, California: Sage Publications, Inc.
Friend, M., & Bursuck, W. (1996). *Including students with special needs: A practical guide for classroom teachers*. Boston: Allyn & Bacon.
Green, K., & Harvey, D. (1983). Cross-cultural validation of the attitudes toward mainstreaming scale. *Educational and Psychological Measurement, 43*, 1255-1261.
Hambleton, R. K. (2004, June). *Traditional and modern approaches to outcomes measurement.* Paper presented at the Advances in Health Outcomes Measurement, Bethesda, Maryland.
Hambleton, R. K., & Swaminathan. (1985). *Item response theory: Principles and applications*. Norwell, MA: Kluwer-Nijhoff Publishing.
Hayes, K., & Gunn, P. (1988). Attitudes of parents and teachers toward mainstreaming. *Exceptional Child, 35*(1), 31-38.
Hodge, S. R., & Jansma, P. (2000). Physical education majors' attitudes toward teaching students with disabilities. *Teacher Education and Special Education, 23*(3), 211-224.
Larrivee, B., & Cook, L. (1979). Mainstreaming: A study of the variables affecting teacher attitude. *The Journal of Special Education, 13*(3), 315-324.
Linacre, J. M. (1998). Rasch first or factor first? *Rasch Measurement Transactions, 11*(4), 603.
Mahat, M. (2007). Measurement of teachers' attitudes toward inclusive education. *Manuscript submitted for publication*.
Masters, G. N. (1982). A rasch model for partial credit scoring. *Psychometrika, 47*(2), 149-174.
Masters, G. N., Adams, R. J., & Wilson, M. (1990). Charting student progress. In T. Husen & T. N. Postlethwaite (Eds.), *International Encyclopedia of Education: Research and Studies. Supplementary Volume 2* (pp. 628-634). Oxford: Pergamon Press.
Moberg, S., Zumberg, M., & Reinmaa, A. (1997). Inclusive education as perceived by prospective special education teachers in Estonia, Finland, and the United States. *Journal of the Association for Persons with Severe Handicaps, 22*(1), 49-55.
Nash, R., Edwards, H., & Nebauer, M. (1993). Effect of attitudes, subjective norms and perceived control on nurses' intention to assess patients' pain *Journal of Advanced Nursing, 18*(6), 941-947.

Oppenheim, A. N. (1992). *Questionnaire Design, Interviewing and Attitude Measurement*. London: Pinter Publishers.
Reynolds, W. M., & Greco, V. T. (1980). The reliability and factorial validity of a scale for measuring teachers' attitudes towards mainstreaming. *Education and Psychological Measurement, 40*, 463-468.
Roberts, C. M., & Lindsell, J. S. (1997). Children's attitudes and behavioural intentions towards peers with disabilities. *International Journal of Disability, Development and Education, 44*(2), 133-145.
Roberts, C. M., & Smith, P. R. (1999). Attitudes and behaviour of children toward peers with disabilities. *International Journal of Disability, Development and Education, 46*(1), 35-50.
Schmelkin, L. P. (1981). Teachers' and nonteachers' attitudes toward mainstreaming. *Exceptional Children, 48*(1), 42-47.
Schumm, J. S., & Vaughn, S. (1995). Getting ready for inclusion: Is the stage set? *Learning Disabilities Research & Practice, 10*(3), 169-179.
Semmel, M. I., Abernathy, T. V., Butera, G., & Lesar, S. (1991). Teacher perceptions of the regular education initiative. *Exceptional Children, 58*(1), 9-6.
Sideridis, G. D., & Chandler, J. P. (1995). Estimates of reliabilities for the Teacher Integration Attitudes Questionnaire. *Perceptual and Motor Skills, 80*, 1214.
Subban, P., & Sharma, U. (2005). Understanding educator attitudes toward the implementation of inclusive education. *Disability Studies Quarterly, 25*(2), Available online at: www.dsq-sds.org/.
Triandis, H. C., Vassiliou, V., & Nassiakou, M. (1968). Three cross-cultural studies of subjective culture. *Journal of Personality & Social Psychology, 8*(4 pt2), 1-42.
Van Reusen, A. K., Shoho, A. R., & Barker, K. (2000). High school teacher attitudes toward inclusion. *High School Journal, 84*(2), 7-20.
Villa, R. A., Thousand, J. E., Meyers, H., & Nevin, A. (1996). Teacher and administrator perceptions of heterogeneous education. *Exceptional Children, 63*(1), 29-45.
Wilczenski, F. L. (1992). Measuring attitudes toward inclusive education. *Psychology in the schools, 29*, 306-312.
Williams, R. J., & Algozine, B. C. (1977). Differential attitudes toward mainstreaming: An investigation. *The Alberta Journal of Educational Research, 23*, 207-212.
Wills, D., & Jackson, R. (2001). Report card on inclusive education in Australia. In D. Wills (Ed.), *Education for All: UNESCO Report Card on Inclusive Education in Australia*. Canberra: National Council on Intellectual Disability.
Wilson, M. (2003). On choosing a model for measuring. *Methods of Psychological Research Online, 8*(3), 1-22.
Wilson, M. (2005). Subscales and summary scales: issues in health-related outcomes. In J. Lipscomb, C. C. Gotay & C. Snyder (Eds.), *Outcomes Assessment in Cancer: Measures, methods and applications* (pp. 465-479). Cambridge: Cambridge University Press.
Wright, B. D. (1996). Comparing Rasch measurement and factor analysis. *Structural Equation Modelling, 3*(1), 3-24.
Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*(3), 370.
Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago: Meta Press.
Wu, M. L., Adams, R. J., & Wilson, M. R. (1998). ACER ConQuest: Generalised Item Response Modelling Software. Camberwell: The Australian Council for Educational Research.