Pertanika J. Soc. Sci. & Hum. 24 (4): 1349 - 1379 (2016)

PERTANIKA JOURNALS

SOCIAL SCIENCES & HUMANITIES

Journal homepage: http://www.pertanika.upm.edu.my/

Examining Psychometric Properties of Malay Version Children Depression Invento-ry (CDI) and Prevalence of Depression among Secondary School Students

Teh, J.* and Lim, H. L.

School of Educational Studies, Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia

ABSTRACT

This study aimed to examine psychometric properties of a Malay version of the Children's Depression Inventory (CDI) by using Item Response Theory, particularly the Rasch model, among secondary school adolescents. The Survey method was employed in this quantitative study by distributing the self-report Malay version of the CDI (MCDI) questionnaire to a total of 500 secondary school adolescents in two randomly chosen daily, government (public) schools. The criteria for selecting respondents was that they possessed a minimum competency in the Malay language, were aged 14 or 16 years and were enrolled in secondary schools situated in the selected geographic area. Respondents meeting these criteria were then randomly chosen for this study. This validation study revealed that MCDI items pointing toward one dimension based on item MNSQ, PTMEA values, Rasch residual based on principal component analysis, adhered to Rasch model's expectation, good item separation index (7.30) and very high reliability (0.98). Independent T-test result revealed that females and males differed in terms of the MCDI total score. Female respondents were more susceptible to depression as compared to their male counterparts. School form was not a significant predictor for depression as evidenced by insignificant T- test results.

Keywords: Malay version Children Depression Inventory (MCDI), Rasch model, secondary school students; psychometric properties

ARTICLE INFO

Article history:
Received: 7 July 2015
Accepted: 15 June 2016

E-mail addresses:
julieteh85@yahoo.com, julieteh85@gmail.com (Teh, J.),
hllim@usm.my (Lim, H. L.)
* Corresponding author

INTRODUCTION

Depression, a treatable psychiatric disorder (Malaysian Mental Health Association, 2014), has been studied extensively both in overseas and local contexts. Globally, it has received much attention from researchers and health care personnel in

ISSN: 0128-7702

recent decades. According to the World Health Organization's (WHO) Department of Mental Health and Substance Abuse, up to 350 million people are estimated to be affected by depression (Marcus et al., 2012; WHO, 2012). WHO also projected depression to be the second leading cause of disability worldwide. This number significantly places a great burden (Marcus et al., 2012) on the existing global crisis of non-communicable diseases and affects all communities across the globe (Marcus et al., 2012; Ferrari et al., 2013). Considerable efforts have been made in studying the disorder's burden both at an international and individual country level. For example, Ferrari et al. (2013) analysed the burden of two subtypes of depression-Major Depression Disorder (MDD) and Dysthymia and presented severity proportions, burden by country, region, age, sex, and year. The findings from this study indicated that depression disorders were a leading cause of disease burden in the 1990 and 2000 GBD studies, in which the statistic had increased by 37.5% in 20 years due to population growth and ageing.

Based on results from the third National Health and Morbidity Survey (NHMS) (2006), approximately 12% of Malaysians aged between 18 and 60 years old were suffering from some forms of mental illness as compared to 10.6 % in 1996. Of those affected by mental illness, depression was ranked number one and made up 2% while psychosis accounted for 1%, worrying 1.8% and the rest accounted for anxiety disorders and other types of mild mental diseases. In 2008, mental disorders were estimated to contribute 16.8 % of the global burden of disease (WHO, 2008 in WHO, 2011). Mental illness in the local Malaysian context has reached an alarming stage when the latest statistics are compared to those of previous surveys. The statistics involving adolescents aged 15 years and below rose from 13% in 1996, 19.4% in 2006 and 20% (1.0 million) in 2011 (The Star, 2011; NHMS, 2011). Other data reported by the former Health Minister, Datuk Seri LiowTiong Lai (The Star, 2011) also pointed towards the rising number of mental disorders among adolescents in the country. It was found that the number of people seeking medical attention for mental illness was 10.6% in 1996 but this number had increased to 11.2%, with 6.4% having considered suicide, in 2006. A high percentage of those people seeking medical treatment for mental illness was from those aged 16 to 19 years and the elderly aged between 70 and 74 years (The Star, 2011). The same pattern of increasing mental health problems among adolescents was also reported by Dr. Lokman Hakim Sulaiman, Deputy Director of General Health (Public Health). Quoting data from the Healthy Mind Programme survey conducted among Form Four (grade 10^{th}) students in Malaysia he indicated that approximately 7% of 19,919 students from 157 schools showed signs of severe and extremely severe stress, anxiety and depression based on the Depression, Anxiety and Stress Scale (DASS) questionnaire(New Straits Times, 2014).

The scope of research studies has not merely been on examining prevalence rates. Causative factors, age of onset of depressive symptoms and their negative effects on all aspects of life have also been investigated using cross sectional or longitudinal methodologies. Results of longitudinal studies of mental disorders have revealed that depression usually starts during adolescence (Hankin, 2006; Van Lang et al., 2007). The main reason given for this finding is that adolescence is a transition period from childhood to adulthood which demands great coping skills and which is characterised by many cognitive , emotional, psychological and behavioural changes. Such demands and changes might result in this aged group feeling helpless, negative and depressed (Mazza et al., 2010; Mey, 2010). This result is consistent with results of a study by Allen et al. (2007) which showed that the increased prevalence of this mental disorder among adolescents and the life time prevalence rates were approximately 20% to 25%, equivalent to that of the adult population in his study of a United States population (Kessler et al., 1994).

In terms of gender, research studies either at international or local context consistently reported difference in the prevalence of depression. For instance, a study of a sample of 1611 secondary school students conducted in Iran (Daryanavard et al., 2011), found that the female to male ratio of depression occurrence was 1.72 to 1. Similarly, a cross sectional study in Malaysia (Adlina et al., 2007) also reported that female secondary school students were more depressed than their male counterparts. This finding was based on the fact that the females' mean CDI T-score of 54.73 was significantly higher than the males' CDI T-score of 51.92. In contrast, Weiss and Weisz (1988) found no significant gender difference between males' and females' CDI scores based on their one way ANOVA result.

Due to time and cost factors, depression in children and adolescents has over the last few decades been studied and measured using self-report approaches. Using a self-report approach, youths (excluding those with reading difficulty) are typically asked to endorse a series of statements that best describe their behaviour or characterizes their emotions with regard to a specific symptom of depression during the past two week period. Alternatively, the reading of statements in the questionnaire can be assisted by researchers for those who cannot read. Additionally, there are studies that involve multiple informants; they are children, their parents and/or teachers. Using this approach, self-reports are completed by the children while the other informants complete different forms that rate depressive symptoms in the children based on the informants' observations of the child's behaviour in the school setting and/or in the home environment. This research approach of involving teachers and/or parents is paramount in obtaining more accurate information pertaining to a child's mental health status while at the same time establishing the concurrent validity of the

study conducted. In summary, both methods have their own pros and cons and the choice of approach will depend mainly on the researchers' decision.

Among the variety of self-report instruments available, the CDI has been widely used in studying children and adolescent depression (Kovacs, 1992). The development of this instrument was based on the Beck Depression Inventory (BDI) (Kovacs, 1992) for adults. The CDI requires a minimum reading ability level, has a short administration time of about 15 minutes and targets those aged 7 to 16 years (Kovacs, 1992). These features give rise to its popularity among researchers examining depressive symptoms among youth in non-clinical (schools) and clinical settings worldwide. The psychometric properties of the instrument have been extensively studied using Classical Test Theory (CTT) (Goodchild & English, 2004). Classical test theory is the earliest theory of measurement. The major focus of this theory is estimating the reliability of the observed scores of a test. Examinee observed scores will always depend on the selection assessment items from the domain of assessment items over which their ability scores are defined. Examinees will have lower scores on difficult tests and higher scores on easier tests (Lord & Novick, 1968).

The main reason for examining the CDI's internal structure is to determine the instruments' ability to adequately fulfil two major criteria - validity and reliability. In other words, if the instrument is well designed it will produce psychometrical sound results.

The CDI was initially designed for English speaking samples, but now it has been translated into many other languages to suit the needs of samples from different countries around the world. A review of the literature indicates that the translated versions of the CDI possess different factor structures as compared to the original English version which has only five factors. The factor analysis studies of the translated CDI have produced incongruent results; some yielded as many as eight factors (Monreal, 1988) whereas some studies found only two factors (Del Barrio & Carrasco, 2004). To date, only two validation studies have been conducted of the MCDI in Malaysia and both used children and adolescent samples attending psychiatric clinics (Rosliwati et al., 2010; Tan et al., 2013). The techniques employed in both these studies were based on the CTT method to examine the construct validity of the MCDI.

However, construct validity is an ongoing endeavor that requires the collection of evidence from multiple sources and various backgrounds of samples to determine the psychometric merit of the instrument (Messick, 1995). Messick argued that construct validity essentially must be evaluated in terms of the mosaic of evidence that imbues a person's test scores with meaning. As Bond and Fox (2015) stated, the test score (data) must be able to relate the items and persons together in a coherent and integrated way that represents the construct examined.

Over the last two decades, however, the concern in the health and psychological fields of modern test theory such as those

based on Rasch's measurement model has increased (Shea et al., 2009). More and more research studies have shifted from using the classical approach to Item Response Theory (IRT) given that CTT is weak and the explanation power is not as robust as IRT (Hambleton & Jones, 1993). This is due to two reasons. First, item difficulty and item discrimination are both sample dependent (Hambleton & Jones, 1993; Bond & Fox, 2001). This means that the responses endorsed on the questionnaire itself are affected by the nature of samples recruited in the validation studies (Bond & Fox, 2001). The conclusion of whether statements in questionnaires can be understood in the same way by samples that differ in terms of their cognitive ability is doubtful. A good instrument should be working the same way regardless of the samples and their underlying cognitive ability trait. Second, modern test theory, including the Rasch measurement model (one of the models under IRT), has unique features that can overcome the flaws of CTT. This theory has very robust explanation power (Hambleton & Jones, 1993) because it enables the estimation of the location of item difficulty and samples' ability along an interval scale that uses a unit of measurement called logit (Bond & Fox, 2001). In other words, the results obtained are more meaningful because the instrument works exactly the same way for every sample, no matter whether they are from low or high average knowledge level.

The considerable advantages of IRT has prompted researchers to examine various depression scales including the Hospital Anxiety and Depression Scale (HADS) (Gibbons et al., 2011), the Depression, Anxiety and Stress scale (DASS) (Shea, Tennant & Pallant, 2009), the Beck Depression Inventory (BDI) (Lerdal et al., 2014), to name a few. Being a downward extension of the BDI, the CDI (Kovacs, 1992) has not been extensively examined by means of the Rasch modern psychometric approach. Only one previous study could be found that examined the psychometric properties of the original English version of the CDI using the Rasch model and this study was based on a non-clinical sample of adolescents from the United States (Lee et al., 2011). Other than this particular study, there has been no other similar study conducted in the Malaysian context. Therefore, conducting a study to examine the psychometric properties of the MCDI is necessary before the instrument can be applied to youth of different cultures in the Malaysian context.

AIMS OF THE STUDY

This study aimed to examine the psychometric properties of the MCDI in measuring depression of secondary school adolescents in two selected daily government schools. Additionally, the study also attempted to determine if there are statistically significant differences in CDI total scores according to gender and class form (school year level). Specifically, the objectives of this study were:

1. To examine the psychometric properties of the MCDI by investigating the

dimensionality, item fit statistics, point measure correlation, Rasch residual based principal components analysis, response categories functioning (category frequency, average measure and category threshold), reliability and separation indices, and appropriateness of item difficulty based on a person-item map.

2. To investigate the prevalence of depression among adolescents attending daily government (public) secondary schools.

MEASUREMENT THEORY: ITEM RESPONSE THEORY

Item response theory (IRT) is the modern test theory and also known as latent trait theory (Baghael, 2008). The name item response theory is due to the focal point of the theory being on the item. IRT initially was introduced to the psychometric field with the aim of assisting in ability assessment of students. The emergence of IRT has greatly influenced how test items are created, calibrated and evaluated, and the conclusions that can be drawn about students' learning progress based on the test scores obtained. It is known as the study of test and scores of items based on assumptions concerning the mathematical relationship between a person's abilities and item responses. In brief, the theory is mainly concerned with test scoring and the development of test items (An & Yung, 2014). Generally, test items are developed in educational and psychological fields with the intention to measure various types of abilities such as mathematical ability, and traits such as personality or behavioural characteristics including depressive tendencies. In addition to the benefits of IRT to the educational field, it is also widely utilized in other disciplines such as medicine, marketing, health sciences, psychology and business (An & Yung, 2014).

While there are many models associated with IRT (e.g. the Rasch model, the 2 parameters logistic model and the 3 parameters logistic model) the present study uses only the Rasch model.

The Rasch Measurement Model

The Rasch model, named after George Rasch, is a psychometric model that can help transform raw data obtained from the human science assessment into equal interval scales. The equality of interval scales is achieved through log transformations of raw data odds and abstraction is accomplished through probabilistic equations (Bond & Fox, 2001). Another feature of the logit scale is that the distance between units on the scale has a consistent value of meaning (Bond & Fox, 2001). This person-item logit scale describes clearly the location of sampled individuals and items of the instrument along the scale (Bond & Fox, 2001).

The Rasch model has a similarity to another of the models in IRT: one parameter logistic model. This model has the fewest parameters: a person trait or underlying ability and an item difficulty parameter, which estimates the degree to which a trait is measured by an item. The Rasch

model is considered as one parameter model where only the item difficulty parameter is taken into account. However, the collected data must first fit the model as the essential prerequisite, not the other way around. In sum, the Rasch model provides approximations of measures which aid in the understanding of the processes underlying the reason why sampled individuals and items of an instrument behave in a particular way. These approximations of measures aid in solving the problems that are impossible to be solved with any other models.

i. Dimensionality

One of the advantages of the Rasch model is that it builds a hypothetical unidimensional line along which items and persons are located according to the item difficulty and person ability measures (Baghael, 2008). The items that fall close enough to the hypothetical line contribute to the measurement of the single dimension which defines construct theory. Those that fall far from it are measuring another dimension which is irrelevant to the main dimension. Messick (1989) stated that the irrelevant dimensions or sub-dimensions always produce reliable (reproducible) variance in test scores that threaten the construct validity of the instrument. In short, the Rasch model measures one attribute or dimension at a time so that a useful and meaningful composite description can be developed. This focus on an attribute or dimension at a time in measurement is referred to as unidimensionality (Bond & Fox, 2001).

The Rasch assumption of unidimensionality can be measured by the following item statistics: (a) item fit mean square (MNSQ), (b) point measure (PTMEA) correlation and (c) Rasch residual based principal components analysis (PCA).

Item fit mean square (MNSQ) indicates how well each item in the MCDI represented a single underlying dimension, that is depression. MNSQ is calculated by dividing the chi-squared statistic by its degree of freedom and the ratio scale form with expected values of +1.0 (Linacre & Wright, 1994b; Bond & Fox, 2001). Both infit and outfit MNSQ values are always positive (i.e., >0). Items with mean-square values greater than 1.0 indicate underfit to the Rasch model, i.e., the data are less predictable than the model expects, whereas items with mean-square values less than 1.0 indicate overfit to the Rasch model, i.e., the data are more predictable than the model expects (Linacre & Wright, 1994b). These statistics are derived for every item and provide information on the consistency of the responses to each item in MCDI. Point measure (PTMEA) correlation is important to determine if items in the MCDI function as intended. The PTMEA values ranges from -1 to +1, with negative values indicating that items are improperly scored or they do not function as intended (Linacre, 2005a; in Lee et al., 2011). Rasch residual based principal components analysis (PCA) is a statistical test to investigate the dimensionality of the MCDI. Under the Rasch model conditions, the residuals associated with any item are presumed to have no relationship to any other

item in the questionnaire. Consequently, each item's residuals comprise a unique factor meaning that the values in the diagonal of the correlation matrix would be set to zero (Linacre, 1998).

ii. Response Categories

The test of response categories investigates whether each response category in the MCDI (namely, 0 for no symptoms, 1 for mild symptoms, 2 for serve symptoms) is adequately used or function as intended. In the Rasch model, it is achieved by examining the category use statistics (for instance, frequencies of category and the average measures) and also thresholds of each.

Categories frequencies denotes the total number of respondents who endorse a particular response category across each item (Bond & Fox, 2001). According to Linacre (1999a), each category in the MCDI should have a minimum of 10 responses (Bond & Fox, 2001). The average measure, on the other hand, denotes the empirical mean of the measure across categories. It is defined by Linacre (1995) as "the average of the ability estimates for all persons in the sample who chose that particular response category, with the average calculated across all observations in that category" (cited in Bond & Fox, 2001). The average measure values are expected to increase monotonically across categories because the high category endorsement pattern is from a high mean. Category thresholds are boundaries between categories, the measures where adjacent categories are equally probable and they should also increase monotonically (Bond & Fox, 2001).

iii. Reliability and separation indices

The Rasch measurement model provides indices which assist researchers to examine the adequacy of MCDI items that stretch along the continuum as opposed to clumps of them and also sufficient spread of ability among persons (Bond & Fox, 2001). Person reliability index indicates the replicability of person ordering expected if this research sample of persons were made to answer another set of items that measure depression (Wright & Masters, 1982; in Bond & Fox, 2001). Item reliability index, on the other hand, refers to the replicability of item placements along the pathway if this pool of MCDI items were given to another sample which has equivalent ability levels (Bond & Fox, 2001).

Separation indices estimate the spread of items, or persons, on the measured variables. They are estimated by the ratio of the person or item adjusted standard deviation to the root-mean-square standard error (RMSE), the error standard deviation (Linacre, 2014). These indices provide information of separation in standard error units, with adequate separation in persons and items, measured with an estimate of at least 2.0.

iv. Person-item map

Wright map is easily interpreted when researchers are familiar with what constitute the left and right columns. The left-hand column locates the person ability measures

along the variable. The persons often have a normal distribution (Linacre, 2014). The right-hand column, on the other hand, locates the item difficulty measures along the variable. By ordering item difficulty and person ability on the same linear continuum, researchers can truly understand where a students falls on the scale. If two students received identical raw scores on the MCDI, researchers might conclude they have the same ability. However, by examining the test items in hierarchical order, the researchers might conclude that one individual has a higher ability because he or she may have correctly answered more challenging items. In addition to establishing students' ability, the item hierarchy also allows researchers to determine whether the MCDI is appropriate for the secondary students.

CHILDREN DEPRESSION INVENTORY (CDI)

The Beck Depression Inventory (BDI) (Beck,1960) was used as a model for developing the Children's Depression Inventory (CDI) (Kovacs, 1992). The former scale is a clinically-based, 21 items, self-rated instrument used to screen depression symptoms among adults. Due to the BDI already being in existence and mainly targeting adult group, there was a need to create another similar scale for children and youth. In addition, depression in children and adolescent are easily overlooked and difficult to diagnose due to different symptoms attributed to age and developmental functions (Harrington et al., 1996). Cognitive, social, emotional and biological changes that transpire over time from childhood to adolescence were proposed as causing phenomenological manifestations of depression symptoms (Cicchetti & Toth, 1998; Weiss & Garber, 2003). Researchers also suggested that "younger children may not have developed the requisite cognitive, social, emotional or biological capacities to experience certain typical adult depression symptoms and secondly, the causes and/or consequences of depression may change across different developmental periods" (cited in Hankin, 2006). Hence, Maria Kovacs, an American clinical psychologist, developed the CDI by changing the wordings of BDI items to make CDI items specifically appropriate for youth aged 7 to 17 years old (Weiss et al., 1991; in Huang & Dong, 2013) and first published it August 1979 (Kovacs, 1992). Items developed can be grouped into 5 factors, including "negative mood", "interpersonal problems", "ineffectiveness", "anhedonia" and "negative self-esteem" (Kovacs, 1992). Descriptions of the subscales of the self – report CDI form are illustrated in Table 1.

Based on the description and review of the literature discussed above, the following is a diagram of the theoretical framework created and which serves as a guideline for this study. The diagram illustration is attached in Figure 1. All 27 items of the MCDI were thoroughly examined by the IRT approach, Rasch model, for its psychometric properties such as dimensionality, response categories, person and item reliability and separation indices as well as the appropriateness of item difficulties for the

recruited respondents based on person-item map. Additionally, the prevalence of depression among the secondary school adolescents was investigated along with significant differences of depression between gender and forms.

Previous prevalence studies and research related to the examination of psychometric properties of CDI instrument

CDI has been frequently used by researchers around the world to study depression in children and adolescents in both clinical and community samples (Friedberg & McClure, 2002; Cole & Martin, 2005; Lee et al., 2011). Specifically, its usage in research settings is based on the underlying assumption of the CDI, that is the depressive symptoms experienced by children especially adolescents can be described through the similar symptoms of adults (American Psychiatric Association, 2000), for instance sadness, guilt, anhedonia (inability to experience pleasure from activities usually found enjoyable) (Dorland, 2009), chronic fatigue, disturbed sleep, social withdrawal and suicidal ideation (American Psychiatric Association, 2000).

Regardless of its popularity of being used in both clinical and research studies and several empirical studies that have proven the dimensionality of the CDI scales, many criticisms arose pertaining to its usage and assumptions due to the incongruent results obtained from factor analysis studies. The assessment of factor structure underlying the CDI with community and clinical samples varied distinctly. These contrasting findings of the internal structure of the CDI scale can be traced back to research studies undertaken previously. Huang and Dong (2013) conducted a meta-analysis of

Table 1
CDI subscales and their definitions

Number	Factors	Description
1	Negative mood	This subscale reflects feeling sad, feeling like crying, worrying about bad things, being bothered or upset by things and being unable to make up one's mind.
2	Interpersonal problems	This subscale reflects problems and difficulties in interactions with people, including trouble getting along with people, social avoidance and social isolation.
3	Ineffectiveness	This subscale reflects negative evaluation of one's ability and school performance.
4	Anhedonia	This subscale reflects "endogenous depression", including impaired ability to experience pleasure, loss of energy, problems with sleep and appetite, and a sense of isolation.
5	Negative Self-Esteem	This subscale reflects low self-esteem, self-dislike, feelings of being unloved, and a tendency to have thoughts of suicide.

(Cited from Children's Depression Inventory Manual, 2003 in Adlina et al., 2007); Kovacs, 2004 in Pscyhological Assessments Australia, n.d.).

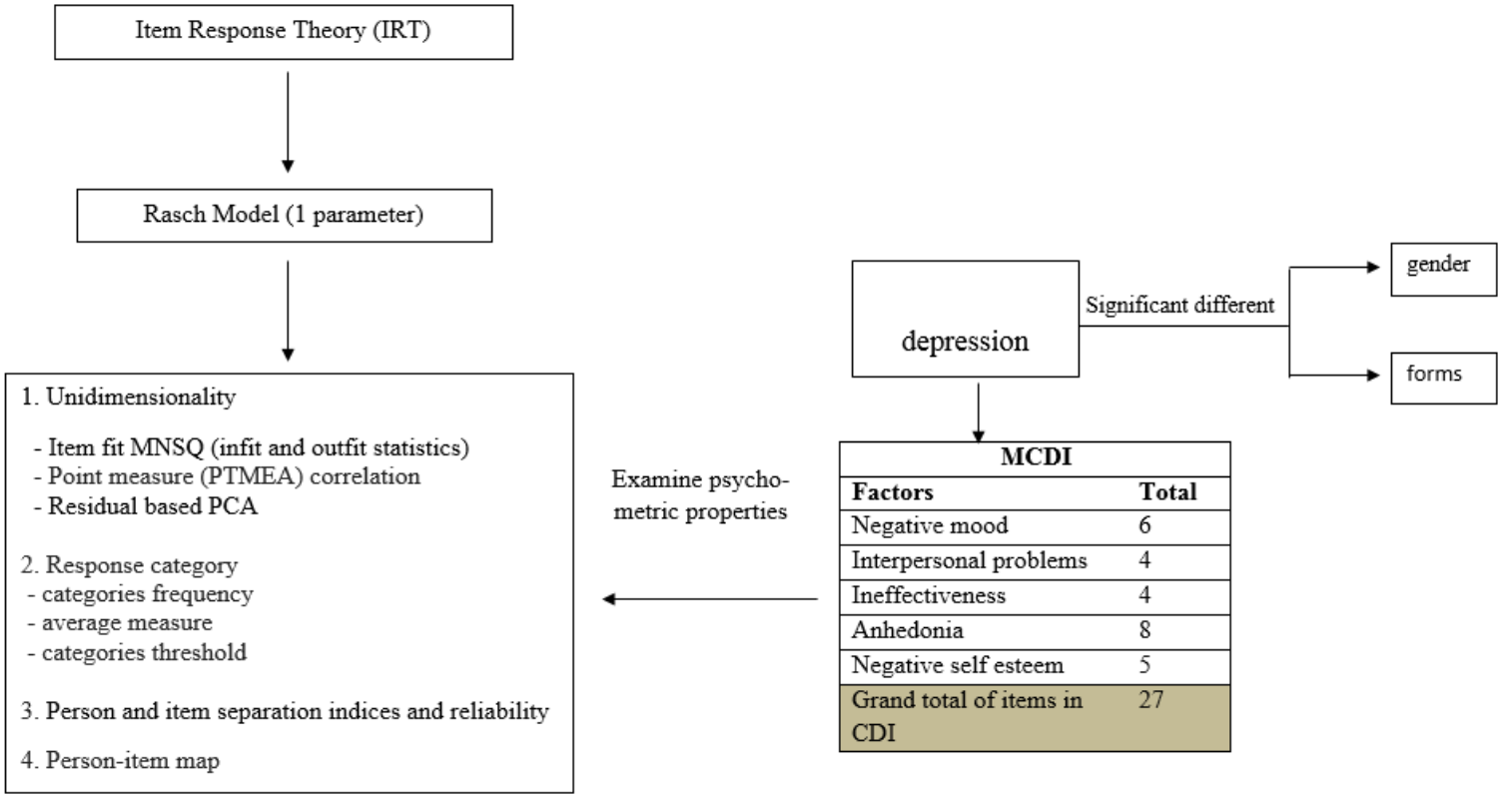


Figure 1. Theoretical framework of the study

the dimensionality of the CDI, involving 24 studies, comprised of 35 independent samples with 18099 participants. They found that the English version of the CDI scale consists of 5 factors whereas non-English versions best comprised only 4 factors. Some researchers found a slightly higher number of factors and argued that the CDI actually consists of 6 factors (Craighead et al., 1998; Gomez et al., 2012), which is different from the 5 factors reported by the original author (Kovacs, 1992). Another study reported as many as 8 factors (Monreal, 1988). Thus, the conclusion from all these CDI structure studies is that the number of factors ranged from 2 to 8.

Despite the shortcomings detected in the scale, especially its internal structure, the popularity of the CDI for use in investigating children and adolescent depression cannot be denied. Many researchers have investigated the psychometric properties such as reliability in the form of test-retest reliability, Cronbach's alpha internal consistency estimates by using Classical Test Theory (CTT) in their research studies (Kovacs, 1985).

The Malay version sample form was validated by Rosliwati et al. (2010) in a study of children and adolescents attending outpatient clinics in Kota Baru, Kelantan state of Malaysia. The psychometric properties reported was adequate as evidenced by Cronhach's alpha value of 0.83, sensitivity of 93% and specificity of 96% based on the area under the ROC (Receiver Operating Characteristic) curve. The study also suggested a cut-off score of 18 to differentiate clinical samples from normal adolescents in the community.

In another validation study of the MCDI conducted in Malaysia by researchers Tan et al. in 2013 who recruited adolescents aged between 12 to 17 years from 6 hospitals in different regions of Peninsular Malaysia, it was found that both reliability and validity values were acceptable. Internal consistency in terms of Cronbach's alpha value was 0.849, similar to findings from another study by Rosliwati et al. (2010). The sensitivity and specificity were 55.6 and 54.4 respectively while the recommended cut-off score based on ROC curve was 17, which was one point less than findings reported by Rosliwati et al. (2010). The scale had good discriminant validity as the scale was able to differentiate between normal and clinical adolescents, with a higher total mean score than those positively diagnosed with depression.

Previous prevalence studies and research related to the examination of the psychometric properties of the CDI from a Classical Test Theory (CTT) approach, as well as several research findings have been described above. The popularity of CTT in validation studies may be due to the reasons of weak assumptions and easily met criteria of the test theory as compared to those from the Item Response Theory (IRT) (Hambleton & Jones, 1993). Despite the easier interpretation and less stringent assumptions of CTT, the true score obtained in the test or questionnaire depends solely on the content of the instrument and the sampled individuals, therefore there is no absolute characteristics of the persons and items.

Under such circumstances, researchers are confronted with problems should they want to compare research findings between groups, not to mention results obtained from different instruments, even where those instruments were similarly operationalized to measure the same construct of depression.

METHODOLOGY

From a total of 25 secondary schools in Kulim district, a simple random sampling technique was used to choose 2 target schools. At each school chosen, this study employed purposive sampling method, involving only Form 2 (8th grade) and Form 4 (10th grade) as research samples. Form 3 and Form 5 students are not encouraged by the Ministry of Education to be involved in research study as they involve with national examinations. At the class level, a simple random sampling technique was again used to determine which classrooms of Form 2 and Form 4 would be involved and given the self-report questionnaire. The subjects in this study were 500 secondary school adolescents (250 from each school) aged 14 (Form 2) and 16 (Form 4) years of both genders.

The breakdown of respondents according to gender and school form are shown in Table 2. According to race, Chinese students constituted the highest number of participates followed by Malay and Indian. There was a total of 288 female respondents.

The Children's Depression Inventory (CDI), Malay language version, was used in this study. The Malay version was translated and validated through factor analysis by Rosliwati et al. (2010). The Malay version of the CDI is a self-report questionnaire consisting of 27 items with 3 response scales of 0 (no symptom), 1 (mild symptom) and 2 (severe symptoms), on which respondents rated themselves based on what they thought and experienced over the past 2 weeks. The scale is made up of five factors: 'Negative Mood' (6 items), 'Interpersonal Problems' (4 items), 'Ineffectiveness' (4 items), 'Anhedonia' (8

Table 2
Respondents by school form, gender and race (N=500)

School form			Race			Total
			Chinese	Indian	Malay	
2	Gender	Female	52	57	34	143
		Male	40	39	28	107
		Total	92	96	62	250
4	Gender	Female	70	13	62	145
		Male	61	12	32	105
		Total	131	25	94	250
Total	Gender	Female	122	70	96	288
		Male	101	51	60	212
		Total	223	121	156	500

items) and 'Negative Self Esteem' (5 items). The total score of depressive symptoms is the sum of all the separate item scores and ranges from 0 – 54, where the maximum score of 54 indicates the more clinically severe level of depression (Kovacs, 1992). However, while a cut-off score of 19 or 20 is generally accepted on the CDI, it is not an absolute, could be identified as severely depressed (Kovacs, 1992). This high cut off-score used in the non-clinical sample aimed to minimize false positives (i.e. children and adolescents who are falsely classified as depressed by means of the CDI) (Kovacs, 1992).

Researchers briefed the respondents on the nature and purpose of the study in their classrooms during the regular school day. Respondents were also informed of their rights to withdraw from the study at any point in time should they feel they were not comfortable while filling in the survey form, with no harm or risk associated with their withdrawal. This study obtained ethical approval from the Universiti Sains Malaysia (USM) and the Malaysia Ministry of Education (MOE), Division of Planning and Policy Research. After permission was granted by MOE, the application process proceeded to Kedah State Education Department. Permission to use the MCDI was obtained from the original author, Dr Kovacs and Multi Health System (MHS) publisher via email. Then, with the two official letters from MOE and State Education Department, researcher went to the randomly selected daily government (public) secondary schools and obtained the school headmasters' permission prior conducted both the pilot study and the actual field study. The IBM Statistical Package for the Social Sciences (SPSS) Statistics version 21.0 and Winsteps version 3.66 were used in this study. SPSS was used for descriptive statistics analysis and Winsteps for IRT-Rasch analyses.

RESULTS

All items in MCDI had acceptable infit and outfit statistics, either above 0.60 or below 1.4 threshold values (for Likert scale) for respondents from both school forms (See Table 5). These results fell within the acceptable threshold values, suggesting there existed no redundancy and heterogeneity of items in the Malay version of the CDI for the respondents.

Regarding point measure (PTMEA) correlation, all 27 items in the MCDI for both Form 2 and Form 4 respondents exhibited positive and moderate to strong PTMEAs and ranged from 0.26 to 0.55. All items were acting as expected with regard to the underlying construct but not multidimensional.

Figure 3 shows 26.9% of the variance was accounted for by the unidimensional model which closely matched the modelled value of 27.5%. Additionally, the unexplained variance in first contrast had an eigen value of 1.8 and accounted for 4.9% for the unmodelled data. These results further suggested that the MCDI scale for both Form 2 and Form 4 respondents could be considered unidimensional since the eigen value of the first contrast recorded less

than 5% of the unmodelled variance.

In general, each and every evidence of item statistics [item fit mean square (MNSQ), point measure (PTMEA) correlation as well as Rasch residual based principal components analysis (PCA)] supporting the fact that the MCDI items appeared to be unidimensional for both Form 2 and Form 4 respondents in this study.

From the analysis, it was noted that there were five items in which the response scale had a response of less than 10. They were item 3, 5, 9, 26 and 27. The scale of 2 (indication of clear symptoms) in all these five items had responses of 3 (item 3), 8 (item 5), 7 (item 9), 8 (item 26) and 4 (item 27), following the sequence of items mentioned above. Other than these five items, the remaining items in MCDI had more than the recommended minimum of 10 responses (Linacre, 1999a in Bond and Fox, 2001) for each 0 (no symptoms), 1 (mild symptoms) and 2 (severe symptoms) scale categories (refer appendix 1).

Average measure analysis showed that generally all measure values increased monotonically (ranged from -2.29 to + 2.29) across the response categories of 0, 1 and 2. Category threshold analysis findings revealed that thresholds (-1.11 and 1.11) and all three response category scales of 0, 1 and 2 across all items in CDI had their respective distinct peaks and intersections, further suggesting that they were good functioning category rating scales for both Form 2 and Form 4 respondents.

Additionally, each category's outfit MNSQ statistics were less than 2.0, indicating they provided information rather than noise in the measurement process

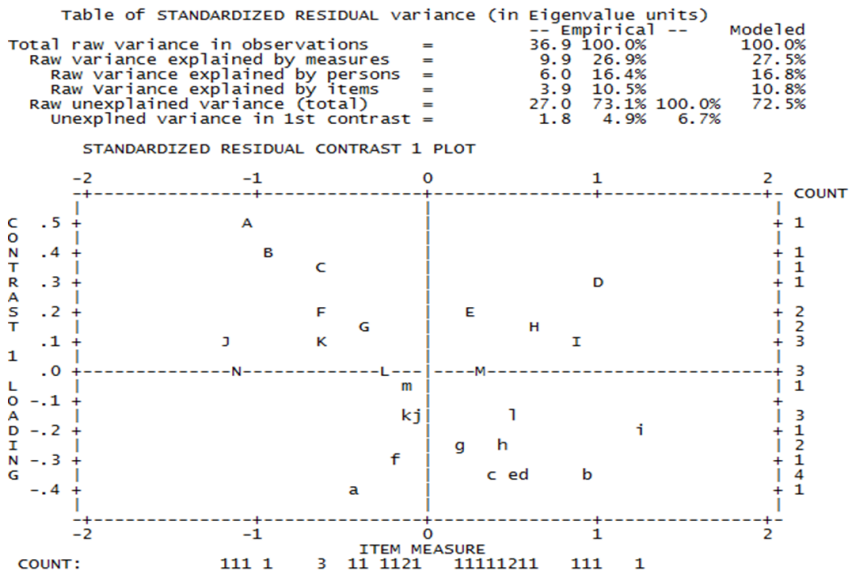


Figure 3. Plot of standardized Rasch residual based principal components analysis (PCA) for respondents (N=500).

(Linacre, 1999a). The results indicated that the three-category rating scale is appropriate for the MCDI.

Person separation reliability value for respondents is given in Table 3. The respondents of these two school forms had an index of 1.90. Both school forms respondents had a person separation value less than 2.00 threshold and concurrently, the reliability estimate was slightly lower than 0.80 threshold. These two values indicate that the MCDI inadequately differentiate between individuals in the study.

Similar to the person separation reliability estimate, results pertaining to the items is presented in Table 4. The respondents had an item reliability index of 0.98, much above the 0.80 threshold. In addition, the sample also had item separation index which was very much greater than the 2.00 threshold, thus, these measures indicated good separation between items used in the translated version of the CDI.

Person latent trait (depressive symptoms) and item difficulty (known as item measure) were arranged following the sequence of highest to lowest; hence, respondents with a higher level of depression and items gauging a more severe degree of depression were located at the top of the Wright map (See Appendix 2). Item difficulty distribution was found higher than the person ability (depressive symptoms level) distribution. Item difficulty covered a range of approximately – 1.17 to + 1.26 logits (within 2 standard deviations) while person ability covered a range of approximate - 6.85 to + 1.57 logits (coverage of more than 2 standard deviations).

The Wright map (See Appendix 2) showed that for both the respondents of Form 2 and Form 4, there was not a good deal of overlap (or not symmetrical at vertical axis) between the ranges of all item difficulties with the respondents' range of latent trait parameters. This further indicated that approximately half of the MCDI items which measured the most severe degree of depression were too severe for the samples targeted. Respondents who had higher levels of depression overlapped with MCDI items measuring less severe levels of depression. In addition, from the Wright map, it also illustrated that those respondents with mild to low levels of depression had no overlap with any of the MCDI items. This finding indicates a floor effect for the MCDI items presented to the respondents in the study.

The least likely to be endorsed factor was interpersonal problem. This was checked by just 02. % of the sample for item 27 and 0.6% of the sample for items 5 and 26. The most likely to be endorsed factor was ineffectiveness. It was checked by 34.4% of the sample for item 24 and 26.4% of the sample for items 15 and 23. Generally, interpersonal problem has the lowest calibration, followed by negative self-esteem, negative mood, anhedonia and ineffectiveness, in ascending order. The person mean is approximately 1.5 logits lower than the item mean. This indicates that the majority of the sample was not very severely depressed.

The distribution of total MCDI score (raw score) transformed to comparable T scores is presented in Appendix 3. The highlighted two rows with the raw score of 20 and above for females and 25 and above for males, equivalent to a T score of above 65 and indicates positive for depression.

In addition, there is evidence that female respondents significantly obtained much higher total MCDI scores as compared to their male counterparts at almost each different total MCDI score (raw score) level, as is presented in Appendix 4. Besides that, independent sample T-tests revealed that female respondents had higher mean total MCDI scores, 13.56 (*SD*=6.904) compared to males 12.03 (*SD*=6.303) with a *t* value of 2.547 (*df*=498, *p*=.011). This result indicates that the difference in MCDI total score between genders was statistical significant with *p* value less than 0.05 (See Appendix 5). Thus, it also can be concluded that female respondents from both school forms obtained significantly different total MCDI scores than male respondents.

The descriptive statistic showed respondents of different school forms differed in terms of the total MCDI score (raw score). In terms of mean total MCDI (raw) score, the opposite pattern was observed with Form 4 respondents recording a higher mean score (13.20) compared to Form 2 respondents (12.62). However, when the analysis was subjected to independent t-test, results revealed a non-significant *t* value of -0.976, *df*= 498, *p*>0.05 (See Appendix 6). The conclusion can be derived from the findings is that the respondents of Form 2 statistically obtained no difference in CDI total score as compared to those Form 4.

DISCUSSION

The acceptable fit statistics of this study were not consistent with another three studies which utilized fit reference values of 0.6 to 1.5. Firstly, Lee, Krishnan and Park (2011) found two items with poor infit and five items with poor outfit. Secondly, Sauer et al. (2013) reported two items with

Table 3
Person summary statistics for respondents (N=500)

	Average ZSTD (*SD*)					
Measure	Infit	Outfit	Adjusted *SD*	RMSE	Separation	Reliability
-1.62	1.02 (0.00)	1.00 (0.00)	0.83	0.44	1.90	0.78

Table 4
Item summary statistics for respondents (N=500)

	Average ZSTD (*SD*)					
Measure	Infit	Outfit	Adjusted *SD*	RMSE	Separation	Reliability
0.00	1.00 (0.00)	1.00 (0.00)	0.67	0.09	7.30	0.98

Note: RMSE= root mean square error; ZSTD= z-standardized fit statistic.

outfit slightly exceeding the recommended threshold of 1.50 and Thirdly, Shea et al. (2009) reported item five in the depression scale showed (4.76) very high misfit in the first analysis. In terms of PTMEA, this study finding was consistent with Lee et al. (2011) whose values of PTMEA results were positive and also ranged between 0.29 and 0.61. The Rasch based PCA finding of this study, however, was indeed much lower compared to that of Lee et al. (2011) which was 54.7 % (8^{th} grader) and 59.6 % (12^{th} grader). Compared with another study by Sheaet. al (2009), they found that the depression scale failed to fit the Rasch model (appeared multidimensional), however, results improved with the revised scale by removing item five. Although MCDI items pointed toward one dimension, the low variance obtained in this study compared to others, was most probably attributed to by either person measure S.D. or item difficulty S.D. or both were small (Linacre, 2011).

In PCA also, the eigen value served an important function in dimensionality of the items in the MCDI questionnaire. This study's result was similar to those obtained by Lee et al. (2011) as their study had a first construct eigen value of 2.4 (reference value adopted was less than 3.0) for both 8^{th} and 12^{th} graders and the unmodelled variances for lower and higher graders were 4.0% and 3.5% respectively (reference value adopted was less than 5%).

From the perspective of categories response, validation studies using the Rasch model by Sauer et al. (2013), Shea et al. (2009) and Lee et. al (2011), have also reported that the category thresholds monotonically ordered with reasonable step calibration values, indicating that answer format of the scale is appropriate and is consistent with this study's finding.

The 500 respondents were inadequately differentiated in this study as evidenced by a separation value of 1.90 (slightly less than the desired value of 2.0) and consistent with findings from the study conducted in 8^{th} (PSI=1.93) and 12^{th} graders (PSI=1.69) by Lee et al. (2011). However, this study found moderate reliability of 0.78 for person separation index (PSI), similar to 0.79 (8^{th} grader) and 0.74 (12^{th} grader) in the Lee et al. (2011) study, but slightly lower compared to 0.84 in the Sauer et al. (2013) study and 0.89 in the Shea et al. (2009) study.

However, in terms of item perspective, the finding of this study revealed good separation value that was 7.30 with very high reliability of 0.98. This similar phenomenon of off-target person separation index but good separation of CDI items was also reported in the Rasch validation study by researchers Lee et al. (2011) with a value of 5.54 (item separation index) and 0.97 (reliability). A very high item reliability (1.00) was also reported by Sauer et. al (2013). In the context of all the results obtained in this study, this high reliability value of CDI items was informative as it demonstrated that all items were clear in measuring varying sensitivity levels of depression; just that when items were presented to individuals with depression symptoms of low to moderate severity

levels, items would not reliably reflect respondents' true level of depression (evidenced by low PSI). Hence, it would seem that absence of good overlap between CDI items and respondents recruited in this study who were presumably a nonclinically depressed sample was the most probable reason causing this phenomenon.

The uneven spread of MCDI items along the y-axis which did not line up with the respondents in the Wright map, suggested that approximately half of the MCDI items were unable to provide useful information about the depression as they were not good in detecting milder forms of the disorder. This finding is similar to the findings for non-clinical secondary school samples in the United States (Lee et al., 2011) and the 5035 non-clinical adult sample in Germany reported by Sauer et al.

Table 5
Item statistics of Malay version CDI for respondents (N=500)

Item	Item Difficulty	Infit MNSQ	Outfit MNSQ	PTMEA Correlation
1	0.52	0.88	0.84	0.44
2	-1.11	1.15	1.12	0.26
3	1.00	0.88	0.85	0.50
4	-0.19	0.97	0.98	0.45
5	0.89	1.03	1.04	0.29
6	0.25	1.00	0.99	0.41
7	0.53	0.81	0.67	0.49
8	0.28	1.03	1.15	0.36
9	0.92	0.85	0.79	0.45
10	0.17	0.93	0.95	0.40
11	-0.65	0.97	0.94	0.49
12	0.42	1.05	1.07	0.34
13	-0.36	1.03	1.03	0.39
14	-0.11	0.99	0.97	0.43
15	-1.04	1.15	1.20	0.40
16	-0.27	1.15	1.24	0.34
17	-0.65	1.09	1.12	0.41
18	-0.13	1.08	1.08	0.38
19	-0.63	1.21	1.22	0.31
20	-0.46	0.85	0.83	0.55
21	-0.05	0.97	1.03	0.43
22	0.38	1.04	1.05	0.36
23	-0.97	0.99	0.97	0.49
24	-1.17	0.96	0.95	0.50
25	0.55	0.87	0.84	0.48
26	0.63	1.09	1.11	0.31
27	1.26	0.94	0.87	0.36

(2013). The most probable reason for the consistently poor person-item map found in this study and others mentioned above is that the majority of respondents recruited were mentally healthy (absence of depressive symptoms) at the time of the studies. From a clinical point of view, the floor effect in the present non-clinical sample does not seem problematic because the CDI instrument is designed for clinical use (Kovacs, 1992) and thus it should be used with more severe levels of depression (Sauer et al., 2013).

From the gender perspective, more female respondents were found to experience depressive symptoms compared to their male counterparts in this study, a surplus of 26 persons. Females also had significantly higher total scores compared to male respondents. In the aspect of mean total score of the MCDI, this study found that female respondents scored significantly higher than their male counterparts with the values being 13.56 and 12.03 respectively. Independent sample T-test results revealed there was a statistically significant difference in CDI total scores between gender (t=2.547, df=498, p=.011). This finding was consistent with findings from the four studies by Adlina et al. (2007), Teoh (2008 & 2010), Yaacob et al. (2009), and Ghazali and Azhar (2012). Ghazali and Azhar found female secondary school students aged between 13 to 16 years old in Sarawak reported a higher prevalence of depression compared to males, the percentage of female to male was 56.9% and 43.1% respectively but independent T-test result was insignificant [$t(384)=0.13$, $p=0.89$]. The study by Yaacob et al. also found females (n=728) had higher depression mean scores, 16.85 in MCDI compared to males which was only 15.51 (n=679), a score difference of merely 1.34. Results of the study by Yaacob et al. (2009) revealed that a statistically significant difference existed in CDI depression scores between gender with a t value of - 4.35 ($p \leq .00$). A similar pattern was also detected in the findings reported by Adlina et al. in 2007 which revealed females were more depressed than males. Their study involved 2048 secondary school respondents. Similarly, Teoh in her two studies in 2008 and 2010 reported females in lower secondary school were more prone to develop depression with the rate of 14% and 20.6% respectively. The reason that females were more susceptible to depression in this study might be due to two reasons, (1) more female respondents recruited in this study which gave higher prevalence percentage and (2) pubertal development started much earlier in females than their male peers and they were found to be more likely to become depressed (reported by Ge et al., 1996, 2001; Graber et al., 1997).

Independent sample T-tests were also conducted in this study to look for the presence of significant differences between school forms. From Table 10 (See Appendix 6), results showed that Form 2 students had higher MCDI total scores (raw scores) at 18 different score levels (marked by asterisk *) as compared to only 13 different score levels in Form 4 students. However, in SPSS mean comparison, Form 2 had a lower mean total MCDI raw score (12.62)

as compared to Form 4 respondents (13.20). When it came to the independent T-test, results revealed a *t* value of -0.976, *df*= 498, *p* =0.33, indicating that the difference of mean total MCDI raw score between school forms was statistically not significant. Comparison of this study's finding with other research could not be established due to the lack of available evidence in this particular area. In summary, based on the two independent T-test results, it can be concluded that the female respondents were more prone to depression compared to male respondents and that school forms was not a good indicator in predicting depression symptoms.

CONCLUSION

The findings of this study showed that the MCDI demonstrates sufficient evidence of psychometric properties. In addition, findings of this study showed that a minority (13.8%) of the respondents were positive for depression based on the MCDI total score (T-score) of 65 and above. Independent T-test result revealed that females and males differed in terms of their CDI total score. Thus, this translated version instrument is useful as a first line tool in screening for depressive symptoms among students before more detailed clinical diagnostic assessment procedures are administered by professional psychiatrists and psychologists. The school counsellor (provided that they are trained to be familiar with the administration and interpretation of results) can benefit from the information gathered by providing timely counselling to the affected students and make early referral to psychiatrist and psychologists for further depression diagnostic confirmation and follow up management. At a broader perspective, the results of the study pertaining to depression prevalence serve to remind schools, state educational departments and the Malaysia Ministry of Education (MOE) stakeholders to work collaboratively and closely with Malaysia Ministry of Health (MOH) in designing and maintaining appropriate school programmes for monitoring the mental health status of students by all means, before the situation reaches the stage which is beyond anyone's capability for any intervention.

REFERENCES

Adlina, S., Suthahar, A., Ramli, M., Edariah, A. B., Soe, S. A., Mohd, A. F., & Karuthan, C. (2007). Pilot study on depression among secondary school students in Selangor. *The Medical journal of Malaysia*, *62*(3), 218-222.

Allen, N. B., Hetrick, S. E., Simmons, J. G., & Hickie, I. B. (2007). Early intervention for depressive disorders in young people: the opportunity and the (lack of) evidence. *Medical Journal of Australia*, *187*(7 Suppl), S15-S17. Retrieved from https://www.mja.com.au/journal/2007/187/7/early-intervention-depressive-disorders-young-people-opportunity-and-lack

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual (4th edition)*. Washington, DC.

Baker, F. B. (2001). *The basics of Item Response Theory (2nd edition)*. United States of America: ERIC Clearninghouse on Assessment and Evaluation. Retrieved from http://files.eric.ed.gov/fulltext/ED458219.pdf.

Baghael, P. (2008). The Rasch Model as a construct validation tool: Transaction of the Rasch measurement. *American Educational Research Association*, *22*(1), 1145—1146.

Balan, R., & Chan, A. (2013).Stress and mental health problem on secondary school student in Sabah. *Malaysian Journal of Youth Studies*, *8,* 167 -176. Retrieved from http://www.ippbm.gov.my/dokumen/jurnal/Jurnal_IPPBM_Volume8%284_july%29.pdf

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences.* Mahwah NJ: Lawrence Erlbaum Assoc.

Cicchetti, D., & Toth, S. L. (1998).The development of depression in children and adolescents. *American Psycholist*, *53*(2)*,* 221-241.

Cole, D. A., & Martin, N. C. (2005). The longitudinal structure of the Children's Depression Inventory: testing a latent trait-state model. *Psychological Assessment*, *17*(2), 144. Retrieved from http://data.psych.udel.edu/abelcher/Shared%20Documents/1%20Assessment%20 (48)/c%20Diagnosis%20and%20assessment%20of%20specific%20problems/Cole, %20Martin%202005.pdf

Craighead, W. E., Smucker, M. R., Craighead, L. W., & Ilardi, S. S. (1998). Factor analysis of the Children's Depression Inventory in a community sample. *Psychological Assessment*, *10*(2), 156.

Daryanavard, A., Madani, A., Mahmoodi, M. S., Rahimi, S., Nourooziyan, F., &Hosseinpoor, M. (2011).Prevalence of Depression among High School Students and its Relation to Family Structure. *American Journal of Applied Sciences*, *8*(1), 39.

Del Barrio, V., & Carrasco, M. A. (2004). *CDI. Inventario de depresión infantil.* Madrid: TEA Ediciones.

Dorland, W. A. N. (2009). *Dorland's pocket medical dictionary (28th edition): International edition*. Philadelphia: Saunders, Elsevier.

Ferrari, A. J., Charlson, F. J., Norman, R. E., Patten, S. B., Freedman, G., Murray, C. J., ... & Whiteford, H. A.. (2013) Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010. *PLoS Med* , *10*(11): e1001547. doi:10.1371/journal.pmed.1001547

Ge, X., Conger, R. D., & Elder Jr, G. H. (2001). Pubertal transition, stressful life events, and the emergence of gender differences in adolescent depressive symptoms. *Developmental psychology*, *37*(3), 404.

Ge, X., Conger, R. D., & Elder, G. H. (1996). Coming of age too early: Pubertal influences on girls' vulnerability to psychological distress. *Child Development*, *67*(6), 3386-3400.

Ghazali, S. R., & Azhar, A. (2012, January).Depressive symptoms in association with socio-demographic factors, family functioning and classroom environment in a Malaysian secondary school population. *International Journal of Psychology*, *47*(1), 54-54.

Gibbons, C. J., Mills, R. J., Thornton, E. W., Ealing, J., Mitchell, J. D., Shaw, P. J., ... & Young, C. A. (2011). Rasch analysis of the hospital anxiety and depression scale (HADS) for use in motor neurone disease. *Health Qual Life Outcomes*, *9*(82).

Gomez, R., Vance, A., & Gomez, A. (2012). Children's Depression Inventory: invariance across children and adolescents with and without depressive disorders. *Psychological Assessment*, *24*(1), 1.

Goodchild, S., & English, L. D. (2004). Researching Mathematics Classrooms: A Critical examination of methodology. Westport, CT: Greenwood Publishing Group Inc

Graber, J. A., Lewinsohn, P. M., Seeley, J. R., & Brooks-Gunn, J. (1997). Is psychopathology associated with the timing of pubertal development?. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(12), 1768-1776.

Hambleton, R. K., & Jones, R. W. (1993). An NCME Instructional Module on Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. Retrieved from http://ncme.org/linkservid/66968080-1320-5CAE-6E4E546A2E4FA9E1/showMeta/0/

Hankin, B. L. (2006). Adolescent depression: Description, causes, and interventions. *Epilepsy and Behavior*, *8*, 102–114.

Harrington, R., Rutter, M., & Fombonne, E. (1996). Developmental pathways in depression: Multiple meanings, antecedents, and endpoints. *Development and Psychopathology*, *8*(4), 601-616.

Huang, C., & Dong, N. (2014). Dimensionality of the Children's Depression Inventory: Meta-analysis of pattern matrices. *Journal of Child and Family Studies, 23*(7), 1182-1192.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., ... & Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Survey. *Archives of general psychiatry*, *51*(1), 8-9.

Kok, J. K., & Goh, L. Y. (2011, October). Young People and Suicide Issue. In *International Conference on Humanities, Society and Culture, IPEDR* (Vol. 20). Retrieved from http://www.ipedr.com/vol20/7-ICHSC2011-M00018.pdf

Kovacs, M. (1992). *Children's Depression Inventory*. North Tonawanda, NY: Multi-Health Systems, Inc.

Kovacs, M. (2004). Children's Depression Inventory (CDI) Profile report for John Sample. Toronto: Multi-Health Systems Inc (MHS) in *Pscyhological Assessments Australia.* (n.d.). Retrieved from http://www.psychassessments.com.au/products/22/prod22_report1.pdf

Lee, Y. S., Krishnan, A., & Park, Y. S. (2011). Psychometric Properties of the Children's Depression Inventory: An Item Response Theory Analysis Across Age in a Nonclinical, Longitudinal, Adolescent Sample. *Measurement and Evaluation in Counseling and Development*, 0748175611428329.

Lerdal, A., Kottorp, A., Gay, C. L., Grov, E. K., & Lee, K. A. (2014). Rasch analysis of the Beck Depression Inventory-II in stroke survivors: A cross-sectional study. *Journal of Affective Disorders*, *158*, 48-52.

Linacre, J. M. (1999a). Investigating rating scale category utility. *Journal of Outcome Measurement, 3*(2), 103-122.

Linacre, J. M. (1995). Categorical misfit statistics. *Rasch Measurement Transactions, 9*(3), 450-451. Retrieved from http://www.rasch.org/rmt/rmt93j.htm

Linacre, J. M. (1998). Structure in Rasch residuals: Why principal components analysis (PCA)? *Rasch Measurement Transactions*, *12*(2), 636.

Linacre, J. M. (2011). *Old Rasch Forum - Rasch on the Run: 2011. 426. RPCA = help*. Retrieved from http://www.rasch.org/forum2011.htm

Linacre, J. M. (2014). Winsteps® (Version 3.81.0) [Computer Software]. Beaverton, Oregon: Winsteps.com.

Linacre, J. M., & Wright, B. D. (1994b). Reasonable mean-square fit values. *Rasch Measurement Transactions*, *8*(3), 370. Retrieved from http://www.rasch.org/rmt/rmt83b.htm

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison- Wesley.

Malaysian Mental Health Association. (2014). *Depression is common and treatable*. Retrieved from http://mmha.org.my/wp-content/uploads/2014/08/MMHA-Bulletin-August-2014-FINAL-final-ang-Publisher_Revised2_CMYK-1.pdf

Malaysian Psychiatric Association. (2009). *National Suicide Registry Malaysia - Annual Report 2007*. Retrieved from http://www.psychiatry-malaysia.org/article.php?aid=1029

Marcus, M., Yasamy, M. T., Ommeren, M. V., Chisholm, D., & Saxena, S. (2012). *Depression A Global Public Health Concern*. Retrieved from http://www.who.int/mental_health/management/depression/who_paper_depression_wfmh_2012.pdf

Mash, E. J., & Wolfe, D. A. (1999). *Abnormal child psychology.* Belmont, CA: Wadsworth.

Mazza, J. J., Fleming, C. B., Abbott, R. D., Haggerty, K. P., & Catalano, R. F. (2010). Identifying trajectories of adolescents' depressive phenomena: An examination of early risk factors. *Journal of Youth and Adolescence, 39*(6), 579-593.

Mey, S. C. (2010). Psychological Profile of Adolescents in Malaysia. *International Journal of Educational & Psychological Assessment*, *4*.

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement, (3rd ed.)*. Washington, D.C.: American Council on Education.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons'responses and performances as scientific inquiry into score meaning. *American Psychologist, 50,* 741–749.

Monreal, P. (1988). *Estudio de la sintomatología depresiva infantil en la provincia de Gerona. Análisis de algunos factores individuales y psicosociales*. (Unpublished Doctoral Dissertation). Universitat Autònoma de Barcelona, Spain.

National Health and Morbidity Survey (NHMS). (2006). *Fact sheet.* Malaysia: Ministry of Health.

National Health and Morbidity Survey (NHMS). (2011). *Fact sheet.* Malaysia: Ministry of Health. Retrieved from http://www.moh.gov.my/index.php/file_manager/dl_item/624746305a584e305833426b5a69394f51305176546b684e553138794d44457858305a4251315266553068465256 51756347526d

National Suicide Registry Malaysia (NSRM). (2009). *National Suicide Registry Malaysia Annual Report 2009*. Malaysia: Ministry of Health Malaysia, Clinical Research Centre (CRC). Retrieved from http://www.crc.gov.my/wp-content/uploads/documents/report/NSRM_report_2009.pdf

New Straits Times. (2014). *20% of children in Malaysia are depressed*. Retrieved from http://www.malaysiandigest.com/frontpage/29-4-tile/483487-20-of-children-in-malaysia-are-depressed.html

Rosliwati, M. Y., Rohayah, M., Jamil, B.Y. M., & Zaharah, S. (2010). Validation of the Malay Version of Children Depression Inventory (CDI) among children and adolescents attending outpatient clinics in Kota Baru, Kelantan. *Malaysian Journal of Psychiatry*, *17*(1). Retrieved from http://www.mjpsychiatry.org/index.php/mjp/article/view/24/24

Sauer, S., Ziegler, M., & Schmitt, M. (2013). Rasch analysis of a simplified Beck Depression Inventory. *Personality and Individual Differences*, *54*(4), 530-535.

Shea, T. L., Tennant, A., & Pallant, J. F. (2009). Rasch model analysis of the Depression, Anxiety and Stress Scales (DASS). *BMC psychiatry*, *9*(1), 21.

Tan, S. M., Loh, S. F., Bujnag, M. A., Haniff, J., Rahman, A., Nazri, F., ... & Iryani, T. (2013). Validation of the Malay version of children's depression inventory. *International Medical Journal*, *20*(2), 188-191.

Teoh, H. J. (2008). Interpersonal and environmental factors that predict depression in secondary school girls. *Malaysian Journal of Medicine and Health Sciences*, *4*(2), 83-95.

Teoh, H. J. (2010). A survey of urban child and adolescent mental health problems in an urban Malaysian population. *Malaysian Journal of Psychiatry*, *19*(1).

The Star. (2011). *Promote good mental health*. Retrieved from http://www.thestar.com.my/story/?file=%2f2011%2f2%2f11%2ffocus%2f8041496&sec=focus

The Star. (2012). *Suicide rate on the rise in Malaysia*. Retrieved from http://www.thestar.com.my/News/Nation/2012/06/05/Suicide-rate-on-the-rise-in-Malaysia/

The Star. (2013). *Students advised to seek help early*. Retrieved from http://www.thestar.com.my/News/Community/2013/10/23/Students-advised-to-seek-help-early-Those-who-fail-to-cope-with-life-may-resort-to-suicide/

Van Lang, N. D., Ferdinand, R. F., & Verhulst, F. C. (2007). Predictors of future depression in early and late adolescence. *Journal of Affective Disorders*, *97*(1), 137-144.

Weiss, B., & Garber, J. (2003).Developmental differences in the phenomenology of depression. *Developmental Psychopathology*, *15*, 403-430.

Weiss, B., & Weisz, J. R. (1988). Factor structure of self-reported depression: clinic-referred children versus adolescents. *Journal of Abnormal Psychology*, *97*(4), 492.

WHO. (2008). *Suicide and Suicide Prevention in Asia*. Retrieved from http://www.who.int/mental_health/resources/suicide_prevention_asia.pdf

WHO. (2011). *Mental Health Atlas 2011 - Department of Mental Health and Substance Abuse*. World Health Organization. Retrieved from http://www.who.int/mental_health/evidence/atlas/profiles/mys_mh_profile.pdf?ua=1

WHO. (2012). *Depression*. Retrieved from http://www.who.int/mediacentre/factsheets/fs369/en/

Wilson, M. (2011). Some notes on the term: "Wright Map". *Rasch Measurement Transactions*, *25*(3), p.1331. Retrieved from http://www.rasch.org/rmt/rmt253b.htm

Yaacob, S. N., Juhari, R., Abu Talib, M., & Uba, I. (2009). Loneliness, stress, self esteem and depression among Malaysian adolescents. *Jurnal Kemanusiaan*, *14*, 85-95.

APPENDIX 1

Table 6
Category frequencies of three response scales

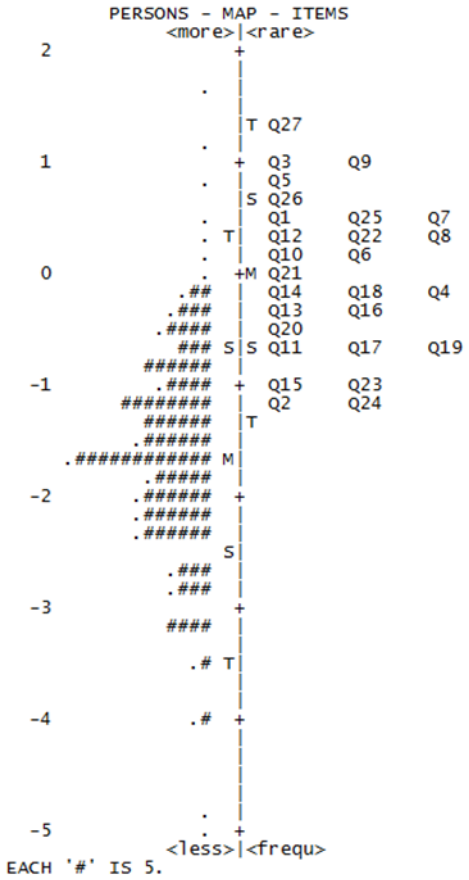
ITEM CATEGORY/OPTION/DISTRACTOR FREQUENCIES: MISFIT ORDER

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE MEASURE	S.E. MEAN	OUTF MNSQ	PTMEA CORR.	ITEM	
16	A 0	0	356	71	-1.89	.05	1.0	-.37	Q16	0
	1	1	93	19	-1.09	.09	.9	.27		1
	2	2	51	10	-1.03	.12	1.6	.21		2
19	B 0	0	224	45	-2.01	.07	1.1	-.32	Q19	0
	1	1	228	46	-1.39	.06	1.2	.24		1
	2	2	48	10	-1.24	.11	1.4	.13		2
15	C 0	0	215	43	-2.13	.07	1.1	-.41	Q15	0
	1	1	193	39	-1.37	.05	.8	.22		1
	2	2	92	18	-1.12	.09	1.5	.25		2
8	D 0	0	353	71	-1.91	.05	1.0	-.40	Q8	0
	1	1	128	26	-1.01	.07	.9	.37		1
	2	2	19	4	-1.16*	.24	2.0	.10		2
2	E 0	0	78	16	-2.32	.15	1.1	-.28	Q2	0
	1	1	392	78	-1.54	.04	1.1	.21		1
	2	2	30	6	-1.38	.20	1.1	.07		2
17	F 0	0	273	55	-2.01	.06	1.1	-.38	Q17	0
	1	1	161	32	-1.37	.06	1.0	.19		1
	2	2	66	13	-.88	.11	1.2	.30		2
26	G 0	0	313	63	-1.90	.06	1.1	-.31	Q26	0
	1	1	179	36	-1.26	.06	1.1	.29		1
	2	2	8	2	-.94	.41	1.4	.09		2
18	H 0	0	358	72	-1.91	.05	1.0	-.40	Q18	0
	1	1	101	20	-1.04	.08	.8	.30		1
	2	2	41	8	-.95	.13	1.4	.21		2
12	I 0	0	355	71	-1.89	.05	1.0	-.36	Q12	0
	1	1	130	26	-1.08	.07	.9	.34		1
	2	2	15	3	-1.10*	.19	1.6	.10		2
22	J 0	0	302	60	-1.94	.06	1.0	-.35	Q22	0
	1	1	186	37	-1.25	.06	1.1	.31		1
	2	2	12	2	-.67	.23	1.1	.15		2
5	K 0	0	403	81	-1.79	.05	1.0	-.28	Q5	0
	1	1	89	18	-1.14	.08	1.0	.24		1
	2	2	8	2	-.39	.38	1.1	.16		2
13	L 0	0	169	34	-2.23	.09	1.0	-.40	Q13	0
	1	1	311	62	-1.37	.05	1.0	.36		1
	2	2	20	4	-1.17	.19	1.1	.10		2
21	M 0	0	268	54	-2.04	.06	1.0	-.41	Q21	0
	1	1	210	42	-1.27	.06	1.1	.32		1
	2	2	22	4	-.59	.18	1.0	.22		2
6	N 0	0	259	52	-2.01	.07	1.0	-.37	Q6	0
	1	1	229	46	-1.33	.05	1.0	.30		1
	2	2	12	2	-.06	.22	.8	.25		2
23	m 0	0	206	41	-2.23	.07	1.0	-.48	Q23	0
	1	1	217	43	-1.36	.05	.8	.26		1
	2	2	77	15	-.93	.09	1.1	.30		2
14	l 0	0	277	55	-2.01	.06	1.0	-.39	Q14	0
	1	1	197	39	-1.31	.05	1.0	.28		1
	2	2	26	5	-.49	.15	.8	.27		2
4	k 0	0	181	36	-2.19	.08	1.0	-.40	Q4	0
	1	1	303	61	-1.41	.05	1.0	.30		1
	2	2	16	3	-.23	.16	.9	.26		2
11	j 0	0	289	58	-2.08	.06	.9	-.49	Q11	0
	1	1	140	28	-1.19	.06	.8	.29		1
	2	2	71	14	-.85	.09	1.0	.32		2

24	i	0	0	137	27	-2.37	.09	1.0	-.43	Q24	0
		1	1	297	59	-1.52	.05	1.0	.16		1
		2	2	66	13	-.76	.09	.9	.34		2
10	h	0	0	418	84	-1.84	.05	.9	-.42	Q10	0
		1	1	50	10	-.75	.09	.4	.30		1
		2	2	32	6	-.61	.16	1.4	.27		2
27	g	0	0	405	81	-1.83	.05	1.0	-.36	Q27	0
		1	1	91	18	-.92	.09	.9	.34		1
		2	2	4	1	-.37	.31	.9	.11		2
3	f	0	0	272	54	-2.11	.06	.9	-.49	Q3	0
		1	1	225	45	-1.13	.05	.8	.47		1
		2	2	3	1	.08	.61	.9	.13		2
1	e	0	0	379	76	-1.90	.05	.9	-.43	Q1	0
		1	1	107	21	-.96	.08	.9	.36		1
		2	2	14	3	-.26	.20	.7	.23		2
25	d	0	0	342	68	-1.98	.05	.9	-.47	Q25	0
		1	1	147	29	-.99	.06	.8	.43		1
		2	2	11	2	-.44	.25	.9	.18		2
9	c	0	0	389	78	-1.89	.05	.9	-.44	Q9	0
		1	1	104	21	-.88	.07	.7	.39		1
		2	2	7	1	-.04	.45	.9	.19		2
20	b	0	0	247	49	-2.18	.06	.9	-.52	Q20	0
		1	1	212	42	-1.26	.05	.8	.34		1
		2	2	41	8	-.50	.12	.8	.34		2
7	a	0	0	394	79	-1.91	.05	.9	-.50	Q7	0
		1	1	91	18	-.75	.06	.5	.42		1
		2	2	15	3	-.30	.27	1.1	.24		2

* Average measure does not ascend with category score

APPENDIX 2



Each "·" is 1-4 persons
M : mean
S : one standard deviation
T : two standard deviation

Figure 2. Distribution of persons and 27 Malay version CDI items

APPENDIX 3

Table 7
Distribution of total CDI score of respondents according to gender (N=500)

Raw score		T – score	Interpretation of overall symptoms/complaints	Frequency		Percentage	
Female	Male			Female	Male	Female	Male
≥23	≥29	Above 70	Very much above average	31	5	10.76	2.36
20-22	25-28	66 to 70	Much above average	27	6	9.38	2.83
17-19	21-24	61 to 65	Above average	28	10	9.72	4.71
14-16	16-20	56 to 60	Slightly above average	44	33	15.28	15.56
7-13	7-15	45 to 55	Average	117	122	40.63	57.54
4-6	3-6	40 to 44	Slightly below average	30	27	10.42	12.74
1-3	0-2	35 to 39	Below average	9	9	3.12	4.26
0	N/A	30 to 34	Much below average	2	0	0.69	0.00
N/A	N/A	Below 30	Very much below average	0	0	0.00	0.00
			Grand total	288	212	100.00	100.00

APPENDIX 4

Table 8
Distribution of Malay version CDI total score for positively screened depression according to gender (n=69)

CDI total (raw) score	Female respondents		Male respondents		Cumulative frequency
	frequency	percentage	frequency	percentage	
20	11	18.97	N/A	N/A	11
21	9	15.52	N/A	N/A	9
22	7	12.07	N/A	N/A	7
23	8	13.80	N/A	N/A	8
24	7	12.07	N/A	N/A	7
25	3	5.17	3	27.27	6
26	5	8.62	2	18.18	7
27	1	1.72	1	9.09	2
29	0	0.00	3	27.27	3
30	2	3.45	1	9.09	3
32	2	3.45	1	9.09	3
36	1	1.72	0	0.00	1
39	1	1.72	0	0.00	1
42	1	1.72	0	0.00	1
Grand total	58	100.00	11	100.00	69

APPENDIX 5

Table 9
Independent sample T-test for gender (N=500)

Group Statistics

	Gender	N	Mean	Std. Deviation	Std. Error Mean
Total score	Female	288	13.56	6.904	.407
	Male	212	12.03	6.303	.433

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Total score	Equal variances assumed	1.451	.229	2.547	498	.011	1.534	0.602	0.351	2.178
	Equal variances not assumed			2.582	475.559	.010	1.534	0.594	0.367	2.702

APPENDIX 6

Table 10
Sample T-test for school form (N=500)

Group Statistics

	Form	N	Mean	Std. Deviation	Std. Error Mean
Total score	2	250	12.62	6.998	.443
	4	250	13.20	6.373	.403

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Total score	Equal variances assumed	1.152	.284	-.976	498	.330	-.584	.599	-1.760	.592
	Equal variances not assumed			-.976	493.707	.330	-.584	.599	-1.760	.592