

ORIGINAL ARTICLE

**VALIDATION OF THE MALAY-TRANSLATED VERSION
OF THE CENTER FOR EPIDEMIOLOGICAL STUDY –
DEPRESSION SCALE (CES-D)**

Nurul Hazrina Mazlan & Affizal Ahmad**

***Forensic Sciences Programme, School of Health Sciences,
16150 Kubang Kerian, Kelantan, Malaysia.**

Abstract

Objective: This study aims to establish the validity and reliability of the Malay-translated version of the CES-D. **Methods:** This study employed a cross-sectional study design. The participants were female inmates (n=90) from local prisons which were selected based on purposive-selective sampling. The analyses include face validation, factor analysis, and reliability testing. A test-retest was conducted within a one-week interval. **Results:** The mean score for depression among the participants is 18.97 (*SD*=6.51). Further descriptive analysis showed that 58.9% of them scored above the mean score, which is considered high. Preliminary construct validation analysis confirmed that factor analysis was appropriate for the Malay-translated version of the CES-D. Furthermore, the factor analysis showed similar factor loadings to the original English version. The total internal consistency of the translated version, which was measured by Cronbach's alpha coefficient, was equal to 0.75. The test-retest reliability of the total score, measured by Pearson's correlation was equal to 0.69. **Conclusions:** Face validity, construct validity, and reliability analysis were found satisfactory for the Malay-translated version of the CES-D. The Malay-translated version of the CES-D was found valid and reliable to be used in future studies, with comparable properties to the original version and to previous studies. *ASEAN Journal of Psychiatry, Vol. 15 (1): January – June 2014: 54-65.*

Keywords: Depression, Female Inmates, CES-D, Validation, Reliability

Introduction

Depression is one of the most prevalent mental health problems worldwide [1]. It affects anyone regardless of gender, age or background. In prison settings, the prevalence of depression is often higher than the general population [2], especially among females [3]. The prolonged effects of depression are notable and these have often been observed in prison setting. The severe effects of prolonged depression include self-harm, self-mutilation, and suicide [4,5]. Therefore, early detection of depressive symptoms is crucial to enable consecutive measures. For this purpose, use of a reliable psychometric instrument to assess depression is very important for accurate

screening of the problems. Consequently, proper diagnosis and treatment of depression could be applied. Appropriate and early treatment is very crucial in cases of prolonged depression in order to prevent unwanted consequences [1,4,5].

The original English version of the Center of Epidemiological Study – Depression scale (CES-D) was designed by Radloff (1977) as a screening instrument intended to measure common symptoms of depression [6]. The items in the instrument enquire certain depressive symptoms, such as poor appetite, sleep disturbance, and loss of concentration. Originally, it was designed for use in the general population [7], with estimated internal

consistency using Cronbach's alpha coefficient equal to .85 for the healthy general population, and .90 among a studied psychiatric population [6]. The original English version of the CES-D contains 20 items using a 4-point Likert-scale in accordance to the frequency of depressive symptoms (0 = rarely or none of the time, to 3 = most or all the time).

The CES-D has been used widely in many studies and has been translated and validated in many languages including Arabic [8], Greek [9], Korean [10], Portuguese [11], Brazilian Portuguese [12,13], Spanish [14], and French [15]. However, no published study on the Malay-translated version of the CES-D was found. So far, Malaysia has no local published psychometric instrument specifically to measure depression and thus, local researchers still use instruments originally developed in other languages. In addition, many previous studies [8,9,10,12,14] across different cultures show distinct factorial structures of the CES-D that are unique to the cultures. The effect of cultural differences on the factorial structures of the CES-D necessitates the availability of the CES-D in the native language of the assessed population. It also emphasizes the needs to explore the underlying construct of the instrument among the assessed population itself [16]. Since the national language of the country is Malay, it requires the CES-D to be translated into Malay and validated. The accurate outcome of the current study is very important to help identify possible depression among Malaysians in any future study.

The main objective of the current study is to validate the Malay-translated version of the Center for Epidemiological Study – Depression scale (CES-D) using face validity and construct validity (Appendix A). Subsequently, the reliability of the translated version is determined in order to identify if it is reliable for use in the future.

Methods

Study design and participants

This study employed a cross-sectional study design. The source population was female prisons in the Peninsular Malaysia. Two

prisons were selected as the sampling frame. The sampling method for all phases was purposive selective sampling based on availability of the participants and selection criteria. The inclusion criteria of the participants are ability to read and write in Malay on their own. Female inmates who cannot communicate or understand Malay were excluded from the study. Female inmates who were on death row, serving life sentence or having diagnosed physical or mental health problems were excluded too.

The number of participants for each phase depends on the type of validation method. In phase one which focuses on face validity, 15 female inmates were recruited as the participants. For construct validity in phase two, the sample size was calculated based on Gorsuch's (1983) suggestion that total numbers of items in an instrument are multiplied to 5 to obtain the required number of participants [17]. For reliability testing, the required sample size was calculated based on the Cronbach's alpha formula. With consideration of estimated 20 percent dropout, the higher of the two resulting calculations was taken as the required sample size.

The final required sample size was 120 participants. However, due to limited number of female inmates available at the two selected prisons, a total of 90 participants were successfully recruited. In test-retest phase (phase three), the required sample size was calculated by a statistical software and the final sample size was 40 participants.

Translation process

For the purpose of validation, the researchers translated the English version of the instrument into Malay. The translation process involved two stages. The first stage was forward translation. In this stage, the original English version of the CES-D was translated into Malay by the researchers. After the translation was checked several times for adequacy of wording and sentence, the second stage was conducted.

The Malay version of the CES-D was translated back to English by a psychologist and a medical expert in the field. This stage is

known as back-translation. Both experts had no prior knowledge of the original version of the CES-D. The English-translated version was then compared to the original English version. The final comparison did not find much difference. The final Malay version was checked through by a language expert from a linguistic centre for any grammatical or language error.

Data collection

The research protocol of this study was reviewed and approved by the Research Ethics Committee (Human) of Universiti Sains Malaysia. The data collection was conducted at each respective prison, situated in the Peninsular Malaysia. Prior to the data collection, a briefing was held to explain the purpose of the current study and to communicate relevant information. Any doubts were clarified before seeking the participants' agreement to participate. The participants were assured that they may withdraw from the study at any time during the data collection process. A respondent information sheet and a consent form were given to each respondent to be read and signed. As the respondents agreed to participate in the study, data collection commenced.

In phase one, face validation was conducted. The selected participants (n=15) were asked to go through the Malay version of the CES-D. After reading through the instrument, they were asked if they fully understood the instrument and its meaning. No participant indicated problems with the translated instrument. In the second phase, construct validation and reliability testing were conducted. For this purpose, the Malay version of the CES-D was distributed to the participants (n=90) with the help of the prison staff. The researcher briefly explained the content and how to answer the instrument before asking the participants to complete the instrument. The participants were encouraged to ask if they had any problem with the instrument. The average time taken to complete the instrument was 7 minutes. These completed instruments were returned to the researcher. At the end of phase two, no

participants had indicated problems with the Malay version of the CES-D.

Phase three, which was the test-retest reliability, was conducted after a one-week interval at both prisons. It is considered as the shortest possible time interval to avoid insufficiency of participants due to limitations, such as length of prison incarceration for participants who had been involved in the second phase. In this phase, the same Malay version of the CES-D was given to 40 participants who had involved in phase two. It is to test if the participants would provide the same answer as in the previous phase. The same procedures were repeated as in phase two.

Analysis

The collected data was computed and analyzed using SPSS version 19.0. Descriptive statistics were used to summarize the demographic information and to obtain the descriptive details of depression among the participants. Then, data analysis for construct validity and reliability testing was run.

To explore the construct validity of the translated instrument, confirmatory factor analysis was performed using principal component analysis with varimax rotation. This procedure was selected based on previous studies [9] [16] in order to assess the factor structure of the translated version of the CES-D. Prior to factor analysis, the preliminary analysis for factor analysis was evaluated.

The preliminary analysis indicates the adequacy of the instrument to proceed with factor analysis [18]. The preliminary analysis is represented by the value of the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy, individual Measure of Sampling Adequacy (MSA) and the Bartlett's test of sphericity. The KMO value is expected to exceed the acceptable limit of 0.50 [19], with the individual MSA higher than 0.50 [20]. Some items might have the individual MSA below 0.50. The individual MSA affect the value of the KMO and items with low individual MSA might be excluded from the analysis [20]. Lastly, the Bartlett's test of sphericity indicates the appropriateness of

factor analysis for the translated instrument [20], thus it is expected to be significant.

The analysis then proceeded with assessment of the factor structure. Since confirmatory factor analysis was conducted using SPSS application, the number of factor was fixed based on previous studies [9,16], which is three factors. The number of factors presented the subscales or content domains of the instrument. Each factor explains certain percentage of variance in the instrument. Assessment of the factor structure include factor loading where items that are highly loaded into each factor were examined and then compared to previous studies.

To assess the reliability of the Malay version of the CES-D, the internal consistency and test-retest reliability of the translated instrument were measured. The internal consistency reliability of the instruments is

represented by Cronbach's alpha coefficient (α). Subsequently, Pearson's correlation coefficient (R) was calculated to evaluate the test-retest reliability. The correlation coefficient was calculated for the total score and individual items of the translated instrument.

Results

Demographic information

The participants' ages ranged from 18 to 53 years (Mean = 28.81 years, $SD = 8.01$ years). The summary of participant's demographic information is shown in Table 1. Regarding ethnicity, the majority of them are Malay (87.8%), and most of the participants are married (52.2%). Many of them have secondary education as the highest education level (80.0%) and almost half of them had no stable job before being incarcerated (42.2%).

Table 1. Summary of participants' demographic information (n= 90)

<i>Demographic Information</i>	<i>N</i>	<i>%</i>
Ethnicity		
Malay	79	87.8
Chinese	6	6.7
Indian	5	5.6
Marital status		
Single	22	24.4
Married	47	52.2
Divorcee	16	17.8
Widow	5	5.6
Highest education		
Primary	12	13.3
Secondary	72	80.0
Tertiary	6	6.7
Employment prior to incarceration		
Permanent job	34	37.8
Had no stable jobs	39	42.2
Unemployed	18	20.0

Descriptive analysis

In the current study, the mean score for Malay-translated version of the CES-D based on the responses given by the participants is 18.97. In addition, the standard deviation is 6.51. The descriptive summary of the Malay-translated version of the CES-D among the participants

is tabulated in Table 2. Based on the mean score, the participants were divided into two groups; participants who scored below the mean score and participants who scored above the mean score. The percentage of those who scored below the mean score (41.1%) is shown less than those who scored above the mean score (58.9%).

Table 2. Descriptive results for the Malay-translated version of the CES-D

Measurements	Frequency (n)	Percentage (%)
Mean score (<i>SD</i>)	18.97 (6.51)	
Below mean score	37	41.1
Above mean score	53	58.9

Factor analysis

The preliminary analysis for factor analysis of the Malay-translated version of the CES-D showed a satisfactory result. The value of the KMO Measure of Sampling Adequacy was .65. The individual MSA lies within .44 to .78. Some items have low individual MSA.

Nevertheless, since the value of KMO is above the acceptable limit, the analysis proceeded with all items regardless of the individual MSA. In addition, Bartlett's test of sphericity was found highly significant ($p < 0.001$). The factor loading of the translated instrument is shown in Table 3.

Table 3. Factor loadings for the Malay-translated version of the CES-D

Items no	Factor 1	Factor 2	Factor 3
1	0.695		
2	0.568		-.0358
3	0.602		
4		0.467	-0.448
5			0.404
6	0.601		
7		-.0761	
8		0.596	
9			0.406
10	0.493		
11	0.649		-0.373
12		0.466	
13	0.531		-0.459
14	0.681		
15			0.434
16		0.701	
17	0.404		
18	0.661		
19	0.540		
20			0.383

As mentioned earlier, three factors were extracted from the Malay-translated version of the CES-D based on the assessment of the screeplot. They explained 42.20 percent of the total variance. As shown in Table 3, items in Factor 1 are item number 1, 2, 3, 6, 10, 11, 13, 14, 17, 18, and 19. The items largely represent somatic complaints and interpersonal relationship problems. Factor 1 explained 23.06 percent of variance. Items in Factor 2 include item number 4, 7, 8, 12, and 16 which represent positive effect. It explained 10.56 percent of variance. For Factor 3, the items are item number 2, 5, 9, 11, 13, 15, and 20, which

mostly represent depressed affects and explained 8.57 percent of total variability. Several items were found not to fit into the factor compared to previous studies [9] [16].

Reliability testing

The internal consistency of the Malay-translated version of the CES-D was found considerably high. The total Cronbach's alpha of the instrument is equal to 0.75. Cronbach's alpha was also calculated for individual factors. For Factor 1, the Cronbach's alpha is 0.83, whereas for Factor 2, the Cronbach's

alpha is equal to 0.66. The Cronbach's alpha for Factor 3 is 0.60.

The test-retest reliability of the total score is found to be satisfactory ($R=.69$). The Pearson's correlation coefficient for test-retest

is shown in Table 4. Individual items however showed a wide range of variability with the lowest value was item number 20 ($R=.009$) and the highest was item number 7 ($R=.069$). Ten items had a correlation value of less than 0.50.

Table 4. Pearson's coefficient (R) for individual item and total score of the Malay-translated version of the CES-D

Items	Pearson's R
Item 1	0.53
Item 2	0.54
Item 3	0.14
Item 4	0.26
Item 5	0.60
Item 6	0.32
Item 7	0.70
Item 8	0.48
Item 9	0.43
Item 10	0.58
Item 11	0.52
Item 12	0.28
Item 13	0.64
Item 14	0.37
Item 15	0.29
Item 16	0.51
Item 17	0.58
Item 18	0.38
Item 19	0.63
Item 20	0.09
Total	0.69

Discussion

Previous studies on translated versions of the CES-D [8 – 10, 12,14] have yielded findings that are unique to their respective cultural identity. This emphasizes the option to generate a translated version of the CES-D which is unique to the Malaysian culture [16]. In response to this need, an initiative was taken to translate the CES-D into Malay and subsequently, validate the translated instrument among Malaysians. To begin with, the validation of the Malay-translated version of the CES-D involved a small scale, focusing on a population of Malaysian female inmates. This population was chosen because of the availability of the participants at a fixed place as well as possible occurrence of depression. Depression, which is expected to be

considerably high among female inmates [2] [3], may assist in the validating process of the Malay CES-D and at the same time, allow for preliminary identification of the problems among the population.

For the purpose of validating the Malay-translated version of the CES-D, two methods of validation processes were selected. These are face validation and construct validation. Face validity is defined as "the degree that respondents judge that the items of an instrument are appropriate to the targeted construct and assessment objectives" [21, p.99]. It is necessary as a preliminary method to ensure that the targeted construct as measured and respondents are able to understand the assessment objectives of the instrument [21]. Therefore, face validity was

conducted based on the participants' level of comprehension after they had gone through the translated instruments. Upon inquiry by the researcher regarding the translated version, the selected participants unanimously agreed that the instrument can be fully understood and carry the original meanings of the instrument. The participants testified that they had no problems in understanding and responding to each item in the translated version.

Subsequently, construct validation was performed to assess the underlying factor structure of the translated instrument. Radloff (1977) suggested that the original CES-D had four factor structures [6]. However, some previous studies identified only three factor structures in their data for the CES-D [9,16]. Similar to the previous studies, three factor structures were found fit to the Malay-translated version of the CES-D. Therefore, three factors were extracted in the current study. Each factor represents certain depressive factors in reference to the original and previous studies [9,16] including somatic complaints, depressed affects, interpersonal relationship problems, and positive affects. In comparison to a previous study, one item was found not to fit into Factor 2 (item number 7) and one item for Factor 3 (item number 15). Item number 7 was expected to fit into Factor 3 whereas item number 15 was expected to fit into Factor 1 [9,16]. However, in this current study, these two items were loaded into different factors. The Malay translation of item number 7 - *I felt that everything I did was an effort* - is considered to be a positive affect rather than a depressed affect. Item number 15 - *People were unfriendly* - in Malay carries a depressed affect rather than an interpersonal relationship problem. Thus, in the Malay translated version, item number 7 is loaded into Factor 2, whereas item number 15 is loaded into Factor 3.

In determining the validity of an instrument, the reliability of the instrument is equally important. Reliability shows the consistency of an instrument in measuring a construct when given to the same person at a separate time or given to a different person with depression [20]. Two types of reliability often measured are internal consistency and test-retest reliability. In most reliability testing, internal

consistency is represented by the value of Cronbach's alpha, whereas test-retest is represented by Pearson's correlation coefficient or intraclass correlation coefficient. The value of Cronbach's alpha is suggested to be in between 0.70 and 0.80 for an instrument to be considered reliable [20]. Also, the Pearson's correlation coefficient of the instrument, which represents the test-retest reliability in the current study, is expected to be higher than 0.50 to be considered moderately or highly reliable.

The original English version of the CES-D has the internal consistencies of 0.85 and 0.90 among two different populations [6]. In the current study, the Malay-translated version of the CES-D produced a total Cronbach's alpha of 0.75, which is within the suggested reliability range. In comparison to the original English version, the Cronbach's alpha of the Malay-translated version of the CES-D however is lower. The difference between the original English and the Malay-translated version of the CES-D might be explained by the different population being tested, the number of participants involved in each study, and the methods of data collection. In a study conducted by Radloff (1977), the selected populations were healthy general population and clinical psychiatric patients. More than 1,000 individuals from the healthy general population were involved in the study, as well as a group of 105 clinical patients. The method of data collection in all the studies was structured interview [6], which may contribute to bias depending on the perspective of the interviewer. Compared to the original study, the current study involved only 90 female inmates and using self-administered method of data collection. These factors may actually differentiate the value of the Cronbach's alpha between the current study and previous studies. In addition, the internal consistency is expected to be higher among clinical patients compared to the other population [6].

Subsequently, it was found in the Malay-translated version of the CES-D that Factor 2 and Factor 3 had Cronbach's alpha lower than .70. However, both values exceeded the acceptable Cronbach's alpha value ($\alpha > .050$). According to Cortina (1993), the value of Cronbach's alpha depends heavily on the

number of items in an instrument [22] and in the case of a psychological construct, the Cronbach's alpha may go below .70 due to the variety of the constructs being measured [23]. In addition, the nature of items in the factor may influence the internal consistency, such as items that exhibit instability [24]. For example, item number 12 in Factor 2 – *I was happy* – exhibits feature of instability between different people as well as over separate time. Another example is item number 2 in Factor 3 – *I did not feel like eating; my appetite was poor* – also exhibits feature of instability between people.

The test-retest reliability of the Malay-translated version of the CES-D exhibits satisfactory results and explainable inconsistency in some items. The total test-retest reliability was considerably good ($R=.69$). Test-retest reliability assesses the consistency of measures between scores of the instruments given to the same person twice [25]. However, it depends on several factors, including the time interval between test and retest [24], and the effect of memory [26]. As mentioned earlier, the individual Pearson's correlation showed inconsistency between test and retest score in some items where the items had unexpectedly low Pearson's correlation. According to Radloff (1977), the CES-D was designed with the current affective symptoms being focused on, which are expected to fluctuate in between different times [6]. Factors such as changes over time, time interval, and particular life events [6] may explain the variability and the low Pearson's correlation in some items. As an example, item number 3 – *I felt that I could not shake off the blues even with the help from my family or friends*. The response to this item might change over time among the inmates population depending on their chance to interact with others, especially family. Item number 20 – *I could not get "going"* – also may differ over time depending on the participants' emotion at the time the measure was taken. Response to depressive items, especially among prison population is expected to be inconsistent over time with consideration to their imprisoned situation. Nevertheless, the total Pearson's correlation was considerably high and reliable.

Several studies on the translated versions of the CES-D were discussed in the current study. First, the researcher found the version of CES-D that was translated and validated in Greek [9]. The samples were patients diagnosed with major depression and a normal control group. Test-retest in this study was one to two days' interval. The study revealed three factors structure of the Greek version of the CES-D with factor loadings almost similar to the current study. The Cronbach's alpha is 0.95 for the Greek version of the CES-D, with the total Pearson's correlation R equal to .71. The individual Pearson's R was also high ($R = 0.45$ to 0.95) [9]. In comparison to the current study, the shorter time interval between test and retest might contribute to the high Pearson's correlation in the study [6] as well as memory effect [26]. In addition, the sample which involved clinical patients is expected to yield high internal consistency and test-retest reliability [6].

The CES-D was also translated into Korean with Cronbach's alpha of 0.89 [10]. The samples were Korean immigrants who lived in Canada and were compared to published scores of American adults. No test-retest was done. In the study, Noh and colleagues found that the Korean tends to be reluctant to express the positive affects in the CES-D, which resulted in high occurrence of depressive symptoms among the sample. The Korean samples were more likely to respond to negative items, therefore the four positive affect items in the original CES-D were revised to negative items. As a result, the revised Korean version of the CES-D produced a higher internal consistency, which is 0.93 [10]. In another study to validate the Arabic version of the CES-D involving 435 Lebanese community adults, only two factors structure were derived [8]. The study suggested that there is no difference in psychological, somatic and interpersonal expression of depression among the Lebanese, thus revealed only two factors that are depressed affect and lack of positive affect in the Arabic version of the CES-D [8].

In Spanish version, the CES-D produced a Cronbach's alpha of 0.89 [14] among a sample of 554 university students. The factors structure was assessed by exploratory factor

analysis. Similar to the original version, four factors structure were revealed in this study.

However, the factor loading did not match the original four factors structure by Radloff (1977). Also, no test-retest was conducted [14]. Two validation studies on Portuguese version of the CES-D were found. The first study involved two samples which are college students (non-clinical population) and addicts (clinical population) [12]. The study yielded a Cronbach's alpha of 0.85. Factor analysis which was conducted produced four factors structure. Based on structure analysis, two items (item number 2 and 4) were suggested to be excluded due to low correlation with others items [12]. Later, another study to validate the Portuguese version of the CES-D was conducted by Batistoni and colleagues (2007) [13]. They used the same Portuguese version which was validated by Silveira and colleagues (1998) [12] but among a different population. The sample in this second study was Brazilian elderly aged 60 and above (n=903). In their study, Batistoni and colleagues found three factors structure, namely negative affects, problems initiating behaviors, and positive affects. In addition, the internal consistency was equal to 0.86 [13]. No test-retest was found in both studies. In all these studies, none involves offenders or prison population.

Several limitations were acknowledged in the current study. The first limitation is the study population, where only female inmates were involved. Certain factors such as limited freedom, imprisonment experiences, and stressful life events [6] needed to be considered. Inmates may experience higher level of depressive circumstances than the free-living people did, thus influences the reliability of the translated instrument [6]. Therefore, the current validated Malay-translated version of the CES-D is more valid and reliable for use among female inmates than Malaysian population in general. The second limitation is the validation method. In the current study, only two validation methods were exercised. The Malay-translated version of the CES-D was not compared to any gold standard and thus, the cut-off point, sensitivity, and specificity of the translated instrument could not be established. This limits the extent

of applicability for Malay-translated version of the CES-D. More validation methods such as criterion and content validity, as well as a more in-depth data collection method are suggested for future studies. The third limitation is the purposive-selective sampling method which was used to select the participants. In prison settings, the applications of probability sampling methods such as random sampling, systematic sampling, and cluster sampling are limited due to restricted access to certain information related to the inmates. In addition, not all inmates were available during data collection processes therefore the researcher had to purposively select those who were available and fulfilled the selection criteria. The fourth limitation is the time interval for test-retest. Test-retest studies are normally done within one to five weeks interval [27]. In this study, the time interval between test and retest took the shortest possible time interval (i.e., one week) to avoid insufficiency of participants due to certain restriction such as length of stay for participants who had involved in the first phase.

The overall results of face validity, factor analysis, and reliability testing of the Malay-translated version of the CES-D were found to be satisfactory. The Malay version of the CES-D was found comparable to the original version and to previous studies, suggesting that the Malay-translated version is valid and reliable to be used as a tool for screening depression among Malaysian population, especially female inmates, in the future. The validated Malay version of the CES-D is expected to contribute to increase employment of screening tool for depression among females, especially in local settings. Application of simple screening tool would increase the possibility for early detection of depression and thus enable for early diagnosis and interventions. Depression is a common disorder among females, especially for those who live in highly stressful environment (i.e., prisons). The availability of validated screening tool in local language would definitely benefit females in helping them and people around them aware of their problems. In addition, a specific study to design local psychometric instruments may be conducted in

the future with reference to the current findings and limitations.

Acknowledgement

A special thank you to Universiti Sains Malaysia, USM Fellowship, and Malaysian Prison Department (*Jabatan Penjara Malaysia*) for supporting this study.

References

1. Wiener JM, Breslin NA. The behavioral sciences in psychiatry. Philadelphia, USA: Williams & Wilkins; 1995.
2. Fazel S, Danesh J. Serious mental disorder in 23,000 prisoners: A systematic review of 62 surveys. *Lancet*. 2002; 359: 545-50.
3. Tye CS, Mullen PE. Mental disorders in female prisoners. *Aust N Z J Psychiatry*. 2006; 40(3): 266–71.
4. Zanarini MC, Frankenburg FR, Deluca CJ, Hennen J, Khera GS, Gunderson JG. The pain of being borderline: Dysphoric states specific to borderline personality disorder. *Harv Rev Psychiatry*. 1998; 6(4): 201-7.
5. Roe-Sepowitz D. Characteristics and predictors of self-mutilation: A study of incarcerated women. *Crim Behav Ment Health*. 2007; 17: 312-21.
6. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Appl Psych Meas*. 1977; 1: 385-401.
7. Knight RG, Williams S, McGee R, Olanman S. Psychometric properties of the Centre for Epidemiologic Studies Depression scale (CES-D) in a sample of women in middle life. *Behav Res Ther* 1996; 35(4): 373-80.
8. Kazarian SS, Taher D. Validation of the Arabic Center for Epidemiologic Studies Depression (CES-D) scale in a Lebanese community sample. *Eur J Psychol Assess*. 2010; 26(1): 68-73.
9. Fountoulakis K, Iacovides A, Kleanthous S, Samolis S, Kaprinis SG, Sitzoglou K, et al. Reliability, validity and psychometric properties of the Greek translation of the Center for Epidemiological Studies-Depression (CES-D) scale. *BMC Psychiatry*. 2001; 1(3): research article.
10. Noh S, Kaspar V, Chen X. Measuring depression in Korean immigrants: assessing validity of the translated Korean version of CES-D scale. *Cross-Cult Res*. 1998; 32: 358-77.
11. Gonçalves B, Fagulha T. The Portuguese version of the Center for Epidemiologic Studies Depression scale (CES-D). *Eur J Psychol Assess*. 2004; 20(4): 339-48.
12. Silveira DX, Jorge MR. Propriedades psicométricas da escala de rastreamento populacional para depressão CES-D em populações clínica e não-clínica de adolescentes e adultos jovens [Psychometric properties of the Epidemiological Screening Scale for Depression (CES-D) in clinical and non-clinical populations of adolescents and young adults]. *Rev Psiq Clín*. 1998; 25(5): Special edition.
13. Batistoni SST, Neri AL, Cupertino APFB. Validity of the Center for Epidemiological Studies Depression scale among Brazilian elderly. *Rev Saúde Pública*. 2007; 41(4): 598-605.
14. Vázquez FL, Blanco V, López M. An adaptation of the Center for Epidemiologic Studies Depression scale for use in non-psychiatric Spanish populations. *Psychiatry Res*. 2007; 149: 247-52.
15. Morin AJS, Moullec G, Mañano C, Layet L, Just J-L, Ninot G. Psychometric properties of the Center

- for Epidemiologic Studies Depression scale (CES-D) in French clinical and nonclinical adults [Propriétés psychométriques du Center for Epidemiologic Studies Depression scale (CES-D) sur un échantillon français d'adultes cliniques et non-cliniques]. *Rev Épidémiol Santé Publique*. 2011; 59(5): 327-40.
16. Gupta R, Yick A. Validation of the CES-D Scale for older Chinese immigrants. *Aging Ment Health*. 2001; 7(20): 56-70.
 17. Gorsuch RL. Factor analysis. 2nd ed. New Jersey, Erlbaum: Hillsdale; 1983.
 18. Tucci AM, Kerr-Corrêa F, Souza-Formigoni MLO. Childhood trauma in substance use disorder and depression: An analysis by gender among a Brazilian clinical sample. *Child Abuse Negl*. 2010; 34: 95-104.
 19. Kaiser HF. An index of factorial simplicity. *Psychometrika*. 1974; 39: 31-6.
 20. Field A. Discovering statistics using SPSS. 3rd ed. London: SAGE Publication Ltd.; 2009.
 21. Hardesty DM, Bearden WO. The use of expert judges in scale development. Implications for improving face validity of measures of unobservable constructs. *J Bus Res*. 2004; 57: 98-107.
 22. Cortina JM. What is coefficient alpha? An examination of theory and applications. *J Appl Psychol*. 1993; 78: 98-104.
 23. Kline P. The handbook of psychological testing. 2nd ed. London: Routledge; 1999.
 24. Wikman A, Wärneryd B. Measurement errors in survey questions: Explaining response variability. *Soc Sci Med*. 1990; 22: 199-212.
 25. Vitoratou S, Ntzoufras I, Smyrnis N, Stefanis NC. Factorial composition of the Aggression Questionnaire: A multi-sample study in Greek adults. *Psychiatry Res*. 2009; 168: 32-9.
 26. Otter ME, Mellenbergh GJ, Gloppe KD. The relation between information-processing variables and test-retest stability for questionnaire items. *J Educ Meas*. 1995; 32(2): 199-216.
 27. Liu Y, Wang M, Tynjälä J, Villberg J, Zhang Z, Kannas L. Test-retest reliability of selected items of Health Behaviour in School-aged Children (HBSC) Survey Questionnaire in Beijing, China. *BMC Med Res Methodol*. 2010; 10: 73-82.

Corresponding author: *Affizal Ahmad, Senior Lecturer in the Forensics Science Program, School of Health Sciences, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia.*

Email: affizal@kb.usm.my

Received: 21 June 2013

Accepted: 2 September 2013

Appendix A

Malay-translated version of the CES-D: Pusat untuk Kajian Epidemiologi Skala Kemurungan

Arahan: Di bawah adalah senarai sebahagian perkara yang anda mungkin rasa atau lakukan. Tandakan (✓) pada kenyataan yang paling tepat dengan diri anda.

Sepanjang minggu lepas...	Jarang / tiada	Kadang- kadang (1 – 2 hari)	Kerap kali (3 – 4 hari)	Pada setiap masa (5 – 7 hari)
Fikiran saya diganggu oleh hal yang selalunya tidak mengganggu saya				
Saya tiada selera untuk makan				
Saya rasa saya tidak dapat menghapuskan perasaan tertekan walaupun dengan bantuan kawan-kawan saya				
Saya rasa saya sebaik orang lain				
Saya mempunyai masalah untuk menumpukan perhatian kepada kerja yang saya lakukan				
Saya rasa tertekan				
Saya rasa semua yang saya lakukan adalah satu usaha				
Saya rasa mempunyai harapan yang baik untuk masa depan saya				
Saya fikir hidup saya telah mengalami kegagalan				
Saya merasa sangat takut				
Tidur saya terganggu				
Saya gembira				
Saya bercakap kurang berbanding kebiasaannya				
Saya berasa kesunyian				
Orang di sekeliling saya tidak mesra				
Saya menikmati hidup saya				
Saya memaki-hamun / menyumpah				
Saya berasa sedih				
Saya rasa orang lain tidak sukakan saya				
Saya tidak dapat meneruskan hidup				