TRANSLATION AND VALIDATION OF THE SELF-COMPASSION SCALE FROM ENGLISH TO BAHASA MELAYU AMONG LOCAL STUDENTS IN A MALAYSIAN PUBLIC UNIVERSITY

ROSS STEPHENSON

FACULTY OF EDUCATION
UIVERSITI MALAYA
KUALA LUMPUR
2021

TRANSLATION AND VALIDATION OF THE SELF-COMPASSION SCALE FROM ENGLISH TO BAHASA MELAYU AMONG LOCAL STUDENTS IN A MALAYSIAN PUBLIC UNIVERSITY

ROSS STEPHENSON

DISSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF EDUCATION (MEASUREMENT AND EVALUATION)

FACULTY OF EDUCATION
UNIVERSITI MALAYA
KUALA LUMPUR

2021

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: Ross Stephenson

Registration/Matric No: POM 170004

Name of Degree: Master of Education (Measurement and Evaluation)

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

TRANSLATION AND VALIDATION OF THE SELF-COMPASSION SCALE FROM ENGLISH TO BAHASA MELAYU AMONG LOCAL STUDENTS IN A MALAYSIAN PUBLIC UNIVERSITY

Field of Study: Education (Measurement and Evaluation)

I do solemnly and sincerely declare that:
(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every right in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 01/02/2021

Subscribed and solemnly declared before,

Witness's Signature Date: 01/02/2021

Name:

Designation:

ABSTRACT

Self-compassion is a growing area of research worldwide and has shown promise to be useful in providing insight and intervention against rising mental health issues. Self-compassion has also been linked to psychological well-being as well as negative psychological outcomes, which makes it an ideal measure for mental health. However, the instrument to measure self-compassion has not been comprehensively translated and validated in Malaysia yet. Previous studies are lacking in the psychometric evidence for a valid self-compassion scale in a local context. This study presents the comprehensive translation and validation of the 26-item self-compassion scale (SCS) by Neff (2003) from English to Bahasa Melayu (BM). The curent researcher first translated the SCS from English to BM using the backtranslation method. Cognitive interviews were conducted as a form of response validity. The first translation was pilot tested with 60 participants and internal consistency was found to be good. The translation was then reviewed and all items were retained but nine items were edited to make the translation more accurate. The SCS-BM was then administered to a total of 302 participants (M age = 21.9, SD age = 4.1) . Confirmatory factor analysis on the SCS-BM supported the original 6-factor model proposed by Neff (2003). Construct validity was demonstrated through a significant positive correlation between the SCS-BM and the SWLS-BM and a significant negative correlation between the SCS-BM and the BDI-BM. Test-retest reliability obtained after a 3-week interval was found to be acceptable. The translated SCS-BM produced by this study is a reliable and valid measure that can be further used to assess self-compassion in Malaysia.

PENTERJEMAHAN DAN PENGESAHAN SKALA IHSAN KENDIRI DARIPADA BAHASA INGGERIS KE BAHASA MELAYU DALAM KALANGAN PELAJAR TEMPATAN DI SEBUAH UNIVERSITI AWAM MALAYSIA

ABSTRAK

Ihsan kendiri adalah bidang penyelidikan yang berkembang di seluruh dunia dan telah menunjukkan potensi untuk dingunakan untuk memahami dengan lebih teliti isu-isu kesihatan mental yang semakin meningkat. Walaubagaimanapun, skala untuk menguji ihsan kendiri belum lagi diterjamahkan dan disahkan secara komprehensif di Malaysia. Kajian ini menerangkan penterjemahan dan pengesahan 26-item *self-compassion scale* (SCS) dari Neff (2003) dalam Bahasa Melayu (ihsan kendiri). SCS masih belum diterjemahkan dan disahkan secara komprehensif untuk penduduk Malaysia. Pertamanya, peneyelidik menerjemahkan SCS dari bahasa Inggeris ke BM menggunakan kaedah *backtranslation*. Wawancara kognitif telah digunakan sebagai bentuk kesahan tindakbalas. Terjemahan pertama diuji perintis dengan 60 peserta dan konsistensi dalaman adalah baik. Terjemahan itu kemudian dikaji semula dan semua item disimpan dan sembilan item diedit untuk meberikan terjemahan yang lebih tepat. Terjemahan terakhir SCS-BM diberikan kepada sejumlah 302 peserta (M umur = 21.9, SD umur = 4.1). *Confirmatory factor analysis* pada SCS-BM menyokong model 6-faktor asal yang dicadangkan oleh Neff (2003). Kesahan konstruk juga ditunjukkan dengan korelasi positif di antara SCS-BM dan SWLS-BM dan juga korelasi negatif di antara SCS-BM dan BDI-BM. Kebolehpercayaan ujian yang diperolehi selepas masa 3-minggu adalah baik. Kesimpulannya, SCS-BM adalah skala ihsan kendiri yang disahkan dan boleh digunakan untuk menilai ihsan kendiri seseorang dalam Malaysia.

ACKNOWLEDGEMENTS

Firstly I want to thank my God for helping me get through this entire masters dissertation. It was difficult trying to juggle a day job and writing a masters thesis but somehow God brought me throught it all. All Glory to Him. Throughout this process I also received a lot of help and support from the key individuals below.

I would like to give my heartfelt thanks to my supervisor, Dr. Harris Shah, for his constant help and assistance throughout my thesis journey. I always appreciate how quick he is to respond to messages and emails and how he guided and pointed me in the right direction when it came to putting this thesis together.In fact, it is thanks to Dr. Harris that I have now found a new love for the field of self-compassion and it has helped me improve my personal well-being in life.

I also want to thank En. Yufri from the Education Faculty admin for being so patient in answering my questions and helpfully explaining the complex administration that goes behind a thesis submission. Although he has so many students who are submitting he always remembers me and is quick to clarify any doubts I have. Together with the rest of the admin in the faculty I am grateful that even during a time of global pademic they are still working to help us get our documents processed.

To my family and friends who have supported me through many sleepless nights, I thank you all. To my friend Sabrena Gabriella Arosh, who was the reason I even joined UM in the first place, thank you for your guidance and tips in my thesis journey. For Lim Wee Lee, my coursemate, friend, and brother-in-Christ, thanks for always being there to communicate and share updated information with me. Finally to my two best friends, Lee Hauii and Ng Hao Yu, my heartfelt thanks for making the stress and burden of a masters thesis more bearable with good times and good company. I am eternally grateful for all these peole who have helped me. Praise God.

CONTENTS

Chapter 1: Introduction

Chapter 2: Literature Review

Chapter 3: Methodology

Chapter 4: Findings

Chapter 5: Discussion and Conclusion

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

BDI	:	Becks Depression Inventory
BM	:	Bahasa Melayu (the Malay language)
CFA	:	Confirmatory Factor Analysis
CFI	:	Comparative fit index
SCS	:	Self-compassion Scale
SMT	:	Social Mentality Theory
SWLS	:	Satisfaction with Life Scale
RMSEA	:	Root Mean Square Error of Approximation
SRMR		Standardized Root Mean Square Residual
χ^2	:	Chi-Square
α	:	Cronbach's alpha
r	:	Correlation coefficient

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

When was the last time you recall being compassionate towards yourself? Chances are it was probably very challenging for you to come up with a single clear example. The notion of self-compassion has only been studied by psychologists in the last decade. Only after being first operationally defined by Kristin Neff in 2003, there has been a sharp increase in research interest in the area of self-compassion. As of 2018, there have been hundreds of studies conducted on self-compassion, more specifically with regards to the Self-Compassion Scale (SCS) developed by Neff (2003), which she has given permission to use freely for research (Appendix A). The full SCS can be referred to in Appendix B. Several key studies will be discussed in detail in the subsequent literature review.

Neff (2003) went on to describe several positive psychological outcomes of self-compassion. Among these positive outcomes, what was most notable was that self-compassion is associated with general wellbeing in many different areas. In addition to that, self-compassion appears to contain all the benefits of self-esteem without any of its associated negative consequences (Neff, 2003). Arimitsu (2014) explained that this is because self-compassion has not shown any significant correlation to narcissism (an unfortunate trait linked to high self-esteem). It is likely that this is the main impetus behind why so many other researchers have adapted the SCS-English into their own local languages so that local researchers can then adapt the scale for use in more research with regards to assessing self-compassion, and to a greater degree, mental health in general. Currently the scale has been translated into

Mandarin, Portuguese, Greek, Spanish, Japanese, Korean, etc. For this current study it will be translated from English into Bahasa Melayu (BM).

Although a majority of the research on the SCS has produced consistently positive results, it has also generated some controversy with regards to the psychometric properties of the SCS. On one hand, most researchers conducting translation studies have confirmed the original 6-factor model structure defined by Neff (Arimitsu, 2014; Castilho & Pinto-Gouveia, 2011; Chen, Yan, & Zhou, 2011; Dundas, Svendsen, Wiker, Granli, & Schanche, 2016) yet there is some contrary evidence to this which suggest a bi-factor model instead (Costa, Marôco, Pinto-Gouveia, Ferreira, & Castilho, 2015; López, et al., 2015). This will be discussed further in the literature review.

Although research on the SCS is relatively new, it has shown much promise with regards to mental health, especially with regards to countering negative self-attitudes. The problem with negative self-attitudes is that they lead to a host of negative psychological problems (Harter, 1999, as cited in Neff, 2003). In the past, people would counter such negative attitudes by raising self-esteem through praise. However, this was found to be largely ineffective and led to a fostering of self-centeredness (Neff, 2003). Instead, Neff (2003) argues that being self-compassionate would result in a recognition of one's own shortcomings, accepting them, and furthermore improving upon them out of a sense of care for one's own well-being as opposed to protecting one's own status of worth (i.e. self-esteem). Self-compassion is thus in perfect alignment with the recent movement in the world of psychology with regards to "positive psychology"; a movement that Diener (2000) describes as encouraging researchers in this field to encourage behaviours which help people lead more satisfying and fulfilled lives. After laying the groundwork, Neff (2003) herself

recommends that the function of self-compassion as a mental health intervention should be explored in future research.

1.2 Rationale of the Study

There are four reasons which form the rationale for this study, namely, the importance of self-compassion in society, the rising mental health issues in Malaysia, the importance of understanding and preserving the wellbeing of students, and the need for local translation.

Heine and Hamamura (2007) noted that Western society often focused on individual self-enhancement, i.e. the pursuit of viewing the self in the most positive way possible. Asians, however, are more prone to self-criticism rather than self-enhancement (Heine & Hamamura, 2007). The authors elaborated that in an Asian cultural context, self-criticism is considered to be adaptive as it highlights personal weakness and encourages individuals to improve and contribute harmoniously back to society. This may seem beneficial on the surface, but on the other hand, Kagitcibasi (2005) argues that because self-criticism often employs harsh self-regulation tactics, it hinders an individual's ability for self-compassion; an individual cannot be self-critical and self-compassionate at the same time, one must give way to the other. Self-compassion provides a far more positive solution as it helps lower depressive thoughts in individuals (Yamaguchi, Kim, & Akutsu, 2014), reduces the experience of anxiety, negative self-evaluation and depression (durmaz & Aydin, 2017). The aforementioned research provides insight into how self-compassion is important to ensure the mental wellbeing of a society, especially in the Asian context.

Next, there is an unfortunate rise of mental health issues in Malaysia and subsequent lack of facilities to cope with its rise. The Malaysian Healthcare Performance Unit (MHPU) released a comprehensive report in 2016 detailing the

prevalence of mental health problems in Malaysia, which they determined to be 29% (adults) and 12% (children). The report highlighted a rise in incidences of depression and suicide. Suicide is especially concerning as many students in Malaysia have unfortunately been resorting to taking their own life (Sivanandam, Zainal, & Tan, 2018) to put an end to their debilitating stress, especially with regards to their education and examinations (Lee, Menon, & Rajaendram, 2018). What is even more worrying is that there are only four out of the 14 states in Malaysia have specialized psychiatric hospitals. It is thus unsurprising that the authors of the MHPU report also highlighted a severe lack of psychiatrists and psychologists to meet the mental health demands of a growing population (MHPU, 2016). All these reasons provide a legitimate concern towards the mental health of Malaysians and provide rationale for any research which could shed light on the growing situation.

Thirdly, the current research will be primarily focusing on university students and this is because students hold an important place in society after they graduate, especially with regards to the positions they will hold for the country (Ying & Tang, 2015). Mey and Yin (2014) elaborate that university students are expected to enhance and contribute in the near future to their nation's socio-economic growth. Considering the weight of this societal expectation placed upon the shoulders of students, it is unsurprising that many Malaysian students fall prey to stress and anxiety as a result of academic related problems (Shamsuddin et al., 2013). This is why, in spite of the high growth in academic achievement among Malaysian students (Elsevier, n.d.), their mental health has been on the decline (Mey & Yin, 2014). If this situation is left unchecked, it will result in more negative consequences as Shamsuddin et al. (2013) revealed in their findings that more senior Malaysian university students were found to have higher levels of depression as compared to those who had just entered

university. This evidence suggests that researcher need to focus efforts on providing a solution that can form basis on an early intervention to help prevent a further deterioration of student mental health. Research by Neff, Hsieh and Dejitterat (2014) hinted at using self-compassion as a solution for this problem. Their findings indicated that self-compassion helps to release students from the debilitating shackles of harsh self-criticism, lowers their anxiety levels, and better prepared them to cope with academic failure (Neff et al., 2007).

Finally, according to the MHPU (2016) report, a staggering 30.3% of the rural population suffer under the burden of mental health problems, and a majority of the rural population of Malaysia speak more BM as compared to English. In addition to that, Shamsuddin et al. (2013) found results that students from these rural areas of Malaysia suffered from greater mental health issues (e.g. depression) as compared to other geographical groups. Public universities that cater to such groups of rural students should therefore have the necessary instruments ready to be deployed in order to asses mental health accordingly.

In summary, the rationale behind this research is that self-compassion has shown to be important especially for countering self-criticism in Asian society, the rise in Malaysian mental health issues especially among Malaysian students indicates that a solution needs to be found, self-compassion has been shown to be beneficial in preserving the mental well-being of students and thus protecting an important group of people in society, and lastly a translated instrument is necessary to cater to the needs of BM speaking students in local universities. All the aforementioned reasons provide a basis for conducting the current research.

1.3 Statement of Problem

Firstly, there are insufficient studies investigating the psychometric properties and usage of the SCS in Malaysia. At the time of writing, the author was able to find only one study by Khatib, Roseliza-Murni, Hoesni, and Manap (2018) which translated the SCS into BM but only examined its reliability. The main disadvantage of the study by Khatib et al., (2018) is that they have provided limited evidence on the psychometric properties of the SCS-BM. This is because their study only utilized the measure of internal consistency, which is not sufficient as a standalone statistical analysis for a translation study. Additionally, their study only examined the validity of the SCS-BM using face and content validity but failed to provide quantitative evidence for the validity of the SCS-BM. The authors themselves admit this and add a suggestion for future researchers to look into conducting a factor analysis of the SCS in Malaysian context (Khatib et al., 2018).

As previously mentioned in Section 1.2, Malaysia faces a rising problem of mental health issues among the youth of the country. This critical problem should be a motivator to researchers to develop methods of assessment and intervention in the local community. Although the SCS is not mandated for use by counsellors and clinicians, it can still be used as a potential tool for counselling interventions and clinical therapies. However, the only way to get this started is by first producing a reliable and valid translation of the SCS from English to Malay.

Therefore, in order to address these two problems, the current study strives to develop a reliable and valid SCS-BM by using confirmatory factor analysis and with a solid translation framework to ensure a measure of self-compassion that is effective and localized for use in the Malaysian population to help assess mental health.

1.4 Purpose of the Study

The overarching aim of this current study is to produce a valid and reliable translation of the SCS-English into BM. This overarching aim can be divided into its component parts, namely, the aim to translate the SCS-English into BM, followed by determining its reliability, validity, and stability over time, and identifying the psychometric properties of the SCS-BM.

1.5 Objectives of the Study

The research objectives of this study are as follows:

1. To determine the internal consistency of the SCS-BM among university students in a local public university.
2. To determine the test-retest reliability of the SCS-BM among university students in a local public university.
3. To determine the factor structure of the SCS-BM via Confirmatory Factor Analysis (CFA) among university students in a local public university.
4. To determine the construct validity of the SCS-BM among university students in a local public university.

1.6 Research Questions

Based on the aforementioned research objectives, the research questions are as follows:

1. What is the internal consistency of the SCS-BM among university students in a local public university?
2. What is the test-retest reliability of the SCS-BM among university students in a local public university?
3. What is the factor structure of the SCS-BM via CFA among university students in a local public university?

4. What is the construct validity of the SCS-BM among university students in a local public university?

1.7 Significance of Study

Taking into account the aforementioned rationale and problem statement, the consequent significance of this study is that if the SCS can be translated, validated, and shown to be reliable and psychometrically stable, then mental health practitioners can use it as part of their mental health assessments and interventions in Malaysia.

In addition to that, there is a lack of validity evidence in Malaysia. The previously mentioned study by Khatib et al. (2018) only looked into the reliability of the SCS scale, it lacks an essential component of validity. It would thus be significant for the current study to include validity in order to better understand the properties of SCS in a Malaysian context.

The SCS is also fairly straightforward and can thus even be administered as a self-assessment tool in areas that lack specialized psychiatric services, i.e. a basic school counsellor could administer the SCS-BM as opposed to finding a licensed and experienced mental health worker. One of the reasons why the SCS needs to be translated and localized is because the national language of Malaysia is BM and localization will help adapt the scale to be more easily understood by the local population. These reasons justify the translation and validation of the SCS so that it can have a significantly positive impact with regards to the mental health of the local Malaysian community.

1.8 Limitations

Firstly, one limitation of this study is that reliability tests are easily influenced by sample size. A study by Bujang, Omar and Baharum (2018) showed how Cronbach's alpha for internal consistency achieve strong results as long as the sample size is at

least 30. Therefore, reliability may not necessarily be a good indicator to use. However, this study has included both tests for internal consistency and test-retest reliability for two reasons. Firstly, according to Maneesriwongul and Dixon (2004) reliability testing is a necessary inclusion for any studies employing backtranslation. According to the authors, measuring internal consistency is an indicator of how reliable the target language version of the translation is. Secondly, both internal consistency and test-retest reliability were included because they are featured in a majority of past translation studies on the SCS (Refer to Table 2.2). Therefore, by conducting these reliability tests in this current study, it provides this study the opportunity to easily refer to and compare with past literature on the psychometric properties of the reliability of the SCS.

Secondly, due to an oversight by the researcher, normality was more heavily relied upon in terms of data cleaning as opposed to a more relevant measure such as the Mahalanobis distance. As defined by Brereton (2015), the Mahalanobis distance is the "distance for one variable is the same as the number of standard deviations; an observation is away from the mean." (p. 143). This is a far more precise method in detecting significant outliers in the dataset as compared to using a histogram. Although the current study did achieve acceptable normality, which provides some confidence in going ahead with the subsequent analysis, the researcher acknowledges the omission of more appropriate measures which should be employed in data cleaning with regards to factor analysis. On a positive note, the factor analysis revealed acceptable fit values which means that even though data cleaning was lacking, acceptable model fit was still obtained.

Thirdly, this study is that it that it relies on a primarily student-based sample size which was recruited via haphazard sampling method. Due to time constraints in

the completion of this study, this study resorted to sampling within the student population of University Malaya. Additionally, at the time of conducting this research, many of the university students that this study intended to sample from had gone on their study break. This meant that sampling would become increasingly difficult due to the lack of students. Hence, haphazard sampling was employed as it was the most suitable given the particular timing of this research. This study acknowledges the bias in haphazard sampling as it lacks randomization.

Fourthly, translation quality was limited due to budget constraints. This current study is not under a grant and the researcher conducting this current study is not under a scholarship. Due to budget constraints, the researcher was unable to afford paying for more qualified translators and instead utilized readily available translators who had a relevant background.

Lastly, test-retest reliability results were affected by attrition. This study initially proposed using a sample of 100 participants from the entire pool of 300 targeted participants to complete the SCS-BM after a 3-week interval. Unfortunately, due to attrition, only 67 were willing to fill up the SCS-BM again. Therefore, caution should be taken when interpreting the test-retest reliability results.

1.9 Delimitations

The first delimitation is that this study has decided to focus only on the validation of the original model of the SCS which is a six-factor CFA model which has been validated by other researchers (Arimitsu, 2014; Castilho & Pinto-Gouveia, 2011; Chen et al., 2011; Dundas et al., 2016; Lee & Lee, 2010).

The next delimitation is that the data obtained on the SCS is measured using a 5-point Likert scale which expresses data in terms of an ordinal scale of measurement. According to Knapp (1990), an ordinal scale is one that has mutually exclusive

categories on the scale which can be ranked in some order. Some researchers emphasize that the data obtained from Likert rating scales should be treated as ordinal data (Pett, 1997, as cited in Jamieson, 2004) because the interval points within the Likert rating scale cannot be presumed to be equidistant from each other and there is no true zero value (Cohen, Manion, & Morrison, 2007; Kuzon, Urbanchek, & McCabe, 1996). For example, in the SCS, the values on the scale range from 1: Almost never, 2: Rarely, 3: Sometimes, 4: Most of the time, and 5: Almost Always. Therefore, the SCS data fits the criteria of being ordinal-level data. In contrast, Blaikie (2003) notes that many researchers treat Likert rating scale data as interval data rather than ordinal data because they assume the interval points are equidistant.

Ordinal data has its limitations in psychometric testing. Firstly, according to Knapp (1990) some researchers state that basic mathematical computations (i.e., addition and subtraction), which can easily be performed on interval data, should not be performed on ordinal data. This restriction in performing basic mathematical calculations on ordinal data leads to the second limitation which is that the data cannot be accurately represented by the mean value or analyzed using typical parametric tests. However, Baker, Hardyk, and Petrinovich (1966) as cited in Knapp (1990) have actually shown empirical evidence that there is little difference in treating ordinal data like interval data.

The current study accepts the aforementioned limitations of using ordinal data, however, will be treating the data obtained from the SCS as interval for the following reasons: past research and sufficient sample size. Firstly, based on the literature review (see Table 2.2), all past researchers treated the SCS 5-point Likert rating scale data as interval data. This is fairly common because according to Field (2009) many scientists themselves treat ordinal data as interval data; this is particularly true within the realm

of the social sciences and even in the field of medicine (Kuzon et al., 1996). The current study accepts that there will be some impact on the preciseness of the measurement but nonetheless, the advantage of treating the data as interval will allow for more straightforward comparisons with past translation studies on the SCS. Proponents of treating Likert rating scale data as interval would be Carifio and Perla (2007, 2008), who explain that the summative analysis of Likert rating scale data results in the data approaching interval or even ratio levels of measurement. On one hand, Kuzon et al. (1996) argues that scales measuring variables at the ordinal level should not use any form of mean calculation and instead should rely on median or mode, on the other hand however Carifio and Perla (2007, 2008) counter this argument by explaining that by taking the mean of multiple items a researcher can thus consider ordinal data as interval data instead. Neff (2003) also treats the SCS scores as interval data and describes using the mean to calculate an individual's self-compassion. Although not ideal, Knapp (1990) explains that researchers who continue to treat ordinal data as interval would need to at least focus on ensuring a sufficient sample size to achieve a normal distribution. This is echoed by Kuzon et al. (1996) who reinforces that parametric tests, which researchers often use, require a normal distribution within the sample.

Therefore, the current study acknowledges that treating Likert rating scale data as interval instead of ordinal may reduce the accuracy in statistically analyzing the SCS. However, by following Knapp's (1990) suggestion, the current researcher will be taking measures to ensure a sufficiently large sample size in order to achieve a normal distribution.

1.10 Summary

Self-compassion appears to hold many benefits in promoting positive psychological wellbeing without negative repercussions. Meanwhile, Malaysian students are at a high risk of developing mental health issues due to various stressors. Producing a valid and reliable translated version of the SCS would therefore be beneficial in helping local researchers better understand and intervene in issues of student mental health.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

The purpose of the following literature review is to provide an insight into the background of the SCS and evaluate several past translation studies which have been done since its introduction into psychology literature. This review will cover a suitable theoretical framework for self-compassion as well as it's conceptual and operational definition. The translation studies have been broadly divided into Asian and European studies respectively in order to clearly see if there are any cultural differences in the translation and administration of the SCS across different samples. Particular focus will be given towards the methods of translation and factor model analysis used in past studies to form the foundation upon which the current study will be based upon.

2.2 Theoretical Framework: Social Mentality Theory

A suitable theoretical framework for self-compassion is found in the Social Mentality Theory (SMT) first proposed by Gilbert (1989) which is defined as the emotional, behavioral and cognitive processes that are directed by an individual's motivation to form reciprocal social relationships (Liotti & Gilbert, 2011). These relationships can be focused on different roles such as caring, sexuality, competition, and cooperation (Gilbert, 2014). Essentially, each role activates different mental components, which can be broadly divided into two categories, namely, the threat (limbic) system and the caring (oxytocin) system. For example, when a mother sees a crying baby, the caring system is activated, producing chemicals to stimulate feelings of care (i.e. oxytocin), emotions are activated and moved by the baby's cries, and cognition is activated to think about ways of administering caring behavior.

Gilbert's (1989) original description of the SMT was primarily focused on how individuals relate to others however, the author continued to build on his theory and Gilbert (2005) as cited in Hermanto and Zuroff (2016) proposed that caregiving and care-seeking social mentalities can be activated both when relating to others (e.g. comforting a crying baby) and also when relating to the self (e.g. comforting yourself after a personal failure). Gilbert (2000) as cited in Hermanto and Zuroff (2016) theorized that social mentalities can be activated by either relationship with the self or with others because both internal (self) or external (others) stimuli can produce similar responses. For example, a person can activate feelings of sexual arousal either in the presence of a sexual partner (external stimulus) or by mental fantasy (internal stimulus) (Hermanto & Zuroff, 2016). In other words, an individual can tap into the original system designed for relating with others and instead use it to relate to the self. Gilbert (2007) refers to this as the "compassionate mind" which is the sum of the different components (i.e. neurological, emotional, cognitive and behavioral) working in tandem to produce either a care or a threat response with regards to stimuli. The understanding that social mentality is both a sum of components and can be activated with regards to the self provides the basis for self-compassion to be understood.

Liotti and Gilbert (2011) propose that self-compassion is rooted in the social mentality theory. Evidence for this can be seen in two ways. Firstly, the components of social mentality, namely, emotion and cognition, theoretically cover the six-factors of self-compassion proposed by Neff (2003), self-kindness (vs. self-judgement), common humanity (vs. isolation) and mindfulness (vs. over-identified). The positive components of Neff (2003)'s concept of self-compassion can be theoretically included under the caring system of social mentality, while the negative components of self-compassion fall under the threat system. Further research by Neff, Kirkpatrick, and

Rude (2007) confirmed this with evidence indicating that self-compassion could actually deactivate the threat system and activate the caring system towards one's self, thus promoting emotional calmness and a sense of security.

Gilbert (2014) also noted the role of oxytocin within the social mentality theory framework, however the role that oxytocin plays is complex. On the surface, oxytocin is primarily associated with the care system as it stimulates an individual to exhibit caring behaviours (e.g. mother caring for a child), however, oxytocin can also lead to increased aggression against out-groups (e.g. a mother protecting her child from a dangerous person) (Gilbert, 2014). Considering this, oxytocin may not necessarily produce positive effects with it comes to self-compassion, especially if the individual has a hostile relationship with themselves (Rockliff et al., 2011). Evidence for this was hinted in research by Chang, Pua and Ng (2019) who discovered that individuals with maladaptive perfectionism had a hostile relationship with their selves based on self-criticism. This self-criticism was a result of the activation of the threat system according to the social mentality theory. However, the authors results indicated that self-compassion was helpful in deactivating the self-criticism from the threat system, and instead activated the care system to produce self-soothing behaviours such as self-acceptance and self-kindness. Self-compassion was concluded to be effective in promoting positive cognitions about the self whilst at the same time reducing negative self-evaluations (Chang et al., 2019).

In summary, according to Neff et al. (2007), the social mentality theory provides a basis for understanding on how self-compassion helps individuals by activating the care (self-soothing) system, which entails feelings of safeness, security and oxytocin, while simultaneously deactivating the threat system, associated with defensiveness, insecurity and the limbic system; this has been visually illustrated in

Figure 2.1. Therefore, the social mentality theory provides a sufficient theoretical framework for the current research on self-compassion.

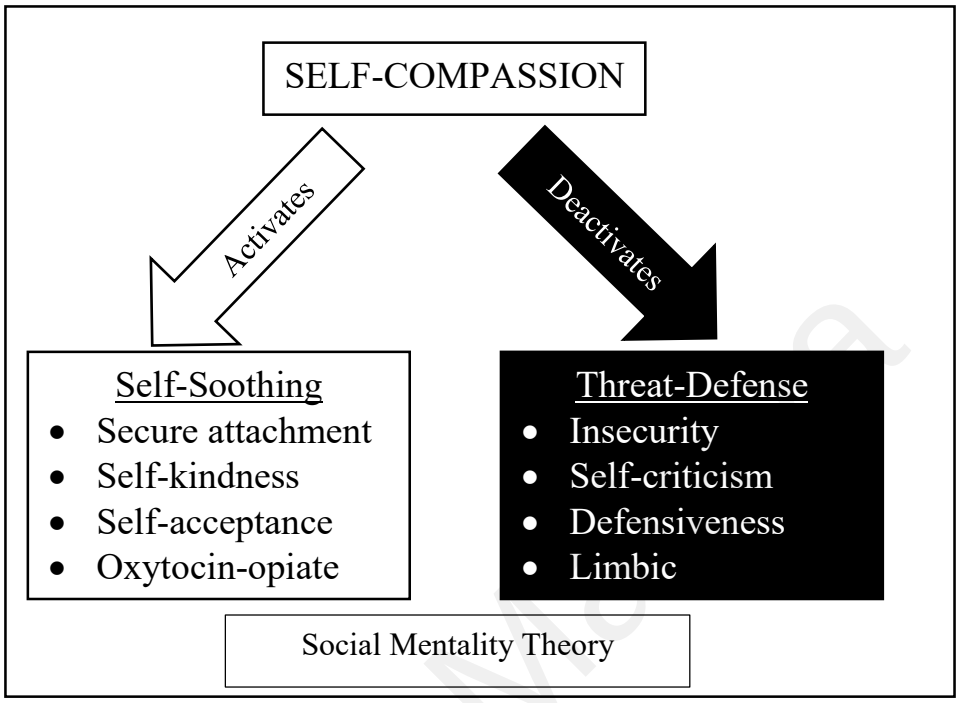


Figure 2.1. Visual representation of Self-compassion in the SMT

2.3 Conceptual Framework: Self-Compassion Scale

Kristin Neff built the concept of self-compassion based on Buddhist philosophy (Neff, 2003). Unlike the Western understanding of compassion, which entails an individual being kind and caring towards others, the Buddhist view of compassion is that it encompasses not only others, but the self as well (Neff, 2003). Essentially, all the positive outward characteristics (kindness, caring, being open and non-judgmental, etc.) we often associate with compassion for others can also be channeled towards the self. She goes on to specify that:

> Self-compassion, therefore, entails three basic components: 1) extending kindness and understanding to oneself rather than harsh self-criticism and judgment; 2) seeing one's experiences as part of the larger human experience rather than as separating and isolating; and 3) holding one's painful thoughts and feelings in balanced awareness rather than over-identifying with them (Neff, 2003, p. 224).

Her final scale consisted of 26-items with 3 pairs of opposite concepts. The first pair is self-kindness vs. self-judgement, which refers to a person being kind and understanding towards themselves as opposed to being harsh and critical. The second pair is common humanity vs. isolation, which refers to people being able to accept that failure is simply part of the human condition and that we are all connected, as opposed to seeing one's self as separated from others. The third pair concerns mindfulness vs. over-identification, which refers to someone being aware and able to recognize their weaknesses but without being consumed or over-identifying with those inadequacies (Neff, 2003).

2.3.1 Self-kindness vs. Self-judgement

According to Neff et al. (2020) "Self-kindness entails being warm, supportive, and understanding toward oneself" (p. 1). In other words, self-kindness includes everything we associate about kindness towards others but this time, by applying it to ourselves. Individuals with self-kindness are able to approach their personal flaws in a more gentle and supportive manner (Tsai, 2015). It is important to note that Neff (2009) conceptualizes self-kindness in the face of some form of disappointment or distress. For example, when we are experiencing distress, we should offer kind words and supportive behaviours to ourselves; in the same way that we would have offered kindness to a friend who was experiencing distress. Self-kindness does not need to arise when an individual is at rest or experiencing a neutral situation. This ties with the Social Mentality Theory by Gilbert (1989) whereby a relevant internal stimulus (i.e. experiencing a personal failure) can trigger an individual's self-soothing system (refer to Figure 2.1) resulting in an expression of self-kindness towards one's self, provided the individual has a good sense of self-compassion.

In contrast to self-kindness, Neff (2003) conceptualizes that an individual who has low self-compassion would express self-judgement, which can be defined as "harshly criticizing oneself for one's failings" (Neff et al., 2020, p. 2). Instead of the gentle and comforting language that one uses when expressing self-kindness (Tsai, 2015), self-judgement is characterized by severe and harsh words which seek to condemn rather than comfort. For example, when a student fails an exam, they may be inclined to blame themselves for not trying hard enough. With reference to the Social Mentality Theory in Figure 2.1, an individual with low self-compassion who experiences a distressing situation would undergo an activation of the threat-defense system which results in an expression of self-criticism.

Therefore, although self-kindness and self-criticism are conceptually distinct, Neff (2003) places them on opposite ends of the construct of self-compassion because they are both part of the same mental system which governs how individuals emotionally respond to failure and pain (i.e. with more self-kindness or more self-judgement towards themselves).

2.3.2 Common Humanity vs. Isolation

With regards to the second pair of concepts, Neff (2003) highlights how Buddhist philosophy accounts for the both others and the self. Neff et al. (2020, p. 1) defines common humanity as "recognizing the shared human experience of imperfection and understanding that all humans fail and make mistakes." It was noted that Neff (2009) conceptualized common humanity in the face of failure. For example, when driving on the road, if an individual with good self-compassion should get into a car accident, they can choose to reassure themselves that car accidents are part and parcel of the human experience of driving. In the absence of failure or mistakes, there is no need for an individual to relate their own personal suffering to the common

human experience of suffering. This ties in with the Social Mentality Theory by Gilbert (1989) whereby a relevant stimulus (i.e. experiencing a car accident) can trigger an individual's self-soothing system (refer to Figure 2.1) and result in feelings of secure attachment, provided the individual has self-compassion. Gilbert (2014) explains how secure attachment is linked to common humanity by drawing upon the influences of the Attachment Theory proposed by Bowlby (1969) in his observation of infants. An infant that receives secure attachment from their primary caregiver will then gain the confidence to become independent and explore the world around them (Bowlby, 1969 as cited in Gilbert, 2014). As an infant explores their surroundings, they come into contact with other human beings and in their development process, they form a mentalization of others and self. The end result of this process is called intersubjectivity, which the American Psychological Association (n.d.) defines as "the sharing of subjective experience between two or more people…" (p. 1). Cortina and Liotti (2010) emphasize that intersubjectivity arises from attachment. They elaborated that individuals with secure attachments should experience greater levels of intersubjectivity because they are able to confidently explore and interact with other social beings. Moreover, Gilbert (2014) explains that having a safe and secure base from which to operate from allows an individual the capacity to feel connected with others. It is this sense of connectedness to others that forms the basis of Neff's (2003) conceptualization of common humanity, whereby individuals realize that they are part of a connected human experience.

In contrast to common humanity, Neff (2003) conceptualizes that an individual with low self-compassion would experience a sense of isolation, which is defined as "feeling alone in the experience of suffering (Neff et al., 2020, p. 1). For example, if an individual with low self-compassion experiences a car crash, they would be more

prone to feeling all alone in their predicament and isolated from the rest of the world. With regard to the Social Mentality Theory in Figure 2.1, an individual with low-self compassion would trigger the threat-defense system and result in an inward expression of insecurity. Feelings of insecurity can lead to fear, and Cortina and Liotti (2010) explain that fear reduces an individual's level of intersubjectivity. Feelings of insecurity that arise from the threat-defense system also lead to a sense of disconnectedness with the world (Gilbert, 2014), which essentially forms the basis of Neff's (2003) concept of isolation.

Therefore, although common humanity and isolation are conceptually distinct, Neff (2003) places them on opposite ends of the construct of self-compassion because they are both part of the same mental system which governs how individuals cognitively understand personal failure and suffering (i.e. as part of the universal human experience or a sense of isolation from the world).

2.3.3 Mindfulness vs. Over-identification

Neff et al. (2020) defines mindfulness as "being aware of one's present moment experience of suffering with equanimity and balance." (p. 1). In other words, mindfulness allows individuals to observe their thoughts without over exaggerating them. For example, if an individual with self-compassion failed to pass a job interview, they would likely be able to take a balanced view and consider that their failure was not only due to personal inadequacies but instead it was the result of a variety of other factors such as the interviewers disposition, and so on. According to the SMT in Figure 2.1, an individual with self-compassion who experiences suffering can activate their self-soothing system which releases oxytocin and furthermore calms the body down via the parasympathetic nervous system (Gilbert, 2014). Mindfulness helps individuals access this self-soothing system by putting them in a more receptive mode of "being"

rather than actively "doing" (Gilbert, 2014). Tsai (2015) describes this state of mindfulness as having a balanced awareness of one's suffering in the moment. In addition to that, mindfulness is often achieved through meditation and Gilbert (2014) noted that people who meditate for long periods of time are able to experience contentedness and peacefulness with the self. According to Scheff (1981) as cited in Neff (2003), being mindful also helps to create a "mental space" that is necessary for one to be able to view themselves in a larger context instead of standing alone. Mindfulness can therefore create both a relaxed state in both the body and mind when an individual is facing a challenging situation.

In contrast to mindfulness, Neff et al. (2020) defines over-identification as being "fused with one's suffering to the point that perspective is lost." (p. 1). Over-identification occurs when an individual becomes so absorbed in their personal suffering that they end up exaggerating their inadequacies instead of taking a balanced perspective (Bennett-Goleman, 2001, as cited in Neff, 2003). For example, an individual with low self-compassion who failed their job interview would only focus on blowing their personal inadequacies out of proportion instead of taking a balanced view and considering other factors that may have contributed to the failure. The SMT in Figure 2.1 explains this by showing how an individual without self-compassion who experiences suffering would activate their threat-defense system which in turns triggers the limbic system that controls the sympathetic nervous system. The sympathetic nervous system is also known as the fight-or-flight system whereby the body prepares itself for stress by increasing heart rate, restricting digestion, and so on. Without mindfulness, a person whose threat-defense system has been activated will be in a constant "doing-mode" as described by Gilbert (2014) which thus prevents them from being able to take a more balanced view of their situation.

Therefore, although mindfulness and over-identification are conceptually distinct, Neff (2003) places them on opposite ends of the construct of self-compassion because they are both part of the same mental system which governs how individuals pay attention to personal failure and suffering (i.e. either in a more mindful or over-identified manner).

2.3.4 Self-compassion as a Global Construct

While each of the items outlined above are distinct, they interplay with each other to contribute towards self-compassion (Neff, 2016). For example, when an individual sees another person suffering, they are usually spurned to show them compassion. In the same way, when a person can recognize that their own individual suffering is no different from other human beings (common humanity), they can then start showing some compassion towards their own feelings (self-kindness), just as they would show it to others. Conversely, a person who lacks self-compassion would not be able to regulate their harsh criticisms (self-judgement) directed at the self, and end up feeling alone in their failures (isolation). In the words of Neff (2003), "These aspects of self-compassion are experienced differently and are conceptually distinct, but they also tend to engender one another" (p. 225).

Neff (2003) proposes that self-compassion is a result of the positive and negative sub-sales of the SCS interacting. Neff (2003) illustrated that just because someone is not being judgmental towards themselves, it does not necessarily mean that they are showing kindness either. With reference to the literature review in Section 2.2, Neff (2003) was quoted saying that one of the principles of self-compassion is "extending kindness and understanding to oneself rather than harsh self-criticism and judgment" (p. 224). This means that self-compassion requires an individual to be taking an active role in being self-compassionate towards themselves, in the face of

suffering, rather than just avoiding negative thoughts. Therefore, having low self-judgement does not mean the individual is showing themselves kindness, it simply means that they are not critical of themselves.

There are two past researchers who obtained findings which supported the idea that the positive and negative sub-scales of the SCS interact to form a global construct of self-compassion. Firstly, Phillips (2019) found that participants who scored high on the three compassionate subscales (self-kindness, common humanity and mindfulness) simultaneously scored low in the uncompassionate subscales (self-judgement, isolation and over-identification), and vice versa. Phillips (2019) further noted that there were no participants who scored both high and low in both compassionate and uncompassionate subscales which suggests that all the six subscales interact together to form a balanced system. In addition to that, recent research by Kim et al. (2020) in the area of brain scans has revealed that compassionate and uncompassionate self-responses can either respectively deactivate or activate activity in the anterior insula, anterior cingulate, and amygdala. Because both compassion and uncompassion trigger the same brain regions, it provides evidence that the three pairs of opposite constructs in Neff's (2003) conceptualization of self-compassion are operating in tandem rather than being independent of each other.

In conclusion, the six-subscales of self-compassion operate in a synergistic fashion to produce a global construct of self-compassion that not only acts as a buffer against negative emotions, but it also promotes a positive sense of self.

2.3.5 Operational Definition of Self-compassion

Neff (2003) operationally defined self-compassion as comprising of three measurable components, namely self-kindness (vs. self-judgement), common humanity (vs. isolation), and mindfulness (vs. over-identified). The three

aforementioned components and their corresponding opposites form the six-subscales that make up the SCS. The SCS consists of 26-items on a 5-point Likert rating scale. An individual's self-compassion is therefore the total score obtained after answering all 26-items in the SCS. The full English SCS has been attached in Appendix B. The six sub-scales, their respective items, and an example of an item have been provided in Table 2.1 below.

Table 2.1

Sub-scales of SCS with Item Numbers and Example

Sub-scale	Item Number	Example
1. Self-kindness	5, 12, 19, 23, 26	#5 - I try to be loving towards myself when I'm feeling emotional pain.
2. Self-judgement	1, 8, 11, 16, 21	#1 - I'm disapproving and judgmental about my own flaws and inadequacies.
3. Common Humanity	3, 7, 10, 15	#3 - When things are going badly for me, I see the difficulties as part of life that everyone goes through.
4. Isolation	4, 13, 18, 25	#4 - When I think about my inadequacies, it tends to make me feel more separate and cut off from the rest of the world.
5. Mindfulness	9, 14, 17, 22	#9 - When something upsets me I try to keep my emotions in balance.
6. Over-identified	2, 6, 20, 24	#2 - When I'm feeling down I tend to obsess and fixate on everything that's wrong.

The SCS is a 5-point Likert rating scale where 1 denotes Almost Never and 5 denotes Almost Always. The original author has not provided descriptors for the remaining scale levels. This scale captures discrete interval data and sub-scale scores are determined by calculating the mean of all the sub-scale items. For a total score of self-compassion, the negative sub-scale items need to be reverse coded, namely Self-judgement, Isolation, and Over-identified (i.e. 1 = 5, 2 = 4, 3 = 3. 4 = 2, 5 = 1). The positive sub-scale items, namely, Self-kindness, Common Humanity, and

Mindfulness) can be summed up as they are. The researcher should calculate the mean score of each individual sub-scale. The resulting 6-subscale mean scores should then be added up and divided by six; this is the grand mean of all the 6-scubscale scores. This final score is the total self-compassion score of the individual. The highest possible value of total self-compassion is five and the lowest possible value of total self-compassion is one. Neff (2003) does not provide any suggestions for cut-off points for the SCS scale.

2.4 Past Studies on the SCS

The SCS has been translated into a variety of languages such as Chinese, Japanese, Korean, Portuguese, Dutch, Norwegian, Spanish, Hungarian, German, and many others (Neff et al., 2018). There has been substantially more research in translating the SCS into a variety of European languages, whereas only three Asian language translations have been officially published on Neff's website. Several of these studies will be explored in greater detail, with regards to their translation methods, and their results on validity, reliability, and factor structure of the SCS.

2.4.1 Asian Translation Studies

It is interesting to note that the first few studies validating translations of the SCS were primarily from Asian countries such as China, Korea, and Japan. Even though Neff hails from America, Asian researchers picked up on the SCS and sought to apply it into their local context very quickly. On the other hand, considering the scale is based on Buddhist philosophy, it seems prudent that it was translated and validated in Eastern societies, albeit in a slightly roundabout fashion after first being validated in a Western context. One of the early validation studies was conducted by Arimitsu (2014) on a Japanese sample. The author first used a backtranslation method to translate the SCS into Japanese after which a Confirmatory Factor Analysis (CFA)

was run with data from 366 participants and revealed that all 26-items had an acceptable fit (CFI = .86). In addition to that, the SCS-J produced a strong internal consistency (Cronbach's α = .84). Further testing with 101 participants on the SCS-J with a 3-week internal revealed good test-retest reliability (r = .83). Lastly, the author checked for construct validity by correlating the SCS-J with Narcissistic Personality Inventory (NPI), the Subjective Happiness Scale (SHS), the State-Trait Anxiety Inventory-Trait form (STAI), and the Beck Depression Inventory (BDI). Correlations revealed the SCS-J to have both convergent (SHS), and discriminant validity (NPI, STAI, BDI) with the other respective scales. Overall, Arimitsu (2014) concluded that the SCS-J was valid, reliable, and produced the same 6-factor correlated model as the original SCS. These findings were similarly obtained by other Asian researchers, namely Chen et al. (2011) who also utilized a backtranslation method with their Chinese sample, obtained a strong internal consistency (r = .84) and a good test-retest reliability (r = .89) as well as supporting the 6-factor model for their SCS-Mandarin. In addition to that, a study in Korea also found similar results. Lee and Lee (2010) investigated the reliability and validity of the SCS-Korean with a sample of 405 Korean women and discovered that it too continued to exhibit good validity and reliability after translation as well as having a good fit for the 6-factor model.

2.4.2 Malaysian Translation Study

There is currently only one SCS translation which has been conducted in Malaysia. Using a backtranslation method, Khatib et al. (2018) were able to produce a reliable SCS in BM, however at the time of writing, the current researcher was unable to contact them. With the help of three psychology experts, the authors claimed that 12 items in the scale needed to be fixed, however the authors did not provide any details on exactly what was wrong with those items or how those items were eventually

fixed. In addition to that, the study is rather limited as it only looked one psychometric property, which was the internal consistency of the scale, α = 0.80, however this finding is consistent with past research (Khatib, et al., 2018). The current study will be extending the investigation of the SCS further by using a factor analysis, test-retest reliability, as well as exploring construct validity (i.e, convergent and divergent).

2.4.3 European Translation Studies

In more recent years, more European researchers have validated the SCS in languages such as Portuguese, Italian, German, Spanish, and others. One of the earliest translations in Europe was done in Portuguese by Castilho and Pinto-Gouveia (2011). The authors were primarily concerned with translating and investigating the psychometric properties of the SCS within the Portuguese community. They employed the standard backtranslation method, similarly applied in Asian studies, and the final SCS-P achieved a strong internal consistency of .89 as well as a good test-retest reliability of .78, with an interval of 4 weeks in between testing. What sets this study apart from others is that the authors validated the SCS-P in two very different Portuguese speaking samples. The first sample consisted of a large mixed group of university students and members of local community groups (N=1128) while the second sample consisted of clinical participants (N=316) who had either Axis I or Axis II disorders as determined via clinical assessment. The SCS-P was found to be valid and reliable in both samples and CFA results confirmed the original six-factor structure of SCS as well as the single higher-order factor as well. In addition to that, Costa et al. (2015) also examined the SCS-P with a clinical sample and produced the same CFA results. Both these studies add to the growing evidence that the factor structure of SCS remains consistent across different samples (both clinical and non-clinical).

2.4.4 Summary of Translation Studies

In summary, both Asian and European translation studies of the SCS have many similarities. Firstly, backtranslation is the preferred method of translating the scale into the local language. Secondly, all these studies obtained strong reliability for the SCS after it was translated. Thirdly, almost all the studies had a minimum sample size of about 300 people (the Malaysian study is the only exception). Fourthly, out of the seven quoted studies which ran factor analysis, a total of five studies discovered a six-factor model while only two studies found a bi-factor model instead. Due to all these similarities, it is fairly apparent that the SCS does not have any major issues when being translated across different cultures.

2.4.5 Novel Features of a European Study

The Norwegian translation (SCS-D, D for Dutch) of the SCS by Dundas et al. (2016) has two novel features that set it apart from all other SCS translation studies. Firstly, this study has the smallest sample size in comparison to all other European studies, it reports a bivariate correlation, and is one of the only studies to employ a mediation analysis to study the relationship between SCS and mental health.

With regards to sample size, Dundas et al. (2016) recruited only 277 participants for their study while almost all other European translation studies on the SCS exceeded 300 participants. Dundas et al. (2016)'s results supported the SCS factor structure in spite of having a smaller sample size compared to other studies, which suggests that the SCS can be validated with a comparable sample size.

Dundas et al. (2016) also reported that the 6-factor structure of the SCS-D had good fit in comparison to a 3-factor model and also noted a bivariate correlation between 3 dimensions of the opposite factors in SCS. Dundas et al. (2016) elaborated that there is no mutual exclusivity between the positive and negative subscales of the

SCS. In other words, there is a connection between the positive and negative subscales of the SCS. Some researchers refer to the 3-factor or bi-factor model in an attempt to separate the positive and negative subscales of the SCS, but Dundas et al.'s (2016) results suggest that those alternative models are unsupported.

Secondly, what makes Dundas et al.'s (2016) is noteworthy because it is one of the only studies which employed mediation to analyze the relationship between self-compassion and depression. In comparison, with reference to Table 2.2, the other studies chose to investigate the relationship between SCS and mental health outcomes in the framework of convergent and divergent validity with relevant scales. Dundas et al. (2016) discovered that the link between self-condemnation and depressive symptoms was weaker in people who had higher levels self-compassion, thus concluding that the relationship between self-condemnation and depressive symptoms was moderated by self-compassion. Similarly, a further mediation analysis found evidence to suggest that individuals with high self-compassion decrease self-condemnation, which in turn decreases depressive symptoms. No mediation effect was found within the group that had low self-compassion (Dundas, et al., 2016). In other words, self-compassion can be considered as a resilience factor as opposed to a risk factor because individuals with higher self-compassion are able to reduce negative mental health outcomes, but at the same time, those with low self-compassion do not exacerbate negative mental health outcomes (Dundas et al., 2016)

2.4.6 Bi-Factor Model

In spite of the previous evidences for a 6-factor model, Costa et al. (2015) considered a bi-factor model of SCS instead. This bi-factor model separately grouped the positively framed SCS factors (self-kindness, common humanity, mindfulness) vs. the negatively framed SCS factors (self-judgement, isolation, over-identification).

López et al. (2015) also discovered evidence for a bi-factor model of SCS. Their study included 1736 adult local community participants from the Netherlands who were required to complete the Dutch translation of the SCS. Their study was unable to confirm the six-factor model via CFA analysis, whereby they obtained poor fit for CFI = 0.896, TLI = 0.879 and RMSEA = 0.095. Investigating further, the authors discovered via an EFA that factors were loading onto 2-factors instead of six. Results showed 12 of the negatively formulated items loading onto Factor 1 (negative: self-criticism) while the remaining 12 positively formulated items loading onto Factor 2 (positive: self-compassion); this explained 45.4% of the variance. Therefore, López et al. (2015) made an argument in favour of testing SCS using a bi-factor model instead of the usual six-factor. According to them, "when analyzing all items simultaneously…we found that the positively and negatively formulated items formed separated factors, indicating that self-kindness, common humanity and mindfulness combine into one factor, and self-judgment, isolation, and over-identification into another factor (López et al., 2015, p. 8). The authors reasoned theoretically that the SCS was actually measuring self-compassion (self-kindness, common humanity and mindfulness) and self-criticism (self-judgment, isolation, and over-identification). They elaborated that the SCS should be thus considered as a bipolar construct measuring along a scale with high self-compassion at one end, and high self-criticism on the other end. The bi-factor model suggests that the negative and positive sub-scales exist in silo, whereas, Neff (2003) argues that both positive and negative subscales interact with each other to ultimately produce self-compassion.

In spite of this evidence of a bi-factor model, another question must first be asked; Why were the researchers unable to confirm the six-factor model of SCS in their Dutch sample in comparison to other samples? López et al. (2015) briefly

mention that they utilized a 24-item SCS-Dutch which had previously been translated but two items had been removed due to translation issues. Other studies almost never removed any items, which leads to a slight doubt in the accuracy of the Dutch translation which may in turn, be responsible for the lack of a six-factor model confirmation. Amidst the wealth of translation studies which confirmed the six-factor model, only two studies were unable to replicate this result. Hence, there is a greater burden of proof on the shoulders of researchers who find different results to further investigate and supply evidence to support the bi-factor model.

2.4.7 Diverse Sample and ESEM Study

To address the controversy of the bi-factor model, Neff et al. (2018) rose up to the occasion and in response, examined the SCS using 20 diverse samples via a CFA and Exploratory Structural Equation Modelling (ESEM) analysis. One of the critiques towards the study by López et al. (2015) is that the researchers neglected to perform an additional CFA to comprehensively confirm the emergence of the bi-factor model in their EFA. In addition to that, according to Neff et al. (2018) the bi-factor model may be explained by the possibility of overlap in meaning between the factors due to similarity between items. Cross loadings thus occur and because of the stringent parameters of the CFA which forces items to load onto factors, but fails to account for cross loading intricacies, the "negative" and "positive" factors appear to group together distinctly. Neff et al. (2018) states that the "ESEM framework appears to be especially appropriate for the SCS because it can simultaneously model both the specific and overall relationship of items using a bi-factor analytic approach as well as their interaction as a system with an ESEM approach" (p. 12). This is because in an ESEM analysis, the associations between factors are expressed as a function of item cross-loadings, hence overcoming the drawback of a CFA. Their results showed excellent

fit for both a single-factor ESEM and a six-factor correlated ESEM across their 20 diverse samples. Neff et al. (2018) concluded that SCS measures self-compassion as a function of 6 distinct factors which are also interacting to form a global construct. Therefore, this study will hence be going forward with the psychometric analysis of a six-factor and single higher order factor model. These aforementioned studies have been briefly summarized in Table 2.2. For a more comprehensive list of past studies on the SCS kindly look at Neff et al. (2018).

Table 2.2

Summary of Translation Studies

No	Authors(s)	Nation	N	Translation Method	Accepted Model	Internal Consistency (α)	Test-retest (*r*)
1	Arimitsu (2014)	Japan	366	Backtranslation	Six-factor first order CFA	.84	.83
2	Chen et al. (2011)	China (undergraduate students	660	Backtranslation	Six-factor model (EFA & CFA)	.84	.89
3	Lee & Lee (2010)	Korea (community females)	405	N/A	Six-factor first order CFA	.87	N/A
4	Castilho and Pinto-Gouveia (2011)	Portugal Community: Clinical:	1128 316	Backtranslation	1. Six-factor first-order CFA 2. Six-factor higher-order CFA (with one G-factor)	.89	.78
5	Dundas et al. (2016)	Norwegian (Undergrad students)	277	Backtranslation	Six-factor first-order CFA	.89	N/A
6	Costa et al. (2015)	Portugal (Community and Clinical)	361	Backtranslation	Bi-factor first-order CFA	Both factors = .91	N/A
7	López et al. (2015)	Netherlands (community)	1643	Backtranslation	Bi-factor first-order EFA	a) Positive = .86 b) Negative = .90	N/A
8	Khatib et al. (2018)	Malaysia (adolescent students)	125	Backtranslation	No factor analysis	.80	

The studies included in Table 2.2 have been selected by this current study because they each possess something unique. Arimitsu (2014), Chen et al. (2011), as well as Lee and Lee (2010) were chosen because they represent the earliest studies done on the SCS in an Asian context. Castilho and Pinto-Gouveia (2011) and Costa et al. (2015) are noteworthy because they included both a community and a clinical sample in their studies. In terms of sample size, Castilho and Pinto-Gouveia (2011) has the biggest (N = 1128 + 316) while on the other hand, Dundas et al. (2016) had the smallest (N = 277), yet both were able to find a six-order factor model and high internal consistency and reliability. Lastly, López et al. (2015) and Costa et al. (2015) were included because they were the only two studies which reported a bi-factor model. Finally, the study by Khatib et al. (2018) was included as it is (at the time of writing) the only translation study of the SCS in Malaysia. For a more comprehensive list of studies kindly refer to Neff et al. (2018).

2.4.8 Stability Across Samples

In spite of pilot testing however, there is still a concern with regards to the stability of the SCS across different samples. Considering that Zeng, Wei, Tian, and Liu (2016) used an SCS which was already validated in Mandarin (albeit with a different sample of Chinese participants), if the reason they failed to obtain similar results as other studies was because of a lack of pilot testing with the exact target sample (Chinese Buddhist), then future researchers would need to painstakingly pilot test the SCS every time they intend to administer it to a new sample. In addition to this, some criticism has been aimed at the over-reliance on student samples with regards to examining the psychometric properties of the SCS (López et al., 2015).

Rising up to address this concern was Neff et al. (2018), who conducted a very recent study to examine the factor structure of the SCS using an impressive total of 20

diverse samples; consisting of students, community and even clinical samples. With the help of 16 other authors, Neff defended the psychometric properties of her scale by confirming that it was both valid and reliable even when tested across many different samples. Their results indicated excellent fit for a 6-factor model (representing the constituent subscale scores of the SCS) as well as prominent single overarching factor (representing an overall self-compassion score from the scale). The results of their study put to rest concerns that other researchers have brought up regarding the integrity of the psychometric properties of the SCS across different samples.

2.4.9 Construct Validity

Theoretically, and in past studies, the SCS has been linked to both psychological well-being and psychological problems (Castilho & Pinto-Gouveia, 2011; Chang et al., 2019; Dundas et al., 2016; Neff, 2009; Neff, 2012; Neff et al., 2014). In order to assess this the researcher utilized construct validity which can be measured in two main ways: convergent and divergent validity. Past researchers utilized a variety of additional scales in order to establish the construct validity of the SCS after translation. The SWLS is suitable to be used for convergent validity because it measures subjective well-being (Diener, 1984), which parallels the conceptualization of self-compassion. As a result of that, the SWLS has been utilized by researchers to assess the covergent validity of the SCS (Benda & Reichová, 2016; Deniz, Kesici, & Sümer (2008); Kotsou & Leys (2016). On the other hand, the BDI represents a suitable scale to test for divergent validity because it accounts for negative items that represent depression, which is the opposite of the concept of self-compassion. Past researchers have utilized the BDI as a measure to assess the divergent validity of the SCS (Benda & Reichová, 2016; Garcia-Campayo et al. (2014); Lee and Lee, 2010).

The SWLS and BDI are valid, reliability and well established in past research. Moreover, they have been translated into many different languages, making them an ideal choice due to their accessibility. In addition to that, Swami and Chamorro-Premuzic (2009) have confirmed that the SWLS-BM is valid and reliable to be used within the Malay speaking Malaysian population, while at the same time, Mukhtar and Oei (2008) confirmed that the BDI-BM is equally valid and reliable to be used within a Malay speaking Malaysian population.

Table 2.3

Studies Testing Construct Validity

No	Authors(s)	Construct Validity	Scale Used	Results
1	Benda & Reichová (2016)	Convergent Divergent	SWLS BDI	Significant Significant
2	Kotsou & Leys (2016)	Convergent	SWLS	Significant
3	Deniz, Kesici, & Sümer (2008)	Convergent	SWLS	Significant
4	Lee & Lee (2010)	Divergent	BDI	Significant
5	Garcia-Campayo et al. (2014)	Divergent	BDI	Significant

The current researcher will be keeping in alignment with past researcher and shall thus employ the usage of both the SWLS and the BDI to test the convergent and divergent validity of the SCS respectively; both of these combined will provide the researcher with an understanding of the construct validity of the SCS once it has been translated into BM.

2.5 Backtranslation Framework

The conceptual framework of this current study rests predominantly upon the translation process. As can be seen from Table 2.2, past studies on the SCS rely heavily on the standard backtranslation method. Most of the authors cite Brislin (1970) as reference for their backtranslation methodology. Table 2.4 below is a summary of the four main processes involved in backtranslation as listed by Tyupa (2011).

Table 2.4

Summary of Backtranslation Process

Steps	Process
1. Forward Translation	Original scale is forward translated into target language.
2. Backtranslation	Translated scale is then translated back from target language to the original language.
3. Backtranslation Review	The back translated scale is then compared to the original scale and inconsistences are resolved to ensure both scales are congruent.

In the case of the SCS, its original language is English, which is first translated into a target language, i.e. BM. Brislin (1970) suggests that a bilingual translator be employed to do this. Next, this translation is then backtranslated back into English. According to Brislin (1970), a separate translator should be hired in order to prevent any carry over effects from the first translation. This second translator is thus seeing the scale for the first time, and will not be biased by its previous translation. According to Maneesriwongul and Dixon (2004), translators should be bilingual and the goal of their translation should be to achieve "conceptual equivalence." This means that the translators should be both fluent in English and BM as well as being familiar with the area of psychology so that they can more accurately translate the conceptual meaning from one language to another with regards to the SCS. Lastly, the backtranslation and original SCS-English should be then compared to check for any inconsistencies in meaning of translated items. Any inconsistencies or ambiguities in understanding the backtranslated version will be addressed by this study.

2.5.1 Notable Translation Method

Although almost a majority of researchers utilized a standard backtranslation method when translating the SCS from English to the target local language, there was a study by de Souza and Hutz (2016) which employed a more exhaustive method for

translation involving a total of nine steps, including, but not limited to: two forward translations by Brazilian scientific scholars, two focus groups, refinement by bilingual experts, a backtranslation, and finally a check by Kristin Neff herself. The authors went out of their way to translate the scale in this way to ensure cultural validation, and it is unsurprising that they obtained a very strong internal consistency ($\alpha = .92$). While some researchers may praise the thoroughness of such a translation procedure, others may bemoan its overly lengthy and cost inefficient process. Hence, the current study will seek to find a balance between ensuring quality translation and maintaining cost efficiency.

2.5.2 Weakness of Backtranslation

In spite of the popularity of the backtranslation method in SCS studies, there are some critiques on this method of translation that should be addressed. According to Behr (2016), backtranslation does not necessarily result in translation accuracy due to the lack of cultural adaptation in the backtranslation process; translators are more likely to view a translation from a purely linguistic rather than a cultural perspective. The author elaborated that "…translation could lack naturalness, be difficult to comprehend or may simply be downright wrong" (p. 2). In the same vein, there are some concepts or phrases which are difficult to express in a conceptually similar manner in a different language (Maneesriwongul & Dixon, 2004). In order to address this Behr (2016) calls for an emphasis on translation quality, rather than just going through the motions of the process of backtranslation itself. Unfortunately, however, hiring multiple expert bilingual translators to enhance backtranslation quality is both costly and time consuming (Maneesriwongul & Dixon, 2004).

According to Maneesriwongul and Dixon (2004), backtranslation alone is not sufficient; based on a comprehensive review of past literature, they suggest that backtranslation should be paired with reliability testing. To quote the authors, "assessment of internal consistency gives an indication of reliability of the target language version" (Maneesriwongul & Dixon, 2004, p. 9).

In addition to that, Behr (2016) and Maneesriwongul and Dixon (2004) note the inclusion of a qualitative component for some backtranslation studies. Behr (2016) elaborated more on this qualitative component taking the form of a type of interview called cognitive interviews. These cognitive interviews were used to gather more in-depth feedback regarding the quality of translations. The current study will therefore be adopting cognitive interviews to help overcome the aforementioned weaknesses of the backtranslation method.

2.5.3 Cognitive Interviews

In order to maintain translation quality while balancing between the costs of a lengthy translation procedure, the current study intends to utilize cognitive interviews to preserve translation integrity. Cognitive interviews are a type of interview which provides validity evidence for responses to survey items by looking at how target participants comprehend individual items in a scale (Padilla & Benítez, 2014; Willis, 1999). This allows a researcher to check if participants really understand the items after they have been translated. As mentioned in Section 2.3.2 a weakness of typical backtranslation methods is that the translations are primarily decided upon by linguistic experts but the layperson does not always conform to expert-translations. Cognitive interviews therefore provide a cost-efficient addition to the research process in order to ensure the translated SCS-BM items will be of good quality. In addition to that, the previous seven translation studies mentioned in Table 2.2 as well as a list of

35 other studies on the SCS (Neff et al., 2018) did not utilize cognitive interviews to check participant comprehension of translated items; hence, the current study will be the first to explore this.

2.6 Summary

Considering the literature review on SCS, it becomes apparent that SCS is not only a viable but a useful tool to be translated and used in a local context. Upon reviewing both the Asian and European translation studies, it appears that there are no significant differences in the SCS when translated and applied to culturally different samples. Additionally, in spite of some controversy regarding its scale properties, the literature review strongly suggests that the SCS maintains its psychometric integrity, even in the face of diverse samples. With the addition of cognitive interviews to boost the backtranslation process, the current study intends to produce a valid and reliable SCS-BM. Furthermore, translation of the SCS into BM will add to the theoretical body of literature documenting the psychometric properties of the various translations of SCS. Moreover, Malaysian researchers and mental health workers will benefit from having a valid and reliable scale by which to measure self-compassion.

CHAPTER 3

METHODOLOGY

3.1 Introduction

The following section contains an outline of the methodology of this study. This includes the research design and its corresponding paradigm. The instrumentation used and necessary improvements made by the current study. Information about sampling will also be provided. Lastly, the entire procedure for the research will be outlined and this will be divided into two sections, namely, Pilot Study and Actual Study.

3.2 Research Paradigm

This study utilizes a quantitative design; hence it aligns with a positivist research paradigm. According to the positivist paradigm, there exists an objective reality underlying behavior that can be measured and tested. With regards to the current study, the assumption is that there is an objective measurement of the latent variable of self-compassion which can be measured and tested via the SCS.

The purpose of this paradigm is to provide generalizability, prediction and also causal explanations for behavior (Lincoln & Guba, 1985). Similarly, the current research seeks to generalize and make predictions regarding the psychometric properties of the SCS. There is no casual effect being tested as this is a translation and validation study.

Lastly, Lincoln and Guba (1985) listed that the positivist paradigm requires researchers to consider elements of theory, operational definitions, control, experimentation, statistical analysis, deduction, and etc. with regards to its approach to research. Correspondingly, the current research will be based upon positivist principles with regards to measuring and testing the reliability and validity of the SCS with regards to theoretical framework suggested by Neff, the operational definition of

self-compassion, as well as subsequent factor analysis to confirm the reality of self-compassion.

3.3 Research Design

The current study will be utilizing a quantitative research design with backtranslation as the instrument development method and surveys for testing. It is quantitative due to the numerical nature of the data collection and analysis which will be performed in this study. Part of the study will be correlational as it involves reliability testing. Another part of the study will be using factor analysis as it involves CFA. The component of cognitive interviews will be conducted in a structured manner, hence, should also be considered as form of quantitative data because it will not be open ended.

3.4 Instrumentation

This study will be using a total of three scales, namely, the SCS-English (Neff, 2003), the Satisfaction with Life Scale (SWLS) and Becks Depression Inventory (BDI). Sections 3.4.2 and 3.4.2 will go into detail into providing justification for the use of the SWLS and BDI as measures of construct validity respectively.

3.4.1 SCS Properties

Permission to use this scale for research is provided by Neff (See Appendix A). The SCS consists of 26-items in a self-report questionnaire that measure the six components of the latent construct of self-compassion (Neff, 2003). The first two opposite components are Self-Kindness (e.g. Item 12: When I'm going through a very hard time, I give myself the caring and tenderness I need) vs. Self-judgement (e.g. Item 1: I'm disapproving and judgmental about my own flaws and inadequacies). The second pair of opposite components are Common Humanity (e.g. Item 15: I try to see my failings as part of the human condition) vs. Isolation (e.g. Item 4: When I think

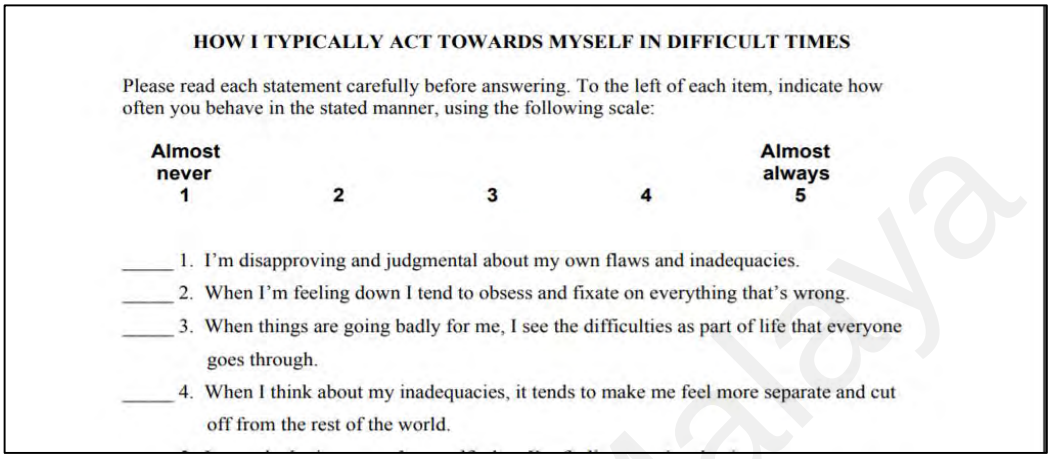
about my inadequacies, it tends to make me feel more separate and cut off from the rest of the world). The final pair of opposite pair of components are Mindfulness (e.g. Item 9: When something upsets me I try to keep my emotions in balance) vs. Over-identified (E.g Item 20: When something upsets me I get carried away with my feelings).

The three positive sub-scales are: Self-kindness, Common Humanity and Mindfulness, while the respective negative sub-scales are: Self-judgement, Isolation and Over-identified. The three negative subscales need to be reverse scored before calculating the mean (i.e. 1=5, 2=4, 3=3, 4=2, 5=1). Responses for each item can range from 1 (Almost Never) to 5 (Almost Always). Neff (2003) advises researchers to calculate scores by first calculating the mean for each sub-scale individually and then later calculating the grand mean for a total self-compassion score. Therefore, the maximum Self-compassion score a person can achieve is 5 and the minimum self-compassion score a person can achieve is 1. The SCS-English displayed a strong internal reliability of Cronbach's α =.92, while the individual sub-scales also exhibited good internal reliability with a range of Cronbach's α from .75 to .81. In addition to that, test-retest reliability for the whole scale was also strong after a three-week interval, with Cronbach's α =.93 for the entire scale and Cronbach's α ranging from .80 to .88 for the six-subscales.

3.4.1.1 Improvements to the SCS

In the original English SCS this study noticed that Neff only provided two labels out of the 5-points in the Likert rating scale. The current research feels this is not appropriate because according to Bertram (2012) each level of the scale requires a specific descriptor. For example, in many 5-point Likert rating scales, the descriptors are as follows: 1 – Strongly Disagree, 2 –

Disagree, 3 – Neither Agree or Disagree, 4 – Agree, 5 – Strongly Agree. In contrast however, the original SCS-English contains only descriptors for 5: Almost Never and 1: Almost Always. Figure 3.1 illustrates this problem visually.



HOW I TYPICALLY ACT TOWARDS MYSELF IN DIFFICULT TIMES

Please read each statement carefully before answering. To the left of each item, indicate how often you behave in the stated manner, using the following scale:

Almost never				Almost always
1	2	3	4	5

_____ 1. I'm disapproving and judgmental about my own flaws and inadequacies.
_____ 2. When I'm feeling down I tend to obsess and fixate on everything that's wrong.
_____ 3. When things are going badly for me, I see the difficulties as part of life that everyone goes through.
_____ 4. When I think about my inadequacies, it tends to make me feel more separate and cut off from the rest of the world.

Figure 3.1. Screenshot of the SCS-English

As can be seen from Figure 3.1, there is a clear ambiguity for the points 2, 3, and 4 because they are not clearly labelled. In addition to this, the literature review did not indicate if other researchers noticed or made any additions to the scale after translation. The current study has thus corrected this and the additional descriptors which follow the typical Likert rating scale format can be seen in Appendix B, namely, 1 – Almost Never, 2 – Rarely, 3 – Sometimes, 4 – Most of the time, 5 – All the time. These additional descriptors are adapted from a research done by Awang, Afthanorhan, and Mamat (2016).

3.4.2 SWLS

The SWLS has been utilized by researchers to assess the construct validity of the SCS (Benda & Reichová, 2016; Deniz, Kesici, & Sümer (2008); Kotsou and Leys (2016). In addition to that, the SWLS has already been translated into BM and can thus easily be incorporated into the current study without a further need for translation. The SWLS-BM consists of five statements to measure an individual's satisfaction with

their current life. Respondents can choose to answer from a 7-point Likert scale ranging from 1: Strongly Disagree to 7: Strongly Agree. The SWLS-BM is scored by simply summing up the score of each item; the higher the total score, the more satisfied the individual is with their life. The SWLS-BM can be found in Appendix P. The SWLS was translated into BM by Swami and Chamorro-Premuzic (2009) and they discovered that it exhibited good psychometric properties such as a good internal consistency of Cronbach's $\alpha = 0.83$. In addition to that, the authors also ascertained, via confirmatory factor analysis, that the SWLS-BM possessed a unidimensional factor structure. The authors also conducted multi group analyses which revealed that the unidimensional model remained constant across ethnic groups and sex; this is consistent with the original SWLS model. Swami and Chamorro-Premuzic (2009) concluded that the SWLS-BM is a valid and reliable scale for measuring subjective well-being among Malay speaking groups in Malaysia.

The SWLS is a suitable measure for construct validity because it measures the concept of life satisfaction, which can be considered as convergent with the concept of self-compassion, hence it falls under the type of construct validity known as convergent validity. According to Diener (1984) the SWLS is based upon the concept of subjective wellbeing (SWB), and SWB consists of three components, namely: positive affect, negative affect, and life satisfaction. With regards to life satisfaction, Pavot and Diener (2008) claim that it represents a form of cognitive evaluation about the circumstances of one's life. This is similar to the concept of mindfulness in the SCS which refers to the cognitive observation of one's thoughts without over exaggerating them (Neff, 2003). In addition to that, the self-kindness factor in the SCS also represents a form of cognitive evaluation as an individual with self-kindness will evaluate themselves and then choose express compassion towards themselves (Neff,

2003). Therefore, life satisfaction, mindfulness, and self-kindness share a similarity in that they are all conceptually understood to refer to a cognitive process in which an individual is evaluating their thoughts. In the case of the SWLS, this evaluation is to determine one's satisfaction, whereas in the SCS, this evaluation is to ensure that negative do not get blown out of proportion. Thus, the SWLS provides a suitable measure of convergent validity to the SCS because they both measure similar cognitive evaluation processes.

3.4.3 BDI

Past researchers have utilized the BDI as a measure to assess the construct validity of the SCS (Benda & Reichová, 2016; Garcia-Campayo et al. (2014); Lee and Lee, 2010). In addition to that, the BDI has already been translated into BM and can thus easily be incorporated into the current study without a further need for translation. The BDI-BM consists of 20 statements designed to measure factors of depression (e.g., sadness, pessimism, suicidal thoughts, etc.). For each statement, respondents will choose to answer from a 4-point Likert scale. The score on each item is summed up to create the total score of an individual's depression; the higher the score the more severe that person's depression. The BDI-BM used in this study can be found in Appendix P. The psychometric properties of the BDI-BM were assessed by Mukhtar and Oei (2008), who sampled among students, community, general medical patients and patients suffering from major depressive disorders. The authors discovered that the factor structure of the BDI-BM loaded appropriately onto two factors, namely, Cognitive/Affective and Somatic/Vegetative; this is consistent with Beck's model. Internal consistency for the entire scale was found to be strong, with Cronbach's α = .91. In addition to that, the authors found evidence for concurrent and discriminant validity when running correlations between the BDI-BM and other relevant scales.

Hence, researchers Mukhtar and Oei (2008) concluded that the BDI-BM is a valid and reliable measure for levels of depression among the Malay speaking community of Malaysia.

The BDI is a suitable measure for construct validity because it measures the concept of depression, which can be considered as divergent with the concept of self-compassion, hence it falls under the type of construct validity known as divergent validity. The BDI is broadly divided into two subscales: cognitive-affective and somatic-performance (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). The cognitive-affective component in the BDI is similar to the cognitive evaluation aspect of the SCS, in that they both represent a cognitive process in understanding depression and self-compassion respectively. Beck et al. (1961) lists out many negative items which represent overt behavioural manifestations of depression, such as: self-hate and self-accusation, which are a cognitive evaluation that results in an individual becoming very critical towards themselves. This is similar to the factor of self-judgement in the SCS which also refers to an individually cognitively evaluating themselves in a way that results in their own self-directed harsh criticism (Neff, 2003). In addition to that, the BDI also measures more and social withdrawal, which is behavioural rather than cognitive. Similarly, the SCS measures isolation, which Neff (2003) describes as an individual feeling alone and separated from everyone else due to the failures. The items in the SCS provide a total score on self-compassion, while the BDI on the other hand provides a total score on depression. Therefore, the BDI is a suitable scale to be used for divergent validity because it provides a score that is conceptually opposite to the SCS.

3.5 Procedure

Firstly, a backtranslation will be conducted to translate the SCS-English into SCS-BM. Secondly, cognitive interviews and subsequent textual analysis will be conducted on a sample of participants who have completed the SCS-BM. An informed consent form and the SCS-BM items will be printed and given to participants to sign and complete respectively. All responses will be recorded and transcribed and those transcriptions will be analyzed accordingly. The SCS-BM will then be administered via a paper survey to 60 participants. They will only receive this link after signing the informed consent form. Data will be collected and analyzed using Cronbach's α in order to determine the internal consistency of the overall scale and its six sub-scales. Based on the results of the cognitive interviews and reliability analysis, the SCS-BM will be refined further if necessary. The finalized SCS-BM, the BDI-BM, and the SWLS-BM, will then be administered to at least 300 participants. subsequent data will be collected and analyzed via a CFA, a test-retest reliability, and construct validity.

3.5.1 Ethical Concerns

The researcher did not foresee any sever ethical issues arising from this current study. All participants were required to sign an informed consent (Refer to Appendix C) before participating in the study. The informed consent provided information on the aim of the study, the procedure, participant rights, confidentiality of information and the contact details of the researcher and supervisor were provided should participants have any concerns of questions. Participants had the right to withdraw their participation at any time during the study, but none of them chose to do so. All data was kept confidential and was stored only on the researchers' personal computer only to be used for data analysis.

3.6 Pilot Test 1

Pilot Test 1 included the translation process of the SCS from English to BM. This involved backtranslation by an experienced translator. The translator appointed for this study was Pn Siti Raihani Mohd Saaid, who holds a certificate in translation from the Malaysian Translation and Book Institute (*Institut Terjemahan & Buku Malaysia, ITBM)*, kindly refer to Appendix R for the letter of appointment. The translators' agreement was communicated verbally to the researcher's supervisor. This translation was followed by a review of items by the researcher. Subsequently, cognitive interviews were conducted to assess the quality of translated items via response validity. A final review of items was conducted to bring Pilot Test 1 to completion.

3.6.1 Translation Process

A backtranslation of the SCS scale will be done to translate the English-SCS into a BM-SCS. Keeping in mind that the SCS is a questionnaire in the field of psychology, therefore the translations should be handled by individuals who are experienced in the area of psychology and fluent in both English and BM. Next, this study then compared the backtranslation to the original English version to determine consistency in meaning. This will help determine if the items being translated are still carrying the same meaning as the original English version. Finally, cognitive interviews will be conducted to evaluate if translated items are understood by participants. Any uncertainties in translated items will be edited accordingly.

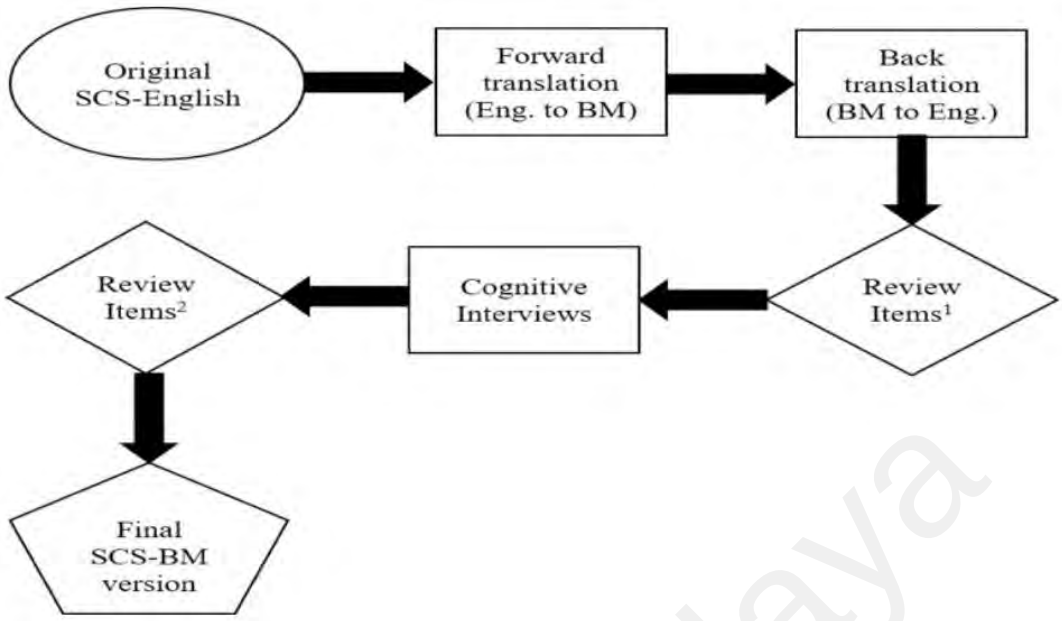


Figure 3.2. Outline of Backtranslation Process in this study (adapted from Tyupa, 2011).

3.6.2 Sampling

This study intends to sample participants who are at least 18 years of age and older. In addition to that, participants need to be fluent in their understanding of BM. The sampling method used is convenience sampling. This study conveniently sampled among the already present BM-fluent students in the Masters in Education Psychology program at University Malaya. This is advantageous because these master's students are familiar with concepts in psychology which should allow them to provide more in-depth responses during the cognitive interviews.

Due to the qualitative nature of cognitive interviews, there is no hard and fast rule as to how many participants to sample. Instead, this study will sample enough participants until theoretical saturation is achieved. Theoretical saturation is defined by Lewis-Beck (2004) as a "phase of qualitative data analysis in which the researcher has continued sampling and analyzing data until no new data appear" (para 1); this

study estimates theoretical saturation to be reached somewhere between five to ten participants.

3.6.3 Cognitive Interviews: Response Validity

In order to determine response validity, six cognitive interviews will be conducted with participants who are fluent in understanding BM. A cognitive interview is a type of structured interview that seeks to explicitly understand the thought process of a person when answering an item (Padilla & Benítez, 2014). The benefit of a cognitive interview is that it would help indicate whether or not participants were processing and understanding the translated BM items in the same way that they should have understood the items in the original English version. Therefore, participants will be interviewed on their understanding of all the 26-items in the SCS-BM, one at a time. Hence, this is considered a form of validity evidence as it will help determine if participants are understanding and responding to items accurately.

The cognitive interviews will be conducted and analyzed according to a method outlined by Padilla and Benítez (2014). As can be seen from Table 3.1, this is a structured form of interviewing because it utilizes specific probing questions to elicit understanding from the participants. The data collected is considered as qualitative data because it deals with participants verbal responses. The cognitive interview is testing the participant's understanding of the BM-SCS items, hence the interview itself shall also be conducted by this study in BM. The examples below are provided in Table 3.1 for both English and BM (italics).

Table 3.1

Types of Cognitive Interview Probes

Type of probe	Example / *Contoh*
1. General probe	"What were you thinking about when you answered this question?" (p. 142) *Apakah yang kamu fikir tentang apabila menjawab item ini?*
2. Comprehension probe	"What does the term 'health' mean to you?" (p. 142) *Apakah yang kamu faham daripada perkataan 'kesihatan'?*

After conducting the cognitive interview, the responses will be transcribed and analyzed. Qualitative analysis primarily involves grouping similar statements or phrases by participants in response to each item. The similarities will be grouped to form conceptual themes. These themes provide insight into how participants understand each item. Differences between participants are compared on each item and finally this study can conclude how each item has been understood overall by participants. The textual analysis will be based on an approach outlined by Willson and Miller (2014) in order to determine response validity. Table 3.2 below shows the outline of the analytic steps.

Table 3.2

Analytic Steps for Cognitive Interview Analysis

Analytic Step	Description
1. Summarizing	Identification of potential themes
2. Comparing across respondents	Identify "What the item captures"
3. Comparing across respondents	Response process differences across respondents
4. Concluding	Explanation of item performance

The main purpose of this analysis is to determine whether or not participants understand the BM-SCS items in the same way that the English-SCS items are

understood. this study will evaluate this by considering how participants responses fall within the boundaries of the theoretical framework suggested by Neff (2003). Each of the 26 English items in the SCS correspond to a specific aspect of self-compassion, hence this study will need to play close attention to determine if participants' responses for each item lean towards an understanding that is congruent with the respective aspect of self-compassion that the item is testing for. Hence, at every step of analysis, the overarching focus is to determine congruity of understanding between the two versions of the SCS. If participants express clear understanding of items in alignment with how the original SCS items are meant to be understood, then those items can be considered valid.

Focus should be given to differences in understanding between respondents towards any items as well as any phrases that participants found to be ambiguous or difficult to understand. Discrepancies or ambiguities in understanding will be addressed by this study in order to better refine the BM-SCS translation so that items carry both the same meaning as the original English-SCS and also to ensure that items are easily and properly understood by participants.

3.6.4 Data Collection

For Part 1 of this study, there are two distinct points of data collection, namely the participants who volunteer to complete the SCS-BM and their subsequent cognitive analysis interviews. In order to maintain ethical standards during the data collection process, an informed consent form (Appendix C) will be prepared, informing participants that their data will be kept private and confidential, as well as their right to withdraw at any given moment. Data from their responses will be keyed into an excel file and saved onto the researcher's computer. Permission will be obtained for the audio recording of participant's responses. All responses will be transcribed

verbatim and participants' names will not be identified; a code will be used instead, i.e. Participant 1.

3.7 Pilot Test 2

After completing the translation, cognitive interviews and review process in Pilot Test 1, the produced scale, i.e. the SCS-BM was now ready for Pilot Test 2 which involved assessing the scale's internal consistency. Items with the lowest reliability score were reviewed accordingly before moving onto the next step in the research.

3.7.1 Sampling

This study utilized haphazard sampling to administer the SCS-BM to 60 participants who are Malaysians above the age of 18 and are fluent in BM. Participants will be haphazardly sampled from the university students of University Malaya. Surveys were administered at the university library as that is a central area which most students can be found at, especially for the study break period. According to Field (2009), using the central limit theorem, a sample size of at least 30 is necessary to achieve a normal distribution, hence to err on the side of caution, this study has doubled the minimum amount to reach a total of 60 participants; this should be sufficient for running the subsequent reliability analysis.

3.7.2 Data Collection

The SCS-BM will be printed and given to participants to fill up. Respondents will be required to agree to an informed consent form before answering the items. The data collected from the SCS-BM is discrete and will be measured at the interval level. All data will be collected and saved in the researcher's computer. Reverse coding will be done accordingly. All participant data is private and will not be identified. Both the sub-scale total scores and the overall total scores for each participant will be summed and entered into the SPSS software for subsequent reliability analysis.

3.7.3 Analytical Procedure

The primary aim of Pilot Test 2 of this study is to determine the internal consistency of the SCS-BM. This will be achieved via analysis using Cronbach's α. According to Field (2009) Cronbach's α is a common tool in measuring the internal consistency for scales. Additionally, Cronbach's α has been used by all past studies in this field. this study will carry out the analysis using the SPSS software which comes with a Cronbach's α analysis option. Based on past research and Field's (2009) suggestions, a minimum overall internal consistency of at least α = .7 should be achieved. If the overall BM-SCS has a reliability that is lower than this, the entire scale would need to be revised. This study will also look at the individual items in order to determine what needs to be edited accordingly to improve reliability. Items that are reducing the reliability of the overall sub-scale can be identified using the "Reliability if item removed" function in SPSS. Once identified the researcher can choose to delete the item or edit it to improve its accuracy.

3.8 Factor Analysis

This section will describe the sampling method and target sample as well as the data collection. The six-factor model of the SCS using CFA which has been validated by past research and relevant fit indices has also been provided.

3.8.1 Sampling

A minimum of 300 participants will be haphazardly sampled from the university students of University Malaya. According to Kass and Tinsley (1979) as cited in Field (2009), it is advisable to sample around 5 to 10 participants per item. The SCS consists of 26 items and multiplying that by 10, the total suggested sample size should be 260. However, Tabachnick and Fidell (2007) as cited in Field (2009) add that having a sample of at least 300 is suitable for factor analysis. Hence the current

study will be aiming to sample at least 300 participants. With regards to the sampling method, as Neff et al. (2018) has already shown, the SCS has shown consistent factor models across a variety of samples and sampling types. Therefore, for the sake of efficiency, the current study will be relying on haphazard sampling. Surveys were passed out at the university library area. All participants must be at least 18 years of age and fluent in understanding BM. No other criteria is needed.

3.8.2 Data Collection

The SCS-BM, the SWLS-BM, and the BDI-BM will be printed and administered altogether to participants (Refer to Appendix P). Demographic questions will be included as well, namely: age, sex, ethnicity, religion. Participants emails will be collected as well to aid in the process of contacting them again should they be selected for the test-retest portion. Data collected from these three scales is discrete and will be measured at the interval level. All data will be kept private and stored only on the researcher's computer. Participant data is not identifiable. In order to ensure reliability, the test-retest approach will be used for the SCS-BM scale. There is no hard and fast rule with regards to the sample size for test-retest reliability, however, based on past research cited in Table 2.2 the current study suggests to use a minimum of 100 participants. These 100 participants will be randomly selected from the original 300 respondents and they will be asked to complete the SCS-BM scale again after a time interval of 3 weeks.

3.8.3 Confirmatory Factor Analysis

Data will first be analyzed using a CFA via AMOS software. A CFA is a multivariate statistical analysis that determines if scale items match their intended construct (Statistics Solutions, 2013). Figure 3.3 below lists the CFA model that will be tested in this study. This was the original six factor model of the SCS discovered by Neff

(2003). This is also known as a first order factor model for a CFA. Furthermore, this first order six-factor model of SCS has been confirmed by Arimitsu (2014), Chen et al. (2011), Lee and Lee (2010), Castilho and Pinto-Gouveia (2011), Dundas et al. (2016).

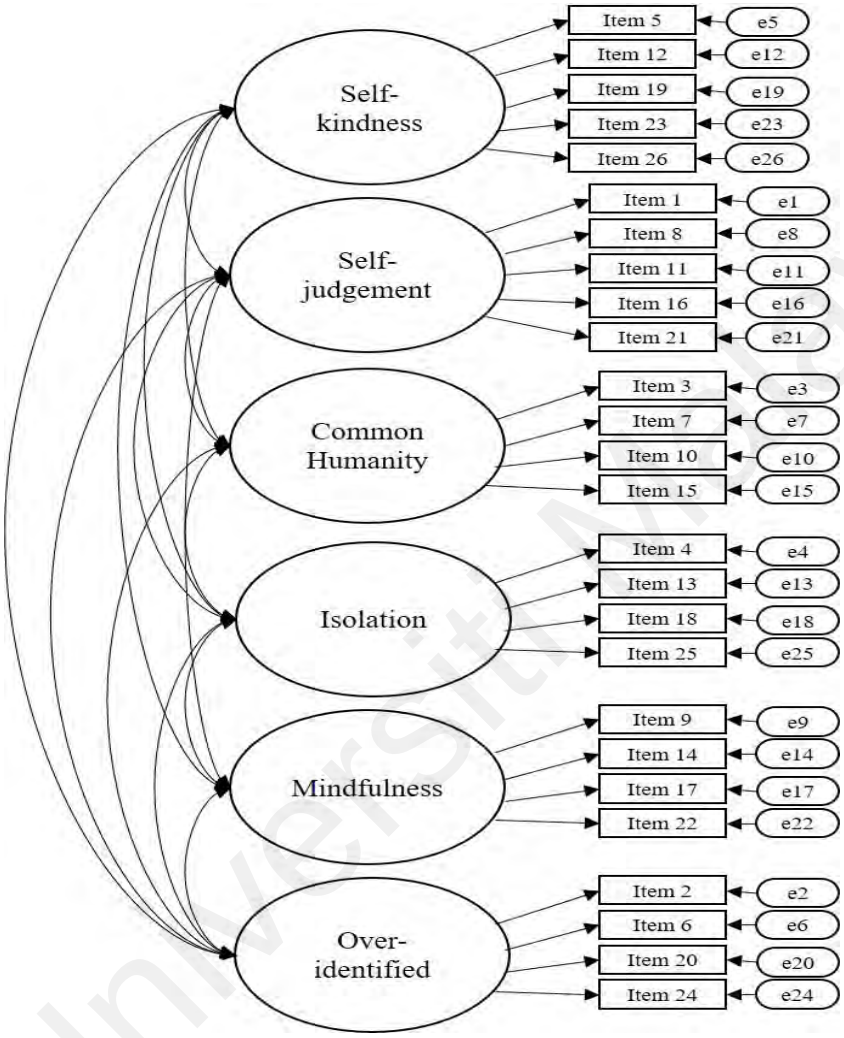


Figure 3.3. First order 6-factor CFA Model of SCS

3.8.4 Fit Indices

In order to determine the fit of the model, this study will be focusing on three fit indices, namely the χ^2, the CFI, RMSEA, and the SRMR. Below is Table 3.3, detailing the definition of each index and its respective values which indicate a good fit as provided by Hooper, Coughlan, and Mullen (2008).

Table 3.3

List of Fit Indices

No	Fit Indices	Name	Definition	Good fit value
1	χ^2	Chi-Square	• Evaluates the overall fit of the model by determining the discrepancy between the fitted covariance matrices and the sample. • The Null Hypothesis (H_0) is that the data fits the model.	$\chi 2 \leq 2$ p-value > 0.05
2	CFI	Comparative Fit Index	• Compares the fit of a null (independent model) to the target model. • Least affected by sample size.	CFI ≥ 0.90
3	RMSEA	Root mean square error of approximation	• Reports how well the model fits the population covariance matrix. • Favours parsimony	RMSEA < 0.08
4	SRMR	Standardized Root Mean Square Residual	• This reports the square-root of the difference between the hypothesized model, the sample covariance matrix, and residuals.	SRMR < 0.08

Although Hooper et al. (2008) list eight fit indices in total, Kline (2005) argues that the Chi-Square test, the CFI, and the RMSEA, and SRMR are the most essential indices to be included when evaluating a model, hence, only those four will be focused on for this study. In addition to this, past research cited in Table 2.2 also commonly report these four fit indices (χ^2, CFI, RMSEA, and SRMR) in their reports, hence the current study will be including them in order to better gauge the fit of the CFA model and also to provide straightforward comparison within the past literature.

3.9 Test-retest Reliability

Next, the test-retest reliability will be determined using Pearson's r correlation. According to Field (2009) test-retest reliability refers to the quality of a measure to be interpreted consistently over time. Based on past research cited in Table 2.2, a minimum reliability score of $r = 0.8$ should be achieved.

3.10 Construct Validity

In order to determine construct validity a convergent and divergent approach will be used. This study chose appropriate scales based on two criteria: 1) The scale should have been used in past translation studies of the SCS and 2) The scale must have a translated BM version so that it can be easily included into this current study.

In order to assess convergent validity, the Satisfaction with Life Scale (SWLS) by Diener, Emmons, Larsen, and Griffin (1985) is used. This scale has been used in past SCS translation research by a number of researchers (See Table 2.4). All these researchers found a positive correlation between the SCS and the SWLS. Permission to use the SWLS is attached in Appendix N.

In order to assess divergent validity, the Becks Depression Inventory (BDI) by Beck et al. (1961) was used. This scale was chosen because it has also been used by researchers in many of the past translations of the SCS (See Table 2.4), all of whom found a negative correlation between the SCS and the BDI. This study was unable to obtain direct permission to use the BDI-BM from Mukhtar and Oei (2008), however, proof of attempt at contacting the author of the translation has been attached in Appendix O. According to The University of Kentucky (n.d.) a material can be used under the provisions of fair right usage if the material is used for educational or non-profit purposes.

3.11 Summary

This chapter details the entire research process from start to finish. Firstly, a quantitative design utilizing backtranslation as a framework for translating the SCS from English to BM. Additionally, justifying the usage of the BDI and SWLS as complementary scales to assess construct validity. Next, the Pilot Test 1 to produce the SCS-BM followed by utilizing backtranslation followed by reviewing items and a cognitive interview to improve translation quality. After that, Pilots Test 2 was to identify its internal consistency. Finally, a CFA is conducted to assess the psychometric properties of the SCS in terms of several key fit indices as well as convergent and divergent validity.

CHAPTER 4

FINDINGS

4. 1 Introduction

This chapter details the findings obtained from the various analysis conducted in this study. It will be divided into two sections, firstly, pilot test 1 (cognitive interviews and reliability), secondly, pilot test 2 (factor analysis and construct validity).

4.2 Pilot Test 1: Translation of SCS-English into BM

Firstly, the SCS-English underwent forward translation from English to BM. This forward translation was completed by a certified language translator. After this, the forward translation was then back translated from BM back into English. This backtranslation was completed by a professional counselling lecturer fluent in both English and BM. Refer to Appendix D for a full table of the translation for each item in the SCS-BM. With reference to Figure 3.2, this study conducted the first round of item review and resolved five items below which were ambiguous.

4.2.1 Item 2

Item 2 falls under the Over-identified subscale of the SCS. Below is Table 4.1 displaying the original SCS-English for Item 2 and its following translations and correction.

Table 4.1

Item 2 Translation and Correction

Original SCS-English	Backtranslation (English)	Forward Translation (BM)	Corrected (BM)
Item 2: When I'm feeling down I tend to obsess and fixate on everything that's wrong	When feeling depressed, I tend to be obsessive and correct all that does not feel right	Apabila berasa murung, saya cenderung untuk obsess dan membetulkan semua perkara yang tidak kena.	Apabila berasa sedih, saya cenderung untuk obses dan terlalu memikir semua yang tidak kena.

During the review, it was noticed that in Item 2, the backtranslation carried the meaning of correction in the latter part of the item, whereas the original item was referring to being judgmental. The BM translation displayed this with the phrase "*…membetulkan semua perkara yang tidak betul"*, which carries the meaning of desiring to correct everything that is wrong. With reference to Section 2.5, Over-identified items should carry the meaning of someone who is obsessed with their personal suffering. Therefore,

In order to resolve this, this study edited it to read *"…memikir semua yang tidak kena",* which carries the meaning of fixating on everything that is wrong; this is in alignment with what an over-identified item should mean.

4.2.2 Item 9 and Item 17

Item 9 and Item 17 falls under the Mindfulness subscale of the SCS. Below is a table displaying the original SCS-English for Item 9 and Item 17 and its following translations and correction.

Table 4.2

Item 9 & 17 Translation and Correction.

Original SCS-English	Backtranslation (English)	Forward Translation (BM)	Corrected (BM)
Item 9: When something upsets me I try to keep my emotions in balance	When some things let me down, I try to maintain my emotion.	Apabila sesuatu mengecewakan saya, saya cuba untuk mengekalkan emosi saya.	Apabila sesuatu mengecewakan saya, saya cuba untuk menyeimbangkan emosi saya.
Item 17: When I fail at something important to me I try to keep things in perspective	When I fail to achieve something which is important to me, I try to make the situation under control.	Apabila gagal mencapai sesuatu yang penting untuk diri, saya cuba agar keadaan menjadi terkawal.	Apabila gagal mencapai sesuatu yang penting untuk diri, saya cuba membuat perbandingan yang adil.

Firstly, for Item 9, it was noticed that the backtranslation and the original were not quite congruent. The original specified keeping emotions in balance whereas the backtranslation stated "maintain my emotions", which translated into BM as *"...mengekalkan emosi saya."* The word "maintain" on its own is ambiguous as it does not inform the reader whether they should maintain negative/emotions emotions or balance their negative and positive emotions. With reference to section 2.5, mindfulness is about taking a step back and keeping this in a balanced perspective. In order to resolve this, one word in the BM translation was changed *"...menyeimbangkan emosi saya";* the word *menyeimbangkan* carries the meaning of balance as opposed to *mengekalkan* which means to maintain.

Secondly, the backtranslation of Item 17 was also not congruent with the original. The original item referring to a characteristic of mindfulness which is being able to take a step back and keep thoughts in a balanced perspective (Refer to Section 2.5), however the backtranslation said "make the situation under control", which was translated into BM as *"...saya cuba agar keadaan menjadi terkawal"*. This phrasing is not consistent with the characteristic of mindfulness, which is more about mentally reframing a situation rather than physically taking action to keep things under control. Therefore, this study edited the BM translation into *"...saya cuba membuat perbandingan yang adil"*, which carries the meaning of making a fair comparison when facing failure. This carries the meaning of mindfulness more accurately.

4.2.3 Item 16

Item 16 falls under the Self-judgement subscale of the SCS. Below is a table displaying the original SCS-English for Item 16 and its following translations and correction.

Table 4.3

Item 16 Translation and Correction

Original SCS-English	Backtranslation (English)	Forward Translation (BM)	Corrected (BM)
Item 16: When I see aspects of myself that I don't like, I get down on myself	When I look parts of myself that I do not like, I am frustrated with myself.	Apabila saya melihat aspek kendiri yang tidak saya gemari, saya kecewa dengan diri sendiri.	Apabila saya melihat aspek kendiri yang tidak saya gemari, saya menyalahkan diri sendiri.

This study noted that the backtranslation of Item 16 did not carry the same meaning as the original. This item should express Self-judgement which refers to someone being harsh and critical on themselves (refer to Section 2.3). The backtranslation version said, "…I am frustrated with myself" which was translated from the BM *"...saya kecewa dengan diri sendiri",* however, being frustrated is not congruent with being harsh and critical. Item 16 is challenging to translate because the phrase "…get down on myself" is an expression that has no BM equivalent. According to the online Macmillan dictionary, the phrase "…to get down on myself" is synonymous with criticizing and blaming (Macmillan Dictionary, n.d.). Hence, the BM translation was edited to read as *"...saya menyalahkan diri sendiri"*, which carries the meaning of blaming oneself and more accurately translates the original item of "…I get down on myself."

4.2.4 Item 23

Item 23 falls under the Self-kindness subscale of the SCS. Below is a table displaying the original SCS-English for Item 23 and its following translations and correction.

Table 4.4

Item 23 Translation and Correction

Original SCS-English	Backtranslation (English)	Forward Translation (BM)	Corrected (BM)
Item 23: I'm tolerant of my own flaws and inadequacies	I can accept all my weaknesses and inadequacies.	Saya boleh menerima segala kelemahan dan kekurangan sendiri.	Saya bertolak ansur dengan segala kelemahan dan kekurangan sendiri.

The backtranslation for Item 23 is not congruent with the original item; the backtranslation uses the phrase "I can accept…" but the original one says "I am tolerant…". Item 23 should be expressing self-kindness, which is a characteristic of being compassionate towards oneself as opposed to being harsh and critical (refer to Section 2.3). Although the word "accept" (backtranslation) might appear as a form of kindness at first glance, however it carries a different meaning from the word "tolerant" (original). According to the Cambridge Dictionary (n.d.), the word tolerant is defined as: "willing to accept behaviour and beliefs that are different from your own, although you might not agree with or approve of them". The key point here is an individual who is expressing tolerance does not necessarily agree with something, whereas the word "accept" indicates a person is in agreement with something. Hence, in order to ensure Item 23 was closer to the original meaning this study edited the BM translation to read as, *"Saya bertolak ansur…"*, this more accurately carries the meaning of being tolerant.

4.2.5 Summary of Translation Findings

In summary, this study reviewed all the 26 translated SCS items by comparing the backtranslation to the original. Incongruencies in meaning were determine for Items 2, 9, 16 17, and 23. Items were edited to produce a translation that more

accurately captured the meaning of the subscale and the meaning of the respective original items. Upon confirming the SCS-BM, this study then continued with Pilot Test 1.

4.3 Cognitive Interview Analysis

A cognitive interview analysis was conducted to determine response validity for each of the 26 items in the SCS-BM. A total of six participants were recruited for this interview. Below is a list of two items which participants had difficulty understanding accurately.

4.3.1 Item 8

Table 4.5

Item 8 Translation and Correction

Item 8 in English	Item 8 in BM	Corrected Item 8 in BM
When times are really difficult, I tend to be tough on myself	Apabila keadaan menjadi benar-benar sukar, saya cenderung untuk berkeras terhadap diri sendiri.	Apabila keadaan menjadi benar-benar sukar, saya cenderung untuk lebih kritikal terhadap diri sendiri.

Item 8 falls under the factor of Self-judgment. During the cognitive interview analysis, it was noticed that participants had a problem understanding the BM word *berkeras* which is a word that was translated from the English phrase "tough on myself". Below are some qualitative data excerpts from participants in response to being asked what they understood about the meaning of Item 8.

> *"Keras terhadap diri sendiri ni perlu ada sebab kalau bukan kita sendiri yang push diri kita untuk capai sesuatu, orang lain mungkin tak boleh tolong."* (Participant 2)

“"Tak salahkan Tuhan, tak salahkan orang lain. Oh, saya akan perbaiki diri saya." (Participant 3)

“…maksudnya menjadikan macam seolah kita lebih, lebih lebih bersifat juanglah…. Maknanya, you are a fighter, in this case.” (Participant 4)

“Actually, saya kurang faham tapi ini…Maksudnya, bila keadaan tu benar-benar sukar, saya akan cuba paksa diri sendiri untuk terus jadi kuat…Untuk never give up lah.” (Participant 5)

Five of the participants communicated the same idea that *berkeras* means pushing themselves to do better. They felt that the item was referring to the spirit of never giving up, trying to improve themselves and becoming stronger. Unfortunately, these responses do not capture the idea of self-judgement, which is the focus of Item 8. Only Participant 6 expressed an accurate understanding for self-judgement.

"Berkeras", bagi saya ada dua makna. Satu, mungkin boleh lebih menghakimi ataupun lebih punish. (Participant 6)

Participant 6 was the only participant who understood that *berkeras* was in reference to self-judgement. Because of this, the current study felt it was prudent to change the phrase *berkeras* to *lebih kritikal* so that participants would understand the item more accurately. The challenge in translating Item 8 is because there is not direct BM translation for the English phrase, “tough on myself”. In order to capture the same meaning, this study opted for *lebih kritikal* in order to communicate that this item is about being critical about oneself.

4.3.2 Item 14

Table 4.6

Item 14 Translation and Correction

Item 14 in English	Item 14 in BM	Corrected Item 14 in BM
When something painful happens, I try to take a balanced view of the situation	Apabila berlaku sesuatu yang menyakitkan, saya cuba melihat situasi itu dari sudut yang berimbang.	Apabila berlaku sesuatu yang menyakitkan, saya cuba melihat situasi itu dari sudut yang seimbang.

Item 14 falls under the factor of Mindfulness. Participants understood what the item meant however two participants briefly expressed that the word *berimbang* may be incorrect. Their quoted transcripts are as follows.

> *"What does it mean by "berimbang. Apabila berlaku sesuatu yang menyakitkan, saya cuba melihat dari sudut yang seimbang? Pada saya, mungkin sudut yang seimbanglah. Bukan "berimbang? Ya. This is the first time "berimbang". I don't think there is anything such a word. It's like, "sudut yang seimbang", it's a balance angle, right?"* (Participant 3)

> *Ya, betul…kena seimbangkan"* (Participant 4)

Both participants expressed the word *seimbang* to be a more accurate term to use. An online translator was used and it was discovered that *berimbang* is an Indonesian word for "balanced" while *seimbang* is a BM word for "balanced". Hence, using both the data from the participants and the online translator verification, this study changed the word *berimbang* to *seimbang* to provide a more accurate translation for the word "balanced".

4.3.3 Summary

In short, the cognitive interviews revealed that participants understood almost all the items in the SCS-BM without issue. Their understanding was found to be congruent with how Neff conceptualized the different factors that make up the SCS. Only Item 8 and 14 required some changes due to translation ambiguity.

4.4 Pilot Test 2: Internal Consistency

A total of 61 participants were recruited. These participants were all University Malaya students over the age of 18 and fluent in BM. The internal consistency for the entire SCS-BM scale of 26 items was high (α = .89). Table 4.7 displays the individual Cronbach alpha scores for each subscale. Refer to Appendix E for the SPSS output of the entire SCS and each subscale's reliability.

Table 4.7

Internal Consistency of SCS-BM Subscales

Subscale	Internal Consistency (α)
1. Self-kindness	.81
2. Self-judgement	.59
3. Common Humanity	.83
4. Isolation	.65
5. Mindfulness	.79
6. Over-identified	.71

Cut off scores for Cronbach alpha are taken from Taber (2017) who lists that scores between 0.64 to 0.85 are moderate to reliable. Scores ranging from 0.85 to 0.93 are considered strong. Following this, the overall SCS-BM has a strong reliability with its score of .89. Only the Self-judgement scale (.59) needs to be looked at in more detail because its score falls below the acceptable cut-off point.

4.4.1 Self-judgement Subscale

The subscale with the lowest Cronbach alpha was self-judgement. This study looked at the total item correlation for the self-judgement scale (Refer to Appendix E) to determine which item was affecting its internal consistency.

Table 4.8

Item Total Correlation for Self-judgement Subscale

Item	Item Total Correlation	Cronbach's Alpha if Item is Deleted
Item 1	.192	.607
Item 8	.403	.503
Item 11	.420	.494
Item 16	.374	.519
Item 21	.342	.537

According to Table 4.8, Item 1 had the lowest item total correlation ($\alpha = .19$) in the self-judgement subscale. In addition to that, deleting Item 1 would increase the Cronbach alpha to $\alpha = .60$ from its current $\alpha = .59$. This study inspected Item 1 in more detail to determine if it should be deleted or if it could be corrected and retained. Below is Table 4.9 comparing Item 1's original English translation against the BM translation.

A note should be made regarding the controversy of using correlations in item analysis. Correlational analysis is primarily designed to analyze the strength of the relationship between variables (Field, 2009) and it is also often influenced by the number of items and sample size of the study. Because of that, caution should be exercised when using correlational analysis to assess item quality as having a large sample can give false positive results. The researcher acknowledges the drawbacks of using correlation analysis, however, in this study, correlational analysis (Table 4.8) serves only as a simple guide to identify which items might require some improvement, particularly in terms of translation. As will be seen in Table 4.9 below, the translation of said Item 1 has been improved for the SCS-BM.

Table 4.9

Item 1 Translation and Correction

Item 1 in English	Item 1 in BM	Corrected Item 1 in BM
I'm disapproving and judgmental about my own flaws and inadequacies	Saya tidak dapat menerima dan gemar menghakimi kelemahan dan kekurangan sendiri.	Saya selalu menghakimi kekurangan diri.

The most likely problem in the construction of Item 1 is its double-barreled nature. According to Lavrakas (2008), a double-barreled question is one that asks respondents about multiple constructs in a single item. In the case of Item 1, it can be seen in the English version that the question is asking about self-disapproval and judgmental attitude, as well as flaws and inadequacies. While these concepts are highly similar, Lavrakas (2008) and Hillerns and Lei (2015) all agree that the conjunction "and" is an indicator that the item is attempting to measure across multiple topics. This should be avoided so as not to confuse respondents. Hence this study has corrected Item 1 in Table 4.9 to become more direct and straightforward in BM, i.e. *Saya selalu menghakimi kekurangan diri.* If this were to be translated into English, it would be "I am judgmental towards my inadequacies". This change echoes advice from Hillerns and Lei (2015) who advocate for simplified and direct phrasing to be used when constructing items so that respondents are clear about what the item is asking of them.

4.5 Final Review of Items

With reference to Figure 3.2, this study then conducted the final review of items in the SCS-BM. During the final review, only one item was identified as requiring a more refined translation.

Table 4.10

Item 18 Translation and Correction

Item 18 in English	Item 18 in BM	Corrected Item 18 in BM
When I'm really struggling, I tend to feel like other people must be having an easier time of it	Apabila saya berjuang bersungguh-sungguh, saya cenderung merasakan yang situasi orang lain tidak sesukar saya.	Apabila saya bersusah payah, saya cenderung merasakan bahawa situasi orang lain tidak sesukar saya.

Item 18 falls under the factor of Isolation. The phrase "struggling" in the English translation was translated into *berjuang.* However, upon reviewing this, it was noticed that this BM translation carries a meaning more similar to that of fighting or battle (Glosbe Online Dictionary, n.d). Hence, it is incongruent with the meaning of the phrase "really struggling". In order to better capture the original item's meaning, this study edited the translation to *bersusah payah* which, according to the online Glosbe Online Dictionary (n.d.), carries the meaning of struggle.

4.6 Summary of Pilot Study 1

In summary, the translation process was successful. The internal consistency analysis revealed an item with low correlation that was subsequently corrected. The cognitive interviews revealed two ambiguous items that were subsequently corrected. The first and last review of the items revealed several items with translation issues which were all subsequently corrected as well. All 26 translated items were retained with only nine items edited. Table 4.11 below clearly lists all the different edited times and their respective reasons for being edited, namely: inaccurate translation, ambiguous meanings, or low correlation.

Table 4.11

Edited Items

Item	When Was the Item Identified	`Reason for Edit
Item 2	Review 1	Translation
Item 9	Review 1	Translation
Item 16	Review 1	Translation
Item 17	Review 1	Translation
Item 23	Review 1	Translation
Item 8	Cognitive Interview	Ambiguous
Item 14	Cognitive Interview	Ambiguous
Item 1	Internal Consistency	Low Correlation
Item 18	Review 2	Translation

4.7 Demographic Data

A confirmatory factor analysis was conducted using the AMOS software in order to test the 6-factor model of self-compassion. A final total of 302 participants' responses were used in this study. A total of 308 participants were initially recruited but six participants scores were removed for the following reasons: Four participants were not from Malaysia, one answered that they were not fluent in BM, and one was below the minimum age of 18 years. The minimum age was 18 years, while the maximum age was 53 years, and the mean age was 22 years. For gender, there were 72 males (24%) and 230 females (76%). For religion, there were 26 Buddhist (8%), 11 Hindus (4%), 54 Christians (18%), and 209 Muslims (69%), and two atheists (1%). For race there were 63 Chinese (21%), 15 Indians (5%), 198 Peninsula Malays (65%), 21 East Malays (7%) (Kadazan, Bidayuh, Iban, Suluk, Murut, Bajau, Bajar, and Bidayuh), two Eurasians (1%) and two who chose not to specify their race (1%). Refer to Appendix Q for a pie chart breakdown of all the demographic data.

4.8 Normality

This study assessed the normal distribution of the SCS-BM scores by utilizing several measures of normality. Firstly, for a visual representation of normality, a histogram

distribution of the SCS-BM was produced. Next, for a numerical representation, this study will be relying on the Kolmogorov-Smirnov statistic, skewness, and kurtosis values. Total sample size is N =302. Refer to Appendix F for full SPSS output of normality.

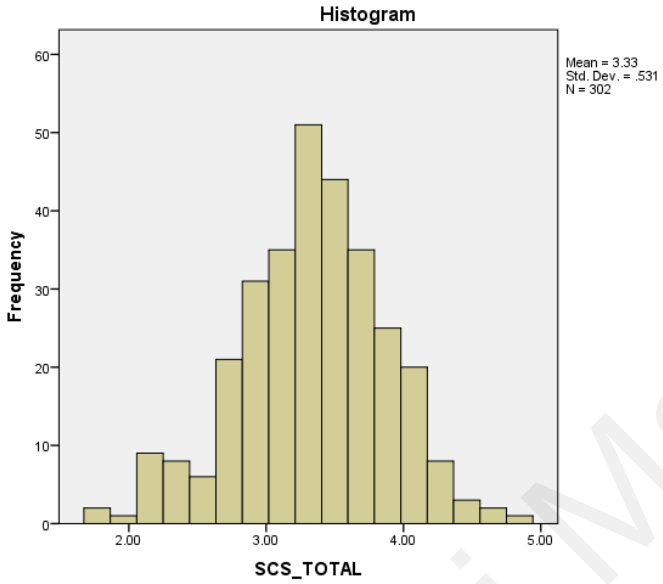


Figure 4.1. Histogram Distribution of SCS-BM

Table 4.12

Kolmogorov- Smirnov Statistic

	Kolmogorov-Smirnov		
	Statistic	df	Sig.
SCS-BM	.052	302	.045

Table 4.13

Skewness and Kurtosis

	Skewness	Kurtosis
SCS-BM	-.206	.148

Referring to Figure 4.1, the histogram distribution provides visual evidence for a fairly normal distribution of the SCS-BM scores. Next, with reference to Table 4.12, the Kolmogorov-Smirnov (K-S) test statistic revealed the sample was not normally

distributed (p = 0.045) However, this was just slightly below the cutoff point of 0.05 for normality. The K-S test was chosen over the Shapiro-Wilk test because according to Elliot and Woodward (2007) the Shapiro-Wilk test is more suitable for sample sizes of 50 and below, hence, due to the large sample size in this study (N = 302), the K-S test is more suitable. Lastly, both skewness and kurtosis were expected to be within -1/+1 for acceptable normality (Field, 2009). With reference to Table 4.13, Skewness was -.206, which is a slight negative skew, indicating a slightly higher number of scores in the right side of the distribution. In addition to that, Kurtosis was .148 which indicates a slight positive kurtosis, meaning that the normal distribution has a slightly higher peak. However, both skewness and kurtosis were within the cutoff range for a normal distribution.

In conclusion, although the K-S test was just slightly below the normal range cutoff, the other indicators, namely the histogram, skewness, and kurtosis all indicated that the SCS-BM scores are normally distributed. Therefore, this study proceeded to accept that normality was assumed for the distribution of the SCS-BM scores.

4.9 Internal Consistency

In section 4.4, internal consistency was run for the pilot study with only 61 participants. In this section 4.10, internal consistency was conducted with the full 302 sample of participants. The internal consistency for the entire SCS-BM scale of 26 items was high ($\alpha = .89$). Table 4.14 displays the overall Cronbach alpha scores for the entire SCS-BM and Table 4.15 displays the Cronbach's alpha for each subscale. Refer to Appendix I for the SPSS output of the entire SCS and each subscale's internal consistency

Table 4.14

Internal Consistency of overall SCS-BM

Cronbach's Alpha for SCS-BM (α)	N
.897	26

Table 4.15

Internal Consistency SCS-BM Subscales

Subscale	Internal Consistency (α)
1. Self-kindness	.785
2. Self-judgement	.730
3. Common Humanity	.775
4. Isolation	.723
5. Mindfulness	.720
6. Over-identified	.769

Cut off scores for Cronbach alpha are taken from Taber (2017) who lists that scores between 0.64 to 0.85 are moderate to reliable. Scores ranging from 0.85 to 0.93 are considered strong. With reference to Table 4.14, the overall SCS-BM has a strong internal consistency with its score of α = .89. Furthermore, with reference to Table 4.15, each subscale of the SCS has moderate internal consistency, ranging from α = .72 to α =.78.

4.10 Confirmatory Factor Analysis

The 6-factor model proposed by Neff was tested using the AMOS software. The 6-facor model is shown in Figure 4.2 below. Each factor is first placed inside an oval shape, next, connected to each factor are the various items which correspond to their respective factors; these are put into rectangles. Finally, the error terms are created.

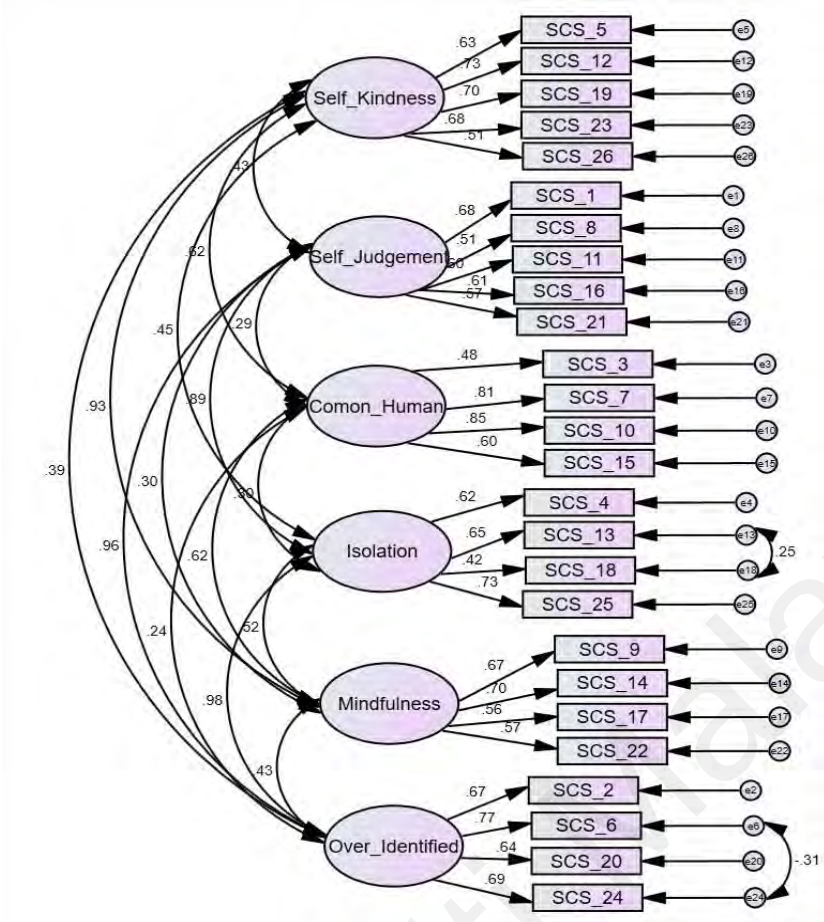


Figure 4.2. The 6-Factor CFA Model

4.10.1 Fit Indices

The Chi-square value indicated $\chi 2$ (282, N=302) = 1.95 p<0.001. Although the χ^2 was within the cut off range, the p-values was significant which indicates a lack of fit. The Comparative Fit Index indicated model fit with CFI = 908 which is above the minimum cut off range of .90. The Root Mean Square Error of Approximation indicated acceptable fit with RMSEA = .056 which was within the cut off score of < .08. Lastly, the Standardized Root Mean Square Residual indicated acceptable fit with SRMR = .061 which was also within the cut off score of < .08. Table 4.14 below summarizes the results of the CFA. Refer to Appendix G for full AMOS output for each respective fit index.

Table 4.16

6-Factor CFA Fit Indices

Fit Indices	Score	Fit Statistic
Chi-Square (χ^2)	1.95, p<0.001	Lacks fit
CFI	.908	Acceptable fit
RMSEA	.056	Acceptable fit
SRMR	.061	Acceptable fit

4.10.2 Regression Weights

With reference to Figure 4.3, it was observed that all items loaded significant onto their constructs. This is shown in CFA via AMOS with the triple-asterik (***) symbol. With reference to Table 2.1, take the example of Self-kindness which should be accounted for in Neff's (2003) SCS by Items #5, #12, #19, #23, and #26. In Figure 4.3 it can be seen that all those corresponding items have loaded significantly onto the construct of Self-kindness. This indicates that Items #5, #12, #19, #23, and #26 significantly measure the construct of Self-kindness.

		Estimate	S.E.	C.R.	P	Label
SCS_5	<--- Self_Kindness	1.000				
SCS_12	<--- Self_Kindness	1.235	.120	10.285	***	
SCS_19	<--- Self_Kindness	1.153	.116	9.912	***	
SCS_23	<--- Self_Kindness	1.121	.115	9.778	***	
SCS_26	<--- Self_Kindness	.715	.094	7.623	***	
SCS_1	<--- Self_Judgement	1.000				
SCS_8	<--- Self_Judgement	.804	.100	8.061	***	
SCS_11	<--- Self_Judgement	.899	.095	9.481	***	
SCS_16	<--- Self_Judgement	1.029	.107	9.655	***	
SCS_21	<--- Self_Judgement	.991	.111	8.931	***	
SCS_3	<--- Comon_Human	1.000				
SCS_7	<--- Comon_Human	2.027	.255	7.949	***	
SCS_10	<--- Comon_Human	2.166	.270	8.027	***	
SCS_15	<--- Comon_Human	1.316	.188	6.987	***	
SCS_4	<--- Isolation	1.000				
SCS_13	<--- Isolation	1.159	.123	9.420	***	
SCS_18	<--- Isolation	.674	.104	6.462	***	
SCS_25	<--- Isolation	1.309	.127	10.319	***	
SCS_9	<--- Mindfulness	1.000				
SCS_14	<--- Mindfulness	.996	.094	10.577	***	
SCS_17	<--- Mindfulness	.836	.097	8.615	***	
SCS_22	<--- Mindfulness	.893	.101	8.815	***	
SCS_2	<--- Over_Identified	1.000				
SCS_6	<--- Over_Identified	1.113	.095	11.734	***	
SCS_20	<--- Over_Identified	1.043	.104	10.041	***	
SCS_24	<--- Over_Identified	1.076	.101	10.611	***	

Figure 4.3. Regression weights for all factors and items

4.10.3 Correlated Errors

After running the initial CFA, it was noticed there was some correlated error among error terms within the same item. Figure 6 below shows a brief summary of some of the covariances between errors and items, the full list can be found in Appendix H.

			M.I.	Par Change
e24	<-->	Isolation	5.320	.056
e6	<-->	Self_Judgement	4.135	.040
e6	<-->	e24	12.962	-.139
e22	<-->	Over_Identified	10.814	-.065
e22	<-->	Comon_Human	5.673	-.039
e22	<-->	Self_Kindness	6.924	.048
e22	<-->	e24	4.656	-.084
e22	<-->	e6	5.028	-.073
e9	<-->	Over_Identified	11.565	.059
e9	<-->	Self_Judgement	5.054	-.040
e9	<-->	e2	9.530	.099
e9	<-->	e22	8.414	-.083
e25	<-->	Mindfulness	4.250	-.044
e25	<-->	e24	13.060	.159
e25	<-->	e2	6.446	-.105
e18	<-->	e14	4.255	.069
e13	<-->	Over_Identified	4.497	.049
e13	<-->	Self_Judgement	4.954	-.054
e13	<-->	e18	14.727	.187

Figure 4.4. Covariances Between Factors and Error Terms

From Figure 6, it was noted that error terms e6 <-->e24 and e13 <-->e18 both had the highest error covariance among all the error terms, 12.96 and 14.72 respectively. Both e6 and e24 are error terms under the same item, namely Over-identified, and e13 and e18 are under the same item as well, namely, Isolation. Therefore, this study proceeded to correlate both pairs of error terms as they were within their own respective items and this would thus improve model fit.

4.10.4 Drawbacks of CFA

Although CFA has been the primary statistical analysis in past translation research on the SCS, it has certain drawbacks. The CFA is based on Classical Test Theory (CTT) which, according to Bell (1982), focuses more on the item than the person. In the case of CFA, the analysis focuses on determining only one thing, that is, the factor loadings on each item with regards to the specified model. Therefore, it is limited because it cannot assess the person-fit data in the sample obtained. Considering the possible cultural differences in the conceptualization of the SCS in a Malaysian context, the CFA is unable to account for whether or not Malaysian students are truly able to understand five-point Likert rating scale for the SCS-BM. It is assumed based on past research; however, it has not been tested by this current research.

In addition to that, according to Brown (2006) CFA is influenced by sample size. Brown (2006) provides an example on how the chi-square fit index is sensitive to large samples and may produce a false negative, as is the case in the current research where the chi-square result indicated a lack of fit. Therefore, this sensitivity to sample size is affecting the validity of the model as it may not be accurately showing the fit index with regards to participants scores. This drawback of the CFA's sensitivity to sample size indicates that alternative measurement models should be considered when analyzing the validity of scales.

The current researcher acknowledges the drawbacks of using a CFA with regards to the inability to assess person-fit statistics and sensitivity to sample size. However, because a majority of past research utilized the CFA in their analysis of the translated SCS, the results of this study can therefore be easily compared to the past literature.

4.11 Test-retest Reliability

The SCS-BM was administered a second time via email after a time interval of three weeks. The test-retest reliability was found to be r = .79, N = 67, p<0.001 for the overall SCS-BM. Refer to Appendix K for the SPSS output for test-retest reliability. This indicates a positive correlation of medium strength. Due to attrition only 67 participants volunteered to participate in the test-retest reliability component via email instead of the expected 100 participants.

Table 4.17

Test-retest Reliability

		SCS-BM Re-test
SCS-BM Test	Pearson's Correlation (*r*)	.79
	Sig.	.000
	N	67

4.11.1 Test-Retest Reliability of Sub-scales

In addition to the aforementioned test-retest reliability of the overall SCS-BM after a 3-week interval, the reliability scores for each sub-scale have also been listed accordingly. Refer to Appendix L for full SPSS output on the sub-scales test-retest reliability.

Table 4.18

Test-Retest Reliability of SCS-BM Sub-scales

No	Sub-scale	*r*
1	Self-kindness	.67
2	Self-judgement	.66
3	Common Humanity	.48
4	Isolation	.75
5	Mindfulness	.69
6	Over-identified	.74

Each sub-scale had a positive moderate to strong test-retest reliability correlation. This indicates that the sub-scales of the SCS-BM are able to consistently

measure their respective factors over a time period of 3 weeks. The only sub-scale that had a low reliability value was common humanity, $r(67) = .48$, $p < .001$.

4.12 Construct Validity

In order to test for construct validity, the SCS-BM was correlated with two other relevant scales. With reference to Table 4.19, the SCS-BM had a moderate negative correlation with the BDI-BM, $r(302) = -.62$, $p < .001$. Secondly, the SCS-BM had a weak positive correlation with the SWLS-BM, $r(302) = .49$, $p < .001$. Refer to Appendix J for full SPSS output.

Table 4.19

Correlations between SCS, BDI, and SWLS

		BDI	SWLS
SCS	Pearson Correlation	-.55	.49
	Sig. (1-tailed)	.000	.000
	N	302	302

4.13 Summary

In summary, the translation, review, and cognitive interview processes, resulted in the resolving of eight items which were found to either have some poor translation or ambiguous meanings after translation. The SCS-BM displayed strong internal consistency for the overall scale. Additionally, the CFA analysis showed an acceptable model fit according to three of the four fit indices used. The findings indicate that the SCS-BM is valid and reliable.

CHAPTER 5

DISCUSSION AND CONCLUSION

5.1 Introduction

The results obtained from this study will be discussed in Chapter 5. Firstly, the results from chapter four will be briefly summarized. After that, the results will be interpreted and discussed in light of past literature with regards to answering the research questions outlined in this study (Chapter 1.7). Implications with regards to theory and practice will be provided and the author will also suggest future directions for this research. The chapter will conclude with an overall summary for this current research paper.

5.2 Summary of Findings

Firstly, the SCS-BM displayed strong overall internal consistency (α = .89). The Self-judgement subscale was identified as being low in internal consistency, and upon further investigation, Item 1 in the self-judgement scale was edited because it was double-barrelled.

Next, CFA analysis showed an acceptable model fit. Of the four fit indices used, only one did not indicate fit, which was the chi-square goodness of fit test (CMIN/DF = 1.95). On the other hand, three of the fit indices indicated acceptable fit, namely the CFI = .908, the RMSEA = .056), and the SRMR = .061. Finally, the SCS-BM had fair construct validity between the SCS-BM and BDI-BM (r = -.55), and the SCS-BM and SWLS-BM (r = .49).

5.3 Research Question 1: What is the internal consistency of the SCS-BM?

This question was successfully answered by the current study. The overall internal consistency of the SCS-BM was found to be strong (α = .89). According to cut off scores by Taber (2017) this score is considered to represent strong internal consistency.

The measure of internal consistency was included in this study for two-fold reasons. Firstly, internal consistency is a key feature in translation studies (Maneesriwongul & Dixon, 2004). Secondly, internal consistency has been featured in a majority of the past journals documenting the translation of the SCS into various languages. The current study has found results that support research by Arimitsu (2014), Castilho and Pinto-Gouveia (2011), Chen et al. (2011), Costa et al. (2015), Dundas et al. (2016), Lee & Lee (2010), and Khatib et al. (2018). All these studies obtained an internal consistency of at least .80 and above. All past studies reviewed by this study were able to obtain good internal consistency for their respective translations of the SCS into the local language. Hence, it can safely assumed that the backtranslation process was able to produce a Malay version of SCS that is internally consistent.

5.3.1 Drawbacks of Internal Consistency

According to Cohen et al. (2007), Cronbach's alpha is limited to only measuring the internal consistency between items of a given scale. In other words, it does not assess the unidimensionality of the scale. In other words, although the current research can suggest that the SCS-BM has items which are internally consistent, it cannot be assumed that those items are accurately measuring self-compassion. In addition to that, the Cronbach's alpha calculation included the number of items on the scale being tested (Cohen et al., 2007) and therefore, if the number of items increase, so does internal consistency. This suggests that having a lot of items may produce a strong internal consistency, but it does not guarantee that the items are accurately measuring the concept of self-compassion.

Therefore, the current researcher acknowledges the drawbacks of the measure of internal consistency but also weighs it against the benefits of using it. On one hand,

it is a measure that can be easily influenced by number of items and it lacks the ability of determining if items are measuring a unidimensional construct. On the other hand, it is a measure that has been an integral part of the past studies on the SCS and thus serves as a good measure to compare with previous research.

5.4 Research Question 2: What is the test-retest reliability of the SCS-BM?

This question was successfully answered by the current research. The test-retest reliability was found to be $r = .79$, N = 67, $p < .001$ for the SCS-BM after a 3-week interval. Refer to Appendix K for the SPSS output for test-retest reliability. This finding indicates a positive correlation of medium strength. Some caution should be exercised when interpreting these results as only 67 participants volunteered to participate instead of the expected 100 participants.

Determining test-retest reliability was to ensure that the SCS-BM construct is stable over time. The current study has obtained results which support past studies by Mantzios, Wilson and Giannou (2013), Benda and Reichová (2016), Deniz et al. (2008), and Castilho and Pinto-Gouveia (2011). The longest time frame (5 weeks) for the assessing test-retest reliability was by Mantzios et al. (2013) and the shortest time frame was three weeks, by Benda and Reichová (2016) and Deniz et al. (2008); the current study was also followed the three-week interval. The test-retest reliability results ranged from $r = .78$ to $r = .90$, and the current study's results falls within this at $r = .79$. The test-retest reliability does not seem to be affected by differences in sample size. Mantzios et al. (2013) recruited 244 participants while Benda and Reichová (2016) only recruited 58 participants, the former obtained a score of $r = .89$ and the latter obtained a score of $r = .90$.

5.5 Research Question 3: What is the factor structure of the SCS-BM?

This question was successfully answered by the current study. The original 6-factor structure of the SCS as proposed by Neff (2003) was found to have an acceptable fit.

5.5.1 RMSEA

Firstly, the RMSEA which had a value of .056 for the six-factor model of self-compassion. According to Hooper et al. (2008) this value of 0.56 is within the cut off value (RMSEA < 0.08) suggested for acceptable model fit. The RMSEA is often used in translation studies involving the SCS which utilize CFA, because, according to Hooper et al. (2008), it is popular measure which favours parsimony by determining the fit for the simplest model possible. Many other translation studies obtained results that supported the six-factor model, namely, Arimitsu (2014) with an RMSEA = .066, Lee and Lee (2010) with an RMSEA = .060, Garcia-Campayo et al. (2014) with an RMSEA = .060, de Souza and Hutz (2016) with an RMSEA = .071, Benda and Reichová (2016) with an RMSEA = .078, and Kotsou and Leys (2016) with an RMSEA = .050. Regardless of what language the SCS-English was translated into, these aforementioned researchers were able to obtain acceptable RMSEA fit for their CFA models. Therefore, it can be concluded that the RMSEA fit index provides evidence that the six-factor model of SCS sufficiently captures the underlying latent construct of self-compassion across a variety of different translations.

5.5.2 SRMR

Secondly, the SRMR which had a value of .061. An SRMR = 0 indicated perfect fit, while an acceptable cut off score provided by Hooper et al. (2008) suggests that any value < 0.08 is good fit. In addition to that, Chen (2007) states that the SRMR is a suitable fit index because it operates relatively independent of the sample size. Therefore, based on the results obtained, this study concluded that the SRMR provides

evidence that the six-factor SCS model sufficiently accounts for the underlying construct of self-compassion. This finding also supports past researchers who obtained a good SRMR score for the six-factor model, namely, Benda and Reichová (2016) with an SRMR = .066, Garcia-Campayo et al. (2014) with an SRMR = .05, Kotsou and Leys (2016) with an SRMR = .041, and Petrocchi, Ottaviani and Couyoumdjian (2013) with an SRMR = .07. These aforementioned translation studies validated the six-factor model of the SCS using the SRMR across a variety of languages (Czech, Spanish, French, and Italian, respectively). The findings from this current research support past literature. Therefore, it can be concluded that the SRMR fit index provides evidence that the six-factor model of SCS sufficiently captures the underlying latent construct of self-compassion across a variety of different translations.

5.5.3 CFI

The CFI indicated acceptable with a value of .908. According to Hooper et al. (2008), the CFI compares the goodness of fit between two models: the independent or null model (which suggests no correlations between constructs) against the target or hypothesized model of the six-factor SCS, with a cut off value of CFI > .90 for good fit. Therefore, in this current study, the CFI provides evidence that the six-factor SCS model is accurately accounting for the underlying construct of self-compassion. The current research obtained results similar to other past researcher who utilized the CFI, such as, de Souza and Hutz (2016) with a CFI = .93, Garcia-Campayo et al. (2014) with a CFI = .95, Kotsou and Leys (2016) with a CFI - .94, Petrocchi et al. (2013) with a CFI = .90, and Azizi et al. (2013) with a CFI = .90.

Therefore, the current study adds support to the past literature indicating that the CFI index provides evidence that the six-factor model of the SCS is accurately

measuring the underlying construct of self-compassion across a variety of different language translations.

5.5.4 Chi-Square

Finally, the chi-square goodness of fit test measured by $\chi 2$ indicated lack of fit with a value of 1.64, $p < .001$. According to Hooper et al. (2008), chi-square goodness of fit compares the sample covariance matrix with the model-fitted matrix with the appropriate cut off point for good fit of the chi-square test is $\chi 2 < 2$ and $p > .05$. Due to the lack of fit, this indicates that the theoretical model and the data generated model are different. This result is similar to three other past studies, namely, de Souza and Hutz (2016), Kotsou and Leys (2016 and Petrocchi et al. (2013) who similarly obtained a $p > .001$.

The researcher noted that only a few past studies utilized the chi-square goodness of fit as an index to gauge the ability of the six-factor model to accurately account for self-compassion. These aforementioned studies were unable to obtain fit via the chi-square test but on the other hand, they were able to obtain model fit for other fit indices. This suggests that the chi-square test may not be an accurate index to use when running a CFA on a translated version of the SCS.

In addition to that McIntosh, (2006), as cited in Hooper et al. (2008) states that the chi-square test can be easily rejected if there are any deviations form normality within the sample. With reference to Section 4.9, normality was assumed for the SCS-BM scores, however, the K-S test did indicate a slightly non-normal distribution which may have led to issues with the chi-square test.

5.5.4.1 Sample Size Issues

Although the chi-square is one of the earliest methods used for determining goodness of fit, Petrocchi et al. (2013) claims that the $\chi 2$ value can be

easily influenced due to large sample sizes. Bergh (2015) further supports this by stating that the chi-square test is sensitive to both very large or very small samples. Furthermore, other researchers have added that the chi-square test may falsely reject an adequate model when dealing with a large sample size (Gatignon 2010; Singh et al., 2016, as cited in Kyriazos, 2018). This is a tricky issue to resolve because on one hand, fairly large sample sizes are required for running CFA, but at the same time, a large sample size can affect the chi-square goodness of fit test. This is why Kyriazos (2018) says that although the chi-square is reported in alignment with current reporting practice, he advises researchers to rely on other complementary fit measures in order to comprehensively assess model fit. Thus, this study will be relying more on the other fit indices in this current study to better measure model fit due to the limitations of the chi-square test when dealing with a large sample size.

5.6 Research Question 4: What is the construct validity of the SCS-BM?

This question was successfully answered by the current study. The SCS-BM was found to have moderate negative correlation with the BDI-BM (divergent validity) and a weak positive correlation with the SWLS-BM (convergent validity).

5.6.1 Divergent Validity

With regards to the BDI, the current study obtained a divergent validity of r= -.62, $p < .001$ which is a moderate negative correlation. Other researchers who utilized the BDI to measure divergent validity were Lee and Lee (2010) with r = -.50, Garcia-Campayo et al. (2014) with r = -.43, and Benda and Reichová (2016) with r = -.65. The current study adds to the past literature with evidence that the SCS displays divergent validity when correlated with an opposing construct in the form of the BDI (i.e. depression). In other words, if a person scores high on self-compassion, they would most likely score low on depression, and vice versa.

This finding on the divergent validity of the SCS-BM and the BDI-BM can be explained using the Social Mentality Theory (refer to Chapter 2.2) in the context of the Malaysian student sample (refer to Chapter 1.2). Firstly, according to Mey and Yin (2014), the status of mental health among Malaysian students is poor. Academic pressure and financial stress are among some of the primary contributors to poor mental health (Gani, 2016 as cited in Kotera & Ting, 2019). In spite of the rise in academic growth, the implementation to boost mental health has not achieved much success, and Chong et al. (2013 as cited in Kotera and Ting (2019) attributes this reason to the negative social stigma surrounding mental health among students which in turn causes them to avoid seeking for helpful treatments. Kotera and Ting (2019) elaborate that using a direct approach is not beneficial as students end up feeling more ashamed of themselves about engaging in such interventions. Heine and Hamamura (2007) provided some insight into this behavior because they noted that individuals from Asian cultures display higher levels of self-criticism due to the constant pressure of performing well. In addition to that, Taylor et al. (2011) as cited in Gilbert (2014) emphasized that when individuals are over-reliant on achievement, they become vulnerable to depression. Furthermore, self-criticism is also part of the threat-defense system in the Social Mentality Theory. Gilbert (1989) suggests that depression can arise from individuals who feel a sense of uselessness (e.g. students who are unable to perform well), first towards others, and if it is not resolved, it becomes internalized via the threat defense system in the SMT and further reinforces self-criticism, insecurity, shame, self-anger. Gilbert (2014) advocates that the use of compassion is effective in silencing the voice of the inner critic and unlocking the potential to view one's self in a more positive way, thus preventing depression from coming into fruition. He quotes one patient saying, "I realized I could become the source of love and care for myself

that I did not have as a child' (Gilbert, 2014, p. 32). This is where self-compassion comes into play, with reference to Figure 2.1, self-compassion can help to deactivate the threat-defense system and activate the self-soothing system. This allows a person to be more accepting of their personal failures and to view themselves with less self-judgement, and more self-kindness. Gilbert (2014) describes this as a process in which compassion re-works the "threat-memory" from within and thus an individual is able to view themselves more positively. Therefore, self-compassion within the framework of the SMT can explain why the SCS-BM was significantly negatively correlated with the BDI-BM because Malaysian students who have high levels of self-compassion are thus better equipped at deactivating the threat-defense system and thus reducing the likelihood of experiencing depression.

5.6.2 Convergent Validity

With regards to the SWSLS, the current study obtained a convergent validity of $r = .49$, $p < .001$. This means that a participant who scores high in self-compassion will most likely score high in their satisfaction with life. Three past studies obtained similar results for convergent validity using the SWLS, namely Benda and Reichová (2016) with $r = .52$, Kotsou and Leys (2016) with $r = .50$, and Deniz et al. (2008) with $r = .45$. Therefore, the current study adds to past research on the divergent validity of the SCS with regards to the SWLS by providing evidence that individuals who score higher on self-compassion will subsequently score higher satisfaction with life scores.

This finding on the divergent validity of the SCS-BM and the BDI-BM can be explained using the Social Mentality Theory (refer to Chapter 2.2). With reference to Figure 2.1, it is noted that not only does self-compassion deactivate the threat-defense system, but it simultaneously activates the self-soothing system which results in individuals experiencing feelings of security, self-acceptance and self-kindness.

Gilbert (2014) explains that the self-soothing system taps into the release of oxytocin which, traditionally, plays a role in caregiving, but is also an important factor in care-receiving behaviours. Oxytocin also helps people feel safe and increases the feeling of well-being (Carter, 1998, as cited in Gilbert, 2014). Gilbert conceptualizes that SMT is not only activated in the presence of external stimuli but can also be understood with regards to internal stimuli (Gilbert, 2014). Therefore, care-receiving is not only limited to receiving care from an external person, but it also includes receiving care from one's own self, i.e. self-compassion. Gilbert (2014) elaborates that clinical psychology has largely ignores the activation of self-soothing emotions and behaviours because this concept has always been thought to occurs in the absence of a threat. In other words, Gilbert (2014) believes that most researchers focus on reducing threats in belief that this will increase self-soothing however, without realizing that self-soothing can exist as a separate system that requires attention. Depue and Morrone-Strupinsky (2005) as cited in Gilbert (2014) lend credence to this theory by showing that calm, rest, and contentment is not merely linked to the absence of threat, but rather, to a specific and separate positive affect system (i.e. the self-soothing system). Mindfulness has been suggested as the method in which people are able to access this system of well-being by creating a sense of non-striving accepting (colloquially we might call this a sense of "being in the moment"). Compassion is therefore, not simply about soothing the mind from threats and negative emotions, but rather,

> "Compassion involves developing the courage to be open to our anger and rage, not some kind of 'soothing it away' Indeed, to say it again, soothing is useful to act like a safe haven but also in preparation to courageously engage with what we need to." (Gilbert, 2014, p. 33)

Self-compassion should thus be thought of as an active force that promotes well-being. This is in alignment with the conceptualization of self-compassion by Neff (2003). Indeed, many researchers have found supporting evidence that links self-

compassion and life satisfaction together, such as durmaz and Aydin (2017), who state that self-compassion is positively linked to happiness, positive moods, healthy behaviors and life satisfaction. Moreover, Neff et al. (2007) elaborated that individuals who related to themselves with kindness, especially when facing challenging times, were more likely to be associated with greater happiness and life satisfaction. In addition to that, Kotsou and Leys (2016) found a significant interaction effect between self-compassion and positive affect. In fact, Kotsou and Leys (2016) actually performed a moderation analysis between depression, self-compassion and positive affect and discovered that a higher level of self-compassion was still associated with lower levels of depression, even when removing the effect of positive affect. Therefore, self-compassion within the framework of the SMT can explain why the SCS-BM was significantly positively correlated with the SWLS-BM because Malaysian students who have high levels of self-compassion are thus better equipped at activating the self-soothing system and thus increasing the likelihood of experiencing satisfaction with life.

5.7 Cultural Adaptation

According to Behr (2016), the backtranslation process does not guarantee accurate cultural adaptation. This was evidenced in the study by Zeng et al. (2016) which tested a Mandarin translation in a Buddhist sample but failed to support the factor structure of the SCS. The SCS-Mandarin had already been previously validated, but it is possible that it was not culturally adapted for the sample that Zeng et al. (2016) intended to test on. Therefore, in addition to the backtranslation process, the researcher had to address issues of cultural adaptation via careful reviews of items and feedback from the cognitive interviews which helped to refine the translation process into a more localized translation of the items in the SCS-BM. These issues have been highlighted

and discussed in the following four subsections for future researchers to take note when adapting the SCS into a local language. Also, while the CFA analysis can identify low-quality items it cannot explain why those items have problems. On the other hand, the current study has in-depth qualitative information from cognitive interviews to help provide insight into the cultural adaptation issues faced during this translation of items for the SCS-BM.

5.7.1 Morphological Differences between English and BM

During the reviewing process of the translated items, this study noticed there were a lot of issues in translation of the negative items in the SCS to BM. With reference to Table 4.11, out of 13 negative SCS items a total of five negative items required editing, namely Items 1, 2, 8, 16 and 18. These items were corrected either during the review or cognitive interview and internal consistency processes. One of the reasons contributing towards this may perhaps be the morphological differences between English and BM. According to Mat Awal, Abu Bakar, Abdul Hamid and Jalaluddin (2007), there are differences between the affixes in English and BM, e.g. prefix and suffix. In English, affixes can transform positive meanings into negative ones, for example, the positive word "obedient" with the prefix "dis-" becomes it's negative counterpart word "disobedient" (Mat Awal et al., 2007). However, Mat Awal et al. (2007) clarified that this function of affixes does not exist in the Malay language in the same way as English. Look at Table 5.1 below for a comparison.

Table 5.1

Comparison between Morphological Features of Item 11

SCS-English	SCS-BM
I'm intolerant and impatient towards those aspects of my personality I don't like	Saya tidak toleran dan tidak sabar terhadap aspek-aspek personaliti sendiri yang saya tidak gemari.

In the English version, the two negative words are “intolerant” and “impatient” both of which utilize the suffix “in-” to make the word negative. However, because the BM language does not have a direct translation for that, the translation comes out as *tidak toleran* and *tidak sabar*. The word *toleran* means tolerance and *sabar* means patience, but in order to transform these two positive words into negatives the translator was forced to use an additional word *tidak* (which means no or not in this context). In other words, the translation for “impatient” could be literally understood as “not patient”. While the understanding of both phrases may be similar, it highlights the challenge of accurately translating English phrases into BM due to the morphological differences such as affixes.

5.7.2 Malay Culture

According to Goddard (2002), Malay cultural ideals are primarily concerned with behaviours that are gentle, kind and considerate of other people’s feelings. This idea of being polite was reflected in the previous point of affixes because instead of having a direct translation for the negative word “impatient”, the Malay language is forced to employ a two-word “no + patient” (*tidak sabar*) format in order to achieve a negative phrase. Goddard (2002) added that the cost of not living up to those aforementioned ideals of gentleness or good behaviour is shame and embarrassment; in BM this feeling is called *malu*. Hence, when Malay participants encounter negatively phrased statements which force them to confront concepts which may not be culturally acceptable, they may perhaps have difficulties in processing or agreeing with those statements. Fortunately, the current study conducted cognitive interviews with six Malay participants. The rich qualitative data might hold some evidence of this cultural issue.

5.7.3 Cognitive Interviews

This study re-looked at the qualitative data collected and focused more specifically on the reasoning participants gave when they disagreed with negative items. Below are several excerpts from conversations between the interviewer and participants who expressed reasoning when disagreeing with negative items in the SCS-BM.

> Interviewer: *Nombor 9. Apabila sesuatu mengecewakan saya, saya cuba menyeimbangkan emosi saya. Apakah yang kamu faham daripada frasa "menyeimbangkan emosi"?*
>
> P1: *Apa yang saya faham is that kita tak boleh nak… Like, merasa sangat kecewa tu kita kena… Because kita banyak orang yang kisah about kita. Banyak orang yang support kita. So, if kita tunjuk sangat kat dia orang yang kita tengah frust, kecewa, nanti kita let them down.* (Participant 1)

In the above excerpt by Participant 1, the Interviewer had asked for their opinion on Item 9 (“When something upsets me I try to keep my emotions in balance”). However, Participant 1 provided a reason saying that people should not show their negative emotions to others because others are often supporting us and if we show them how frustrated or disappointed we are about something, we would be letting them down. This is similar to what Goddard (2002) suggested, which was that Malay culture resolves around being considerate about others and not just one’s self. The SCS however, is about one’s own self-compassion. Hence, there may be an issue in how participants processed negative items because they may be thinking about them in terms of how others would view them rather than how they view and understand themselves.

Interviewer: *Apabila saya gagal mencapai sesuatu yang penting buat saya, saya akan dipengaruhi rasa serba kekurangan.*

P3: *Tak setuju juga. Sebab gagal dan berjaya adalah lumrah kehidupan. So, kalau kita gagal disebabkan kekurangan kita, kita cuba improve diri kita. Perbaiki mana yang kita lemah…Maksudnya kalau Tuhan tak bagi kita part ni, mungkin Tuhan bagi kita part lain.* (Participant 3)

In this excerpt by Participant 3, the Interviewer had mentioned Item 6 ("When I fail at something important to me I become consumed by feelings of inadequacy."). Participant 3 responded by saying that they did not agree with the statement because they believed a person should just improve on their weaknesses and added that perhaps that person is weak in some areas because God has not gifted them in that area, but instead God has blessed that person in a different area. This participant's answer suggests that religion may be a contributing factor in how negative items are processed. Instead of focusing on how they personally deal with feelings of inadequacy, Participant 3 outright rejects the possibility of feeling inadequate by claiming that perhaps it is God who has made them adequate in other areas. Goddard (2002) states that Malay culture in Malaysia is heavily influenced by the Islamic religion and hence, this is likely why Participant 3 brought in the element of God when confronted with the concept of inadequacy. Goddard (2002) claims that Malays are brought up from their childhood to value refined behaviour (known as *halus* in BM), therefore, it is reasonable to believe that appearing inadequate is probably a behaviour that is looked down upon in Malay culture, hence, it is avoided and, in this case, a religious element is brought in to provide a defense.

5.7.4 Double-barreled Questions

Extending further from the previous issue of translation is the issue of double-barreled questions. As mentioned previously, Lavrakas (2008) state that double-barreled questions can be misleading for readers because they involve multiple concepts in a single statement. Although the internal consistency results only indicated that one item from the self-judgement scale needed to be edited due to its double-barreled nature, it was noted that there were actually several other items which fall into this construct. Below is Table 5.2 which consists of other negative SCS items and their corresponding double-barreled phrases. All these negative items share something in common, which Hillerns and Lei (2015) warned against when constructing items; they each have a conjunction linking two concepts within the same item. Although in all the cases the conjunction "and" links two similar concepts, nonetheless, multiple concepts can be an issue for participants trying to accurately understand what an item is saying.

Table 5.2

Items and their Corresponding Double-barreled Phrases

Item	Double-barreled Phrase
Item 2: When I'm feeling down I tend to obsess and fixate on everything that's wrong	Obsess and fixate
Item 4: When I think about my inadequacies, it tends to make me feel more separate and cut off from the rest of the world	Separate and cut off
Item 11: I'm intolerant and impatient towards those aspects of my personality I don't like.	Intolerant and impatient
Item 23: I'm tolerant of my own flaws and inadequacies.	Flaws and inadequacies

5.11 Implications of the Study

Firstly, this current study supports past research on other translations of the SCS. Internal consistency of the translated SCS-BM is high, the 6-factor model has acceptable model fit, and the SCS-BM showcases convergent and divergent validity. All these measures are in alignment with past studies who found similar findings when translating and validating the SCS in their local languages, Therefore, the theoretical implication of this current study is that it provides additional evidence for the validity of the 6-factor model of SCS originally proposed by Neff (2003).

Secondly, this current study builds upon and extends the research by Khatib et al. (2018) which primarily assessed the reliability of the SCS in Malay. In Khatib et al.'s (2018) study, the only statistical analysis conducted was to measure the internal consistency of the overall scale and its sub-scales. The current study not only matches this measure of internal consistency but further extends it by providing test-retest reliability scores for both the overall scale and its sub-scales after a 3-week period. Moreover, the current study extended the past Malaysian study further by investigating the construct validity in the form of a CFA to better asses the psychometric properties of the SCS-BM. In addition to that, the current study also considered two measures of construct validity. Therefore, this research supports the past findings discovered by Khatib et al. (2018) and extends the research on the BM translation of SCS further by providing a more comprehensive psychometric assessment of the scale with regards to both reliability and validity.

Finally, the practical implication of this study is that it has produced a valid and reliable SCS-BM for other researchers and individuals to use. Refer to Appendix M for the final SCS-BM. At the time of writing, to the best knowledge of the researcher, the current study is the first study in Malaysia to produce a valid and

reliable SCS-BM. This practical implication thus fulfils the primary research objective of this study, which was to produce an accurately translated SCS-BM and now that it has been produced, the SCS-BM can be administered and used in situations that require the evaluation of self-compassion in the BM language, e.g. local counselling or psychological assessment sessions.

5.12 Suggestions for Future Research

Firstly, researcher may want to consider replacing the chi-square test of goodness of fit with another fit index. Although it is often included traditionally in translation studies, most researchers are unable to find model fit using the chi-square, for reasons such as large sample size, in spite of achieving good model fit with other indexes using the same sample. Therefore, in order to prevent misleading results when evaluating model fit using CFA in SCS validation studies, this study suggests replacing the chi-square measure with another suitable fit index, e.g. GFI or NFI.

Secondly, considering the cultural and morphological differences between English and BM, this study suggests that future studies could look into improving the backtranslation process of the SCS so that it is better suited for a local context in Malaysia. Issues such as the negative phrasing of items and double-barreled questions need to be addressed in greater detail in light of improving translation quality. A future researcher who has sufficient funds may consider hiring more professional translation experts, preferably those in the area of psychology, to further refine the SCS-BM translation.

Lastly, the researcher suggests utilizing Item Response Theory (IRT) for future analysis of the SCS. Reason being that IRT can overcome the aforementioned drawbacks of CTT and internal consistency. One of the primary advantages of IRT analysis is that it is largely sample independent. Unlike the CFA fit indices which can

be influenced by sample size, IRT can assess item fit regardless of the sample of respondents and person-fit regardless of the sample of items (Cohen et al., 2007). In addition to that, while CFA can only measure the factor loading for each item, IRT analysis can assess the item discrimination of each item (Cohen et al., 2007), which then allows researcher to more accurately pinpoint which items in the SCS are not discriminating the self-compassion of participants. Future researchers should therefore look into utilizing IRT to assess the validity of the SCS-BM scale in a local population.

5.13 Conclusion

In conclusion, the results of this study indicate that the SCS-BM is a valid and reliable measure for self-compassion. The SCS-BM has both a high internal consistency and a moderate test-retest reliability for the overall scale. Confirmatory factor analysis indicated that the original 6-factor model proposed by Neff (2003) was found to have adequate fit to the data. This study adds to the growing body of translations of the SCS into local languages and now other Malaysian researchers can continue to build upon the research on self-compassion in Malaysia by using the SCS-BM scale produced by this research.

References

Arimitsu, K. (2014). Development and validation of the Japanese version of the Self-Compassion Scale. *Shinrigaku Kenkyu, 85*(1). 50-59. doi: 10.4992/jjpsy.85.50

Awang, Z., Afthanorhan, A., & Mamat, M. (2016). The Likert scale analysis using parametric based Structural Equation Modeling (SEM). *Computational Methods in Social Sciences,* 13-21. Retrieved from https://pdfs.semanticscholar.org/3df3/9fe4762c84b624db953557f0f7f52a407536.pdf

Azizi, A., Mohammadkhani, P., Foroughi, A. A., Lotfi, S., & Bahramkhani, M. (2013). The validity and reliability of the Iranian version of the Self-Compassion Scale. *Iranian Journal of Clinical Psychology*, *2*(3), 17-23. Retrieved from https://self-compassion.org/wp-content/uploads/publications/SCS-Iranian.pdf

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archive General Psychology, 4*, 53-63.

Behr, D. (2016). Assessing the use of back translation: the shortcomings of back translation as a quality testing method. *International Journal of Social Research Methodology*, *20*(6), 573–584. doi:10.1080/13645579.2016.1252188

Bell, R. C. (1982). Person-fit and person reliability. *Education Research and Perspectives*, *9*(1), 105 – 113. Retrieved from www.rasch.org/erp8.htm

Benda, J., & Reichová, A. (2016). Psychometric characteristics of the Czech version of the self-compassion scale (SCS-CZ). *Ceskoslovenská psychologie, 60*(2), 120-136.

Bergh, D. (2015). Chi squared test of fit and sample size—a comparison between a random sample approach and a Chi square value adjustment method. *Journal of Applied Measures*, *16*(2), 204–217.

Berjuang Bersusah payah [Def 1 & Def 3]. (n.d.). Glosbe Online Dictionary. Retrieved August 28, 2019 from https://en.glosbe.com/ms/en/Berjuang

Bersusah payah [Def 2 & Def 5]. (n.d.). Glosbe Online Dictionary. Retrieved August 28, 2019 from https://en.glosbe.com/ms/en/bersusah%20payah

Bertram, D. (2012). *Likert scale*. CPSC 681 – Topic Report. Retrieved 12 June 2012 from http://poincare.matf.bg.ac.rs/~kristina//topicdane-likert.pdf

Blaikie, N. (2003). *Analyzing quantitative data: From description to explanation.* Sage Publications.

Brereton, R. G. (2015). The Mahalanobis distance and its relationship to principal component scores. *Journal of Chemometrics, 29*(3), 143 – 145. doi: 10.1002/cem.2692

Brislin., R. W. (1970). Back-translation for cross cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185-216.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* The Guilford Press.

Bujang, M. A., Omar, E. D., & Baharum, N. A. (2018). A review on sample size determination for Cronbach's alpha test: a simple guide for researchers. *Malaysian Journal of Medical Sciences, 25*(6), 85-99. doi: 10.21315/mjms2018.25.6.9

Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences, 3*(3), 106–116. doi:10.3844/jssp.2007.106.116

Carifio, J., & Perla, R. J. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education, 42*(12), 1150–1152. doi:10.1111/j.1365-2923.2008.03172.x

Castilho, P., & Pinto-Gouveia, J. (2011). Self-compassion: Validation of the Portuguese version of the Self-Compassion Scale and its relation with early negative experiences, social comparison and psychopathology. *Psychologica*, *54*, 203-231. doi: 10.14195/1647-8606_54_8

Chang, D., Pua, X. C., & Ng, S. S. L. (2019). Relationship between self-compassion, perfectionism, and body image satisfaction among university students in Malaysia (Bachelors thesis), University Tunku Abdul Rahman, Kuala Lumpur. Retrieved from http://eprints.utar.edu.my/3217/1/fyp_PY_2019_CD_%2D_1505207.pdf

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464-504. doi: 10.1080/10705510701301834

Chen, J., Yan, L., & Zhou, L. (2011). Reliability and validity of Chinese version of Self-Compassion Scale. *Chinese Journal of Clinical Psychology*, *19*(6), 734-736.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education* (6th ed.). Routledge.

Cortina, M., & Liotti, G. (2010). Attachment is about safety and protection. Intersubjectivity is about social understanding and sharing. *Psychoanalytic Psychology*, *27*(4), 410 – 441. doi: 10.1037/a0019510

Costa, J., Marôco, J., Pinto-Gouveia, J., Ferreira, C., & Castilho, P. (2015). Validation of the psychometric properties of the Self-Compassion Scale. *Clinical Psychology & Psychotherapy*, *23*, 460-468. doi: 10.1002/cpp.1974

de Souza, L. K., & Hutz, C. S. (2016). Adaptation of the Self-Compassion Scale for use in brazil: Evidences of construct validity. *Trends in Psychology*, *24*(1), 159-172. doi: 10.9788/TP2016.1-11

Deniz, M. E., Kesici, S., & Sümer, A. S. (2008). The validity and reliability of the Turkish version of the Self-Compassion Scale. *Social Behaviour and Personality, 36*(9), 1151-1160. Retrieved from https://self-compassion.org/wp-content/uploads/publications/Turkish_self_compassion_scale.pdf

Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, *95*(3), 542-575. http://dx.doi.org/10.1037/0033-2909.95.3.542

Diener, E. (2000). Subjective well-being: The science of happiness and a proposal for a national index. *American Psychologist*, *55*, 34–43. doi: 10.1037/0003-066X.55.1.34

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment, 49*(1), 71-75. doi: 10.1207/s15327752jpa4901_13

Dundas, I., Svendsen, J. L., Wiker, A. S., Granli, K. V., & Schanche, E. (2016). Self-compassion and depressive symptoms in a Norwegian student sample. *Nordic Psychology*, *68*, 58–72. doi: 10.1080/19012276.2015.1071203

Elliot, A. C., & Woodward, W. A. (2007). *Statistical analysis quick reference guidebook with SPSS examples*. London: Sage Publications

Elsevier. (n.d.). *In an age of globalization, will the Asian academic emergence continue in countries like Malaysia and beyond?* https://www.elsevier.com/research-intelligence/campaigns/malaysia-research-excellence-and-beyond

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). Singapore, Singapore: SAGE Publications Asia-Pacific Pte. Ltd.

Garcia-Campayo, J., Navarro-Gil, M., Andrés, E., Montero-Marin, J., López-Artal, L., & Demarzo, M. M. P. (2014). Validation of the Spanish versions of the long (26 items) and short (12 items) forms of the Self-Compassion Scale (SCS). *Health and Quality of Life Outcomes*, *12*(4). Retrieved from http://www.hqlo.com/content/12/1/4

Get down on [Def. 1]. (n.d.). *In Macmillan Dictionary,* Retrieved August 1, 2019, from https://www.macmillandictionary.com/dictionary/british/get-down-on

Gilbert, P. (1989). *Human nature and suffering*. Lawrence Erlbaum Associates Limited.

Gilbert, P. (2007). *The compassionate mind foundation and compassion focused therapy.* http://www.compassionfocusedtherapy.com/new-page/

Gilbert, P. (2014). The origins and nature of compassion focused therapy. *British Journal of Clinical Psychology, 53*, 6 – 41. doi: 10.1111/bjc.12043

Goddard, C. (2002). Directive speech acts in Malay (Bahasa Melayu): an ethnopragmatic perspective. *Cahiers de praxématique*. 113 – 143. Retrieved from http://journals.openedition.org/praxematique/582

Heine, S. J., & Hamamura, T. (2007). In search of East Asian Self-enhancement. *Personality and Social Psychology Review*, *11*(4), 4-27. doi: 10.1177/1088868306294587

Hermanto, N., & Zuroff, D. C. (2016). The social mentality theory of self-compassion and self-reassurance: The interactive effect of care-seeking and caregiving. *The Journal of Social Psychology.* doi: 10.1080/00224545.2015.1135779

durmaz, D., & Aydin, A. (2017). The relationship between self-compassion and multidimensional perfectionism levels and influencing factors in nursing students. *Journal of Psychiatric Nursing,* 86 – 94. doi: 10.14744/phd.2017.40469

Hillerns, C., & Lei, P. P. (2015). A quest to eliminate double-barreled survey questions. *Center for Health Policy and Research Publications*, 188. Retrieved from https://escholarship.umassmed.edu/cgi/viewcontent.cgi?article=1186&context=healthpolicy_pp

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural Equation Modelling: Guidelines for determining model fit. *E-Journal of Business Research Methods,* 53-59. Retrieved from http://www.ejbrm.com/issue/download.html?idArticle=183

Intersubjectivity [Def. 1]. (n.d). *In American Psychological Association,* Retrieved May 5, 2020, from https://dictionary.apa.org/intersubjectivity

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education, 38*(12), 1217 – 1218. doi: 10.1111/j.1365-2929.2004.02012.x.

Kagitcibasi, C. (2005). Autonomy and relatedness in cultural context: Implications for self and family. *Journal of Cross Cultural Psychology, 36*(403), 403 – 422. doi: 10.1177/0022022105275959

Khatib, N. A. M., Roseliza-Murni, A. R., Hoesni, S. M., & Manap, J. (2018). Kesahan dan Kebolehpercayaan Instrumen Skala Belas Kasihan Kendiri (SBKK) Versi Bahasa Melayu. *Jurnal Psikologi Malaysia, 32*(3), 35-46. Retrieved from http://spaj.ukm.my/ppppm/jpm/article/view/409/348

Kim, J. J., Parker, S. L., Doty, J. R., Cunnington, R., Gilbert, P., & Kirby, J. N. (2020). Neurophysiological and behavioural markers of compassion. *Scientific Reports, 10*(1), 6789 – 6798. doi: 10.1038/s41598-020-63846-3

Kline, R. B. (2005). *Methodology in the social sciences. Principles and practice of structural equation modeling* (3rd ed.). New York, NY, US: Guilford Press.

Knapp, T. R. (1990). Treating ordinal scales as interval scales: An attempt to resolve the controversy. *Nursing Research, 39*(2), 121 – 123. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.460.246&rep=rep1&type=pdf

Kotsou, I., & Leys, C. (2016). Self-Compassion Scale (SCS): Psychometric properties of the French translation and its relations with psychological well-being, affect and depression. *PLoS ONE, 11*(4), 1-11. doi: 10.1371/journal.pone.0152880

Kuzon, W. M. Jr., Urbanchek, M. G., & McCabe, S. (1996). The seven deadly sins of statistical analysis. *Annals of Plastic Surgery, 37*(3), 265 – 272. doi: 10.1097/00000637-199609000-00006

Kyriazos, T. A. (2018). Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, *9*, 2207-2230. doi: 10.4236/psych.2018.98126.

Lavrakas, P. J. (2008). Double-barreled Question. In *Encyclopedia of Survey Research Methods.* Retrieved from https://methods.sagepub.com/reference/encyclopedia-of-survey-research-methods/n145.xml

Lee, W. K., & Lee, K. (2010). The validation study of the Korean version of Self-Compassion Scale with adult women in community. *Journal of Korean Neuropsychiatric Association*, *49*(2), 193-200. doi: 10.17315/kjhp.2008.13.4.012

Lee, C., Menon, S., & Rajaendram, R. (2018, October, 12). Too many teens suffering from stress. *The Star Online*. Retrieved from https://www.thestar.com.my/news/nation/2018/10/12/too-many-teens-suffering-from-stress-examoriented-culture-not-only-affecting-students-but-teachers-a/

Lewis-Beck, M. S. (2004). Research methods. In Bryman, A. & Liao, T. F. (Eds.) In *The SAGE Encyclopedia of Social Science Research Methods*. Retrieved from http://sk.sagepub.com/reference/socialscience/n1011.xml

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage Publications

Liotti, G., & Gilbert, P. (2011). Mentalizing, motivation, and social mentalities: Theoretical considerations and implications for psychotherapy. *Psychology and Psychotherapy: Theory, Research and Practice, 84*. 9 – 25. doi: 10.1348/147608310X520094

López, A., Sanderman, R., Smink, A., Zhang, Y., van Sonderen, E., & Ranchor, A. (2015). A reconsideration of the Self-Compassion Scale's total score: self-compassion versus self-criticism. *PloS One*, *10*(7), 1-12. doi: 10.1371/journal.pone.0132940

Malaysian Healthcare Performance Unit (2016). *Malaysian mental healthcare performance: technical report 2016*, 1-67. Retrieved from http://www.moh.gov.my/index.php/file_manager/dl_item/554756755a584a69615852686269387 94d4445334c3078686347397959573476545756756447467349456 86c5957783061474e68636d5567554756795a6d3979625746755932556 7556d567762334a30494449774d5459756347526d

Maneesriwongul, W., & Dixon, J. K. (2004). Instrument translation process: a methods review. *Journal of Advanced Nursing 48*(2), 175–186. doi: 10.1111/j.1365-2648.2004.03185.x

Mantzios, M., Wilson, J. C., & Giannou, K. (2013). Psychometric properties of the Greek versions of the Self-Compassion and Mindful Attention and Awareness Scales. *Mindfulness, 6*(1), 123 – 132. doi: 10.1007/s12671-013-0237-3

Mat Awal, N., Abu Bakar, K., Abdul Hamid, N. Z., & Jalaluddin, N. H. (2007). Morphological differences between Bahasa Melayu and English: constraints in students' understanding. *Faculty of Communication and Modern Languages, Universiti Utara Malaysia*. Retrieved from http://repo.uum.edu.my/3264/1/N6.pdf

Mey, S. C., & Yin, C. J. (2014). Mental health and wellbeing of the undergraduate students in a research university: A Malaysian experience. *Social Indicators Research, 122*, 539 – 551. doi: 10.1007/s11205-014-0704-9

Mukhtar, F., & Oei, T. P. S. (2008). Exploratory and confirmatory factor validation and psychometric properties of the Beck Depression Inventory for Malays (BDI-Malay) in Malaysia. *Malaysian Journal of Psychiatry Ejournal, 17*(1). Retrieved from https://www.mjpsychiatry.org/index.php/mjp/article/view/27

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity, 2*, 223–250. doi: 10.1080/15298860390209035

Neff, K. D. (2009). The role of self-compassion in development: A healthier way to relate to oneself. *Human Development*, *52*, 211 – 214. doi: 10.1159/000215071

Neff, K. D. (2012). The science of self-compassion. In C. Germer & R. Siegel (Eds.), *Compassion and Wisdom in Psychotherapy* (pp. 79-92). Guilford Press.

Neff, K. D., Bluth, K., Tóth-Király, I., Davidson, O., Knox, M. C., Williamson Z., & Costigan, A. (2020): Development and validation of the Self-Compassion Scale for youth. J*ournal of Personality Assessment*, 1 – 14. doi: 10.1080/00223891.2020.1729774

Neff, K. D., Hsieh, Y. P., & Dejitterat, K. (2014). Self-compassion, achievement goals and coping with academic failure. *Self and Identity, 4*(3), 263 – 287. doi: 10.1080/13576500444000317

Neff, K. D., Kirkpatrick, K. L., & Rude, S. S. (2007). Self-compassion and adaptive psychological functioning. *Journal of Research in Personality*, *41*, 139 – 154. doi: 10.1016/j.jrp.2006.03.004

Neff, K. D., Tóth-Király, I., Yarnell, L. M., Arimitsu, K., Castilho, P., Ghorbani, N., … Mantzios, M. (2018). Examining the factor structure of the self-compassion scale in 20 diverse samples: support for use of a total score and six subscale scores. Examining the factor structure of the Self-Compassion Scale in 20 diverse samples: Support for use of a total score and six subscale scores. *Psychological Assessments*, *31*(1). doi: 10.1037/pas0000629

Padilla, J. -L., & Benítez, I. (2014). Validity evidence based on response process. *Psicothema, 26*(1), 136-144. doi: 10.7334/psicothema2013.259

Pavot, W., & Diener, E. (2008). The Satisfaction With Life Scale and the emerging construct of life satisfaction. *Journal of Positive Psychology*, *3*, 137–152.

Petrocchi, N., Ottaviani, C., & Couyoumdjian, A. (2013). Dimensionality of self-compassion: translation and construct validation of the self-compassion scale in an Italian sample. *Journal of Mental Health, 23*(2), 1-6. doi: 10.3109/09638237.2013.841869

Phillips, W. J. (2019). Self-compassion mindsets: The components of the self-compassion scale operate as a balanced system within individuals. *Current Psychology*. doi: 10. 1007/s12144-019-00452-1

Rockliff, H., Karl, A., McEwan, K., Gilbert, J., Matos, M., & Gilbert, P. (2011). Effects of intranasal oxytocin on compassion focused imagery. *Emotion*, *11*, 1388–1396. doi:10.1037/a0023861

Shamsuddin, K., Fadzil, F., Ismail, W. S. W., Shah, S. A., Omar, K., Muhammad, N. A., … Mahadevan, R. (2013). Correlates of depression, anxiety and stress among Malaysian university students. *Asian Journal of Psychiatry, 6*, 318 – 323. doi: 10.1016/j.ajp.2013.01.014

Sivanandam, H., Zainal, H., & Tan, T. (2018, March, 29). Ministry addresses suicide tendency among students. *The Star Online*. Retrieved from https://www.thestar.com.my/news/nation/2018/03/29/ministry-addresses-suicide-tendency-among-students/

Statistics Solutions. (2013). *Confirmatory Factor Analysis*. Retrieved from http://www.statisticssolutions.com/academic-solutions/resources/directory-of-statistical-analyses/confirmatory-factor-analysis/

Swami, V., & Chamorro-Premuzic, T. (2009). Psychometric evaluation of the Malay Satisfaction with Life Scale. *Social Indicators Research, 92*, 25-33. doi: 10.1007/s11205-008-9295-7

Taber, K. S. (2017). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education,* 1-24. doi: 10.1007/s11165-016-9602-2

The University of Kentucky (n.d.). *Copyright FAQ.* Retrieved from https://psychology.as.uky.edu/copyright-faq

Tolerant [Def. 1]. (n.d.). *In Cambridge Dictionary,* Retrieved August 1, 2019, from https://dictionary.cambridge.org/dictionary/english/tolerant

Tsai, M. Y. (2015). Construction and factorial validation of the Chinese version of the Self-Compassion Scale for gifted students. *Psychology Research*, *5*(11), 634 – 644. doi: 10.17265/2159-5542/2015.11.003

Tyupa, S. (2011). A theoretical framework for back-translation as a quality assessment tool. *New Voices in Translation Studies, 7*, 35 – 46. Retrieved from https://core.ac.uk/download/pdf/53121361.pdf

Willis, G. B. (1999). *Cognitive interviewing: A how to guide.* Retrieved from http://www.chime.ucla.edu/publications/docs/cognitive%20interviewing%20guide.pd

Willson, S., & Miller, K. (2014 in press). Data collection. In Miller, K., Chepp, V., Willson, S., & Padilla, J. L. (Eds.), *Cognitive Interviewing Methodology: A Sociological Approach for Survey Question Evaluation*. New York, NJ: John Wiley and Sons.

Yamaguchi, A., Kim, M. S., & Akutsu, S. (2014). The effects of self-construals, self-criticism, and self-compassion on depressive symptoms. *Personality and Individual Differences, 68*, 65 – 70. doi: 10.1016/j.paid.2014.03.013

Ying, K., & Tang, S. W. (2015). Exploring factors associated with self-compassion among local students in a public university (Master's thesis), Universiti Sains Malaysia, Penang. Retrieved from http://eprints.usm.my/30097/1/EXPLORING_FACTORS_ASSOCIATED_WITH_SELFCOMPASSION_AMONG_LOCAL_STUDENTS_IN_A_MALAYSIAN_PUBLIC_UNIVERSITY.pdf

Zeng, X., Wei, J., Tian, P. S. O., & Liu, X. (2016). The self-compassion scale is not validated in a Buddhist sample. *Journal of Religion and Health*. doi: 10.1007/s10943-016-0205-z