Journal of Applied Structural Equation Modeling

Journal of Applied Structural Equation Modeling: 1(1), i-xiii, June 2017

EDITORIAL

A REVIEW OF THE METHODOLOGICAL MISCONCEPTIONS AND GUIDELINES RELATED TO THE APPLICATION OF STRUCTURAL EQUATION MODELING: A MALAYSIAN SCENARIO

Mumtaz Ali Memon[a*], Hiram Ting[b], T. Ramayah[c], Francis Chuah[d] and Jun-Hwa Cheah[e]

[a]*Centre of Social Innovation, Universiti Teknologi PETRONAS, Perak, Malaysia*
[b]*Sarawak Research Society, Sarawak, Malaysia*
[c]*School of Management, Universiti Sains Malaysia, Penang, Malaysia*
[d]*School of Business & Management, Universiti Utara Malaysia, Kedah, Malaysia*
[e]*International Business School, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia*

[*a]*mumtazutp@gmail.com*, [b]*hiramparousia@gmail.com*, [c]*ramayah@gmail.com*,
[d]*francischuah@uum.edu.my*, [e]*junhwa@ibs.utm.my*

ABSTRACT

Although structural equation modeling (SEM) is a powerful statistical technique, understanding its methodological assumptions before data analyses is essential to attaining more robust results. In this editorial, we outline four major methodological issues which are related to the application of SEM in Malaysia along with their respective guidelines. These issues include 1) probability and non-probability sampling, 2) pre-testing and pilot study, 3) CB-SEM and PLS-SEM, and 4) exploratory and confirmatory factor analysis. We also recommend the steps that the local research community, especially the postgraduate students, should consider taking to keep themselves up-to-date with methodological advances and to make informed decisions about the use of SEM. This humble effort will help to clarify the confusion and doubts many lecturers and postgraduate students in Malaysia might have, and provide directions to what they should do in a practical manner.

Keywords: *PLS-SEM; CB-SEM; structural equation modeling; probability sampling; non-probability sampling; pre-testing; pilot study; CFA; EFA; Methodology*

INTRODUCTION

With the amount of effort dedicated to investigating and understanding human behavior, which itself is complex, it is of no surprise that the methodological development in social science and business research occurs more rapidly than ever. These new findings, meticulous procedures and well thought-out recommendations can be found in the articles published in some of the top journals, such as Organization Research Methods, Psychological Methods, Journal of Management Studies and Journal of the Academy of Marketing Science (see Aguinis & Edwards, 2014; Green, Tonidandel, & Cortina, 2016; Henseler, Ringle, & Sarstedt, 2015; Holland, Shore, & Cortina, 2016; Hulland, Baumgartner, & Smith, 2017; McNeish, 2017). The explanations and guidelines provided by the authors, many of whom are the most prolific and seasoned researchers, are extremely useful to the researchers and the postgraduate students because they point out what is lacking in the existing literature and the customary practices, and offer solutions as well as alternatives to reinforce the rigor and practicality of various research methods and techniques.

Notwithstanding its convenient availability and accessibility in the increasingly networked world, it strikes us that many lecturers and supervisors in Malaysia remain unfamiliar with the current methodological advances. They still rely on what they learned in the past, insist on what they are acquainted with and instruct their students to do things rigidly, thus disregarding what is right, current, documented and more appropriate to specific needs and situations. Most of these instructions often seem easy to be carried out, but they would likely lead to unnecessary procedures, misleading results and misinformed decisions. Although what was published and adopted in the past can still be of use today, dwelling on it in the expense of what is current and recommended certainly has no place in the rigor of research in any form (Guide & Ketokivi, 2015). Consequently, many postgraduate students at the local universities either become fixated to one approach for a long time or get into trouble with some vexatious matters, such as model misspecification, systematic errors, inappropriate use of statistical software and techniques as well as mishandling of data in the course of their studies. Such regrettable situations become worsened when they only want their problems to be fixed rather than to understand the subject matter and learn to better themselves.

Countless times we have been approached by lecturers and students to help address issues pertaining to the research methodology and the use of structural equation modeling (SEM), be it in the workshops, conferences or social networking groups, such as mySEM and Doctorate Support Group (DSG) on Facebook. The manner we respond has always been the same and that is to firstly, understand what their research is about, and secondly, provide some probable guidelines or solutions with the support from the literature. We do the same in the social networking groups although admittedly the amount of information and explanation we can provide there is often limited. In many occasions, we also see resourceful comments from other members, including some experts with impressive track record of publications, and that really helps not only the person who posts, but also the readers. We are glad to see some who taught and commented something incorrect in the past gradually change their stance and have started to preach the right thing. Nevertheless, we are still puzzled that some erroneous teachings continue to be taught and spread online, and many postgraduate students still buy the idea, fall into it and embrace it. It worries us that these ill-considered instructions are still widely disseminated in the methodology classes, workshops, and proposal defense sessions at the local universities in Malaysia.

The purpose of this Editorial is not to merely regurgitate what is already in the literature. It is a continual effort to alarm the current situations of the research community in Malaysia, put together four major methodological issues which are related to the application of SEM in Malaysia along with their respective guidelines, and recommend the steps the researchers and postgraduate students should consider taking to keep themselves abreast with the methodological advances and make informed decisions about the use of the analytical techniques (see also Ramayah, Cheah, Chuah, Ting, & Memon, 2016). For the record, we are neither methodologists nor pioneers in any methods. Instead, we are keen learners who keep close track to all the recent development of methodology to the best of our ability, including that of the use of SEM, either through personal contact with the experts or by reading the articles published in the top social science and business journals. Therefore, it is our hope that this paper would complement the extant literature on the subject matter and articulate what should be considered and adopted before the use of SEM to eradicate the confusion and doubts many postgraduate students in Malaysia might have.

PROBABILITY AND NON-PROBABILITY SAMPLING

"Even the most sophisticated statistical methods cannot offset badly designed samples".
Sarstedt, Bengart, Shaltoni, and Lehmann (2017, p. 10)

Although probability sampling is believed as ideal in research, the vast majority of studies in social science research actually draw upon non-probability samples (Rowley, 2014). In probability sampling, each individual has a known chance of being selected, whereas in non-probability sampling, the chances of being selected are unknown. As such, the inclusion may be considered based on purposive selection, opportunity, or expert judgment (Burns, Duffett, Kho, Meade, Adhikari, & Sinuff, 2008). In reality, non-probability sampling is more frequently employed and more likely appropriate in fieldwork research (Bryman & Bell, 2015). Specifically, studies with humans as subjects are less likely to involve random samples (Polit & Beck, 2010). Yet, the myth of probability sampling being the prerequisite to doing something of a PhD level persists among the Malaysian academics. It is unfortunate that many postgraduate students often receive suggestions from their supervisors or examiners during proposal defense to adopt probability sampling procedures without considering its underlying assumptions. The significance of sampling in general and non-probability sampling in particular, therefore, continues to be overlooked, misunderstood and misappropriated.

A carefully controlled non-probability sampling can provide valid and meaningful results (Cooper & Schindler, 2011). Choosing one sampling strategy over another has nothing to do with the level of research quality. Although probability sampling is ideal in terms of its sampling generalizability, in many occasions it is not appropriate and necessary. Calder, Phillips, and Tybout (1981) have long argued that representativeness is inappropriate when the goal of research is rigorous theory testing. As such, a non-probability sample is deemed more fitting when the purpose is to test the proposed theoretical assumptions (Hulland et al., 2017). On the contrary, probability sampling is preferred when the research emphasizes the generalizability of findings from the sample respondents to the target population. Therefore, a social science study which uses theories to explain a phenomenon in a different context is more likely about theory generalization, rather than sampling generalization. The extension of knowledge (theory and concept) is what warrants a PhD and is most often what the editors of the top journals will preliminary and primarily look for.

When selecting probability sampling and its techniques, it is important to note that it only works when the sampling frame is available. In other words, the complete list of all the subjects in the target population must be obtained (Cooper & Schindler, 2011; Saunders, Lewis, & Thornhill, 2016). Without the sampling frame, random and stratified sampling techniques, for instance, and the use of formula to calculate the sample sizes are categorically inappropriate. Even if a good sampling frame can be obtained to recruit random samples, achieving a 100% response rate remains a big challenge (Rowley, 2014); it is insurmountable in any studies which involve diverse and dispersed populations. In practice, the completeness of the list obtained (e.g. whether it is the latest list) in Malaysia and the manner data collection is administered (e.g. giving every subject in a wide geographical area equal chance) can easily compromise the underlying assumptions of probability sampling. Unfortunately, these assumptions are more than often ignored and thus unfulfilled. As such, the trend of using probability sampling for the sake of sampling generalization when the fundamentals of the said sampling are violated really puzzles us. Such flawed belief, insistence and practice are made more obvious by the fact that the results as well as the implications of the research are not yielding any practical meaningfulness to the target population (Seddon & Scheepers, 2012). We would like to echo what is already in the literature that the use of non-probability samples is appropriate when the selection of sampling strategy suits the sampling objectives as well as the scope of research, the research goal is to attain theory generalization, and the complete sampling frame is not available in a given context. To be familiar with the use of both probability and non-probability sampling techniques, we recommend articles by Hulland et al. (2017), Sarstedt et al. (2017), Rowley (2014), and Calder et al. (1981).

PRE-TESTING AND PILOT STUDY

"Whether we are social researchers or epidemiologists, designing surveys or clinical trials, we aim to provide results that are valid, reliable, sensitive, unbiased and complete".
Collins (2003, p. 229)

Despite the availability of detailed guidelines about how to properly design questionnaires (see Brace, 2008; Dillman, 2011; Rowley, 2014), it is often difficult for researchers and postgraduate students to identify and curb all the potential issues that may arise during data collection (Hulland et al., 2017). Pre-testing is therefore imperative for survey questionnaire to confirm that there is no ambiguity in the questions and the respondents could understand the questions the way they are designed and intended (Sekaran, 2003). Evidently, the pre-testing process "rectifies any inadequacies, in time, before administering the instrument orally or through a questionnaire to respondents, and thus reduce biases" (Sekaran, 2003, p. 249). The lack of understanding about pre-testing will likely lead to poor data quality, and consequently the deletion of items and/or cases during measurement model assessment. Surprisingly, this occurs too frequently among the postgraduate students during data analysis process. Ironically, many postgraduate students are only told to do the pilot study and are even compelled to show the pilot study's results during their proposal defense when what they want to achieve is actually done by pre-testing, much to our chagrin. Although pre-testing is often understood as a pilot study (Baker, 1994), both serve distinctive purposes.

A pilot study is a small-scaled version or trial run—a key step to ensuring a full-fledged study will be carried out successfully (Polit, Beck, & Hungler, 2001; Teijlingen & Hundley, 2002). It is also referred to as the dress rehearsal (Moser & Kalton, 1992). Teijlingen and Hundley (2002) emphasized several objectives for conducting a pilot study, such as a) *testing adequacy of research instruments*, b) *assessment of the feasibility of a full-scale project*, c) *assessing whether the research protocol is realistic and workable*, d) *revealing reveal logistics issues*, e) *collecting preliminary data*, f) *ensuring whether the sampling frame and technique are effective*, g) *determining sample size*, h) *convincing funding bodies that the major study is feasible and worth funding and so on.* Given what a pilot study is, it is therefore inappropriate to use the pilot study sample as the main study sample. The researchers and postgraduate students should also be mindful about sampling too many respondents during a pilot study because they are typically the more responsive and cooperative ones. As such a pilot study using questionnaire-based survey might not be feasible or necessary when the target population is relatively small or the respondents are relatively difficult to get.

There are several rules for determining the sample size for a pilot study. For example, Cooper and Schindler (2011) suggested a sample between 25 and 100 individuals (Cooper & Schindler, 2011). It is also said that a range from 10 to 30 individuals are enough for a pilot test (Hill, 1998; Isaac & Michael, 1995). Moreover, several scholars suggested that the sample size should be 10 percent of the sample projected for the main study (Connelly, 2008). Furthermore, the sample size could also be decided based on the type of analysis at the preliminary stage (Cooper & Schindler, 2011). Traditionally, coefficient alpha (α) is calculated to check the internal consistency reliability of the measures. Hence a sample of 30 individuals is usually advocated. This number originates from the Central Limit Theorem which makes a distributional assumption of the sample size of 30 or more to ensure the mean of any samples from the target population will be approximately equal to that of the population. Regrettably, when the results show poor coefficient alpha values, many postgraduate students tend to proceed with main data collection and that usually culminates in a low-quality data set. Even if the results demonstrate good coefficient alpha values, it is crucial to understand that larger sample size, greater number of items and the manner a pilot study is conducted (e.g. conducting a pilot study on the undergraduate students before the class ends) can falsely inflate the values. If the small sample size is mentioned as one of the reasons to unsatisfactory results, it becomes necessary to revise and retest the items in a rigorous manner until the satisfactory results are obtained.

This brings in and highlights the pre-testing process. The main focus of most pre-tests is to address problems that, if not resolved, would accelerate measurement error (Blair & Conrad, 2011). Kumar, Talib, and Ramayah (2013) asserted that the purpose of pre-testing a questionnaire is to ensure whether a) *the wording of the questions is correct*, b) *the sequence of questions is correct*, c) *the respondents have clearly understood all the questions*, d) *additional questions are needed or some questions should be eliminated*, and, e) *the instructions are clear and adequate.* All developed scales, or items, be it adopted or adapted, should be pre-tested to confirm whether the questions work accurately in a new setting with the new respondents (Kumar et al., 2013). Cognitive interview—a typical semi-structured in-depth interview that focuses on respondents' thought processes associated with answering survey questions—is frequently used as a method of pre-testing (Neuert & Lenzner, 2016). Such interviews can be conducted by two methods: *debriefing* and *protocol.* During the debriefing method, the researcher would carefully observe the respondent when he/she fills out the questionnaire. Once completed, the researcher would ask him/her to reveal any problems with the questions (Shelby, Hunt, Sparkman, & Wilcox, 1982). Conversely, in the protocol method, the respondent is asked to think aloud while filling out the survey questionnaire, and the researcher would then make careful notes of statements (Shelby et al., 1982). Additionally, a *card sorting* method can also be an effective way to examine whether the questions or items belonging to the same constructs are understood as intended (Collins, 2003).

Similar to the pilot study, several guidelines exist in the published literature for pre-testing. For example, Kumar et al. (2013) mentioned that at least 50 people should be asked to participate. Other researchers suggested that a sample of 30 participants is reasonable for pre-testing questionnaires (Perneger, Courvoisier, Hudelson, & Gayet-Ageron, 2015). Willis (2005) recommended a sample size between 5 and 15 individuals for the large-scale surveys. Ferber and Verdoorn (1962) suggested that a sample of 12 individuals is sufficient. It is clear that there is no clear-cut rule for the pre-test sample size. Practically speaking, it should be decided based on the length and complexity of the questionnaire. A long and complex questionnaire might require a larger sample size than a short and simple questionnaire (Shelby et al., 1982). Unlike the pilot study, a pre-test requires no statistical analysis. Moreover, pre-testing should be conducted using actual respondents—broadly representative of the target population to be chosen for main data collection (Cooper & Schindler, 2011; Kumar et al., 2013). Upon successful processing of the first round, a second round of pre-test with the revised version of the questionnaire is highly recommended. We would encourage the postgraduate students to read articles by Hulland et al. (2017), Rowley (2014), Grimm (2010), Collins (2003), and Drennan (2003), amongst others, as they will give you a better understanding of the subject matter, in addition to being good citations.

CB-SEM or PLS-SEM?

"Statistical approaches are like tools in a mechanic's toolbox. Although in any situation there may be more than one tool that can at least "sort of" get the job done, the good mechanic knows the right tool for the right problem"
Hair, Babin, and Krey (2017a, p. 10)

The selection between Covariance based Structural Equation Modeling (CB-SEM) and Variance based Structural Equation Modeling, later coined as Partial Least Squares Structural Equation Modeling (PLS-SEM) is another issue which appears to remain unresolved in Malaysia despite the abundance of literature on the subject. CB-SEM and PLS-SEM are two widely accepted second generation data analysis approaches (Avkiran, 2017; Richter, Sinkovics, Ringle, & Schlägel, 2016). Each of these approaches is suitable for different research contexts (Hair, Hult, Ringle, & Sarstedt, 2017b). Thus, it is imperative for the postgraduate students to understand the differences and the appropriateness of each approach in order to make a correct choice to address their research questions. Therefore, the characteristics and objectives of the research should be the key distinguishing factors between CB-SEM and PLS-SEM (Hair et al., 2017b; Hair, Ringle, & Sarstedt, 2011). For the record, Analysis of Moment Structures (AMOS) (Arbuckle, 2015) is a software that performs SEM; it is not SEM per se. This also applies to SmartPLS (Ringle, Wende, & Becker, 2015). Both software, which are most widely used by Malaysian researchers compared to others, such as MPlus, Lisrel, WarpPLS and ADANCO, operate with different approaches/techniques to serve different purposes. In the following paragraphs, we would like to respond to two frequently asked questions, that is, firstly, when to use PLS-SEM, and secondly, when to use CB-SEM.

When to use PLS-SEM?

Firstly, PLS-SEM can be applied for exploratory research—when "theory is less developed" (Hair et al., 2017b, p. 15). Specifically, when the primary focus of the research is to predict and explain the key target constructs and/or identify the key driver constructs (Hair et al., 2017b; Rigdon, 2012). Secondly, when formative constructs are part of a model (Hair, Hult, Ringle, & Sarstedt, 2014), PLS-SEM would be the preferable choice. A formative construct is like a regression model, where the indicators (or items) are anticipated to cause a latent construct (Hair et al., 2011). Hence, for models with formative constructs, or combination of both reflective and formative constructs, PLS-SEM has the edge over CB-SEM. Also, it facilitates both modes (regression and correlation weights) in the measurement model more efficiently (Hair et al., 2017b; Hair, Hult, Ringle, Sarstedt, & Thiele, 2017c).

Thirdly, PLS-SEM can handle complex cause-effect structural models (Richter et al., 2016; Rigdon, 2012, 2014). For models with many constructs and indicators, PLS-SEM is a suitable analytical method (Hair et al., 2017b). Additionally, data characteristics, such as small sample size and non-normal data, can be some of the reasons to choose PLS-SEM. Hair et al. (2017b) advocated that the complexity of a structural model does not require large sample size because "PLS algorithm does not compute all the relationships at the same time" (p. 24). As far as the data distribution is concerned, PLS-SEM is labeled as soft-modeling because of its greater flexibility to accommodate distributional assumptions (Hair et al., 2017b; Wold, 1980). Hence, when multivariate normality assumption is a concern, PLS-SEM would be a better option for analysis (Hair et al., 2017b). Nevertheless, there is a caveat on the aforementioned. PLS-SEM might cease to perform effectively if the sample size is too small and the data is extremely not normal. A small sample size must be justifiable by matching it against the population. As such, the postgraduate students are not advised to justify the use of PLS-SEM solely based on small sample size (Rigdon, 2016). Besides, notwithstanding what bootstrapping does, they must learn to screen and clean the data before performing data analysis, such as checking multivariate kurtosis and removing influential outliers. Regarding the reasons for using PLS-SEM, we would recommend the reading of articles by Hair et al. (2017b), Richter et al. (2016), and Rigdon (2016).

When to use CB-SEM?

The main aim of CB-SEM is to assess the fit between theoretical covariance matrix and the observed covariance matrix—how well a proposed theoretical model represents the reality of the context under study. CB-SEM is commonly applied for confirmatory or explanatory research. It is a preferred method when the goal is theory testing, theory confirmation, or the comparison of alternative theories (Hair et al., 2017b). Besides, unlike PLS-SEM, CB-SEM can handle non-recursive models. Therefore, CB-SEM should be used for models with circular relationships or loops of relationships between latent variables (Hair et al., 2017b; Hair et al., 2011).

CB-SEM, which uses maximum likelihood estimation, is inflexible in terms of data distribution. For datasets with ideal data distribution, CB-SEM is the preferred method. Nevertheless it is apparent that Malaysian researchers, including the postgraduate students who are under their supervision or guidance, tend to proceed with CB-SEM analysis even when distributional assumptions are seriously violated. Byrne (2016) recommended that prior to any data analysis in AMOS, it is important to check that the data achieves multivariate normality. Failure to do so can severely affect the end results. Therefore, multivariate kurtosis (e.g., Mardia's coefficient) instead of univariate normality must be taken care of before any analyses. It brings the whole CB-SEM analysis into disrepute if the normality issue is neglected or not addressed, and no explanation is rendered to tackle multivariate normality.

It is also a major concern that many are resort to applying Modification Indices (MI) to improve the model to achieve its fit indices. Notably, they like to co-vary the error terms and any possible links in pursuit of satisfactory model fit (Hair et al., 2017b). Notwithstanding, correlating error terms for the sake of good fit indicates that the measurement is flawed (Hair et al., 2017a). This brings theory testing or confirmation into disrepute because it is unlikely to find theories to support the revised model, and CB-SEM is "not ideal for exploratory research" (Hair et al., 2017a, p. 11). In recent years, there has been a strong call by quantitative methodologists (e.g. Goodboy & Kline, 2017; Green et al., 2016; Hulland et al., 2017) to abandon such practice. Goodboy and Kline (2017) argued, "Such re-specifications are likely to reflect mere sampling error instead of the truth or the population model. Indeed, blind model re-specification is likely to lead the researcher away from the true model, not toward it" (p.73). Likewise, Green et al. (2016) emphasized that correlating error terms is almost always bad practice. However, error terms can be allowed to correlate when the errors are attached to indicators that represent the same variable at different time points or when some indicators are components of other indicators (Cortina, 2002; Green et al., 2016). Therefore, if additional specification is deemed necessary, such as error terms covariation, CB-SEM would be the choice. For a detailed discussion of the issues related to CB-SEM, we would suggest the articles by Hair et al. (2017a), Green et al. (2016), Fan, Chen, Shirkey, John, Wu, Park, and Shao (2016), and Goodboy and Kline (2017).

EFA or CFA?

"Every methodology has its limits, so do not fall into the trap of believing that a given methodology will be necessary or appropriate for every application; conversely, do not reject out of hand any methodology because it is flawed or limited, for a methodology may perform adequately for some purposes despite its flaws".
Green, Tonidandel, and Cortina (2016, p. 3)

Factor analysis plays an important role in determining the appropriateness of measures used in the study. It provides confidence on the validity of item measuring a specific construct. However, the choice of factor analysis remains elusive among Malaysian academics and postgraduate students when it comes to understanding and using exploratory factor analysis (EFA), confirmatory factor analysis (CFA), or both sequentially. While EFA is theory generation method, CFA is a theory testing method (Henson & Roberts, 2006). EFA can be employed when little is known regarding the factor structure and number of factors (Green et al., 2016). As such, this method is mainly adopted during the scale development process and used to specify construct dimensions (Pallant, 2007; Reise, Waller, & Comrey, 2000; Thompson, 2004). Apart from these purposes, it is unnecessary to use both CFA and EFA in most cases (Hulland et al., 2017), given the fact that so many scales have been established and validated to cater multitude of research objectives over the years. Conversely, CFA is more appropriate with a well-established scale and a priori knowledge of the factor structure (Green et al., 2016). Unlike EFA, CFA is driven by theoretical expectations regarding the structure of the data (Henson & Roberts, 2006). Therefore, the postgraduate students should proceed with CFA if the scale is well established and adopted from past literature with explicit theoretical grounding. Harping on the use of EFA with an established scale by arguing that the factor structure would differ due to contextual differences, such as culture, is altogether uncalled for. The concern that an adopted scale from the western sources, for example, would mean differently in Malaysia is in fact related to the instrument design of the researchers and the comprehensibility of the items to the target population (Rowley, 2014). This is dealt with not by performing EFA but by conducting pre-testing as mentioned earlier.

Moreover, using both EFA and CFA on the same data set seems to be a common practice in Malaysia. Henson and Roberts (2006) have long argued, "It is not terribly informative, and can be potentially misleading, to follow an EFA with a CFA on the same data set" (p. 400). Green et al. (2016) further elucidate that "conducting both EFA and CFA on the same dataset confirms nothing else except demonstrating that the two modeling approaches on the same data converge" (p.15). Therefore, it is recommended that the "factor structure from an EFA should be confirmed with CFA on a different data set" (Green et al., 2016, p. 18). In an event where the target population is relatively small, such as an organizational study or a research on a particular group of individuals, it is thus impractical to collect data two times to run EFA and CFA due to contextual differences and the assessment of unidimensionality. While the manner EFA is used and misused is another concern which we do not wish to delve into in this Editorial (such as the use of principal component analysis and Eigenvalue of 1 to determine the number of components), collecting data with an established scale for performing EFA is a total waste of effort, time and respondents. Even if the population is large, the postgraduate students should just carry out CFA so long as the questionnaire is well designed (adopted or adapted) with the support from theory and literature, and it has gone through meticulous pre-testing process.

Furthermore, the notion of conducting CFA separately for each construct rather than as a part of a complete model (referred to as partial CFA) is also incorrect. Unfortunately, this is another misleading advice and practice since partial CFA not only masks any lack of fit, but also fails to detect discriminant validity even when there is a strong correlation between latent constructs (Hair et al., 2017a). As such, it is strongly recommended that the postgraduate students should always include all latent constructs in a single CFA and report a full pattern/structure matrix while assessing model fit, convergent validity, and discriminant validity (Hair et al., 2017a; Henson & Roberts, 2006).

Another questionable approach among local researchers and postgraduate students is the tendency to delete items without understanding and/or addressing the theoretical grounding of such a decision. Consequently, we are often approached by the postgraduate students inquiring about the supporting literature to justify the deletion of a large number of items. Unfortunately, the deletion of items simply based on the statistics actually jeopardizes content validity. As a rule of thumb, one cannot delete more than 20 percent of the total items in the model, be it CB-SEM or PLS-SEM (Hair et al., 2017a; Hair, Black, Babin, & Anderson, 2010). In other words, a total of eight items (maximum) can be excluded if the original instrument consists of 40 items. It is recommended that the postgraduate students should explicitly report how many items are discarded in the process, and justify such exclusion would not impair the content. We would recommend some good articles on the subject matter, and they are authored by Hair et al. (2017a), Fan et al. (2016), Green et al. (2016), Brown (2015), Fabrigar and Wegener (2011) and Winter, Dodou, and Wieringa (2009).

CONCLUDING NOTES

These four methodological misconceptions are found to be the most prevailing conundrums among the postgraduate students in Malaysia when it comes to the application of SEM. Obviously there are other issues which are also wrongly understood and adopted, for example the assessment of multivariate normality, the use of negative-worded statements, and the justification for including a moderator or a mediator in the model. We would visit these topics as Editorials in the future. But again, if a person, be it a researcher, an academic, a supervisor or a postgraduate student, would read some of the articles published in the top journals in social sciences and business as mentioned earlier, he or she would be kept abreast of most of these methodological developments.

To conclude this Editorial, we would like to provide some guidelines to deciding what to believe in when a postgraduate student is stuck in the middle of nowhere about the methods and the application of SEM, and is confused with different advices and comments. These guidelines are not the answers to any of the methodological misconceptions and the misuse of SEM; rather they serve as the pointers to those who are willing to learn and do the right thing to progress in their respective research. There are many baffling and even bizarre articles with erroneous teaching out there, and many still seem to believe them and continue to uphold them. We hope the novice researchers, especially the postgraduate students, whom our Editorial is mainly directed to, would take heed of these guidelines.

Firstly, it is always advisable to check the publication of any alleged expert. If someone conducts a workshop on social science methodology and the use of SEM with clear insistence on certain approaches or practices, it is best to review their publication on the subject matter. It is easy to say things unchecked in a classroom or workshop; it is also easy to write a book without stringent reviewing process. However, having papers peer-reviewed by the international research communities and published at the very least shows the credibility of their claims. Given the advancement of communication technology, the portfolio of any researchers or experts can be easily accessed online. There is no secret about it. Like it or not, it is a simple homework all the postgraduate students should do, and it can be done at our fingertips. To follow a person along with his or her teaching blindly without exercising some discretion and contemplating it critically is not an attribute we expect from any postgraduate student.

Secondly, the postgraduate students have to know the top journals in their own fields. Hence, they should look for articles in those journals, read them and refer to them as methodological guides. Before long they would know who the big names in methodology and their respective fields are. As stated earlier there are many confusing and contradicting articles published in refereed and indexed journals out there. While some of these journals are actually not indexed at all, some are paid journals (note: paid journal here refers to journal which promises quick publication with little or no review at a payment) which publish just about anything that is written. In the absence of rigorous editorial and double-blind review process, whatever that is claimed on the papers must be read with great cautions. Why not read the paper by Green et al. (2016) which is published in Organizational Research Methods to understand why EFA is not needed? Also, why not spend the time to read the paper by Hulland et al. (2017) published in Journal of the Academy of Marketing Science so as to justify the need to conduct a pre-test?

Thirdly, if a postgraduate student aspires to be a true researcher, we would recommend him or her to start writing papers that aim at some good journals. There is nothing wrong to begin with a low tier journal in order to build confidence in writing and publishing. Yet there will be a time when we should start aiming for better journals and submit our papers there. It might sound like a cliché but rejection is never a failure. In most cases, the editors and/or the reviewers will give us constructive comments why our paper is rejected or how it could be made better. If the paper requires minor or major corrections, the comments given by the reviewers will help not only to improve the paper but also to make us a better researcher. Rejection can be hurtful, and major correction can be difficult to swallow but these experts from different places in the world are in fact reviewing our papers and providing comments for free. Hence, if there are any deficiencies in methodology or inadequacies in analysis and interpretation, they will likely be pointed out. This compels us to learn, and in many cases, unlearn and relearn at the same time. Even though it requires commitment, persistence, hard work and sometimes a bit of luck to get through the editors and the reviewers of these journals, it is still a bargain. The onus of learning and progressing as well as disseminating the current knowledge about research methods and the application of SEM is on us.

> *"Our goal as scientists is to do good science and to report fully and accurately; it is not to push an ideological agenda. Sometimes we may not like what we find because it contradicts what we were expecting or hoping to find… research must be done, reported, and synthesized in a responsible and honest way".*
> *Antonakis (2017, p. 16)*

REFERENCES

Aguinis, H., & Edwards, J. R. (2014). Methodological Wishes for the Next Decade and How to Make Wishes Come True. *Journal of Management Studies, 51*(1), 143-174. doi: 10.1111/joms.12058

Antonakis, J. (2017). On doing better science: From thrill of discovery to policy implications. *The Leadership Quarterly, 28*(1), 5-21. doi: https://doi.org/10.1016/j.leaqua.2017.01.006

Arbuckle, J. L. (2015). New York Patent No.: IBM.

Avkiran, N. K. (2017). An in-depth discussion and illustration of partial least squares structural equation modeling in health care. *Health Care Management Science*, 1-8. doi: 10.1007/s10729-017-9393-7

Baker, T. (1994). *Doing Social Research* (2nd ed.). New York: McGraw-Hill.

Blair, J., & Conrad, F. G. (2011). Sample size for cognitive interview pretesting. *Public Opinion Quarterly, 75*(4), 636-658. doi: https://doi.org/10.1093/poq/nfr035

Brace, I. (2008). *Questionnaire Design: How to Plan, Structure and Write Survey Material for Effective Market Research*: Kogan Page Publishers.

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). London: The Guilford Press.
Bryman, A., & Bell, E. (2015). *Business research methods* (4th ed.). United Kingdom: OUP Oxford.
Burns, K. E., Duffett, M., Kho, M. E., Meade, M. O., Adhikari, N. K., & Sinuff, T. (2008). A guide for the design and conduct of self-administered surveys of clinicians. *Canadian Medical Association Journal, 179*(3), 245-252. doi: 10.1503/cmaj.080372
Byrne, B. M. (2016). *Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming* (3rd ed.). New York: Routledge.
Calder, B. J., Phillips, L. W., & Tybout, A. M. (1981). Designing Research for Application. *Journal of Consumer Research, 8*(2), 197-207. doi: http://www.jstor.org/stable/2488831
Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research, 12*, 229-238. doi: http://www.jstor.org/stable/4038871
Connelly, L. M. (2008). Pilot studies. *Medsurg Nursing, 17*(6), 411-412.
Cooper, D. R., & Schindler, P. S. (2011). *Business Research Methods* (11th ed.). New York: McGraw-Hill.
Cortina, J. M. (2002). Big Things Have Small Beginnings: An Assortment of "Minor" Methodological Misunderstandings. *Journal of Management, 28*(3), 339-362. doi: https://doi.org/10.1177/014920630202800305
Dillman, D. A. (2011). *Mail and Internet Surveys: The Tailored Design Method* John Wiley & Sons.
Drennan, J. (2003). Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *Journal of Advanced Nursing, 42*(1), 57-63. doi: 10.1046/j.1365-2648.2003.02579.x
Fabrigar, L. R., & Wegener, D. T. (2011). *Exploratory Factor Analysis.* New York: Oxford University Press.
Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S. R., Park, H., & Shao, C. (2016). Applications of structural equation modeling (SEM) in ecological studies: an updated review. *Ecological Processes, 5*(19), 1-12. doi: 10.1186/s13717-016-0063-3
Ferber, R., & Verdoorn, P. J. (1962). *Research Methods in Economics and Business.* New York: The Macmillan Company.
Goodboy, A. K., & Kline, R. B. (2017). Statistical and practical concerns with published communication research featuring structural equation modeling. *Communication Research Reports, 34*(1), 68-77. doi: http://dx.doi.org/10.1080/08824096.2016.1214121
Green, J. P., Tonidandel, S., & Cortina, J. M. (2016). Getting Through the Gate: Statistical and Methodological Issues Raised in the Reviewing Process. *Organizational Research Methods, 19*(3), 402-432. doi: https://doi.org/10.1177/1094428116631417
Grimm, P. (2010). Social desirability bias. In J. N. Sheth & N. K. Malhotra (Eds.), *Wiley International Encyclopedia of Marketing*: John Wiley & Sons Ltd.
Guide, V. D. R., & Ketokivi, M. (2015). Notes from the Editors: Redefining some methodological criteria for the journal. *Journal of Operations Management, 37*, v-viii. doi: https://doi.org/10.1016/S0272-6963(15)00056-X
Hair, J. F., Babin, B. J., & Krey, N. (2017a). Covariance-Based Structural Equation Modeling in the Journal of Advertising: Review and Recommendations. *Journal of Advertising, 46*(1), 163-177. doi: http://dx.doi.org/10.1080/00913367.2017.1281777
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (Seven ed.).
Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2014). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM).* Los Angeles: USA: SAGE.
Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017b). *A primer on partial least squares structural equation modeling (PLS-SEM)* (2nd ed.). Thousand Oaks, CA: Sage.
Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., & Thiele, K. O. (2017c). Mirror, mirror on the wall: a comparative evaluation of composite-based structural equation modeling

methods. *Journal of the Academy of Marketing Science*, 1-17. doi: 10.1007/s11747-017-0517-x

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice, 19*(2), 139-151.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115-135. doi: 10.1007/s11747-014-0403-8

Henson, R. K., & Roberts, J. K. (2006). Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice. *Educational and Psychological Measurement, 66*(3), 393-416. doi: https://doi.org/10.1177/0013164405282485

Hill, R. (1998). What sample size is "enough" in internet survey research? *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century, 6*(3-4), 1-10.

Holland, S., Shore, D. B., & Cortina, J. M. (2016). Review and Recommendations for Integrating Mediation and Moderation. *Organizational Research Methods*, 1-35. doi: 10.1177/1094428116658958

Hulland, J., Baumgartner, H., & Smith, K. M. (2017). Marketing survey research best practices: evidence and recommendations from a review of JAMS articles. *Journal of the Academy of Marketing Science*, 1-17. doi: 10.1007/s11747-017-0532-y

Isaac, S., & Michael, W. B. (1995). *Handbook in research and evaluation*. San Diego, CA: Educational and Industrial Testing Services.

Kumar, M., Talib, S. A., & Ramayah, T. (2013). *Business Research Methods*. Selangor, Malayisa: Oxford University Press.

McNeish, D. (2017). Thanks Coefficient Alpha, We'll Take It From Here. *Psychological Methods*, 1-23. doi: 10.1037/met0000144

Moser, C. A., & Kalton, G. (1992). *Survey Methods in Social Investigation* (2nd ed.). Aldershot UK: Gower.

Neuert, C. E., & Lenzner, T. (2016). Incorporating eye tracking into cognitive interviewing to pretest survey questions. *International Journal of Social Research Methodology, 19*(5), 501-519. doi: http://dx.doi.org/10.1080/13645579.2015.1049448

Pallant, J. (2007). *SPSS survival manual: A step by tep guide to data analysis using SPSS for Windows* (3rd ed.). New York: Open University Press.

Perneger, T. V., Courvoisier, D. S., Hudelson, P. M., & Gayet-Ageron, A. (2015). Sample size for pre-tests of questionnaires. *Quality of Life Research, 24*(1), 147-151. doi: 10.1007/s11136-014-0752-2

Polit, D. F., & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies, 47*(11), 1451-1458. doi: https://doi.org/10.1016/j.ijnurstu.2010.06.004

Polit, D. F., Beck, C. T., & Hungler, B. P. (2001). *Essentials of nursing research: methods, appraisal, and utilization* (5th ed.). Philadelphia: : Lippincott Williams and Wilkins.

Ramayah, T., Cheah, J., Chuah, F., Ting, H., & Memon, M. A. (2016). *Partial Least Squares Structural Equation Modeling (PLS-SEM) using SmartPLS 3.0: An Updated Guide and Practical Guide to Statistical Analysis*. Kuala Lumpur, Malaysia: Pearson.

Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment, 12*, 287-297. doi: 10.1037//1040-3590.12.3.287

Richter, N. F., Sinkovics, R. R., Ringle, C. M., & Schlägel, C. (2016). A critical look at the use of SEM in international business research. *International Marketing Review, 33*(3), 376-404. doi: https://doi.org/10.1108/IMR-04-2014-0148

Rigdon, E. E. (2012). Rethinking Partial Least Squares Path Modeling: In Praise of Simple Methods. *Long Range Planning, 45*(6), 341-358.

Rigdon, E. E. (2014). Rethinking partial least squares path modeling: breaking chains and forging ahead. *Long Range Planning, 47*(3), 161-167. doi: https://doi.org/10.1016/j.lrp.2014.02.003

Rigdon, E. E. (2016). Choosing PLS path modeling as analytical method in European management research: A realist perspective. *European Management Journal, 34*(6), 598-605. doi: https://doi.org/10.1016/j.emj.2016.05.006

Ringle, C. M., Wende, S., & Becker, J. M. (2015). *SmartPLS 3*. SmartPLS GmbH. Boenningstedt. Retrieved from http://www.smartpls.com

Rowley, J. (2014). Designing and using research questionnaires. *Management Research Review, 37*(3), 308-330.

Sarstedt, M., Bengart, P., Shaltoni, A. M., & Lehmann, S. (2017). The use of sampling methods in advertising research: a gap between theory and practice. *International Journal of Advertising*, 1-14. doi: 10.1080/02650487.2017.1348329

Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research Methods for Business Students* (7th ed.). Essex, England: Pearson Education Limited.

Seddon, P. B., & Scheepers, R. (2012). Towards the improved treatment of generalization of knowledge claims in IS research: drawing general conclusions from samples. *European Journal of Information Systems, 21*(1), 6-21.

Sekaran, U. (2003). *Research methods for business: A skill building approach* (4th ed.). United State of America: John Willey and Sons, Inc.

Shelby, D., Hunt, R. D., Sparkman, J., & Wilcox, J. B. (1982). The Pretest in Survey Research: Issues and Preliminary Findings. *Journal of Marketing Research, 19*(2), 269-273. doi: http://www.jstor.org/stable/3151627

Teijlingen, E. v., & Hundley, V. (2002). The importance of pilot studies. *Nursing Standard, 16*(40), 33-36. doi: http://hdl.handle.net/2164/157

Thompson, B. (2004). *Factor Analysis: Understanding Concepts and Applications*. Washington DC: American Psychological Association.

Willis, G. B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Thousand Oaks, CA: Sage.

Winter, J. C. F. d., Dodou, D., & Wieringa, P. A. (2009). Exploratory Factor Analysis with Small Sample Sizes. *Multivariate behavioral research, 44*(2), 147-181. doi: http://dx.doi.org/10.1080/00273170902794206

Wold, H. (1980). Model construction and evaluation when theoretical knowledge is scarce: Theory and application of PLS. In J. Kmenta & J. B. Ramsey (Eds.), *Evaluation of econometric models* (pp. 47-74). New York: Academic Press.