การประยุกต์ใช้การวิเคราะห์องค์ประกอบเชิงยืนยัน
พหูระดับความเพียงในการทำกำรบ้านวิชาสถิติสุริกิจ
ของนักศึกษามหาวิทยาลัยราชภัฏ 1

กันต์ฤทธิ์ คลังพล 2
ทรงกลม ไตรภิรมณ์ 3
ศิริญา กฤษณานวลี 4

บทคัดย่อ
การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อนำเสนอการวิเคราะห์องค์ประกอบเชิงยืนยันพหูระดับโดยศึกษาองค์ประกอบเชิงยืนยันพหูระดับความเพียงในการทำกำรบ้านวิชาสถิติสุริกิจของนักศึกษาที่ประกอบไปด้วยตัวแปรสังเกตได้ 3 ตัวแปร คือ การปฏิบัติตนกฎระเบียบห้องอาจารย์
กำหนดไว้ ความมุ่งหมายในการทำกำรบ้าน และการมีสมาธิในการทำกำรบ้านวิชาสถิติสุริกิจ
gลุ่มตัวอย่างเป็นนักศึกษาชั้นปีที่ 3 จำนวน 316 คน จำแนกตามมหาวิทยาลัยราชภัฏ
3 แห่ง จำนวน 10 แห่งเรียน ผลการวิเคราะห์องค์ประกอบเชิงยืนยันพหูระดับความเพียง
ในการทำกำรบ้าน พบว่า ไม่แสดงการวิเคราะห์องค์ประกอบเชิงยืนยันพหูระดับความเพียงในการทำ
การบ้านสอดคล้องกับค่าของประสิทธิคุณค่าช่วง (χ² = 0.254, df = 1, χ²/df = 0.254,
p-value = 0.614, CFI = 1.000, TLI = 1.017, RMSEA = 0.000, SRMR_w = 0.000,
SRMR_b = 0.019) โดยค่าดัชนีประสิทธิคุณเกณฑ์ระดับนักศึกษามีค่าอยู่ระหว่าง 0.347–0.607
และค่าดัชนีประสิทธิคุณเกณฑ์ระดับห้องเรียนมีค่าอยู่ระหว่าง 0.870–0.894

1 ได้รับทุนสนับสนุนจากหน่วยงานวิจัย หัวข้อ “หนึ่ง 70 ปี จุฬาลงกรณ์มหาวิทยาลัย” จากหน่วยงานวิจัย หัวข้อ “หนึ่ง 70 ปี จุฬาลงกรณ์มหาวิทยาลัย” 2 นิสิตปริญญามหาบัณฑิต สาขาพยากรณ์การวิจัยการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
3 อาจารย์ที่ปรัชญาทรัพยากรพื้นฐาน สาขาวิชาทรัพยากรและอุตุโยบายการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
4 อาจารย์ที่ปรัชญาทรัพยากรพื้นฐาน สาขาวิชาทรัพยากรและอุตุโยบายการศึกษา คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

393
Applying Multilevel Confirmatory Factor Analysis Technique to Study Homework Effort of Rajabhat University's Students

Kanreutai Klangphaho
Duangkarnol Traiwichitkhun
Sirichai Kanchanawasi

ABSTRACT

This research aimed to propose and validate the multilevel confirmatory factor analysis of students' homework effort with 3 indicators: students' compliance with the instructor's rules and regulations, the perseverance they put into their homework assignments, and their concentration on the homework. The subjects were 316 students drawn from 10 classes from three campuses of Rajabhat University. The proposed multilevel confirmatory factor model of homework effort fit well with the empirical data set ($\chi^2 = 0.254$, $df = 1$, $\chi^2/df = 0.254$, $p$-value = 0.614, $CFI = 1.000$, $TLI = 1.017$, $RMSEA = 0.000$, $SRMR_\mu = 0.000$, $SRMR_\gamma = 0.019$). The coefficient of determination of the student-level effects was 0.341-0.607 and the coefficient of determination of the classroom-level effects was 0.810-0.894.
Introduction

The attitude of students at all levels towards their homework assignments has been one of the most popular topics of discussion among teachers, parents, and educators (Simplicio, 2005; Marzano, 2007). The syntheses of research conducted by Cooper (1989), and Cooper, Robinson, & Patall (2006) showed that doing homework helped increase students’ learning, and enhance students’ self-discipline in managing the completion of their homework before the due date. However, the research studies on homework during the period of 1987 to 2006, regardless of their types, had design flaws. The researchers used homework as a learning and teaching tool, and only assigned the experimental groups homework, while the control groups were not given any homework assignments (Kohn, 2006; Cooper, Robinson, & Patall, 2006). Homework is complex because there are different groups of people, e.g. teachers, students, and parents, involved. Also, it serves a variety of purposes, e.g. achievement, improvement, self-regulation; engages tasks of different quality levels e.g. routine tasks versus complex tasks, and affects lesson organization, e.g. discussing, checking, and grading homework. Therefore, research studies on homework should incorporate new methodologies, such as multilevel modeling so that homework-related research studies will be put on the right track (Trautwein & Koller, 2003).

According to Trautwein et al. (2006a), researchers have provided some guidance of how to conduct research studies on homework by using the Multilevel Homework Model, which combines elements of expectancy-value theory (Eccles & Wigfield, 2002; Wigfield & Eccles, 2000), research on learning and instruction (Boekaerts, 1999; Brophy & Good, 1986; Weinert & Helmske, 1995), and self-determination theory (Deci & Ryan, 2002). Stable personal characteristics, namely basic cognitive abilities and conscientiousness, (Costa & McCrae, 1992) are also included. In many studies, researchers used homework as a basic example of problems between teachers and students that affected students’ achievement in their studies. Therefore, it is important for all studies to relate homework to students’ success in order to look at its effects at the classroom-level and the student-level (Trautwein et al., 2002; Trautwein & Koller, 2003; Trautwein et al., 2006a; Trautwein et al., 2006b; Trautwein & Ludtke, 2007; Trautwein, 2007; Trautwein & Ludtke, 2009).
The multilevel analysis can solve the technical problems of the conventional method in the areas of aggregation bias, misestimated standard error and heterogeneity of regression, but it does not give importance to the causal structural relationship between variables (Raudenbush & Bryk, 2002; Farmer, 2000). The Structural Equation Model (SEM), on the other hand, was created to show the relationship between latent variables, and between latent variables and observed variables (Diamantopoulos & Siguaw, 2000). However, its limitation lies in its lack of focus on the natural structure of hierarchical data (Muthén & Satorra, 1989; Muthén, 1994). The multilevel analysis and Structural Equation Model have been developed into the Multilevel Structural Equation Model that can analyze the relationship between hierarchical latent variables. This technique is then suitable for the analysis of homework-related variables that are multilevel and complex. This can solve the weaknesses of the traditional techniques.

In this research study, the researchers, then, proposed a Multilevel Confirmatory Factor Analysis Model of students’ homework effort in the business statistics course.

The Present Study

This study followed the framework of previous research by Trautwein et al. (2006). Based on past studies, this research aimed to develop and validate the Multilevel Confirmatory Factor Analysis Model of Homework Effort for a business statistics course. It was expected that the model would fit the empirical data set.

Methodology

Sample

The sample group comprised undergraduate students from the faculty of Business Management, Rajabhat University. The simple random sampling technique was used to select the sample group from the population. Three out of nine Rajabhat University’s central region campuses were chosen. They included 10 classes with the average classroom size of 31.6. There were more than 15 students enrolling in each class. The total number of the students who participated in the study was 316. This corresponded with the rule requiring that the number of the students in the sample group be larger than the number of the studied observed variables (Muthén, 1989).
Procedure

The researcher contacted the instructors of the statistics course at each of the three campuses by herself and collected the data by distributing a questionnaire to the students to complete. The time allowed to answer the questions on the questionnaire was limited to 20 minutes.

Instruments

The instrument used in this study was a 5-point Likert Scale questionnaire. It measured students' effort in completing their homework in the statistics course. There were 3 observed variables incorporated in the questionnaire: students' compliance with the instructors' rules and regulations, the perseverance they put into their homework assignments, and their concentration on the homework. Twenty-one items were created and modified based on the work of Trautwein et al. (2006a). The Item-Objective Congruency Index (IOC) was used to analyze the question items. Two items with an IOC value of less than 0.6 were eliminated from the questionnaire. The questionnaire with the remaining 19 items that gained an IOC value and the reliability coefficient value of more than 0.6 was tried out with a sample group of 100 students. Cronbach's Alpha was used to evaluate the reliability of the questionnaire. The reliability coefficient value was 0.87 (10 items). Considering each variable, the reliability values of the three variables: students' compliance with the rules and regulations (3 items), perseverance put into the homework (5 items), and concentrations (2 items) were 0.82, 0.79 and 0.61 respectively.

Statistical analyses

Analyzing multilevel confirmatory factor analysis procedures

Multilevel confirmatory factor analysis (MCFA), one of the multilevel SEM techniques, was originally devised to test the factor structure of the measurement used in a study by means of which participants could be categorized into different groups (e.g., Hox, 1998; Zimprich, Perren, & Hornung, 2005; Sun & Wilson, 2008). Multilevel confirmatory factor analysis model may be described as combining one
separate factor analysis model which accounts for the structure of observations on individuals within groups, and another factor analysis model which accounts for the structure of observed group means. Multilevel model thus implies a covariance structure model that is formulated in terms of a conventional factor analysis model on both "between-group" and "within-group" levels. (Muthén 1989, 1994)

Muthén (1994; 2007) developed the MCFA procedure. Figure 1 illustrates two-level confirmatory model with three observed indicators ($y_{1w}-y_{3w}$) depicted by squares. These indicators are the observed respondent ratings for the three items in a scale. The lower half of Figure 1, labeled "within", is consistent with a traditional confirmatory factor analysis on disaggregate data. As shown in this figure, the three observed variables load onto a single latent factor ($\eta_{w}$) at the "within" level. There are also three random errors ($\varepsilon_{1w}-\varepsilon_{3w}$) associated with each variables at this level. The upper half of Figure 1, labeled "between", shows three indicators represented by the circled $y_{iB}-y_{3B}$. They are not observed/raw data, but rather represent the group means aggregated for each observed indicator ($y_{1w}-y_{3w}$). The circled $y_{iB}-y_{3B}$ are represented of point (•) in "within level" that aggregated from each observed indicator ($y_{1w}-y_{3w}$). These group means load onto the aggregate latent variable ($\eta_{B}$) and they are associated with their respective random error terms ($\varepsilon_{1B}-\varepsilon_{3B}$). The full model connects the disaggregate and corresponding aggregate indicators. Thus, the observed values of the original indicators ($y_{1w}-y_{3w}$) are considered to be a function of both the within and between-level latent constructs ($\eta_{w}$ and $\eta_{B}$, respectively). The two-level confirmatory model consists of a simultaneous analysis of both of within and between-group covariance matrices. (Dyer et al., 2005).
In Figure 1, "the between and within components are explained by a single latent factor, however, this need not be the case. For example, one could test a model that proposes a single factor at the aggregate level and two factors at the disaggregate level, or many other similar non-isomorphic structures. If the hypothesized factor structure proposes more than one factor at a given level, the model may also include covariances among those same-level factors (by definition in this type of model, no covariances are allowed among factors at different levels). Similarly, the model may suggest that some indicators are valid at one level only, indicating a fuzzy composition model. Furthermore, the model may show some important covariates (e.g., age, pretest) that might be included in the model, relate to the focal latent construct at only one level. Estimation of these models yields both indicators of model fit, and parameter estimates of the factor loadings, factor variances, and uniquenesses (residuals). Thus, although our illustration presents only a very simple case, the MCFA technique in general promises some flexibility in the type of model that can be specified and tested. An advantage of the MCFA is that the individual- and class-level factor structures are calculated in one step by separating the total covariance into two parts—one between groups and one within groups (i.e. individuals)" (Dyer et al., 2005)
Six indices were used to assess the goodness of fit of the measurement model: the MCFA model to the empirical data. These indices included the Chi-square ($\chi^2$) index, the goodness-of-fit index (GFI), the nonnormed fit index (NNFI), the comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root mean residual (SRMR). The MCFA models were tested with Muthén's maximum likelihood (MUML), which includes robust standard errors and adjustment to the $\chi^2$ test statistic due to unbalanced group sizes. MUML procedure leads to correct model inference asymptotically when level-2 sample size goes to infinity and the coefficient of variation of the level-1 sample sizes goes to zero (Yuan H. K. & Hayashi K., 2005). The six above-mentioned fit indices were chosen for this study because no single fit index is considered to be the definitive marker of a model with “good” fit; each index serves a different purpose and should be interpreted in combination with the other indices. The $\chi^2$ index is an absolute index that tests for lack of fit resulting from overidentifying restrictions placed on a model. A nonsignificant p value (e.g., $p > 0.05$) is desired, but the $\chi^2$ index is usually inflated by the number of restrictions imposed on a model and sample size. Values of 1 for the GFI and the NNFI indicate perfect model fit; however, some researchers have suggested cutoff values greater than .95 to indicate good model fit. The following fit index cutoff values suggested by Hu and Bentler (1999) were used for determining goodness of fit: $CFI > 0.95$, $RMSEA < 0.06$, and $SRMR < 0.08$.

**Missing values**

We analyzed using a special feature of Mplus, having several options for the estimation of models with missing data. Mplus provides maximum likelihood estimation under MCAR (missing completely at random) and MAR (missing at random; Little & Rubin, 2002) for continuous, censored, binary, ordered categorical (ordinal), unordered categorical (nominal), counts, or combinations of these variable types. (Muthén & Muthén, 2007)
Descriptive Results

The data matrix featured 316 students from 10 classrooms. The correlation coefficient among the three observed variables at the student and the classroom levels (Table 1) revealed that there were a significant positive relationship among the variables at the student level ($r = 0.413$–$0.693$, $p < 0.01$). The highest correlation coefficient value was the relationship between the variable “students’ compliance with the rules and regulations” and “their perseverance in doing homework” ($r = 0.693$). The correlation coefficients between students’ concentration and the perseverance they put into doing homework, and between students’ concentration and their compliance with the rules and regulations were 0.456 and 0.413 respectively.

The correlation coefficient of the variables at the classroom level ranged between 0.781 and 0.880. The relationship between students’ compliance with the rules and regulations and their effort in doing homework had the highest value of 0.880. The correlation coefficients between students’ concentration and their compliance with the rules and regulations, and their concentration and the effort they put into their homework were 0.838 and 0.781 respectively.

Table 1 Within, and Between Correlation Matrices

<table>
<thead>
<tr>
<th></th>
<th>compliance</th>
<th>perseverance</th>
<th>concentration</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>within-sample matrix</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>compliance</td>
<td>1.000</td>
<td></td>
<td></td>
<td>0.7978</td>
</tr>
<tr>
<td>perseverance</td>
<td>0.693**</td>
<td>1.000</td>
<td></td>
<td>0.8314</td>
</tr>
<tr>
<td>concentration</td>
<td>0.413**</td>
<td>0.456**</td>
<td>1.000</td>
<td>0.7316</td>
</tr>
<tr>
<td>between-sample matrix</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>compliance</td>
<td>1.000</td>
<td></td>
<td></td>
<td>0.2908</td>
</tr>
<tr>
<td>perseverance</td>
<td>0.880**</td>
<td>1.000</td>
<td></td>
<td>0.2743</td>
</tr>
<tr>
<td>concentration</td>
<td>0.838**</td>
<td>0.781**</td>
<td>1.000</td>
<td>0.2605</td>
</tr>
<tr>
<td>ICC</td>
<td>0.108</td>
<td>0.087</td>
<td>0.078</td>
<td></td>
</tr>
</tbody>
</table>

Note. ICC = intraclass correlation, ***$p<.001$, **$p<.01$, *$p<.05$.  

401
The analysis of the elements of Multilevel Confirmatory Factor Analysis requires two-level variance. Intraclass correlation (ICC) is used to test whether the variables show variations across students within groups, or between groups or between classroom levels. If the ICC is more than 0.05, it means that there are high correlations among variables at least at two levels which is suitable to be tested by means of Multilevel Factor Analysis. However, if the ICC is less than 0.05, this means that there is no variance at the classroom level. It is, therefore, not necessary to analyze the data using Multilevel Factor Analysis (Snijders and Bosker, 1999). Based on Table 1, the ICC value of each observed variable ranged between 0.087 and 0.108, indicating that it was appropriate to use the Multilevel Factor Analysis with this set of data. The observed variables that had the highest variance were students' compliance with the rules and regulations set by teachers (ICC = 0.108), followed by the perseverance put into their homework (ICC = 0.087), and their concentration (ICC = 0.078).

**Multilevel confirmatory factor analysis's homework effort result**

The result of the multilevel confirmatory analysis's homework effort (see in Table 2 and Figure 2) showed that the model fit to the empirical data well and the construct validity was confirmed ($\chi^2 = 0.254$, df = 1, p-value = 0.164, CFI = 1.000, TLI = 1.017, RMSEA = 0.000, SRMRW = 0.000, SRMRB = 0.019 and $\chi^2$/df = 0.254). Moreover, the indices of the Comparative Fit Index (CFI) and the Tucker–Lewis Index (TLI) were approximately equal to 1.000, the root mean square error of approximation (RMSEA) was less than 0.06, and the standardized root mean residual (SRMR) was less than 0.08 (Hu & Bentler, 1999), and $\chi^2$/df was less than 2. They confirmed a very good fit of the model.
Table 2 Validation results to a two-level confirmatory factor analysis model of homework effort

<table>
<thead>
<tr>
<th>variable</th>
<th>Within groups: W</th>
<th>Between groups: B</th>
<th>intercepts (average group mean)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SE</td>
<td>z</td>
</tr>
<tr>
<td>compliance</td>
<td>0.652</td>
<td>0.051</td>
<td>12.682</td>
</tr>
<tr>
<td>perseverance</td>
<td>0.584</td>
<td>0.043</td>
<td>13.684</td>
</tr>
<tr>
<td>concentration</td>
<td>0.779</td>
<td>0.028</td>
<td>27.358</td>
</tr>
</tbody>
</table>

Note. Average cluster size (c) = 31.6, |z| > 2.56; p < .01, \( \chi^2 / df = 0.254 \), p-value = 0.614, CFI = 1.000, TLI = 1.017, RMSEA = 0.000, SRMRW = 0.000, SRMRR = 0.019

The intercepts of the average group means were between 3.469 to 3.973, indicating that at the classroom level, students' effort in doing homework as shown by each variable ranged from medium to high levels. The variable "perseverance" had the highest intercepts of 3.973, while "concentration" had the least intercepts of 3.469. Considering the factor loadings of each observed variables in the multilevel confirmatory factor homework effort model at the student level, it was found that the Standardized Factor Loading Coefficient (β) of the observed variables at the student level gained similar values of between 0.584 and 0.779 (p < .01). It implied that the three indicators of students' effort in doing homework, were equally important. The Standardized Factor Loading Coefficient (β) of the variable "concentration" had the highest value of 0.779, followed by students' compliance with the rules and regulations (β = 0.652) and their perseverance to do homework (β = 0.584).

Factor loadings of each observed variables in the multilevel confirmatory factor homework effort model at the classroom level revealed that the Standardized Factor Loading Coefficient (β) of the observed variables at the classroom level ranged within a narrow interval of 0.900 and 0.945 (p < .01). This implied that the three observed variables were equally important in measuring students' effort in doing homework. The Standardized Factor Loading Coefficient (β) of students' compliance
Applying Multilevel Confirmatory Factor Analysis Technique to Study Homework
Effort of Rajabhat University’s Students

to rules and regulations obtained the highest value of 0.945, followed by their perseverance to do homework ($\beta = 0.912$) and their concentration ($\beta = 0.900$).

Regarding the variance of students’ effort to do homework in the course, which was considered the latent variable in this study, the Coefficient of Determination ($R^2$) was between 0.341 and 0.607. Each of three the observed variables shared the variance of the latent variable at the student level at the percentage of 0.810–0.894. The Coefficient of Determination ($R^2$) value implied that the latent variable shared the variance with at the classroom level better than that at the student level.

In conclusion, the multilevel confirmatory factor homework effort Model 2 consisted of three observed variables: students’ compliance with the rules and regulations, their perseverance and their concentration, possessed structural validity at both the student and the classroom levels.

**Figure 2** Multilevel confirmatory factor homework effort model
Discussion

The study objective was to develop and validate the Multilevel Confirmatory Students’ Homework Effort Model with 3 indicators. The research study revealed that students’ effort in doing homework in the business statistic course had significant variance both at the student and the classroom level, and therefore, were to be analyzed by means of Multilevel Confirmatory Factor Analysis. This corresponded with students’ effort-related studies that also had variance of homework effort at both levels (Trautwein et al., 2006; Luedtke et al., 2007; Trautwein & Lüdtke, 2007; Trautwein & Lüdtke, 2009). Homework-related research studies should pay careful attention to the data with two-level variance to avoid incorrect research conclusions, since the variables related to homework are, by nature, multilevel and hierarchical nested data.

Considering the correlations between the three observed variables at the student and the classroom levels, the two variables that had the highest value of correlation at both levels were students’ compliance with instructors’ rules and regulations and their perseverance in doing homework. It means that the students who strictly complied with the rules and regulations set by their instructors tend to put great effort in doing homework. Therefore, instructors should have their students follow rules and regulations. This would also be beneficial for the students because what they learned from the business statistics course is considered basic knowledge of other courses.

Homework is a prime example of a research area in educational psychology that calls for a multilevel approach (Trautwein & Koller, 2003). Teachers typically assign the same tasks to all the students in a class, and expect them to complete these tasks outside classroom hours. It is therefore likely that multiple factors at different levels contribute to how much effort students put into their homework (Trautwein et al., 2006). Multilevel Confirmatory Factor Analysis can also be used to test whether the structure of any construct differ across levels of analysis. It was used to validate the Multilevel Students’ Homework Effort Model revealed that the model possessed structural validity or perfectly fit the empirical data. It was able to confirm that the variable “homework effort” could be used with the multilevel model, and the factor loadings of the student level were less than that of the classroom level. At the student
level, the observed variables that had the highest value were students’ perseverance, followed by their compliance with the rules and regulations, and their concentration respectively. Regarding the classroom level, the three observed variables obtained approximately the same values.

Regarding the ability to explain variance at the student level, the three variables (students’ compliance with the instructor’s rules and regulations, the perseverance they put into their homework assignments, and their concentration on the homework) shared the variance of the latent variable of homework effort at the percentage of 34.1–60.7 whereas the number ranged between 81.0 and 89.4 at the classroom level. This meant that at the student level, the observed variable that could best share the variance of the latent variable was students’ perseverance, their compliance with the rules and regulations, and their concentration respectively. At the classroom level, there was no significant difference in the ability of the three observed variables to explain the latent variable. In addition, the observed variables could explain the latent variable “homework effort” better at the classroom level than at the student level.

Homework researchers have long been aware of the need to differentiate between the classroom and student levels (Cooper, 1989; Corno, 1980). Because of homework is complex because there are different groups of people, e.g. teachers, students, and parents, involved. Also, it serves a variety of purposes, e.g. achievement, improvement, self-regulation, engages tasks of different quality levels e.g. routine tasks versus complex tasks, and affects lesson organization, e.g. discussing, checking, and grading homework. (Trautwein & Koller, 2003)

In conclusion, the validation of the multilevel confirmatory factor homework effort model confirmed that the model that incorporated the three observed variables had structural validity and could be analyzed at the two levels. An addition limitation of study is number of sample group at classroom level was 10 classes that acceptance criteria. (Johnsrud & Rosser, 2002) although the number of groups recommended was about 20 to 100 (Hox & Mass, 2001; Hox & Krefl, 1994; Hox, 1993). However, the researchers increased the number of participating students in each class to lessen the impact of errors on the estimation of parameter and standard error. The future research should deal with at least 20 groups of students.
References


