Information Security Awareness Measurement with Confirmatory Factor Analysis

Puspita Kencana Sari, Candiwan, Nurvita Trianasari
Faculty of Economic & Business, Telkom University
Jl. Telekomunikasi 1 Bandung, Indonesia

Abstract -- **One of information security management elements is information security awareness program. Usually, this programs only involve the employees within the organization. Some organizations also consider security awareness for some parties outside the organization like providers, vendors, and contractors. This paper add consumers as variable to be considered in information security awareness program as there are also some threats for organization through them. Information security awareness will be measured from user's knowledge and behavior of five information security focus areas in telecommunication, especially related with smartphone users as one segment of telecommunication provider. In other researches, information security measurement from outside an organization is focused in Internet use by end-user. In telecommunication industry, information security threats for consumers not only from Internet, but also by phone call or texting. This research used CFA for data analysis method. The result showed that the indicators of knowledge dimension is not good enough to measure dimensions of knowledge where most of them are not significant. Meanwhile, behavior dimension gave different result with high significant value for measuring security awareness level. Indicators of behaviour dimension is good for measuring dimension of behaviour because only one indicator that isn't significant.**

Keywords: information security; awareness; measurement; confirmatory factor analysis

I. INTRODUCTION

Basic goal of information security is to ensure business continuity and to minimize business damage by preventing and minimizing the impact of security incidents [7]. When an organization apply information security management, it should maintain three basic components [10]. First, confidentiality of sensitive information by protecting it from unauthorised disclosure or intelligible interception. Second, integrity, by safeguarding the accuracy and completeness of information. And the third, availability, by ensuring that information and vital services are available to authorised users when required.

Information security management is derived from potential threats of each organization. Those threats are identified from circumstances or activites that can cause loss or harm for organization, such as financial loss, absence of data or resources, or even loss of company credibility[10]. One most important part of information security management is information security awareness programs. The programs are to ensure that all employees obey the information security policies and procedures that has been established by the organization. Kruger and Kearney said that "*the initial aim or objective of Information security awareness was to ensure that computer users are aware of the risks associated with using information technology as well as understanding and abiding by the policies and procedures that are in place*"[6].

As Symantec reported that Telecommunication sector is In the second rank (10%) after retail (27%) that has risk in data breaches that could lead to identity theft (Top 10 sectors by number of Identities Exposed)[12]. It also reported that Indonesia is in the eight rank of countries with highest cost per capita of a data breach. Indonesia Computer Emergency Response Team (ID-CERT) surveyed, whose some of the respondents are from telecommunication provider, that 53,1% of incident reported from March to April, 2013 is about Network Incident; 15,4% is Intellectual Property Rights; 12,1% is malware incident; and 11,4% is spam[4]. In 2012, number of network incident reported reach 76,53%. Therefore, Internet service provider, including telecommunication provider, should increase their preventive actions to reduce this incident[4].

Internet development drives the increament of its users. APJII released Indonesia Internet Profile in December 2012 that said 65,70% internet user in Indonesia use Smartphone as their devices. Number of smartphone users in Indonesia is 23,8 million people in 2012 and predicted to reach 71,6 million people in 2015. It's also driven by the cheapening of gadgets and services provided by telecommunication providers. But unfortunately, use of mobile technology also increase the threats of information security. In 2010, Yayasan Layanan Konsumen Indonesia (YLKI) recorded that 17,1% from 590 consumen complaint is about telecommunication service, where it is the first rank in

978-1-4799-3704-2/14/$31.00 ©2014 IEEE

that period. About 46,7% from that complaints is about stealing pulse from the customers. This incident indirectly can threaten the credibility of telecommunication provider, which is one aspect of information security management.

Although end users have adapted the technology, they often have a lack of awareness towards the right practice or they possess knowledge but they often do not practice it in proper ways[1]. Mobile users often save their personal and financial information in their phone. It makes them execellent become malware and phishing targets. In November 2010, a virus spread to a million mobile phones in China, the virus was sold to mobile users as an anti-virus application, but in fact turned the mobile phones into zombies and began sending spam SMS to people on phone book[8]. According to Symantec security report that top-three of mobile threats in 2012 are 32% steal information, 25% traditional threats, and 13% send content. Steal information including steal device data, banking trojan, Ddos Utility, Hacktool. Traditional threats including downloader, backdoor. Send content including sends premium SMS and spam[12]. In this smartphone era, there are new threats developed such as vishing attack and smishing attack. Vishing attack is phishing by verbal message, while smishing attack exploit SMS message, compromised text message can contain email and website addresses that can lead the innocent user to malware site[8].

In many literatures, the objects of information security awareness program focus on employees within the organization. In other security standard, such as BMIS from ISACA, define people element of information security management consist of employees, contractors, vendors, and service providers[10]. Meanwhile, they also define that primary people within BMIS are those who are employed or otherwise associated with the organisation[5]. In additioan that in ISO27001 stated that people who work under organisation's control should aware of information security and also all employees of the organization and, where relevant, contractors shall receive appropiate awareness education and training and regular updates in organizational policies and procedures, as relevant for their job function [2].

In this paper, we include cunsumers as people element in information security management. As Peltier said that "*system owners has responsibility to share appropriate knowledge about the existence and general extent of control measures so that other users can be confident that the system is adequately secure*"[9]. Consumer of some organization is also given access to communication network, that means they can access some organization information. Beside that as statement in BS ISO 27001, that detection, prevention and recovery controls to protect against malware shall be implemented, combined with appropiate user awareness [2]. Futhermore that around 40 % of social network user are attacked by malware and in December 2010, one of the first android botnets (called gemini) was discovered and the code was wrapped inside a legistimate android application whose developers did not realize they were spreading malware and this happens again in March 2011 that google discovered a botnet called "droiddream". As Al-Sehri said that "It is essential to keep the public aware of the security threats and educate them towards using good practices in order to get greater security"[1]. Therefore, in this paper we proposed a measurement of information security awareness from consumers of telecommunication providers, especially smartphone users. By knowing the level of awareness from consumers, organization can established appropriate security policies and procedures to get greater security.

II. RESEARCH FRAMEWORK

The research framework used in this research is adapted from Kruger & Kerney Model. The tool was based on social psychology theory that proposes three components to measures a favourable or unfavourable manner to a particular object. That components are cognition, affect, and behaviour[7]. Those components are used to develop three equivalen dimensions namely knowledge (what does a person know), attitude (how do they feel about the topic), and behaviour (what do they do)[7]. But, in this research we only use two of them; knowledge and behavior. Beside measuring the awareness level, we also Each one of these dimensions was then subdivided into the five focus areas; adhere to security policies, protect personal data, fraud/spam SMS, mobile applications, and report for security incident. This is the framework adopted from Kruger & Kerney's model.

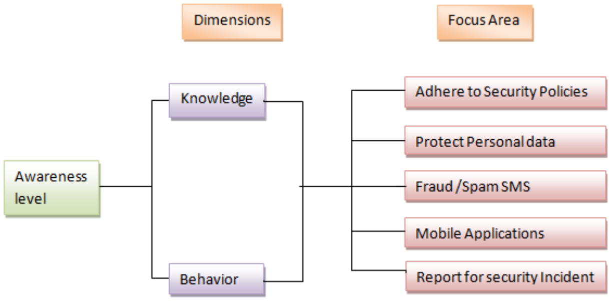

Figure 1. Information Security Awareness measurement framework

These five focus areas were defined from theories, facts and phenomonen about information security in Indonesia related with telecommunication sector. Beside that, some area were defined by an information security expert in telecommunication provider (ISO 27000 auditor). Two problems stated by the expert are adhere to security policies and report for security incident. The first point of awareness in ISO 27001:2013 states that '*persons doing work under the organization's control shall be aware of information security policy*' [2]. Therefore, security policy, as the base of information security management, should be discussed as one the focus area. The next focus area is protect personal data. As we state in the introduction that nowdays people save many information in their smartphone, including personal and confidential data. They use smartphone not only for texting and making phone calls, but also for doing business and other productivity tasks. Accordingly, we put area of protecting personal data as one thing that we should consider in this research. The threats of premium SMS or spaming and mobile applications are based on Symantec Security Reports 2013, ID-CERT and also YLKI complaint report (as stated in introduction). Symantec mentioned that one of top-three of mobile threats is premium SMS or spaming (send content). While YLKI reported that in 2010, the most complaint is about premium SMS. Other mobile threats refered to Symantec Report are traditional threats; such as bakcdoor, malicious code, and so on, that can be caused by mobile application installation in the smartphone. Although some mobile operating systems now have been implementing the sandbox security mechanism that could separate/isolate each programs, such as iOS and Android 4. But, in this research we consider that those kinds of smartphone are not the majority of smartphone used in Indonesia.

III. RESEARCH METHOD

This research used quantitave method where data was gathered using questionnaires. Twenty-two questions were designed to test the knowledge and behaviour of respondents concerning the five main focus areas. Each focus area in each dimensions has two questions, except protect personal data area. Some of the questions were answered on a 3-point scale – true, don't know and false (knowledge dimensions), while others only needed a true or false response (behaviour dimensions). The questionnaire was distributed by online.

Population of this research are people who use smartphone and telecommunication service from Indonesian telecommunication providers. To define sample, this research uses nonprobability sampling with purposive sample technique. For data analysis we use measurement model called Confirmatory Factor Analysis (CFA). It will give a modelling of relationship between latent variabel and observed variabel, where observed variable is reflection of latent variabel (refelective). One of type of CFA model is First Order Confirmatory Analysis model, that is a measurement model where the latent variable is measured only by indicators contained in that variable. In measurement theory, a dimension is a group if same indicators. According to Diamantopolous and Siguaw in Bachrudin (2008), a dimension can be considered as a latent variable[13]. Figure 1 below describe the model that will be used in this research where using two constructs or dimensions as latent variables.

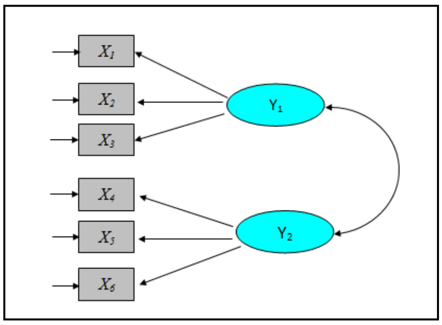

Figure 2. First order Confirmatory Factor Analysis Model with two-Constructs

The variables in this research consist of two dimensions, i.e. knowledge (What do they know about the topic?) and behavior (What do they do?). Each dimensions has five focus areas; (i) adhere to security policies, (ii) protect personal data, (iii) fraud/spam SMS, (iv) mobile applications, and (v) report for security incident. Every focus area has indicators, for instance in protect personal data, the indicators are using password in smartphone and log out from their account after finishing**.** To test the validity of every item in the questionnaire, we used *pearson product moment* correlation where every item that has ng coefficient equal or more than 0,3 is valid. For reliability, testing is used *Alpha Cronbach* method, where the coefficient should be equal or more than 0,5.

Structural Equation Modelling (SEM) is used for evaluating consistency of a theory with empirical data or a statistical technique where the processing simultanously involves errors in measurement, indicator variable or latent variable. While other multivariat techniques are able to test only on relationship. In SEM approach, measurement error in latent variable is taken into account. Latent variable is a variable that can't be measured directly or rated its presence degree. But, manifestation of a latent variable can be observed by recording or measuring various characteristics of the behavior of individuals in certain circumstances, for example through questionnaires.

If the latent variables have been assessed, then the SEM is used to test the hypothesis that indicate the relationship between the latent variables and the relationship between latent variables to the indicators used to measure. The both types of relationship are formulated in two modeling, namely the structural model and the measurement model.

In the structural model described the relationship between latent variables, while the measurement model described the relationship between the latent variable and its indicators. Suppose that there are m endogenous latent variables $[\boldsymbol{\eta}]$ and n exogenous variables with $[\boldsymbol{\xi}]$ structural model below[14]:

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \qquad \text{............ (1)}$$

Confrimatory Factor Model

By using LISREL notation, confirmatory factor model is formulated below[16]:

$$\mathbf{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon} \qquad \text{.............. (2)}$$

$\mathbf{y}$: observed indicators of $\boldsymbol{\eta}$ $(n \times 1)$

$\boldsymbol{\Lambda}_y$: coefficient relating y to $\boldsymbol{\eta}$ $(n \times q)$

$\boldsymbol{\eta}$: vector of latent variables or constructs sized $(p \times 1)$

$\boldsymbol{\varepsilon}$: measurement error for y $(n \times 1)$

In this model, assumed that $\boldsymbol{\eta}$ is not correlated with $\boldsymbol{\varepsilon}$. This model is identical to the model proposed by Bollen (1989) and serve as the basis for analysis of the validity and reliability of measuring instruments.

Generally, SEM procedure consits of 5 steps below:

A. Model Spesification

In this step, we define path diagram that is a combination of measurement and structural model.

B. Model identification

Accordingt to Hair et.a; (1989), the results of model identification can be devided based on their degree of freedom, such as:

- df = 0, the model called *justidentified*
- df > 0, the model called *overidentified*
- df < 0, the model called *underidentified*

In SEM, it is suggested to have overidentified model and avoid underidentified model. Justidentified model (df=0) is can't be tested. While, overidentified model (df>0) is can be tested with various statistic test.

C. Estimation

Model identification phase is a must before estimation proses, because without that phase the result of estimation is meaningles[14][15].

D. Confirmity Test

Some of confirmity test for realibility of this model are Chi-Square χ^2 ,GFI,RMSEA, RMR, etc[16].

E. Respesification

There are 3 modelling strategy for this phase[16]:

a. Confirmatory modeling strategy.

b. Competing models strategy.

c. Model development strategy.

IV. RESULTS AND ANALYSIS

The survey was done for around three weeks, from the 23th December 2013 through 13th Januari 2014. The total number of respondents was 106 users from seeveral cities in Indonesia; Bandung (64%), Jakarta (17%), Surabaya (6%), Palembang (3%) and other cities (10%). Females who use smartphone in this survey are 43 % and males are 57 %. Based on age range, majority of respondents (57%) are from the age group of 20 – 30 years of age then followed by the age group of under 20 years old (18%), 41-50 years old (11%), 31-40 (8%) and over 50 years old (6%).

Regarding the usage of smartphone by the repondents, most of respondent use their smartphone for browsing (80 %), social media (79 %), sms (75%) and email (62 %). However only a few users user their smartpohe for phone call (55%), playing games (44 %) and ohers (5 %). Others include navigation, notes for lecture, e-banking, and productivity applications. This usage is suitable with trend that the use of internet or data is increasing and the use of phone call is decreasing.

Concerning information security breach experience based on the survey, most of respondents have experience around 82 % and they who have no security experience is about 18 %. The detail of this number about security breach experience are as follow; fraud SMS (71%), Spam SMS (53%), fraud call (17%), virus (13%) and others (8%).

The analysis procedure is conducted with significant rate (α) = 0,05 by using LISREL 8.72 as statistical software tool. Model identification proses is aimed to test whether the proposed model resulting a unique estimation or not by using the formula below.

$$df = \frac{p(p+1)}{2} - t = \frac{22(22+1)}{2} - 46 = 207 \qquad \text{............(3)}$$

From that computation, we get $df > 0$ (*over-identified*). Therefore we can conclude that the model generates a unique estimation.

Result of Confirmatory Factor Analysis Model

Figure 3 below describe CFA path resulted by the modelling process.

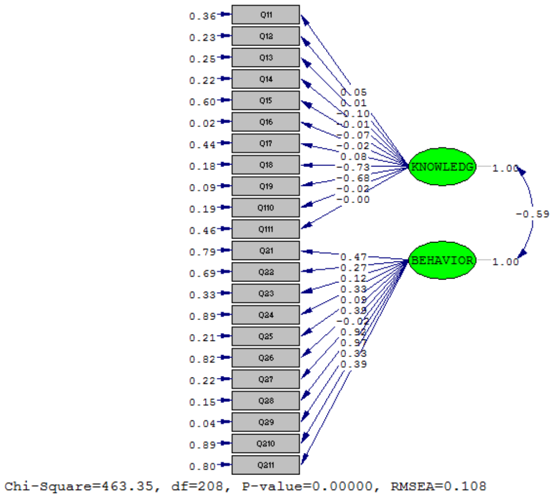

Figure 3. Path Diagram Of LISREL's Output With Estimated Parameter (Unstandardized)

Source : Data processing result with LISREL 8.72

Figure 3 shows standardized loading factor and measurement error for first order CFA model. While Figure 4 below shows that all standardized loading factor on the first level of the measurement model (first-order CFA) for the knowledge dimension is less good, but for the behaviour dimension is good enough for just one item is insignificant. This statement is based on validity criteria where t value loading factor is bigger than critical value $\left(t_{hitung} \geq 1.96\right)$ and the standardized loading factor ≥ 0.50 . Therefore, we can get conclusion that observed variable in first order CFA model can measure its constructs; good enough for Behavior dimension but less for Knowledge dimension.

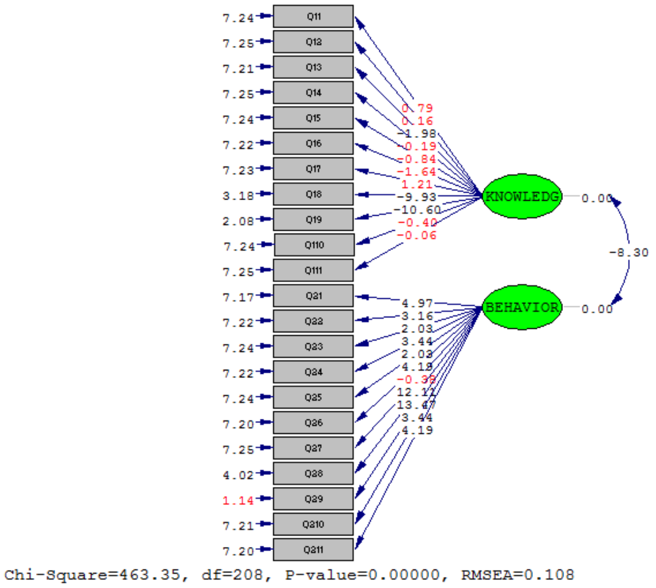

Figure 4. Path Diagram Of LISREL's Output With T-Values Of Statistics Test

Source : Data processing result with LISREL 8.72

Overall Model Analysis

Tabel 1 below describe output of data processing with LISREL 8.72 to evaluate fitness of overall model by inferential and descriptive.

Tabel 1. Confirmity Values Of The Overall Model

Level	Goodness of Fit Statistics	Value	Criteria	Result
Absolute Inferensia	Chi-Square	463,35	$< \chi^2_{tabel}$	Not
	p-value	0.000	> 0.05	Not
Absolute Descriptif	RMSEA	0.108	< 0.05	Not
	RMR	0.045	< 0.05	Fulfilled
	GFI	0.71	> 0.90	Not
	AGFI	0.65	> 0.90	Not

Source : Data processing with LISREL 8.72

Based on confirmity value of overall model on Tabel 1, statistically inferential model is unconfirm. It's showed by chi-square value = 463,35 and p-value = 0.000 that not fulfill the significant level of model acceptance (model should fit with data), where *p-value* ≥ 0.05. But, as stated by Bollen & Long in Wijanto (2008) that model confirmity test is not only depended on chi-square test, but also can use other statistic test[17].

The result of fitness test with absolute descriptive level shows that model is unfit. It's represented by model fitness index value RMSA, GFI, and AGFI are not fulfill the fitness criteria. But, RMR is fulfilled and indicates that the model is fit. Therefore, we conclude that fitness of overall model is not good enough.

Model analysis of Knowledge Dimension

Measurement model (first order CFA) is interpreted as measurement model between endogen latent variabel Knowledge with its indicators (focus area). The estimation result of standardized loading factor parameter (weight value) for Knowledge measurement model from its indicators is showed by Tabel 2 below.

Tabel 2 *Standardized Loading* Indicator Value of *KnowledgesDimension*

Dimension	Item	*Standardized Loading value*
KNOWLEDGE (η_1)	Q11	0.05
	Q12	0.01
	Q13	-0.10
	Q14	-0.01
	Q15	-0.07
	Q16	-0.02
	Q17	0.08
	Q18	-0.73
	Q19	-0.68
	Q110	-0.02
	Q111	-0.00

Source : Data Processing with LISREL 8.72

We can see that Q18 item has the highest value (-0.73). It means that Q18 item has the biggest contribution in measuring Knowledge dimension. Q18 item is discussed about security side in installing and using mobile application. While in the other side, the smallest contribution is given by Q11 item that has value = 0.00. Q11 is an indicator of adhere to security policy.

Model Analysis of Behavior Dimension

First order CFA model is interpreted as measurement model between endogen latent variabel Behavior with its indicators (focus area). Estimation result of standardized loading factor parameter (weight value) for Behavior model can be seen in Tabel 3 below.

Tabel 3. *Standardized Loading* Indicator Value of *Behavior Dimension*

Dimension	Item	*Standardized Loading value*
BEHAVIOR (η_2)	Q21	0.47
	Q22	0.27
	Q23	0.12
	Q24	0.33
	Q25	0.09
	Q26	0.39
	Q27	-0.02
	Q28	0.92
	Q29	0.97
	Q210	0.33
	Q211	0.39

Source : Data Processing Result with LISREL 8.50

As we can see form Tabel 3, that Q29 item has the highest value (0.97). It indicates that Q29 item has the biggest contribution in measuring Behavior dimension. Q29 item is one of mobile application indicators. While the smallest contribution is given by Q27 item which has wight value = 0.02. Q27 item is discussed about threat from premium/spam SMS.

Whereas, correlation between Knowledge and Behavior dimension is 0.59. It indicates that there is a negatif correlation with moderate strength based on theory of Guilford's Emperical Rule[18]

Tabel 4. ***Guilford Correlation Criteria***

Coefficient Interval	Relationship Degree
0.00 - 0.19	Very Low
0.20 - 0.39	Low
0.40 - 0.59	Moderate
0.60 - 0.79	Strong
0.80 - 1.00	Very Strong

V. CONCLUSION

Based on analysis result, we conclude that measurement of Knowledge dimension is not good enough to indicate security awareness level of smartphone users. It can be seen where three from eleven items/indicators have unsignificant value. Different with the Behavior dimensions, only one from eleven items/indicator that has unsignificant value.

The results of this study support our previous research. Our previous research is about measuring awareness level where the total awareness is about 80% (Good). The Knowledge dimension has level in 86% (Good) and Behavior dimension in 73% (average). These results indicate that many events of information security breaches experienced by smartphone users because their behavior that is less concerned about the security of information. Meanwhile, a high level of knowledge dimensions in previous studies may be due to the majority of the indicators that have unsignificant value.

Therefore, for the next research, we suggest to redefine the question items for each security focus areas (indicator), especially in Knowledge dimension. Another option is to redefine the indicators or focus area of both dimensions.

REFERENCES

[1] Al-Sehri, Yasser. Information Security Awareness and Culture. *British Journal of Arts and Social Sciences*. 2012; Vol.6 (No.1): 61-69.

[2] British Standard Institution. *ISO/IEC 27001:2013 Information Tecnology-Security Techniques-Information Security Management Systems-Requirements***.** Switzerland. BSI Standard Limited. 2013

[3] IDCERT. *ID-CERT Annual Report 2012*. Indonesia Computer Emergency Response Team. 2012

[4] IDCERT. *Laporan Dwi Bulan II 2013*. Indonesia Computer Emergency Response Team. 2013

[5] ISACA. *Business Model for Information Security*. USA. 2010.

[6] Kruger, H.A, Kearney, W.D. A Protoype for Assessing Information Security Awareness. *Elsevier Journal: Computers & Security*. 2006; 25 page 289-296.

[7] Kruger, Hennie., et al.,. *A vocabulary Test to Assess Information Security Awareness*. South African Information Security Multi-conference in Port Elizabeth, South Africa. 2010.

[8] Laudon, KC, Traver CG. E-Commerce 2012: Business, Technology, Society. England: Pearson Education Limited. 2012.

[9] Peltier, Thomas R. Information Security Fundamentals, Second Edition. Boca Raton: CRC Press. 2014.

[10] Sari, P.K. *A Concept of Information Security Management for Higher Education*. International Conference on Technology and Operation Management, 3rd. Bandung. 2012: 469-477

[11] Sugiyono. *Statistik Untuk Penelitian*. Bandung: Alfa Beta. 2009.

[12] Symantec. *Information Security Threat Reports*. Symantec Corporation. Volume 18. 2013.

[13] Bachrudin, Achmad & Tobing, Harapan L. Analisis Data untuk Penelitian Survey dengan menggunakan LISREL 8. Jurusan Statistika FMIPA-UNPAD. Bandung. 2003

[14] Bollen, K.A. *Structural Equation with Latent Variables.* New York : Wiley. 1989

[15] Long, J. S. *Confirmatory Factor Analysis: A Preface to Lisrel,* Sage Publications. Ltd, London. 1983

[16] Joreskog, KG & Sorbom.D. LISREL 8. *Structural Equation Modelling with the SIMPLIS command Language,* Hillsdale;Erlbaum. 1993

[17] Wijanto, Setyo H. *Structural Equation Modeling dengan Lisrel 8.8 : Konsep dan Tutorial*. Graha Ilmu. 2008

[18] Guilford ,J.P., *Psychometric Methods*, Tata McGraw-Hill Publishing Company Limited. 1979