

Evaluation Model for Islamic Education Learning in Junior High School and Its Significance to Students' Behaviours

Badrun Kartowagiran^{1,*}, Harsul Maddini²

¹Graduate School, Yogyakarta State University, Yogyakarta, Indonesia

²Palu State Islamic Institute, Palu, Indonesia

*Corresponding author: badrunkw@yahoo.com

Received June 13, 2015; Revised July 07, 2015; Accepted July 15, 2015

Abstract This study was aimed at: (1) developing an evaluation model for the learning of Islam religious education (IRE) in the Junior High School (JHS) and its significance to students' behaviours, and (2) testing the feasibility of the model by looking at the compatibility and significance of the model among the research variables. The study was developmental research using the mixed approach (quantitative and qualitative) of the embedded type in three phases: pre-development, development, and dissemination. The research model consisted of (1) the evaluation procedure and (2) the evaluation instruments in the forms of (a) analysis sheet for the lesson plan, (b) observation sheet, and (c) questionnaire. Validity for the lesson plan analysis and observation instruments was obtained through expert judgment for content validity. The reliability measure of the lesson plan analysis was estimated by the Cronbach's Alpha, while that of the observation sheet by the Genova. The validity evidence for the questionnaire was construct validity using the exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) while its reliability was estimated by the Cronbach's Alpha. The research results showed that (1) the evaluation implementation procedure was systematic; (2) all the research instruments possessed a good measure of validity and reliability; and (3) the model was fit and the influence of each independent variable (Environments of Religiosity, Process, and Result) upon the dependent variable (Impact) was significant. The research result was also completed with a description of the developed model.

Keywords: ERPRI model, evaluation, validity, reliability

Cite This Article: Badrun Kartowagiran, and Harsul Maddini, "Evaluation Model for Islamic Education Learning in Junior High School and Its Significance to Students' Behaviours." *American Journal of Educational Research*, vol. 3, no. 8 (2015): 990-995. doi: 10.12691/education-3-8-7.

1. Introduction

The discussion of evaluation of Islam religious education (IRE) learning in JHS is related to how much the instructional material is eventually internalized by the students which, at a later turn, is expressed in the students' daily behaviours. Thus far, learning outcomes have been seen as scores obtained from tests or examinations. The affective and psycho-motoric learning aspects have not been popular among JHS teachers, including Islamic education teachers. Evaluation of Islamic education learning outcome should be comprehensive covering the aspects of knowledge, skills, and attitudes. A need thus arises to develop an IRE evaluation model that is comprehensive.

Islamic education has a basic frame of thoughts that becomes a reference for all educational practices in all aspects of human life. The basic thought frame is constructed of the three basic concepts of faith, devotion, and good deeds. These basic concepts form the core of Islamic teachings that must be adopted as the base

structure of every system [14]. These three basic teachings have been developed by clerics into three concept studies of faith related to matters of belief in God, laws related to matters of behaviours, and good deeds related to goods behaviours [10].

The three basic concepts are concerned with synergic and harmonious relations to produce compatibility, harmony, and balance in the four relation kinds. The first is the relation between man and his Creator in terms of how far humans have fulfilled their responsibilities to the Creator. The second is the relation among humans in which no one is to act in such a way that it will hurt others. The third is the relation between humans and animals and environment in which no one is to impact destruction. The fourth is the relation between man and himself, in which humans must behave themselves.

These four relation kinds are contained in the Islamic education curriculum organized in the subject-matter components of 1) faith, 2) laws, 3) good deeds, 4) Alquran and Alhadist, and 5) Islam cultural history. These subject-matter components become the curriculum scope of IRE as a broad-field or one-system curriculum. This is different from the curriculum scopes of many Muslim

boarding schools in which subject-matters are organized in separated-subject curricula, such as faith, phonology, laws, Alquran, Alhadist, history, etc. [8].

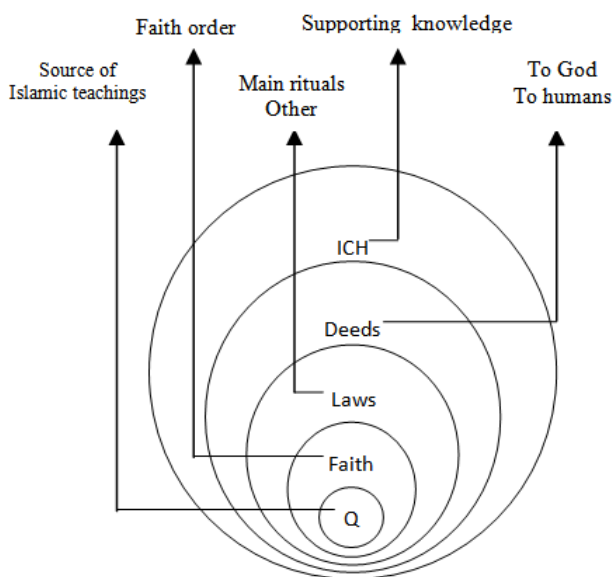


Figure 1. Relations among the five aspects of IRE subject matter

Looking at the conceptions of Islamic education learning, especially in terms of the instructional components and materials, it can be shown that the instructional materials are grouped into two types: main component and supporting component. The main component consists of four aspects: Alquran and Alhadist, faith, law, and deeds; the supporting component consists of one aspect: Islam cultural history. These aspects are inter-related and inter-supporting. Inter-relations among these five aspects can be seen in Figure 1.

The outcome of IRE must be given careful evaluation. Matters that need to be treated are the learning processes, the evaluation conduct, and the application of the learning outcome both in the classroom and outside. Three aspects can be derived from these matters; namely faith (rooting of belief), deeds (internalization of faiths), and worship (application of faiths and deeds). These three aspects will guide the students in the religious practices of their daily life.

Based on the foci of the study as described above, the evaluation model in this study is phased in four steps. These are (1) evaluating the environment of religiosity, (2) evaluating the instructional processes, (3) analyzing the official documents, and (4) looking at the internalization and implementation of IRE values in the students' life in the society.

2. Research Method

This study was developmental research of the mixed approach (quantitative and qualitative) using the embedded type realized in three steps: preliminary investigation, development, and dissemination. The pre-developmental phase consisted of gathering information from the field and conducting reviews of theories. Gathering of information in the field was conducted by observation and interviews with students and teachers of a number of junior high schools in Palu City. An

identification of relevant previous studies and a review of various concepts and theories related to instructional evaluation were given.

The developmental phase was the construction of a model prototype based on information obtained during the pre-developmental phase. The prototype model consisted of a procedure, instrument, and technique to find the effectiveness of the model. The procedure consisted of steps to evaluate the components of the evaluation model. This model has four component, namely: environment of religiosity (ER), process (P), result (R), dan Impact (I), to become the model acronym: ERPRI. The ER had three indicators: manners, software, and hardware; the P had three indicators: planning, conduct, and evaluation; the R had three indicators: psychomotor, affective, and cognitive; and the I had three indicators: faith, attitudes, and behaviours. The indicators affective skill, instructional conduct, and behaviours were elicited by the questionnaire and observation; the indicator lesson planning by the analysis sheet.

Ideally, to uncover religious behaviors, an experience sampling method (ESM) which is a method using experienced persons as samples. ESM is often referred to ambulatory self-report in which the subjects give self-reports along the way or an intensive longitudinal design in which experiences and behaviours are studied in contexts and repeated over times [16].

The ESM is a good research method but is costly and takes a long time. It is on this reason that the ERPRI model developmental study used the cross-sectional design in uncovering the religious behaviours of students. This is in line with the conclusions in the article by McAndrew, M. and Voas, D. [11] who state that cross-sectional survey methods are still capable of diving data about religiosity though in a limited manner.

The dissemination phase was the validation phase of the model prototype that proceeded from focus group discussions (FGD) among experts, to a limited try-out, and then a wider try-out. The FGD participants were experts in instrument development and IRE subject matters. They were asked to give direct inputs by in the document draft (procedure draft and technical plan to validate the model feasibility), and to assess appropriateness by giving scores on the items of those instruments. The written inputs were used to revise the draft, while the scores were used to estimate the validity and reliability of the three instruments: observation sheet, analysis sheet, and questionnaire.

Validity checks for the lesson plan analysis sheet and observation sheet were done by content validity using expert judgment, while the reliability measure of the lesson plan analysis sheet was estimated by a Cronbach's Alpha, the reliability for the observation sheet by GENOVA software program. For the questionnaire, the validity evidence used construct validity by exploratory factor analysis (EFA) on the SPSS software and confirmatory factor analysis (CFA) by LISREL, while its reliability measure was estimated by a Cronbach's Alpha using the SPSS software program. The technique for determining model feasibility in the forms of relation paradigms among the research variables (model components) was the structural equation modeling (SEM) on the LISREL.

3. Research Findings and Development

Validity checks for the lesson-plan analysis sheet and observation sheet were provided by content validity using expert judgment, the reliability measure for the lesson-

plan analysis sheet was estimated by a Cronbach's Alpha, and the reliability for the observation sheet by generalizability theory on the GENOVA. Estimation results for the reliability of the observation sheet can be seen in Table 1.

Table 1. Inter-rater Reliability Indexes of Behaviour Sub-sub-components

No	Sub-sub-componemnts	Reliability Index					
		Generalizability			Phy		
		Stud1	Stud2	Std 3	Std 1	Std 2	Std 3
1.	Behaviour 1	0.81	0.94	0.97	0.81	0.94	0.97
2.	Behaviour 2	0.48	0.79	0.89	0.47	0.77	0.88
3.	Behaviour 3	0.41	0.73	0.85	0.40	0.72	0.84
4.	Behaviour 4	0.51	0.81	0.90	0.51	0.80	0.89
5.	Behaviour 5	0.50	0.78	0.88	0.50	0.78	0.88

These reliability estimates by the generalizability theory were inter-rater reliabilities with two criteria to determine whether or not the instruments were reliable. These were when the generalizability co-efficient and phy co-efficient reached an index of ≥ 0.7 [5].

After being revised, the ERPRI prototype was tried-out to a limited number of 30 JHS students, 4 teachers, and 2 principals. This limited try-out produced the readability measures of the instruments (observation sheet, analysis sheet, and questionnaire). The readability level was used

as a consideration to improve the procedure and instrument drafts, while the responses to the questionnaire were used to provide evidence for the validity and estimate the reliability of the questionnaire. The validity measure was obtained by the EFA technique; the reliability was estimated by a Cronbach's Alpha. Results of these validity evidences and reliability estimates can be seen in Table 2 which shows the number of items that are valid after undergoing calculation by EFA and reliable model sub-components ($\text{Alpha} \geq 0.70$).

Table 2. Number of Valid Items among Sub-sub-components and their Reliability Measures

Component	Sub-component	Item Number	Failing Item	Valid Item	Alpha Reliability
Environment of Religiosity (ER)	Manners (Mn)	8	0	8	0.91
	Software (Sw)	15	5	10	0.71
	Hardware (Hw)	12	4	8	0.90
Process (Proc)	Planning (Pl)	5	0	5	0.82
	Implementation (Im)	15	0	15	0.88
	Evaluation (Ev)	8	0	8	0.96
Result (Rest)	Cognitive (Cog)	10	0	10	0.89
	Affective (Af)	19	5	14	0.86
	Psychomotor (Psy)	6	0	6	0.78
Impact (Imp)	Belief (Blf)	18	5	13	0.86
	Attitude (Att)	17	5	12	0.76
	Behaviour (Bhv)	18	6	12	*)
Total of Items		145	24	121	

*) Using the GENOVA, see Table 1.

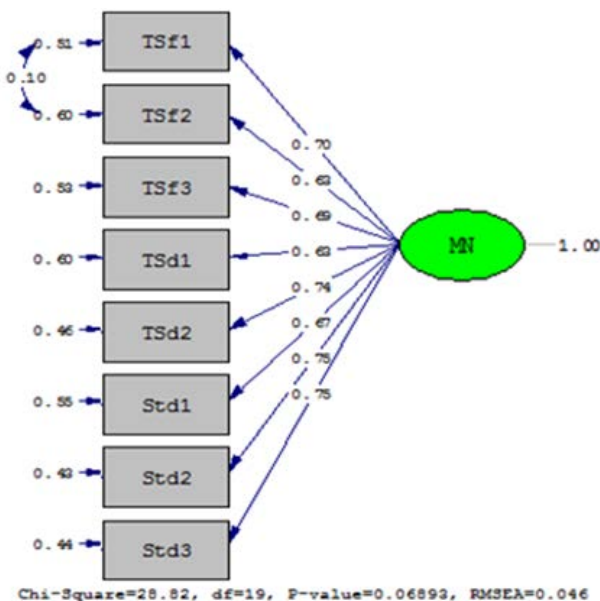


Figure 2. Confirmed results of sub-component MN

The revised questionnaire was then subjected to 250 JHS students, 20 teachers, and 10 principals. The results were used for validity and reliability checks. The validity check was determined by CFA while the reliability estimate was done by the Cronbach-Alpha technique. The validity check was conducted in phases; first on each sub-component by first-order CFA then on each component by second-order CFA. The first order CFA for sub-component Manners was found to be fit $\{(\text{Chi-Square} = 28.82, \text{df} = 19 \text{ or } \text{Chi-Square} < 2\text{df}), (p = 0.06892 \text{ or } p > 0.05), (\text{RMSEA} = 0.046 \text{ or } \text{RMSEA} < 0.08) \text{ and } \text{GFI} = 0.925 \text{ or } \text{GFI} > 0.90)\}$; [9].

By using the same technique, i.e. first-order CFA, confirmation or validation of sub-components Software (Sw) and Hardware (Hw) showed that both were found to be fit. Further, analysis was done on component Environment of Religiosity (ER) with its sub-components Mn, Hw, and Sw. This was done by second-order CFA and the results can be seen in Figure 3. This figure shows that the component ER was fit; meaning that the sub-components Mn, Hw, and Sw were rightly parts of

component ER. Figure 3 also shows that all items in Mn, Hw, and Sw sub-components had a factor loading higher than 0.3 meaning that the items were valid.

The validity checks and reliability estimates for the three other components of Process, Result, and Impact

were done in the same way; i.e. using the second-order CFA. The criteria that were used to select items in the three components were also the same as the criteria for item selection for component ER.

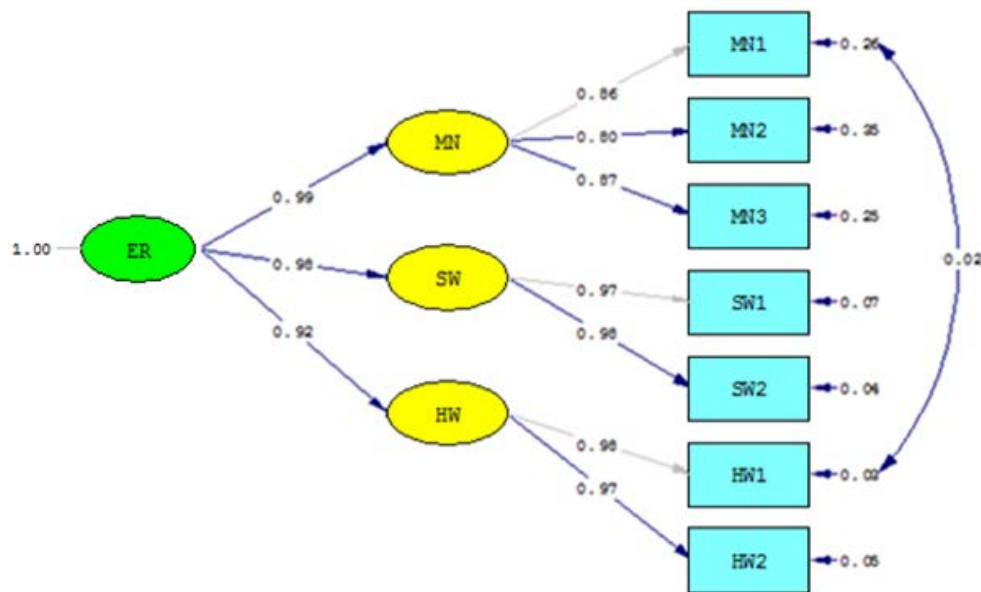


Figure 3. Results of second-order CFA for componen Environment of Religiosity (ER)

The results of the second-order CFA validation for the components Process (Proc), Result (Rest), and Impact (Imp) showed that the three components were fit. This means that planning (Pl), implementation (Im), and evaluation (Ev) were truly parts of the component process; sub-components cognitive (Kog), psychomotor (Psy), and affective (Aff) were truly parts of the component result; and sub-components belief (Blf), attitude (Att), and behavior (Bhv) were truly parts of the component impact.

Table 1 shows that subcomponent religious behavior (behavior) has five indicators, namely behavior 1 (prayers), behavior 2 (fasting and alm giving), behavior 3 (socializing), behavior 4 (piety), and behavior 5 (discipline). Looking at the items in these five indicators, findings of this study are in line with those of El-Menouar [4]. El-Menouar's study mentions five dimensions in the components of Muslim religiosity, namely: (1) basic religiosity, (2) central religious duties, (3) religious experience, (4) religious knowledge, and (5) orthopraxis. This fact indicates that the observation sheet that is used for eliciting behaviours possesses a construct validity measure by the multi trait-multi method technique in which the constructs and the measurement products are the same when used to elicit the same trait.

After undergoing the FGD and limited try-outs, the procedure of the conduct of the ERPRI evaluation model is as follows.

3.1. Evaluating the Environment of Religiosity component which is concerned with three aspects: manners in the interaction among the school academi, the software, and hardware that support the situation and atmosphere of religious life.

3.2. Evaluating the Process component which covers three sub-components of planning, instruction, and evaluation of learning.

3.3. Evaluating the Result component in three sub-components: cognitive, affective, and psychomotoric. The

cognitive evaluation is done by analyzing and synthesizing the results of the daily quizzes, mid-semester test, and end-semester test. The affective evaluation is done by observing the students' discipline, diligence, honesty, and politeness in and out of the classroom. The psychomotoric evaluation is done by analyzing and grading students' progress reports, port-folios, and assignments.

3.4. Evaluating the Impact component of the instructional processes covering the internalization and implementation of religious values in the forms of belief in embracing the religion (faith), taking positions in life matters (attitude), and doing actions in daily life (behavior) elicited through the observation sheet. The behavior sub-component consists of five the indicators of behavior 1, behavior 2, behavior 3, behavior 4, and behavior 5.

Data that were collected in the evaluation conduct were not only quantitative, but also qualitative in nature. Quantitative data were collected through the questionnaire, while qualitative data were collected through the observation sheet and questionnaire guide. Data collection was done by way of the embed technique whereby, to explain something, data were gathered at the same time, qualitative data complementing quantitative data.

Based on the reviews of related theories and research studies, a theoretical frame can be developed in the form of a hypothetical model structure or relation among the components in the ERPRI evaluation model. The hypothetical model structure shows that ER, process, and result give direct and indirect influences on impact. The hypothetical model structure must undergone testing. If the hypothesis is accepted, it can be stated that the model is feasible. ER, process, and result are capable of directly and indirectly influencing the students' religious behaviours.

The technique that was used to test the hypothetical model structure was the structural equation modeling

(SEM). Two requirements are to be fulfilled for the hypothetical model structure to be accepted; i.e. the model must be fit (Chi-square < 2df, p>0.05, and RMSEA<0.08) and the co-efficient path must be significant (t >1.96); or, when displayed in Standardized Solution, the co-efficient

path must be higher than 0.05 [12]. When these two criteria are fulfilled, then the model is effective; i.e. ER, process, and result give direct and indirect impacts on students' religious behaviors. Results of the hypothetical model structure testing can be seen in Figure 4.

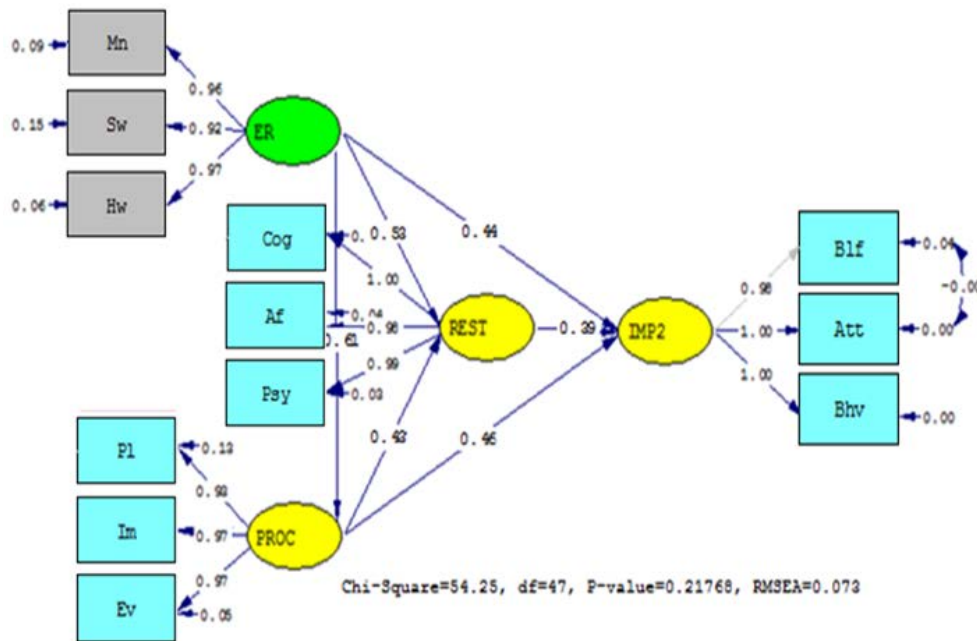


Figure 4. ERPRI evaluation model

Figure 4 shows that the model is fit {(chi square= 54.25 < 2(47), (p-value = 0.21768>0.05), and Root Mean Square Error of Approximation (RMSEA= 0.073 < 0.08) [9]. Figure 4 also shows that all the factor loadings for indicators Mn, Hw, and Sw on the exogenous latent variable ER are significant. This means that Mn, Sw, and

Hw, are the proper variables to measure exogenous latent variable ER. The same happens to the Pl, Im, and Ev for endogenous latent variable Proc; Kog, Aff, Psy for endogenous latent variable Rest; and Blf, Att, and Bhv for endogenous variable Imp.

Table 3. Total Effect among Exogenous Variables and Endogenous Variables

Latent Variabel	r	r ²	Direct Effect (DE)	Indirect Effect (IE)	Total effect (TE)
ER →PROC	0.61	0.36	0,36 (11.15*)	-	0.36
ER →REST	0.53	0.28	0.28 (7.38)	0.07	0.35
ER →IMP	0.44	0.19	0,19 (5.38)	0.11	0.30
PROC →REST	0.43	0.18	0.18 (9.23)	-	0.18
PROC → IMP	0.46	0.21	0.21 (5.35)	0.18	0.39
REST → IMP	0.39	0.15	0.15 (6.74)	-	0.15

*) T value.

Data analysis results (Figure 4) also show that components ER, process, and result have an influence on students' religious behaviors. Complete data analysis results can be seen in Table 3. This table shows that all of the co-efficient paths on direct effects have a value higher than 0.05 (DE > 0,05 or all T values > 1.96). This means that the two requirements for the ERPRI model feasibility are fulfilled. With that, it can be stated that the ERPRI evaluation model is effective.

In addition to the quantitative data, this study also produces qualitative information. All the JHS in Palu City have a mosque in the school and put up school rules and slogans at strategic spots. The mosque is often used for religious practices and even functions as an IRE lab. According to students' responses and observation results, relations among principal and teachers, teachers and teachers, and teachers and students are smooth, warm, and harmonious looking at the status of each individual.

In the learning process, teachers give model examples of how to behave at school in line with the religious teaching and school rules. In their teaching processes, teachers remind the students of good conducts and evaluate the students' behaviours. The majority of the students state that, during the learning processes, they receive encouragement to behave well.

In their daily school life, all students take prayers on time in the mosque and are involved in religious activities such as distributing alms and meat. Communication among students is smooth and warm, and there is never a student brawl. They are polite and well-behaved when communicating with the teachers and school principal.

Results of the present study is in line with those of Brown & Gary [2], Free [7], Brizer, Carroll [3], and Foshee & Hollinger [6] concerning levels of religiosity in relation to respondents' behaviors. The same thing can be said in relation to Waite & Gallagher [15], Popenoe [13],

dan Blankenhorn [1] showing the existence of positive relations between religious practices and marriage quality and stability.

The end product of this study is an evaluation model of IRE in the JHS in relation to students behaviours labeled as ERPRI. This model consists of four components, namely: environment of religiosity, process, result, and impact. The component environment of religiosity depicts the religious atmosphere and environment of the school whether it is conducive for the implementation of IRE for students and teachers. The component process is related to the students' responses towards the learning processes conducted by the teacher. The component result is concerned with learning outcomes achieved by the students in certain periods of time. The component impact is a reflection of the three previous components, hypothesized as an internalization and implementation of the teaching processes done by the IRE teachers in the school.

Data analysis results show that the Environment of Religiosity component, as an exogenous variable, can be measured and described uni-dimensionally, appropriately, and consistently by the three sub-components or the three main factors that build it, namely: manners (MN), software (SW), and hardware (HW). The same thing can be said about the other three components (Process, Result, and Impact) as endogenous variables. Uni-dimensionally, appropriately, and consistently, they can be measured and described by each of their three sub-components or the three main factors that build them; the process component include sub-components planning (Pl), implementation (Im), and evaluation (Ev), component Result include cognitive (Cog), affective (Aff), and psychomotor (Psy), and component impact include belief (Blf), attitude (Att), and behavior (Bhv).

The instruments that are designated for the evaluation of those four components have been tested and found to have good validity and reliability. Evidences for the validity and reliability of the instruments have been obtained in line with standard procedures. Therefore, evaluation results using the instruments can be assured to be proper and reliable so that the ERPRI model can be stated to be eligible for use.

4. Conclusion

Following the data analyses and discussion of the research findings, a number of conclusions are presented as follows:

4.1. The ERPRI model developed for the evaluation of Islamic Religious Education instruction in the Junior High School consists of (a) an IRE evaluation procedure, (b) evaluation instrument, and (c) evidence for model feasibility;

4.2. The procedure for IRE evaluation using this model consists of conducting evaluation on the components of environment (ER), process (Proc), result (Rst), and impact (Imp);

4.3. All of the instruments used in the ERPRI evaluation are in good categories (valid and reliable);

4.4. The model developed in this study is effective, model fit, and significant relations among the model

variables. In this case, (a) the variable Environment of Religiosity has significant direct and indirect influences on the variable Result; (b) the variable Environment of Religiosity has significant direct and indirect influences on the variable Impact; (c) the variable Process has significant direct and indirect influences on the variable Result; (d) the variable Process has significant direct and indirect influences on the variable Impact; and (e) the variable Result has significant direct and indirect influences on the variable Impact;

4.5. The developed model indicates that (a) students' religious behaviours (Impact) are most strongly influenced by learning processes (Process), followed by the religious atmosphere and environment in the school (Environment of Religiosity) and learning outcomes (Result), and (b) learning outcomes (Result) are more strongly influenced by the religious atmosphere and environment in the school (Environment of Religiosity) rather than the learning processes in the school (Process).

This conclusion implies that when it is desired that the students behave well then it must be created religious environment and rules that can motivate students to behave well. In the learning process, teachers must exemplify good behavior, remind students to behave well, and assess students' everyday behavior.

References

- [1] Blankenhorn, D, *Fatherless America*, New York, Basic Books., 1995.
- [2] Brown, D. R., & Gary, L. E, "Religious socialization and educational attainment among African Americans: An empirical assessment", *Journal of Negro education*, Vol. 60, No.3, 411-426. 1991.
- [3] Carroll, S, "Spirituality and purpose in life in alcoholism recovery", *Journal of studies on alcohol*, Vol. 54, No. 3. 297-301, May. 1993.
- [4] El-Menouar, Y, "The Five dimensions of muslim religiosity. Result of empirical study", *Methods, data, and analysis*, Vol. 8, No. 1, 53-78, 2014.
- [5] Feldt, L. S. and Brennan, R. L, "Reliability", *Educational measurement*, edited by Robert L Linn, New York, Macmillan Publishing Company, 1989.
- [6] Foshee, V. A., & Hollinger, B. R, "Maternal religiosity, adolescent social bonding, and adolescent alcohol use", *Journal of early adolescence*, Vol. 16, No. 4. 451-468, Nov. 1996.
- [7] Free, M. D. Jr. "Religiosity, religious conservatism, bonds to school, and juvenile delinquency among three categories of drug users", *Deviant behavior*, Vol. 15, No. 2. 151-170. 1994.
- [8] Hamdan, *Pengembangan dan pembinaan kurikulum (Teori dan praktek kurikulum PAI)*, Banjarmasin, Not published, 2009, 40.
- [9] Joreskog, K. G. & Sorbon, D, *LISREL 8.3: User's reference guide*, Chicago, Scientific Software International, 1996.
- [10] Marzuki, *Prinsip dasar akhlak mulia*, Yogyakarta, Debut Wahana Press & FISE UNY, 2009, 4.
- [11] McAndrew, S. and Voas, D, "Measuring religiosity using surveys". *Question bank*, 4, 1-15, February 2011.
- [12] Pedhazur, E. J, *Multiple regression in behavioral research*, Toronto, Thomson Learning, Inc., 1997.
- [13] Popenoe, D, *Life without father*, New York, Free Press, 1990.
- [14] Taimiyah, I, *Kitab al-iman*, Damaskus, Dar al-Fikr, 1961.
- [15] Waite, L., & Gallagher, M, *The case for marriage: Why married people are happier, healthier, and better off financially*, New York, Doubleday, 2000.
- [16] Zirkel, S., Garcia, J.A., & Murphy, M.C, "Experience-sampling research methods and their potential for educational research". *Educational Researcher*, Vol. 44. No. 1. 7-16. 2015.