Models for Measuring Latent Variables in Asynchronous Learning

Mirisoa Rakotomalala,
Ph d School in Sciences and Technical Engineering and Innovation,
University of Antananarivo,
Antananarivo, Madagascar

Toky Basilide Ravaliminoarimalalason,
Ph d School in Sciences and Technical Engineering and Innovation,
University of Antananarivo,
Antananarivo, Madagascar

Falimanana Randimbindrainibe,
Phd School in Sciences and Technical Engineering and Innovation,
University of Antananarivo,
Antananarivo, Madagascar

Abstract— **Digital technology is currently an important part of many areas of our lives, our society and our culture. This study was conducted in the context of this digital culture, specifically regarding asynchronous learning in open and distance learning. The traces of the activities of learners on the platform were identified and used to measure the direct and indirect effects of learners' engagement and behavior, as latent variables, on their cognition. The results gave two models of measurement of these two latent variables according to the concept of structural equations of research.**

Keywords— Models for measuring, commitment, behavior, learning, asynchronous ;

I. INTRODUCTION

In Open and Distance learning (ODL), the community is formed by a group of learners, gathered for a training activity, to address a concern about a task whose product will be evaluated.

This study is relevant to the analysis of these forums and attempts to propose a structural conceptual equation model, in order to examine the relationship between the learner's attitude and his commitment to his asynchronous learning.

II. ASYNCHRONOUS LEARNING IN ODL

A. Definitions

If distance learning has been around for decades, ODL is a form of education that emerged in the 1990s and overcame the constraints of place and time by relying on digital technologies. By definition, an ODL is a system:

- which is organized, finalized, and recognized as such by the actors;

- which takes the singularity of people in their individual and collective dimensions into account;

- which relies on complementary and plural learning situations in terms of time, place, human and technological pedagogical mediations, and resources.

Distance learning is defined as a mediated training [1]. It uses technological artifacts and media devices.

In this paper, these systems rely on the use of a Moodle platform. This platform has been designed to support a socio-constructivist training framework and integrates tools and learning activities focused on synchronous and asynchronous communications. In addition, with Moodle, learners are at the center of the training.

The forum is a communication tool and an asynchronous collaboration tool in training. This tool has many advantages. It is a tool adapted to the Malagasy context given the poor quality of Internet connection in different regions of Madagascar and especially the presence of load shedding that prevents the synchronous meetings scheduled during the training. This tool is also favorable for collaborative work; it is accessible to everyone, at all time. The tutor is a resource person who helps the learners to solve technical problems or redirects them to the person in charge, answers their questions and takes the role of an animator in a collaborative work if necessary. In a forum, each learner can ask any educational question (course, activity, evaluation) or organizational question (schedule, group appointment), and everyone can answer any questions without waiting for the intervention of the tutor.

B. Asynchronous learning approach in distance learning

Learning consists of acquiring new knowledge, storing it in memory, organizing it and developing automatism. Learning is thus perceived as an individual mental activity based on information to be processed [02]. In 2003, Entwistle identified the different learning approaches among students and favored those that would improve the quality of learning [03]. Those are the in-depth learning approach, indicating the intention to understand oneself, the metacognitive approach that assumes one can step back from one's ways of studying, and the organization of studies to manage time efficiently and work hard. These approaches are related to the student's commitment to training, as well as his motivation and performance.

In the asynchronous training type, the exchange with other learners or with the tutors is carried out via modes of communication that do not require a simultaneous connection. These can be discussion forums or even the exchange of e-mails. Moreover, this mode of training is often based on self-directed learning, with courses, exercises and self-assessments involving a certain autonomy of the learner. Online training in an asynchronous environment means that all the responsibility for learning lies with the learners.

As a learner-centered method, asynchronous learning gives students full responsibility for their online learning experience.

Asynchronous learning requires that the learner be focused, determined and have good time management skills. Successful asynchronous learning requires learners' commitment, motivation and rigorous discipline.

III. METHODOLOGIES

A. Methodology used

Structural equation models have been useful in tackling many of the fundamental problems and have contributed to the evolution of theory in the social and behavioral sciences, psychology and education, and so on. [04]. These methods make it possible to simultaneously test the existence of causal relationships between several explanatory latent variables and several variables to be explained [05].

A manifest variable is a variable for which a measure can be directly collected (observed, measured, etc.). The dominant paradigm in test theory was that represented by Churchill [06], who considers that manifest variables are all supposed to represent their latent variables. In this sense, it is postulated that all the indicators agree in their way of measuring the phenomenon, and all allow the reflection of the same variable. Consequently, it is obligatory to ensure the significance of the latent variable constructed on the basis of these indicators, which must be significantly correlated.

A well-constructed latent variable is a variable whose variation must be accompanied as closely as possible by the variation of all the indicators that compose it.

The traces of learners' activities in the training platform are used to measure the direct and indirect effects of the learners' commitments and their behaviors in the discussion forum on their cognitions, which translate into their pedagogical results.

In this case, the indicators are a reflection of the latent variable which remains the cause of the indicators, and where each indicator is linked to the latent variable by a simple regression equation of the following type:

$$X_i = \lambda_i \xi_1 + \delta_i \qquad (1)$$

With: ξ_1: latent construct, X_i: indicators, λ_i: coefficient representing the effect of the construct on the indicator (loading); and δ_i: measurement error.

B. Model of structural equations

According to Bollen and Long [07] and Schumacker and Lomax [08], the specification, identification, estimation, evaluation and respecification of the model represent the most commonly accepted steps in structural equations modelling.
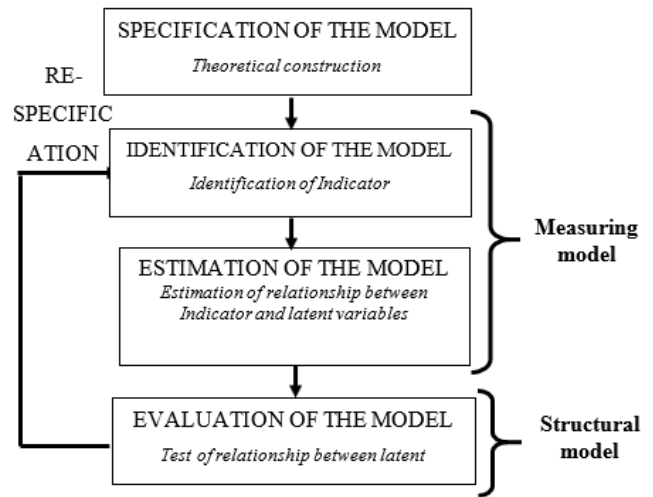

Fig. 1. Steps of structural equations modelling

The specification of the model leads to the design of a "linear relationship schema" that translates into measurement models and structure models. [10]

The identification makes it possible to assign a single solution for each parameter to be estimated. A model leads to express the variance/covariance matrix of the manifest variables ☐ as a function of a set of parameters θ. This model is identified if $\sum(\theta)=\sum(\theta')$ implies $\theta = \theta'$. In practice, the identification of a model supposes the satisfaction of two conditions, namely the condition of order (necessary condition) and the condition of rank (necessary and sufficient condition).

The order condition imperatively depends on the degree of freedom ddl [10] and refers to the correspondence between the parameters to be estimated and the number of variances/covariances of the observed variables [11]. The degree of freedom is expressed in the following way [08]:

$$ddl = \left(P \frac{(P+1)}{2} \right) - N \qquad (2)$$

With: P is the number of measurement indicators of the model, N: the number of coefficients of the model to estimate and ddl is the degree of freedom.

Therefore, three situations are possible: sub-identification (ddl <0), fair identification (ddl = 0) and over-identification of the model (ddl> 0). According to these three situations, the order condition is verified when the model is just identified or over-identified (ddl ≥ 0).

Moreover, the rank condition makes it possible to have a single solution with respect to each of the parameters to be estimated.

The estimation consists of determining the values relative to the various parameters of the model to be tested. ML (Maximum Likelihood) and GLS (Generalized Least Squares)

estimation techniques are the most commonly used methods for estimating a model since they are included in most data processing software [08] [09]. The expression of the Maximum Likelihood method is as follows:

$$F_{ML} = tr(S.\widehat{\Sigma}^{-1}) - p + ln|\widehat{\Sigma}| - ln|S| \qquad (3)$$

If θ is a free parameter vector estimated by the model, we can write the relation 3.03 as follows:

$$F_{ML}(\theta) = tr(S.\Sigma^{-1}(\theta)) - p + ln|\Sigma(\theta)| - ln|S| \quad (4)$$

With: S: the observed variance-covariance matrix, ln: the natural logarithmic function, tr(.): The trace function of a matrix, Σ: the estimated variances/covariances matrix, p: the number of measured or observed variables, and |S|: Determinant of a matrix S.

For the evaluation of the model, its quality can be appreciated through the evaluation of a set of adjustment indices. There are several indices of adjustment that fall into three categories namely absolute indices, incremental indices and parsimony indices.

Absolute indices are qualified as classical indices [09]. They examine the level of correspondence between the proposed model and the observed data [08]. Incremental indices are used to evaluate the contribution of the model studied to a restrictive model, i.e. a comparison between the estimated model and the reference model with zero correlation between the observed data [10]. The parsimony indices control the overestimation of the model. They determine the necessary number of parameters to estimate to reach the specific adjustment level [08]. The goal is to achieve a better mix between maximizing the fit and minimizing the number of estimated coefficients [10].

The respecification of the model represents a phase of reflection that offers the opportunity to reconsider the design of the model and propose possible modifications, taking into account the theoretical framework of the research in question [10]. (Ex: adding or deleting parameters, setting or modifying constraints, etc.)

In terms of the presentation of the results, the latter must incorporate relevant and precise information, making it easy to find solutions to the problems likely to be encountered in the analysis.

IV. RESULTS AND DISCUSSION

The results are divided into two parts. The first part is about the analysis of asynchronous learning activities in the discussion forum in which an introductory module was mobilized. Then, a conceptual model of structural equations of the research will be proposed, and the different stages will be implemented in order to leave the two models of validated measurements from SEM. R is used for statistical processing of data.

A. Analysis of the activity

The goal is to analyze asynchronous learning in a distance education. The analysis preferentially focuses on the learners' forum. The discussion takes place on the platform in an introductory module (GEAL) at the beginning of the training. Learners interact with one another or with the tutor. Three open

and distance courses are involved. The learner needs an internet connection to access and work on the platform.

1) Indicators of the learner's behavior :

The traces of the activities of the participants in an online training can be transformed to represent a measure of behavioral engagement [12], which has considerable implications for how they are organized and simplified to fit well with theoretical constructs. In a current research initiative, the learner's attitude towards a discussion forum was modeled as a continuum of participation in the various interventions available on discussion forums.

From a trace analysis point of view, a preliminary quantitative approach to the more specific statistical study is needed. Participation in number of interventions: consulting forums, consulting discussions, adding or updating a message to an ongoing discussion, and creating new topics are indicators taken into account. All of these indicators characterize the behavior of the learner in a discussion forum.

2) Indicators of the learner's commitment :

The engagement of the learner is measured by the regularity of their participation to the forum, their feeling towards the discussion forum, the completion of the activities from the platform in spite of the existing constraints, and the dedication to stay active in asynchronous learning [13]. In this study, the models explore the different communicational, motivational, and socio-cognitive dimensions of engagement in distance learning contexts. These three dimensions are in continuous and reciprocal interaction according to variable and contingent importance to activities. The qualitative analysis carried out on the contents of the discussion forum gave the indicators to measure, according to the three dimensions above regarding the engagement of the learner. Thus, the collection and analysis of traces of the main activities are based on the messages. These are classified according to the communication message, the motivation indicator and the interaction for collaboration. These three classes constitute the learner's engagement in a discussion forum.

B. Modeling by structural equations

The conceptual model of research presents the different causal links between theoretical constructs and shows that asynchronous learning could play a mediating role in the relationship between learner's engagement and the exploitation of the discussion forum. The causal model consists of two measurement models and structural models. It encompasses the set of independent observable variables (X_i), dependent observable variables (Y_i), an explanatory latent variable (E), two latent variables to be explained (A) and (C), and error terms $(e_i, \delta, \beta$ and $\gamma_i)$. Therefore, it is possible to distinguish two types of equations namely the equations of the measurement model and the equations of the structure model.
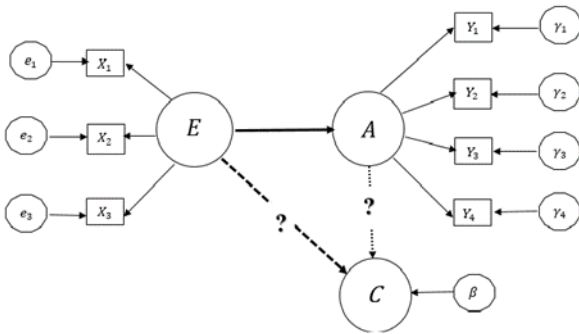
Fig. 2. The research conceptual model

C. Measurement models

Two structural equation measurement models were developed: the learner's attitude in asynchronous learning, and their commitment to distance learning.

These models will be applied to real data from an ODL of Antananarivo IST. These are asynchronous activity traces of 4804 events of 13 learners for 19 days from October 20th and November 07th, 2016.

1) The model for measuring the attitude of the learner

a) First model specification

The latent variable A measurement model: learner attitude towards the discussion forum, can be illustrated by the following equations:

$$Y_1 = \alpha_1 A + \gamma_1$$
$$Y_2 = \alpha_2 A + \gamma_2$$
$$Y_3 = \alpha_3 A + \gamma_3 \quad\quad (5)$$
$$Y_4 = \alpha_4 A + \gamma_4$$

They can be written in the matrix form: Y = αA + γ and form a linear regression.

The covariance matrix between the four variables shows that there are negative covariances between the two variables "Creat_D" and "consu_D".

For Model Estimation, the maximum likelihood estimation was used. The model converges only after 7959 iterations, with a test on a sample of 13 observations.

The degree of freedom ddl = p (p + 1) / 2 –N, with p = 4 and N = 6, is equal to 2, which is verified by the "Degrees of freedom" indicates over-identification of the model.

It is noted that the error on the variable Creat_D is -1112.53 (negative variance) which defines a reason of respecification of the model.

The log likelihood (log1): -164.987 indicates that the model has no likelihood relative to the actual data.

b) Respecification of the model

The two variables "Send_M" and "Creat_D" express the learner's participation in the discussion forum; they will be merged (Envoi_M_2). As a result, there are only three observed variables instead of four as previously seen. The

previous steps were implemented with the same data but with this new model.
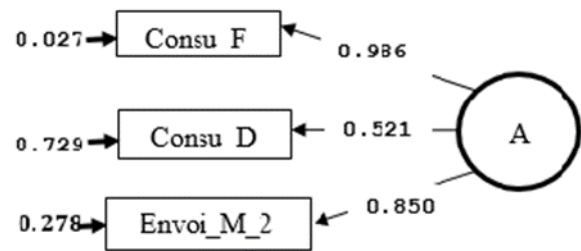

Fig. 3. Measurement model of latent variable A with estimated parameters

The degree of freedom ddl = p (p + 1) / 2 - N becomes zero. The model is just identified. As far as the evaluation of the model is concerned, the minimum of the estimate function (fmin) is equal to 0, indicating that it is a good model compared to the previous model which has fmin = 0.677.

The log likelihood (logl) of the respecified model is -144.312 which is better than that of the initial model -164.987.

The $\chi2$ (chisq) of the model is equal to 0. There is no presumption against the null hypothesis. The actual data do not have much conformity with the underlying statistical model, this may be due to their distribution.

The degree of freedom of the model (df) is 0, as calculated manually. Six parameters (npar) of the model were estimated.

This respecified model can be validated with respect to the first. The chi-square does not prevent this validation, the value of p should never be used to validate a hypothesis from data since it is the opposite that is calculated.

2) The model of measurement of the latent variable E

The learner's engagement measurement model can be translated into the following equations:

$$X_1 = \lambda_1 E + e_1$$
$$X_2 = \lambda_2 E + e_2 \quad\quad (6)$$
$$X_3 = \lambda_3 E + e_3$$

They can be written in the matrix form: X = λE + e which forms a linear regression.

By reviewing the different steps of the learner's engagement measurement model including model specification and model estimation, the following model was obtained:
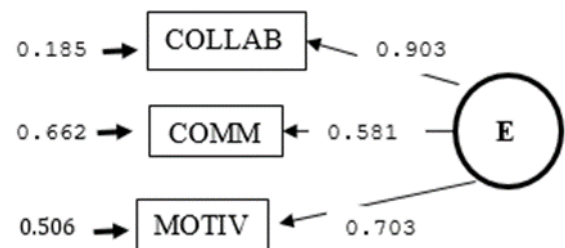

Fig. 4. Measurement model of latent variable A with estimated parameters

The degree of freedom is still zero. Thus, the model is also fair identified.

Validation of the model

```
> fitmeasures(donnees3.fit,
fit.measures=c("fmin", "logl",
"chisq", "df", "npar"))

   fmin   logl  chisq     df   npar

   0.00 -61.47   0.00   0.00   6.00
```

The minimum of the estimation function (fmin) is equal to 0. The log likelihood (log1) is -61.47. An exact likelihood for a null value of this log likelihood is noted, and this value is approached little by little.

logl = -144.312 for the attitude model.

The $\chi2$ (chisq) is still 0.

The estimated parameter number (npar) of the model is 6. These data give the idea of also validating the model described.

V.   CONCLUSION

The interactions between students in distance education aim at deep learning and have been reinforced by a teaching method in collaborative learning. The participation in exchanges within the platform, the richness and the quality of the interactions between learners are recognized like factors that are beneficial to the engagement, the satisfaction and the success of the learners in distance learning.

Still, participation in a discussion forum, reading other learners' interventions, and peer interaction create a sense of belonging to a group. This commitment is considered a latent variable and was measured by a measurement model in a structural equation.

To conclude, two validated measurement models were obtained. The structural modeling of asynchronous learning by manipulating the two models with the real data in three open and remote trainings will allow the validation of the causal relationships between the pedagogical result, the commitment and the attitude of the student facing the discussion forum. The knowledge development effect via the discussion forum will be strongly demonstrated. In other words, engaged, motivated and collaborating distance learners develop their knowledge and skills on an ongoing basis. The impact of forum exploitation on the acquisition of knowledge will be demonstrated empirically.

REFERENCES

[1]   D. Peraya, "La formation à distance : un dispositif de formation et de communication médiatisée. Une approche des processus de médiatisation et de médiation." TICE et développement », n°1, 2005.

[2]   M. Fayol, et D. Gaonac'h, " Le développement de la mémoire ", in A. Blaye, et P. Lemaire, (Edit.) Psychologie du développement cognitif de l'enfant. Paris-Bruxelles : De Boeck Université, 2007, pp. 125-156.

[3]   N. Entwistle, "Concepts and conceptual frameworks underpinning the ETL Project", Edinburgh, School of Education, University of Edinburgh, 2003.

[4]   K. G. Jöreskog, D. Sörbom, "Recent developments in structural equation modling ". Journal of Marketing Research, 19, pp 404-416, 1982.

[5]   Lacroux, « L'analyse des modèles de relations structurelles par la méthode PLS : une approche émergente dans la recherche quantitative en GRH », XXème congrès de l'AGRH, Toulouse, 2010.

[6]   G. A. Churchill, « A paradigm for developing better Measures of marketing constructs », Journal of Marketing Research, vol. 16, pp 64-73, Feb 1979.

[7]   K.A. Bollen, J.S. Long, « Testing structural equation models ». Newbury Park, CA: Sage, 1993.

[8]   R.E. Schumaker, R.G. Lomax, « A beginner's guide to structural equation modelling », Lawrence Erlbaum Associates, 2ème edition, London, 2004.

[9]   L. Chaput, « Modèles contemporains en gestion », Presses de l'Université du Québec, 2006.

[10] P. Roussel, F. Durrieu, E. Campoy, A. El Akremi, « Méthodes d'équations structurelles : Recherches et applications en gestion », Edition Economica, Paris, 2002.

[11] R.H. Hoyle, « Structural equation modelling: Concepts issues and applications », Edition Sage, London, 1995.

[12] B. Poellhuber, N. Roy et I. Bouchoucha, « Relations entre la motivation, l'engagement cognitif et la persévérance dans un MOOC francophone (EDULIB) », Communication au colloque de l'AIPU, Mons, 2014.

[13] K. DE WULF, G. ODEKERKEN-SCHRÖDER, "Assessing the impact of a retailer's relationship efforts on consumer's attitudes and behavior", Journal of Retailing and Consumer Services, 10, 2, pp. 95 – 108, 2003.