

N° d'ordre : 808-2018-SCA-D3

Année universitaire : 2018



**UNIVERSITÉ D'ANTANANARIVO**

**ÉCOLE SUPÉRIEURE POLYTECHNIQUE D'ANTANANARIVO**

**ÉCOLE DOCTORALE EN SCIENCES ET TECHNIQUES DE L'INGÉNIERIE ET DE L'INNOVATION**

**Équipe d'Accueil Doctorale en Sciences Cognitives et Applications**

**Laboratoire de recherche : Sciences Cognitives et Applications**

## **THÈSE de DOCTORAT**

**Pour l'obtention du titre de  
DOCTEUR DE L'UNIVERSITÉ D'ANTANANARIVO**

*Domaine : Sciences de l'Ingénieur*

*Thématique : Sciences et Techniques de l'Ingénierie et de l'Innovation*

*Spécialité : Sciences cognitives et Applications*

*Par* **RAKOTOMALALA Mirisoa**

## **MODELISATION DE L'APPRENTISSAGE ASYNCHRONE DANS UNE FORMATION OUVERTE ET A DISTANCE**

Soutenu le 19 septembre 2018 à Ankatso, devant la Commission d'Examen composée de :

*Président du jury :* Monsieur ANDRIANAHARISON Yvon Dieudonné, Professeur Titulaire  
*Examineurs :* Monsieur RANDRIAMITANTSOA Paul Auguste, Professeur Titulaire  
Monsieur RAKOTOMIRAHLO Soloniaina, Professeur

*Rapporteur Externe :* Monsieur RAKOTONIRINA Christian, Professeur  
*Rapporteur Interne :* Monsieur RASTEFANO Elysée, Professeur

*Directeur de thèse :* Monsieur RANDIMBINDRAINIBE Falimanana, Professeur Titulaire  
*Co-Directeur de thèse :* Monsieur RAVALIMINOARIMALALASON Toky Basilide, Docteur



## **FISAORANA**

Alohan'ny rehetra dia, isaorako ny Tompo mahery, namindra fo sy namela heloka, nanome hery ho ahy, nanatontosana sy nidirana tamin'ny dingana rehetra nilaina tamin'ny fikarohana.

Ny asa aseho ao anatin'ity boky ity dia natao tao amin'ny laboratoara Sciences Cognitives et Applications (SCA)n'ny Sekoly doktoralny Sciences et Techniques de l'Ingénierie et de l'Innovation (ED STII) eo ambany fiahian'ny Sekoly Ambaratonga Ambony Politeknikan'ny Oniversiten'Antananarivo.

Sarotra ny hisaotra ny olona rehetra nanampy nahafahana namita izao tezy izao. Na izany aza anefa dia tsy azo ialana ny fanolorana fisaorana manokana ireto voalaza manaraka ireto :

- Ramatoa RAMIARISON MAVOARILALA Claudine, Profesora, Tale jeneralin'ny fikarohana eo anivon'ny Ministeran'ny Fampianarana Ambaratonga Ambony sy ny Fikarohana Siantifika, sady Prezidan'ny fivondronan'ny Sekoly Doktoralin'i Madagasikara (CEDM),
- Andriamatoa RAMANOELINA Panja, Profesora laharana ambony, Prezidan'ny Oniversiten'Antananarivo,
- Andriamatoa ANDRIANAHARISON Yvon, Profesora laharana ambony, Talen'ny Sekoly Ambony Politeknika Antananarivo (ESPA) sy Andriamatoa RANDRIAMITANTSOA Paul Auguste, Profesora laharana ambony, Talen'ny ED STII, noho ny nandraisan'izy ireo ahy tao amin'ny sekoly izay iadidiany sy ny nanekeny ihany koa ho isan'ny mpitsara ny fanohanana ny voka-pikarohana, na dia teo aza ny fahateran'ny fotoanany nohon ny asa marobe izay sahaniny.

Tiako ihany koa ny manolotra fisaorana mitafotafo sy fankasitrahana feno an'Andriamatoa RANDIMBINDRAINIBE Falimanana, Profesora laharana ambony, Tomponandraikitra ny Equipe d'Accueil Doctorale en Sciences Cognitives et Applications (EAD SCA), tsy nisalasala tamin'ny nanekeny ahy ho isan'ny « doctorant » eo ambany fitarihiny. Ny fahafoizam-potoana nananany isaka ny misy ilàna azy, ny torohevitra maro izay nataony, dia isan'ny nanasokajiaiko azy ho isan'ny Talen'ny Tezy mitsinjo mandrakariva ny mpianatra. Isaorako toy izany koa Andriamatoa RAVALIMINOARIMALALASON Toky Basilide, Dokotera, niara-nifarimbona tamin'ny Talen'ny tezy, nanaraka akaiky ny asako. Ny torohevitra, ny fotoanany, ny talentany

sy ny fahavitrihana siantifika nananany ho fanampiana dia noraisiko ho tena zava-dehibe ary isan'ny hery manosika ahy lalandava nandritry ny fotoana nanaovana ny fikarohana mba ahatratarana ny tanjona.

Tolorana fisaorana manokana ihany koa Andriamatoa RAKOTONIRINA Christian, Profesora eo anivon'ny Ivotoerana Ambony momba ny Teknolojia Antananarivo (IST-T) sy Andriamatoa RASTEFANO Elysée, Profesora ao amin'ny Sekoly Ambony Politeknika Antananarivo, nanaiky ho mpanara-maso ity fikarohana ity.

Manolotra ihany koa ny fisaorana eram-pon eran-tsaina an'Andriamatoa RAKOTOMIRAHO Soloniaina, Profesora ao amin'ny Sekoly Ambony Politeknika Antananarivo nanaiky ho isan'ny mpitsara.

Isaorako ireo mpikambana rehetra ao antin'ny Komity manaraka ny tezy izay nitsara ny tatitra fandrosoan'ny tezy nandritra izay telo taona izay. Misaotra betsaka ihany koa an'ireo mpampianatra mpikaroka rehetra, ny mpandraharahan'ny ED STII sy ny ESPA, fa isan'ny niasa nahatontosa ny tezy izy ireo.

Tsy hijanona ny fisaorana raha tsy mipaka any amin'ireto olona ireto izay nandray anjara mavitrika tamin'ny fanampiana :

- Andriamatoa Jean Lalaina RAKOTOMALALA, Profesora laharana ambony, Tale jeneralin'ny IST-T noho ny tohana isankarazany sy fankaherezana nataony,
- Ramatoa ZAKARIASY Lova, Tale Jeneralin'ny IST Antsiranana amin'izao fotoana, noho ny fiaraha-miasa tamin'ny fotoana nanamboarana ny « module Guide de l'Etudiant pour l'Apprentissage en Ligne » (GEAL) ;
- Ny ekipan'ny « coordination » n'ny FOAD 3, noho ny naneken'izy ireo ny hidirako anaty plateforme ahazoana ny donnée ilaina.

Ary ny farany, nefa tsy ny kely indrindra, dia ny fanolorako fitiavana be ho an'ny vady aman-janako, nanana fahatokisana sy faharetana, ary teo foana nanohana nandritra ny fotoana lava nanaovako ny fikarohana.

## REMERCIEMENTS

Avant tout, je tiens à remercier Dieu tout puissant qui, à travers sa miséricorde et sa grâce, m'a donné la force et le courage de réaliser et de franchir toutes les étapes nécessaires à la recherche et à l'élaboration de cette thèse.

Les travaux présentés dans cette thèse ont été effectués au Laboratoire de Sciences Cognitives et Applications (SCA) de l'École Doctorale Sciences et Techniques de l'Ingénierie et de l'Innovation (STII) sous l'Institution de l'École Supérieure Polytechnique d'Antananarivo (ESPA), de l'Université d'Antananarivo (UA).

Il me sera très difficile de remercier tout le monde car c'est grâce à l'aide de nombreuses personnes que j'ai pu mener cette thèse à son terme. Toutefois, j'aimerais en citer quelques-unes, lesquelles ont joué un rôle incontournable tout au long de mon trajet.

Je tiens à remercier :

- Madame RAMIARISON MAVOARILALA Claudine, Professeur, Directeur Général de la Recherche Scientifique au sein du ministère de l'Enseignement Supérieur et de la Recherche Scientifique, et Présidente du Consortium d'Ecoles Doctorales de Madagascar (CEDM),
- Monsieur RAMANOELINA Panja, Professeur Titulaire, Président de l'Université d'Antananarivo,
- Monsieur ANDRIANAHARISON Yvon, Professeur Titulaire, Directeur de l'École Supérieure Polytechnique d'Antananarivo et Monsieur RANDRIAMITANTSOA Paul Auguste, Professeur Titulaire, Directeur de l'Ecole Doctorale STII, pour m'avoir si bien accueillie dans leur établissement, d'avoir accepté d'être parmi les membres du jury malgré leurs innombrables occupations.

Je tiens aussi à exprimer ma profonde reconnaissance à Monsieur RANDIMBINDRAINIBE Falimanana, Professeur Titulaire, Responsable de l'Equipe d'Accueil Doctorale en Sciences Cognitives et Applications, qui fut pour moi un directeur de thèse attentif et disponible malgré ses nombreuses tâches ; ainsi qu'à Monsieur RAVALIMINOARIMALALASON Toky Basilide, Docteur, Co-directeur de Thèse, dont le soutien et la disponibilité ont été pour moi d'importantes sources de motivation et d'inspiration tout au long de la réalisation de cette

recherche. Sa compétence et sa rigueur scientifique m'ont beaucoup apporté. Ils ont été et resteront des moteurs de mon travail de chercheur.

Je tiens à remercier chaleureusement Monsieur RAKOTONIRINA Christian, Professeur à l'Institut Supérieur de Technologie d'Antananarivo et Monsieur RASTEFANO Elysée, Professeur à l'Ecole Supérieure Polytechnique et à l'Ecole Doctorale en STII, d'avoir accepté d'être rapporteurs de cette thèse.

J'exprime toute ma gratitude à Monsieur RAKOTOMIRAHO Soloniaina, Professeur qui a accepté d'examiner ce travail.

Je remercie également, les membres du comité de suivi de thèse qui ont évalué les différents rapports d'avancement de thèse pendant ces trois années doctorales. Aussi, un grand merci, à tous les Enseignants Chercheurs, Personnels administratifs et techniques de l'École Supérieure Polytechnique d'Antananarivo, sans qui ce travail n'aurait pas pu aboutir. Merci pour votre support et vos encouragements.

Je ne saurais terminer avant de remercier également toutes ces personnes dont la contribution m'a énormément aidée :

- Monsieur Jean Lalaina RAKOTOMALALA, Professeur Titulaire, Directeur Général de l'Institut Supérieur de Technologie d'Antananarivo (IST-T) pour son appui, son écoute et son encouragement ;
- Madame ZAKARIASY Lova, Directrice Générale actuelle de l'Institut Supérieur de Technologie d'Antsiranana (IST-D), pour sa collaboration dans la conception du module Guide de l'Etudiant pour l'Apprentissage en Ligne (GEAL) ;
- Les équipes de coordination des FOAD de m'avoir autorisée à accéder aux différentes informations et données sur la plateforme.

Enfin, les mots les plus simples étant les plus forts, j'adresse toute mon affection à ma famille, qui a fait preuve de confiance, de patience et qui a toujours été là pour m'épauler durant ces longues années de recherche.

## TABLE DES MATIERES

FISAORANA .....	i
REMERCIEMENTS.....	iii
TABLE DES MATIERES .....	v
NOTATIONS & ABREVIATIONS.....	xi
LISTE DES FIGURES.....	xx
LISTE DES TABLEAUX .....	xxii
INTRODUCTION ET PRESENTATION DU PROBLEME.....	1
CHAPITRE 1 - ETAT DE L'ART SUR L'APPRENTISSAGE ASYNCHRONE DANS UNE FORMATION OUVERTE ET A DISTANCE .....	5
1.1 Introduction.....	5
1.2 Formation ouverte et à distance.....	5
1.2.1. Formation à distance.....	6
1.2.2. Formation ouverte .....	7
1.2.3. Diversité de formation à distance à Madagascar .....	7
1.2.4. Formation ouverte et à distance.....	10
1.3 Dispositif de formation .....	11
1.3.1. Apprentissage.....	12
1.3.2. Système d'apprentissage.....	13
1.3.3. La plateforme Moodle et la formation .....	13
1.3.4. Les acteurs et leurs interactions.....	14
1.4 Descriptif global de la formation .....	16
1.4.1. Déroulement de la formation.....	16
1.4.2. Suivi de la formation.....	17
1.4.3. Evaluation des apprenants .....	17
1.4.4. Descriptif détaillé d'un module de formation .....	17
1.4.5. Structure d'un module.....	19
1.4.6. Scénario d'encadrement d'un module et scénario d'apprentissage .....	20

1.4.7. Les dispositifs des trois formations ouvertes et à distance .....	21
1.5 Le module GEAL : Guide de l'Etudiant pour l'Apprentissage en Ligne .....	26
1.6 Science de données et sciences cognitives.....	29
1.6.1. Science de données .....	29
1.6.2. Sciences cognitives .....	30
1.7 Neurosciences cognitives et facilité d'apprentissage .....	31
1.8 Ingénierie pédagogique.....	32
1.8.1. Téléapprentissage.....	32
1.8.2. L'ingénierie des systèmes d'apprentissage et les interactions .....	32
1.8.3. Le design pédagogique, une ingénierie des systèmes d'apprentissage.....	32
1.8.4. Le concept de système d'apprentissage .....	33
1.9 Approche d'apprentissage asynchrone dans une formation à distance .....	34
1.9.1. Contrôle absolu sur sa formation.....	35
1.9.2. L'apprentissage à son propre rythme.....	35
1.9.3. Très pratique .....	36
1.9.4. Moins d'obstacles sociaux.....	36
1.9.5. Possibilité d'interaction malgré les barrières de lieu ou de temps.....	36
1.9.6. Pas de feedback immédiat.....	37
1.9.7. Sans interaction personnelle .....	37
1.9.8. Pas de collaboration directe, ni d'activités en temps réel.....	37
1.9.9. Réduire la motivation.....	38
1.9.10. Nécessité d'un certain degré d'autodiscipline .....	38
1.10 Conclusion.....	38
<b>CHAPITRE 2 – ETAT DE L'ART SUR LES THEORIES DE L'APPRENTISSAGE ET LA MODELISATION DANS UNE FORMATION A DISTANCE .....</b>	<b>40</b>
2.1 Introduction.....	40
2.2 Les théories de l'apprentissage .....	40
2.2.1. Le béhaviorisme .....	40



2.2.2. <i>Le constructivisme</i> .....	42
2.2.3. <i>Le cognitivisme</i> .....	43
2.2.4. <i>Le Socio-constructivisme</i> .....	45
2.2.5. <i>Le connectivisme</i> .....	46
2.3 Analyse des traces numériques d'activités sur la plateforme.....	48
2.3.1. <i>Notion de « trace d'activité »</i> .....	49
2.3.2. <i>Définition de traces numériques d'interaction</i> .....	50
2.3.3. <i>Caractéristiques de l'interaction humaine</i> .....	51
2.3.4. <i>Type et profil d'étudiants</i> .....	52
2.4 Approche morphologique d'un corpus de message.....	53
2.4.1 . <i>Premières analyses à mettre en œuvre</i> .....	53
2.4.2. <i>Caractéristiques quantitatives d'un forum</i> .....	54
2.5 La théorie sur l'analyse de discours : échanges écrits en ligne .....	57
2.5.1. <i>Rigueur d'une analyse de contenu</i> .....	57
2.5.2. <i>Echanges, structures et transactions</i> .....	58
2.5.3. <i>L'interlocution : dimension cognitive et sociale</i> .....	58
2.5.4. <i>Echanger par écrit pour apprendre</i> .....	59
2.5.5. <i>Etapas du processus</i> .....	59
2.5.6. <i>Le codage</i> .....	60
2.5.7. <i>Propriétés des catégories</i> .....	60
2.6. <i>Modélisation</i> .....	61
2.6.1. <i>Modélisation dans les sciences sociales</i> .....	61
2.6.2. <i>Modèle psychocognitif d'encadrement selon Jacques Rodet</i> .....	62
2.6.3. <i>Modélisation dans les sciences cognitives</i> .....	63
2.7. <i>Conclusion</i> .....	73
CHAPITRE 3 – THEORIES DE BASE SUR LE MODELE D'EQUATION STRUCTURELLE .....	74
3.1 Introduction.....	74

<b>3.2 Notion de base en statistique .....</b>	<b>75</b>
3.2.1. Définitions.....	75
3.2.2. Mesure de la relation entre deux variables.....	78
3.2.3. Matrice de corrélation .....	81
3.2.4. Estimation.....	81
3.2.5. Matrice de covariance .....	84
3.2.6. Test de sphéricité de Bartlett .....	87
<b>3.3 Analyse multivariée.....</b>	<b>88</b>
3.3.1. Régression linéaire .....	88
3.3.3. Modèle linéaire multiple.....	90
3.3.4. Loi du Khi-deux .....	91
3.3.5. Loi de Wishart.....	92
3.3.6. Méthode des moindres carrés .....	94
3.3.7. Test de Khi deux.....	95
<b>3.4 Les équations structurelles.....</b>	<b>100</b>
3.4.1. Origine.....	102
3.4.2. Causalité.....	103
3.4.3. L'importance de la théorie .....	103
3.4.4. Quelques notions importantes .....	104
3.4.5. Les étapes .....	104
<b>3.5 Modèles d'équations structurelles .....</b>	<b>105</b>
3.5.1. La modélisation par équations structurelles.....	105
3.5.2. Les fondements conceptuels de la modélisation par les équations structurelles (SEM).....	106
3.5.3. Les étapes d'élaboration d'un modèle d'équations structurelles .....	108
<b>3.6 Conclusion .....</b>	<b>118</b>
<b>CHAPITRE 4 – MODELISATION DE L'APPRENTISSAGE ASYNCHRONE PAR EQUATIONS STRUCTURELLES LINEAIRES .....</b>	<b>120</b>
<b>4.1 Introduction.....</b>	<b>120</b>

4.2 Analyse de l'activité .....	121
4.3 Approche théorique du SEM appliquée à la formation à distance.....	125
4.3.1. <i>L'engagement scolaire et l'approche sociocognitive de la motivation</i> .....	125
4.3.2. <i>Les dimensions de l'engagement scolaire</i> .....	126
4.3.3. <i>L'engagement en formation à distance dans une perspective socio-cognitive</i> .....	127
4.3.4. <i>Modèle d'équations structurelles en formation à distance</i> .....	128
4.4 Mesure de l'engagement .....	130
4.4.1. <i>Mesures de l'engagement scolaire en FAD</i> .....	131
4.4.2. <i>Mesure de l'engagement par les questionnaires</i> .....	131
4.4.3. <i>L'analyse des traces pour mesurer l'engagement comportemental</i> .....	131
4.4.4. <i>L'engagement dans l'apprentissage collaboratif à distance</i> .....	132
4.4.5. <i>Mesure de l'engagement dans l'apprentissage collaboratif à distance</i> .....	134
4.5 La spécification du modèle .....	136
4.5.1. <i>L'engagement de l'apprenant dans une formation à distance</i> .....	137
4.5.2. <i>Attitude dans l'apprentissage asynchrone</i> .....	143
4.5.3. <i>Modélisation d'équations structurelles par le maximum de vraisemblance (LISREL)</i> 145	
4.6 Traitement de données sur R .....	148
4.6.1. <i>Le modèle de mesure de la variable latente A</i> .....	149
4.6.2. <i>Le modèle de mesure de la variable latente E</i> .....	158
4.7 Conclusion .....	164
<b>CHAPITRE 5 – APPLICATION DU MODELE DE L'APPRENTISSAGE ASYNCHRONE .</b>	<b>165</b>
5.1 Introduction.....	165
5.2 Identification du modèle de structure de la recherche .....	165
5.2.1. <i>1<sup>er</sup> cas de modèle</i> .....	166
5.2.2. <i>2<sup>e</sup> cas de modèle</i> .....	169
5.2.3. <i>3<sup>e</sup> cas de modèle</i> .....	172
5.2.4. <i>4<sup>e</sup> cas de modèle</i> .....	174
5.2.5. <i>5<sup>e</sup> cas de modèle</i> .....	178

5.2.6. 6 <sup>e</sup> cas de modèle .....	181
5.2.7. Modèle identifié.....	184
5.3 Le modèle de structure avec les paramètres estimés .....	184
5.4 Estimation du résultat à partir de la variable latente C.....	188
5.4.1. Valeurs des variables latentes A, E et C.....	188
5.4.2. Relation entre le résultat R <sub>r</sub> et le résultat du modèle R <sub>m</sub> .....	189
5.4.3 Résultat pédagogique par rapport à la Cognition selon le modèle .....	191
5.4.3. Valeur seuil selon les distributions du résultat réel et celles du résultat du modèle .....	194
5.5 Prévision de résultat pour l'année universitaire en cours .....	195
5.5.1. Analyse des traces d'activités des apprenants 2 <sup>e</sup> promotion FOAD GSA.....	195
5.5.2. Analyse de contenu des forums de discussions de la 2 <sup>e</sup> promotion FOAD GSA.....	198
5.5.3. Les variables de mesures de A et de E.....	203
5.6 Utilisation du modèle dans un cas général.....	205
5.6.1. Résultat en fonction de chaque variable .....	206
5.6.2. Croissance d'admission en fonction des variables .....	208
5.6.3. Faible variation d'admission même si la valeur des indicateurs de A augmente.....	209
5.6.4. Forts indicateurs pour la variable latente E .....	210
5.7 Discussion.....	211
5.7.1. Collaboration à distance dans un forum structuré.....	211
5.7.2. Intervention tutorale pour aider les apprenants à collaborer et à communiquer .....	212
5.7.3. Motivation, rôle de tous les protagonistes de la formation.....	214
5.8 Conclusion .....	214
CONCLUSION GENERALE.....	216
ANNEXES.....	218
BIBLIOGRAPHIES.....	272
FICHE DE RENSEIGNEMENT .....	289

## NOTATIONS & ABBREVIATIONS

### 1. Minuscules latines

$\hat{r}_p$	Estimateur du coefficient de corrélation
$a_i$	Paramètre à estimer
$e_i$	Indicateur de l'erreur de mesure
$\hat{p}$	Probabilité empirique
$r_{xy}$	Coefficient de corrélation de l'échantillon de deux variables $x$ et $y$
$s^2$	Variance de l'échantillon
$\bar{x}$	Moyenne de l'échantillon
$x_i$	Observation de rang $i$
Cov	Covariance
log	Logarithme à base 10
max	Valeur maximale
min	Valeur minimale
n	Nombre d'observations de l'échantillon
Var	Variance
$k$	Degré de liberté

$\ln$	Fonction logarithmique naturelle
$p$	Nombre d'indicateurs de mesure ou variables observées
$tr(.)$	Fonction trace d'une matrice
$x$	Variable
$y$	Variable
$s$	Ecart-type de l'échantillon

## 2. Majuscules latines

$N$	Nombre d'observations d'une population
$A_i$	Catégorie d'une population
$B_j$	Catégorie d'une population
$ S $	Déterminant d'une matrice $S$
$W_p$	Variable aléatoire de loi de Wishart
$\vec{X}$	Vecteur
$X_i$	Indicateur
$A$	Variable latent

$B$	Variable latent
$E$	Variable latent
$E(X)$	Espérance mathématique
$N$	Nombre de coefficients à estimer du modèle
$P$	Nombre d'indicateurs de mesure du modèle
$Q$	Vecteur colonne
$S$	Variable aléatoire / matrice de variance-covariance observée
$X$	Matrice
$Y$	Matrice

### 3. Minuscules grecques

$\hat{\sigma}_{xy}$	Estimateur de la covariance
$\gamma_i$	Indicateur de l'erreur de mesure
$\delta_i$	Indicateur de l'erreur de mesure
$\alpha_i$	Coefficient représentant l'effet du construit sur l'indicateur

$\beta_i$	Paramètre
$\lambda_i$	Coefficient représentant l'effet du construit sur l'indicateur
$\xi_1$	Construit latent
$\mu_x$	Moyenne de la population de la variable x
$\mu_y$	Moyenne de la population de la variable y
$\rho_{xy}$	Coefficient de corrélation de deux variables x et y
$\hat{\sigma}$	Estimateur de l'écart-type
$\sigma^2$	Variance de la population
$\mu$	Moyenne de la population
$\delta$	Erreur
$\theta$	Paramètre
$\lambda$	Valeur propre
$\epsilon$	Erreur
$\sigma$	Ecart-type de la population



#### 4. Majuscules grecques

$\chi^2$  Chi-deux

$\Sigma$  Variable manifeste / matrice de variances/covariances estimées

$\Gamma$  Fonction gamma

#### 5. Notations spéciales

$1_A(x)$  Fonction indicatrice de  $A$  au point  $x$

♣ Début de démonstration

♦ Fin de démonstration

$\otimes$  Produit de Kronecker

$( )^*$  Opérateur conjugué

$( )^T$  Opérateur transposé

#### 6. Abréviations

ATICE                    Acteurs en Technologie de l'Information et de Communication en  
Enseignement

AGFI                    Adjusted Goodness of Fit Index

CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CLOM	Cours en Ligne Ouvert et Massif
CNTEMAD	Centre National de Télé-Enseignement de Madagascar
CPM-GOMS	Cognitive Perceptual Motor- Goals, Operators, Methods and Selection
CRAC	Centre de Ressources, d' Assistance et de Conseil
CSCL	Computer-Supported Collaborative Learning
Ddl	Degré de liberté
ECC	Effectif Cumulé Croissant
ECD	Effectif Cumulé Décroissant
ELT	Exploitation Logistique et Transports
FAD	Formation A Distance
FCC	Fréquence Cumulée Croissante
FCD	Fréquence Cumulée Décroissante
FOAD	Formation Ouverte et A Distance
GEAL	Guide de l' Etudiant pour l' Apprentissage en Ligne

GFI	Goodness of Fit Index
GLS	Generalized Least Squares
GPL	General Public License
GSA	Génie du Système Automatique
GOMS	Goals, Operators, Methods and Selection rules
HTA	Hierarchical Task Analysis
KLM-GOMS	Keystroke-Level Model - Goals, Operators, Methods and Selection
IST	Institut Supérieur de Technologie
IST-T	Institut Supérieur de Technologie d’AnTananarivo
IST-D	Institut Supérieur de Technologie d’Antsiranana
LARTIC	Laboratoire d'Appui aux Recherches et Technologies de l'Information et de la Communication
LAVAAN	LAtent VAriable ANalysis
LMD	Licence Master Doctorat
MAD	Méthode Analytique de Description
MESupReS	Ministère de l’Enseignement Supérieur et de la Recherche Scientifique

MISA	Méthode d'Ingénierie des Systèmes d'Apprentissage
ML	Maximum Likelihood
MOOC	Massive Open Online Course
MOODLE	Modular Object Oriented Dynamic Environment
NFI	Normed Fit Index
NNFI	Non-Normed Fit Index
PLS	Partial Lest Square
PND	Plan National de Développement
Recre@sup	Réseau des Centres de Ressources pour l'Enseignement Supérieur
RMSEA	Root Mean Square Error of Approximation
RMR	Root Mean square Residual
SA	Système d'Apprentissage
SEM	Structural Equation Models
SRMR	Standardised Root Mean square Residual
TCI	Transit et Commerce International
TLI	Tucker-Lewis Index

UAN	User Action Notation
UNT	Universités Numériques Thématiques
VAE	Validation des Acquis en Entreprise

## LISTE DES FIGURES

<b>Figure 1.01 :</b>	Les constituants d'une séquence.....	19
<b>Figure 1.02 :</b>	Plateforme de la FOAD ELT de l'IST-T .....	23
<b>Figure 1.03 :</b>	Plateforme de la FOAD TCI de l'IST-D .....	24
<b>Figure 1.04 :</b>	Plateforme de la FOAD GSA de l'IST-T .....	25
<b>Figure 1.05 :</b>	Le module GEAL .....	28
<b>Figure 1.06 :</b>	Les différentes disciplines de la science de données .....	30
<b>Figure 2.01 :</b>	Modèle comportemental de Fechner .....	41
<b>Figure 2.02 :</b>	Le connectivisme [2.20] .....	47
<b>Figure 2.03 :</b>	Courbe de Gini ou courbe de concentration .....	56
<b>Figure 3.01 :</b>	Droite de régression.....	90
<b>Figure 3.02 :</b>	La population de la catégorie $A_i$ et de la catégorie $B_j$ .....	99
<b>Figure 3.03 :</b>	Spécification des relations entre les variables latentes et leurs indicateurs..	101
<b>Figure 3.04 :</b>	Représentation du modèle .....	105
<b>Figure 3.05 :</b>	Exemple illustratif d'un modèle causal (indicateurs réfléchitifs).....	107
<b>Figure 3.06 :</b>	Les étapes de la construction du modèle des équations structurelles .....	108
<b>Figure 4.01 :</b>	Le modèle de causalité triadique réciproque de Bandura [4.02] .....	126
<b>Figure 4.02 :</b>	Modèle structurel du SEM.....	128
<b>Figure 4.03 :</b>	Le modèle conceptuel de la recherche.....	129
<b>Figure 4.04 :</b>	Modèle de mesure de l'engagement de l'apprenant .....	137
<b>Figure 4.05 :</b>	Modèle de mesure de l'attitude dans l'apprentissage asynchrone.....	143
<b>Figure 4.06 :</b>	Modèle causal de la recherche .....	145
<b>Figure 4.07 :</b>	Modèle de mesure proposé pour l'attitude de l'apprenant .....	150
<b>Figure 4.08 :</b>	Modèle théorique après respécification .....	153
<b>Figure 4.09 :</b>	Modèle de mesure de la variable latente A avec les paramètres estimés .....	156
<b>Figure 4.10 :</b>	Modèle de mesure à estimer pour l'engagement de l'apprenant .....	159

<b>Figure 4.11 :</b>	Modèle de mesure de E avec les paramètres estimés .....	162
<b>Figure 5.01.</b>	Les six cas possibles pour notre modèle.....	166
<b>Figure 5.02.</b>	Modèle de structure, proposition 1 .....	167
<b>Figure 5.03.</b>	Modèle de structure, proposition 2 .....	169
<b>Figure 5.04.</b>	Modèle de structure, proposition 3 .....	172
<b>Figure 5.05.</b>	Modèle de structure, proposition 4 .....	175
<b>Figure 5.06.</b>	Modèle de structure, proposition 5 .....	178
<b>Figure 5.07.</b>	Modèle de structure, proposition 6 .....	181
<b>Figure 5.08.</b>	Modèle de structure de la recherche .....	187
<b>Figure 5.09.</b>	Comparaison du résultat par rapport à la variable latente C.....	190
<b>Figure 5.10.</b>	Traitement sur Excel.....	192
<b>Figure 5.11.</b>	Distribution du résultat pédagogique.....	194
<b>Figure 5.12.</b>	Extraits de données générées avec les valeurs estimées des variables A, E, C et R 205	
<b>Figure 5.13.</b>	Résultat en fonction de la collaboration .....	206
<b>Figure 5.14.</b>	Résultat en fonction de la communication .....	206
<b>Figure 5.15.</b>	Résultat en fonction de la motivation .....	207
<b>Figure 5.16.</b>	Résultat en fonction de la consultation des forums .....	207
<b>Figure 5.17.</b>	Résultat en fonction de la consultation des discussions .....	207
<b>Figure 5.18.</b>	Résultat en fonction de l'envoi et de la mise à jour de message .....	208
<b>Figure 5.19.</b>	Taux croissant d'admission en fonction des variables .....	208
<b>Figure 5.20.</b>	Taux d'admission avec la variable latente A .....	209
<b>Figure 5.21.</b>	Collaboration et motivation de l'apprenant .....	210

## LISTE DES TABLEAUX

<b>Tableau 1.01:</b> Scénario d’encadrement d’un module.....	21
<b>Tableau 4.01 :</b> Analyse d’une activité d’apprentissage : forum de discussion .....	124
<b>Tableau 4.02 :</b> Les variables opérées dans la recherche .....	147
<b>Tableau 4.03 :</b> Les paramètres à estimer dans la recherche .....	148
<b>Tableau 5.01 :</b> Les trois variables latentes de la recherche .....	189
<b>Tableau 5.02 :</b> Valeur seuil de C.....	191
<b>Tableau 5.03 :</b> Traitement de données sur Excel .....	193
<b>Tableau 5.04 :</b> Analyse de traces dans la première séquence .....	196
<b>Tableau 5.05 :</b> Analyse de traces dans la deuxième séquence .....	197
<b>Tableau 5.06 :</b> Analyse de traces dans la troisième séquence.....	197
<b>Tableau 5.07 :</b> Analyse de contenu des forums de discussions .....	203
<b>Tableau 5.08 :</b> Données après analyse de traces et analyse de contenus .....	203
<b>Tableau 5.09 :</b> Prévision du résultat GSA 2 <sup>e</sup> promotion selon le modèle.....	204



## INTRODUCTION ET PRESENTATION DU PROBLEME

### *Contexte et justification*

Dans le contexte actuel de la politique nationale dans le domaine de l'enseignement supérieur, le Plan National de Développement (PND) 2015-2019 prévoit dans son Axe Stratégique 4, le « programme 4.4. Enseignement supérieur innovant et adapté aux besoins de la société » ; dont le premier sous-objectif est de « 4.4.1.1. Promouvoir l'adéquation, la recherche, le développement, l'innovation et la promotion d'une université numérique ». Aussi, la technologie de l'éducation qui est une de la consistance de la recherche s'inscrit de manière optimale dans la politique nationale en matière d'enseignement supérieur et de recherche.

Ainsi, au fur et à mesure que le système LMD a progressé dans son implémentation, il s'est avéré primordial d'étendre le champ d'action des établissements d'enseignement supérieur suite à la limitation du *numerus clausus* à l'entrée. La solution à la fois pertinente et innovante réside dans l'exploitation de l'espace virtuel, à travers la formation à distance avec un dispositif en ligne. C'est la raison pour laquelle le recours à la formation ouverte et à distance (FOAD) est incontournable pour l'enseignement supérieur malgache.

Distance et technologies ont profondément modifié les modes de médiation et de médiatisation des dispositifs de formation. L'introduction de technologies dans un dispositif de formation y amène tout le potentiel de l'apprentissage, de la communication et de l'organisation du travail à distance. Si la distance implique les technologies, ces dernières permettent en retour d'articuler la distance et la présence. Ces différentes innovations peuvent s'enrichir des travaux en formation à distance sur le suivi de l'apprenant et la centration sur l'activité [1.01]. Elles ont suscité de nouvelles pratiques d'enseignement, dont l'encadrement à distance qui assure les catégories de fonctions tutorales identifiées dans la littérature et synthétisées dans une étude de Said Berouk et Alain Jaillet [1.02].

L'appropriation des méthodes et techniques de l'apprentissage à distance n'est pas évidente pour les étudiants. De plus, la plupart des tuteurs n'aident pas les apprenants à résoudre les problèmes d'ordre technique [1.03]. Pour combler toutes ces lacunes, un module d'entrée intitulé « Guide de l'étudiant pour l'apprentissage en ligne » (GEAL) a été conçu et introduit dans deux formations ouvertes et à distance. Il a pour objectif d'accompagner les apprenants à la prise en main de la

plateforme et de fournir les outils, les conseils et méthodes nécessaires pour réussir dans une formation ouverte et à distance. A partir de toutes ces constatations, l'idée d'exploiter principalement cet instrument a été avancée. Le renforcement de l'apprentissage dans des dispositifs de FOAD est envisagé, et cela à travers une modélisation de l'ingénierie d'apprentissage.

C'est dans ce cadre qu'une recherche a été menée et basée sur l'analyse de l'apprentissage en ligne. Cette recherche relève des sciences cognitives et plus particulièrement du domaine de la science de l'ingénierie. Les approches, les méthodes, les outils et les instruments se contextualisent aux modélisations cognitives utilisant la science de données (Data sciences), la psychologie, la sociologie et l'ingénierie.

Dans cette recherche, la modélisation se concentre sur la gestion de connaissance des apprenants et l'analyse des données sur la plateforme. Autrement dit, comment établir un modèle d'équation structurelle (SEM) d'apprentissage, qui pourrait être appliqué à toute sorte de formation ouverte et à distance, selon l'analyse des traces d'activités d'apprentissage sur une plateforme ?

### ***Problématique***

La problématique porte sur le système d'apprentissage à distance. Nous voulons étudier le comportement de l'apprenant dans un forum de discussion et savoir si dans un enseignement utilisant un dispositif en ligne, l'engagement dans une interaction entre étudiants peut être valorisé.

- Comment soutenir les étudiants dans une FOAD ?
- Quel modèle d'apprentissage asynchrone pourrait constituer un outil d'amélioration de l'efficacité pédagogique ?

### ***Objectifs de la recherche***

La recherche a pour objectif général de présenter un modèle d'apprentissage en ligne. Pour l'atteindre, quelques objectifs spécifiques ont été identifiés :

- Analyser d'abord la trace des activités numérique d'apprentissage asynchrone et le contenu des forums de discussions
- Ensuite, élaborer un modèle d'apprentissage asynchrone fondé sur les relations structurelles linéaires du SEM.

Le système d'apprentissage montre la stratégie d'apprentissage de l'institution de formation et les valeurs auxquelles elle se rattache en matière de support à l'apprentissage. Il permet d'identifier les moyens qui assurent l'apprentissage au sein d'un dispositif de formation à distance. Ainsi, définir une stratégie d'apprentissage dès le départ permet d'assurer le bon déroulement de la formation et d'avoir un résultat satisfaisant. De même, la disposer d'un scénario d'apprentissage adapté à la formation pourrait aider les étudiants à bien suivre une formation ouverte et à distance.

La recherche se joint celle de Jacques Rodet, qui a écrit « *L'ingénierie tutorale* » [1.04] ; mais, concerne l'ingénierie d'apprentissage. Elle est plutôt axée du côté apprenant, et a pour but de modéliser le système d'apprentissage en ligne.

### ***Méthodologies***

Les méthodologies de recherche adoptées ont été basées sur l'analyse des traces numériques laissées dans les forums de discussion sur la plateforme, et sur le modèle d'équation structurelle (SEM) de l'apprentissage asynchrone. Dans ce cadre, plusieurs indicateurs peuvent être construits pour observer les processus sociaux, cognitifs et métacognitifs mis en œuvre. En effet, des variables comme le « volume de participation individuelle », la « fréquence » et même la « densité » des interventions peuvent être mesurées pour rendre compte de la situation étudiée. Pour ce faire, il a été question de procéder à une analyse de contenu susceptible de caractériser l'interactivité virtuelle entre les protagonistes de la formation et répondre aux questions que nous nous sommes posées.

L'asynchrone a été davantage exploitée dans le cadre de cette recherche. Le logiciel scientifique R a été utilisé pour l'analyse statistique des données. Trois formations ouvertes et à distance ont été choisies comme terrain, et constituent les corpus à étudier. La collecte de données s'est déroulée pendant les années 2016, 2017 et 2018. L'intérêt général dans le domaine concerné est l'acquisition d'un modèle d'équation structurelle d'apprentissage asynchrone pour une FOAD. L'élaboration du modèle et la mise en œuvre de l'application se sont déroulées en 2017, tandis que la simulation a été réalisée en 2018. Quatre articles ont été publiés et trois actions de communication ont été effectuées pour le partage des résultats de la recherche.

### ***Résultats attendus***

Le résultat global attendu de la recherche couvre en priorité la stratégie innovante au niveau de l'apprentissage à distance. Les résultats spécifiques sont :

- Un modèle d'équation structurelle d'apprentissage asynchrone dans une FOAD est disponible et,
- L'application partielle de ce modèle dans une formation ouverte et à distance en cours de l'année universitaire 2017/2018 est mise en œuvre.

Cet ouvrage est organisé en trois parties qui sont chacune constituées de deux chapitres :

Un état de l'art sur la modélisation cognitive est exposé dans la première partie. Le premier chapitre est consacré à la formation ouverte et à distance ainsi qu'à la méthodologie d'analyse de traces numériques sur une plateforme de formation. Tandis que le deuxième chapitre concerne l'état de l'art sur les modèles d'apprentissage en ligne.

La deuxième partie développe le modèle d'équation structurelle (SEM) proprement dit de l'apprentissage en ligne selon l'analyse des traces numériques d'activités dans les trois formations, terrains de la recherche. Le troisième chapitre expose la théorie de base sur l'analyse statistique de l'apprentissage étudié et le SEM ; et le quatrième chapitre va sortir deux modèles de mesure d'une équation structurelle de l'apprentissage selon les analyses déjà effectuées.

La troisième partie est consacrée aux applications. A cet effet, l'utilisation des deux modèles conçus pour sortir le modèle de structure sera décrite dans le chapitre cinq. La simulation du modèle dans trois formations ouvertes et à distance permet de valider un module proposé par la recherche et pourra être adapté à toute formation ouverte et à distance. L'annexe complète les théories et les analyses de données non développées dans les parties précédentes. Il contient aussi les publications de notoriété internationale et de notoriété nationale, fruits de travaux effectués durant les trois années de recherche.

Comme l'apprentissage ne peut pas se séparer de la pédagogie, la recherche sur l'apprentissage à distance ne peut pas non plus être effectuée sans des études préalables sur son tutorat en ligne. C'est la raison pour laquelle l'article intitulé « *les pratiques tutorales et attentes des apprenants dans une FOAD* » [1.05] a été publié, permettant ainsi de mettre en évidence les attentes des apprenants sur l'accompagnement des enseignants lors de leur apprentissage en ligne.

# **CHAPITRE 1 - ETAT DE L'ART SUR L'APPRENTISSAGE ASYNCHRONE DANS UNE FORMATION OUVERTE ET A DISTANCE**

## **1.1 Introduction**

Le numérique ne cesse d'évoluer et son usage est inévitable partout, y compris dans le domaine de l'éducation. En effet, les technologies de l'information et de la communication (TIC) offrent d'immenses possibilités pédagogiques pour l'appropriation du savoir. Conscients de l'enjeu des TIC pour le développement du système d'enseignement, plusieurs établissements utilisent la formation à distance comme système de développement du secteur de l'enseignement supérieur à Madagascar. Les uns utilisent des ressources pédagogiques numériques en ligne tout simplement et, les autres recourent aux plateformes numériques et s'en servent comme support de la formation. Deux instituts supérieurs malgaches de technologie utilisent un dispositif en ligne pour leur formation ouverte et à distance (FOAD). Ce dernier a été choisi comme terrain de la recherche.

Le premier chapitre expose les théories de la science de données, de l'ingénierie de connaissance et, de l'apprentissage asynchrone dans une formation ouverte et à distance. C'est une étape indispensable avant d'aborder l'état de l'art sur les modèles d'apprentissage, l'analyse de traces numériques d'activités et la modélisation dans les sciences cognitives qui sont traités dans le second chapitre.

Ce chapitre est un développement de l'article de notoriété nationale citée ci-dessous :

- M. Rakotomalala, T. B. Ravaliminoarimalalason, F. Randimbindrainibe, « *Analyse de l'apprentissage d'un module d'initiation dans une formation ouverte et à distance* », MADA-ETI, [www.madarevues.gov.mg](http://www.madarevues.gov.mg), Vol.1, pp. 18 - 31, mai 2018.

## **1.2 Formation ouverte et à distance**

Qu'est-ce qu'une formation à distance ? Qu'est-ce qu'une formation ouverte ? Ces types de formation existent-ils à Madagascar ? Si oui, lesquels ?

Avant d'aborder la formation ouverte et à distance, il s'avère nécessaire de répondre d'abord à ces questions.

Historiquement, la formation à distance a été développée bien avant le déploiement d'Internet. Elle doit ses origines, notamment dans le monde anglo-saxon, davantage aux mouvements d'éducation populaire et à une politique de démocratisation de l'enseignement supérieur, qu'aux besoins de l'entreprise.

### ***1.2.1. Formation à distance***

#### *Définition 1.01 :*

Un enseignement à distance (EAD), selon le réseau africain de formation à distance (RESAFAD) est « une situation éducative dans laquelle la transmission des connaissances et les activités d'apprentissage se situent en dehors de la relation directe en face à face (ou dite en " présentiel ") entre l'enseignant et l'enseigné ». L'enseignement à distance a largement bénéficié des apports offerts par les technologies de l'information et de la communication.

Selon Daniel Peraya [1.06], la formation à distance est définie comme une formation médiatisée. Celle-ci fait appel à des artefacts technologiques, et à des dispositifs médiatiques. « C'est ainsi que seront mis à disposition des apprenants les ressources pédagogiques, les différents outils de communication et de collaboration, les activités d'apprentissage et les moyens de les réaliser. C'est à travers de tels dispositifs que les apprenants communiqueront et collaboreront, qu'ils se verront aidés dans leurs travaux, que les tâches qu'impose toute formation seront réalisées. »

Dans cette recherche, ce dispositif s'appuie sur l'usage d'une plateforme Moodle. Il a été conçu pour favoriser un cadre de formation socioconstructiviste et intègre des outils et des activités d'apprentissage orientés sur les communications synchrones et asynchrones. Par ailleurs, avec Moodle, l'apprenant est au centre de la formation. Il suit un parcours pédagogique qui favorise l'apprentissage, lequel est perçu comme une activité mentale réalisée individuellement à partir d'informations à traiter [1.07].

### ***1.2.2. Formation ouverte***

La formation ouverte est une stratégie d'apprentissage qui en principe permet à l'apprenant de décider de son parcours d'apprentissage et de le personnaliser. L'apprenant choisit ses objectifs d'apprentissage, les compétences visées et le chemin de son apprentissage. Cette personnalisation doit permettre à l'apprenant d'atteindre ses objectifs dans un temps, lieu et contexte qui lui conviennent.

La formation ouverte possède de nombreux avantages. Elle permet tout d'abord de développer l'autonomie de l'apprenant. Celui-ci apprend à s'organiser selon ses contraintes personnelles ou professionnelles, et à gérer son apprentissage par ses propres moyens. L'apprenant motivé saura toujours retirer des formations en ligne le meilleur qu'elles peuvent offrir : programme de formation 'à la carte', échanges d'expérience, facilité d'apprentissage et progression à son propre rythme.

Souvent les initiatives de formation ouverte sont liées aux dispositifs de formation à distance ou e-learning et soutenues par les technologies de l'information et de la communication. Cette configuration techno-pédagogique est généralement appelée formation ouverte et à distance.

### ***1.2.3. Diversité de formation à distance à Madagascar***

Même si la formation à distance n'est pas encore généralisée dans l'enseignement supérieur à Madagascar, les cas non-exhaustifs suivants reflètent la diversité de formation utilisant des ressources et / ou dispositifs en ligne.

Certains établissements donnent en priorité à leurs étudiants et enseignants l'accessibilité aux ressources pédagogiques, et mettent ainsi en place un dispositif facilitant la diffusion et le partage de ces ressources.

Dans le cas de la Faculté de médecine d'Antananarivo, le Laboratoire d'appui aux recherches et technologies de l'information et de la communication (LARTIC) utilise la plateforme de gestion de ressources Moodle et l'outil de web-conférence pour donner aux étudiants l'accès aux cours dispensés par les enseignants. Le Laboratoire travaille également en collaboration avec les universités numériques thématiques (UNT) comme l'UNF3S et l'UMVF ; les sites miroirs de ces UNT sont installés sur le serveur de LARTIC.

C'est également en vue d'assurer l'accessibilité aux RPN que l'Université d'Antsiranana a mis en place une plateforme Moodle, permettant d'établir un lien vers d'autres UNT et d'orienter les étudiants vers des MOOC susceptibles de les intéresser.

L'Ecole Nationale d'Informatique (ENI) de Fianarantsoa utilise les plateformes Moodle et EMAEVAL (supports pour les évaluations) comme espace numérique de travail de leurs étudiants. Ces plateformes ont comme fonction l'accompagnement pédagogique des étudiants pour un taux de réussite optimal. Malheureusement, l'ENI n'a utilisé ces plateformes que pendant seulement deux années (2012-2013), en raison de la mauvaise qualité de la connexion et d'un problème d'accès à la plateforme. Durant ces deux premières années l'ENT s'est étendue dans toute l'université. A cette époque, l'ENI avait une formation hybride (25% en présentiel et 75% à distance), afin de satisfaire la demande en formation continue.

L'IST d'Antananarivo a implémenté des modules d'appoint pour ses élèves-ingénieurs sur une plateforme Moodle. Ces modules ont été mis en ligne et tutoré à distance dans le but de résoudre les problèmes de salles et de mobilité d'étudiants, lesquelles sont majoritairement des salariés. L'extension des six premiers modules est en cours actuellement, et l'ensemble des UE avec le système VAE qui sont déjà en place devraient pouvoir assurer plus tard une année de formation complète qui serait la première année en formation d'ingénieur des ISTs.

La Faculté de droit, économie et gestion de l'Université de Toliara dispose également d'un système de formation à distance en partenariat avec un Centre privé à Antsirabe qui est le Centre de Ressources, d'Assistance et de Conseil (CRAC). Une convention a été établie entre les deux entités et permet à des étudiants éloignés de Toliara de suivre une formation diplômante encadrée par des tuteurs. Les enseignants de la Faculté DEG sont rémunérés pour fournir des ressources qu'ils envoient par courriel ou sur CD. Ces ressources sont mises en forme par le Centre qui diffuse ensuite des cours photocopiés aux étudiants. Le Centre assure le suivi de ces étudiants à travers un système de tutorat.

L'IHSM Toliara, en partenariat avec des universités belges francophones, a œuvré pour un pôle de formation spécialisé en écotourisme et bio-management. Dix modules ont été produits pour le site web [www.madaecotourismemarin.be](http://www.madaecotourismemarin.be). Des cours en ligne ont été créés et mis en ligne sur la base du volontariat, sans contrepartie pour les enseignants.



Anthropomada.com est un site web dédié à la formation en M1 et M2 en anthropologie sociale dispensée par l'Université de Toamasina. Il met ses ressources numériques à disposition pour tous les étudiants de Madagascar, y compris et surtout ceux qui se trouvent dans les zones enclavées. Des CD Rom contenant les cours ont été édités pour les personnes n'ayant pas accès à l'internet. La formation diplômante est payante mais les cours sont consultables gratuitement pour le grand public. Anthropomada.com est en partenariat avec les universités de Mahajanga et de Toliara. Le site a été développé en collaboration avec le CNAM (France).

Le département Gestion de la faculté DEGS de l'université d'Antananarivo a commencé sa formation à distance en 2014/2015. Les cours à distance ont été fournis sur une plateforme fait maison. L'enseignement est en présentiel sous forme de regroupement.

L'Université Privé de Madagascar (UPRIM) a lancé une formation à distance de même type que celle du CNTEMAD sur plusieurs parcours. Les cours sont donnés aux étudiants sous forme de documents photocopiés et le regroupement se fait chaque samedi après-midi. Ses cibles sont les travailleurs qui veulent avoir une promotion dans leur carrière, les personnes qui veulent avoir un diplôme mais qui ne peuvent pas joindre la formation en salle et les étudiants qui suivent des formations en parallèle. Il y a encore plusieurs universités privées qui offrent actuellement des formations à distance à Madagascar. Les cours sont fournis en numérique, soit sur CD ou flash-disque, soit sur plateforme. L'enseignement dirigé se fait sous forme de regroupement en présentiel, car le tutorat en ligne présente trop de contraintes pour les enseignants : disponibilité, connexion, maîtrise de la prise en main, de la technique de communication, de la conception et de la mise à jour sur la plateforme.

Avec le concours de l'AUF, les instituts supérieurs de technologie (IST) d'Antananarivo et d'Antsiranana ont mis en place des formations ouvertes et à distance dans des parcours en licence professionnelle (L3). Dix-neuf (19) modules en Exploitation Logistique et Transports et treize modules (13) en Transit et Commerce International ont été développés pour ces deux FOAD. La majorité des étudiants sont issus d'autres pays de l'Afrique francophone. Les enseignants concepteurs et tuteurs des modules ont bénéficié d'une série de formations sur la technologie éducative telle que la conception de cours en ligne et l'encadrement à distance. Leurs expériences du numérique ont induit de nouvelles pratiques dans l'enseignement. L'année 2016/2017, une troisième FOAD en Licence Génie du Système Automatisé (GSA) a recruté sa première promotion. Cette formation a comporté 23 modules et a suivi la structure des deux autres FOAD.

C'est dans ces ISTs que la recherche a été menée. La trace numérique des activités et des interventions des apprenants sur la plateforme a été analysée.

#### ***1.2.4. Formation ouverte et à distance***

Si la formation à distance existe depuis des décennies, la FOAD (formation ouverte à distance) est un mode d'enseignement apparu dans les années 1990 qui rompt avec les unités de lieux et de temps en s'appuyant sur les technologies numériques. Dans les dispositifs de FOAD, formateur et apprenant sont éloignés l'un de l'autre.

##### *Définition 1.02 :*

Selon la définition de « Le B.A.BA de la FOAD – Forum Français pour la Formation Ouverte et à Distance (FFFPD) », une *formation ouverte et à distance* est un dispositif :

- Organisé, finalisé, reconnu comme tel par les acteurs ;
- Qui prend en compte la singularité des personnes dans leurs dimensions individuelle et collective ;
- Et repose sur des situations d'apprentissage complémentaires et plurielles en termes de temps, de lieux, de médiations pédagogiques humaines et technologiques, et de ressources.

##### *Définition 1.03 :*

Selon l'AFNOR, une *formation à distance* est un système de formation conçu pour permettre à des individus de se former sans se déplacer sur le lieu de formation et sans la présence physique d'un formateur. La formation à distance recouvre plusieurs modalités (cours par correspondance, e-learning ...) et est incluse dans le concept plus général de Formation Ouverte et à Distance.

### 1.3 Dispositif de formation

*Définition 1.04 :*

On appelle « dispositif » tout système de formation, selon une définition courante. Pour Nicolas Guichon [1.08], le *dispositif* est un objet de formation alliant une logique interne centrée sur l'apprentissage en ligne et une logique externe avec un accompagnement pédagogique.

*Définition 1.05 :*

Daniel Peraya définit de manière très précise ce que représente un dispositif : « *Un dispositif est une instance, un lieu social d'interaction et de coopération possédant ses intentions, son fonctionnement matériel et symbolique enfin, ses modes d'interactions propres. L'économie d'un dispositif – son fonctionnement – déterminée par les intentions, s'appuie sur l'organisation structurée de moyens matériels, technologiques, symboliques et relationnels qui modélisent, à partir de leurs caractéristiques propres, les comportements et les conduites sociales (affectives et relationnelles), cognitives, communicatives des sujets.* » [1.09]. On voit donc très clairement dans la seconde partie de cette définition prise en compte de moyens non technologiques, notamment relationnels, ayant un impact sur les comportements sociaux, psychologiques et cognitifs.

*Définition 1.06 :*

De manière plus ciblée, Simmel et Blandin définissent le *dispositif pédagogique* comme « *un ensemble de moyens matériels et humains, correspondant à une forme de socialisation particulière destiné à faciliter un processus d'apprentissage* » [1.10]. Que ce soit avec les TICEs ou sans, on voit avec ces auteurs qu'on ne peut réduire le dispositif à un simple objet technique (Texte de Julien Da Costa). Da Costa fait aussi remarquer qu'un dispositif correspond à une définition particulière de l'enseignement (il s'inscrit dans un courant, une certaine conception de l'enseignement) et implique une organisation de l'espace et du temps spécifique.

### **1.3.1. Apprentissage**

#### *Définition 1.07 :*

L'apprentissage est un ensemble de mécanismes menant à l'acquisition de savoir-faire, de savoirs ou de connaissances. L'acteur de l'apprentissage est appelé apprenant.

#### 1.3.1.1.Apprentissage à distance

Apprentissage où l'apprenant peut suivre les activités à un endroit autre que celui où le cours est dispensé. L'apprenant ne doit pas se rendre en un lieu précis et peut même réaliser les activités d'apprentissage à un autre moment.

#### 1.3.1.2.Apprentissage synchrone

L'apprentissage synchrone implique d'étudier à l'aide de discussions. Ce style d'apprentissage ne peut se produire qu'en ligne. En étant connecté, on peut rester en contact avec les professeurs et les autres étudiants. Ce système permet aux étudiants de poser des questions à leurs enseignants ou à leurs camarades au moyen d'un système de messagerie instantanée.

#### 1.3.1.3.Apprentissage asynchrone

L'apprentissage asynchrone peut se produire hors ligne ou en ligne. Il implique que les supports pédagogiques et travaux à effectuer sont publiés sur le web, par email ou dans des forums sur une plateforme. Ce style d'apprentissage est intéressant pour les étudiants qui n'ont pas la possibilité d'accéder à une messagerie instantanée. Un des bénéfices de l'apprentissage asynchrone est qu'il permet d'étudier de manière autorégulée.

#### 1.3.1.4.Différences entre apprentissage synchrone et apprentissage asynchrone

Pour l'apprentissage synchrone, il faut suivre le cours dans un lieu, à une date et à une heure bien définis. Il est possible d'interagir avec les autres étudiants et l'enseignant ainsi que de voir immédiatement les réponses aux questions durant le cours.

Pour l'apprentissage asynchrone, il est possible de regarder la vidéo quand on le souhaite mais il n'est pas possible de discuter en live avec les autres étudiants. Ce genre de cours peut être donné hors ligne.

### **1.3.2. Système d'apprentissage**

*Définition 1.08 :*

Le système d'apprentissage indique la stratégie d'apprentissage de l'institution de formation et les valeurs auxquelles elle se rattache en matière de support à l'apprentissage. Il permet d'identifier les moyens qui assurent l'apprentissage au sein d'un dispositif de formation à distance.

« *Un système formel d'apprentissage est un ensemble de moyens matériels et humains, correspondant à une forme de socialisation particulière destinée à faciliter un processus d'apprentissage* » [1.08]. Toutefois, dans un dispositif techno pédagogique, il existe des aspects non technologiques : relationnels, sociaux et cognitifs. En prenant compte de ces aspects dits non technologiques, nous allons essayer de définir quelques éléments du dispositif.

### **1.3.3. La plateforme Moodle et la formation**

*Définition 1.09 :*

**Moodle** est une **plateforme** d'apprentissage en ligne (en anglais : Learning Management System ou LMS) sous licence libre servant à créer des communautés s'instruisant autour de contenus et d'activités pédagogiques.

La formation, corpus de la recherche se fait entièrement à distance via une plateforme Moodle. Les supports de cours et autres ressources pédagogiques des différents modules sont disponibles sur cette plateforme.

Les apprenants réalisent les activités proposées individuellement ou en équipe. Outre l'appropriation des contenus des cours et la consultation des ressources pédagogiques, les apprenants doivent réaliser les exercices et tests d'auto-évaluation ; et exploiter de façon effective les outils de communication et de collaboration intégrés dans la plateforme, qu'ils soient synchrones ou asynchrones.

La formation se déroule sur deux semestres. Au cours du deuxième semestre, les apprenants sont amenés à réaliser en équipe de trois ou de quatre un projet tutoré. Il s'agit d'un problème concret posé par une entreprise dans le domaine de la formation. Les apprenants ont la charge de mener

l'étude et de proposer une ou des solutions à ce problème. Ce projet constitue la base du cadre pratique de leur formation.

A la fin du deuxième semestre, les apprenants doivent également effectuer un stage de fin d'études en entreprise au bout duquel, ils ont à rédiger individuellement un mémoire et à le soutenir devant un jury, soit en présentiel, soit par visioconférence pour les apprenants étrangers.

#### ***1.3.4. Les acteurs et leurs interactions***

Cinq acteurs théoriques interagissent dans un système d'apprentissage : l'apprenant, l'informateur, le concepteur, le formateur et le gestionnaire.

##### *Définition 1.10 :*

Le *concepteur* est l'enseignant auteur / créateur du cours et des activités. Il prépare le scénario pédagogique à partir duquel le scénario d'apprentissage va être élaboré par les tuteurs. Pour la plupart des modules, les concepteurs même deviennent des tuteurs. Ce dernier assure le tutorat du module lors de l'apprentissage.

Le processus d'apprentissage est régi par un acteur appelé « apprenant » qui transforme en connaissances un ensemble d'informations. Par « information », on entend toute donnée, concrète ou abstraite, perceptible par les sens et susceptible d'être transformée en connaissance. La « connaissance » correspond à une information assimilée par une entité cognitive et intégrée par elle à son système cognitif dans un contexte et un usage. La transformation par l'apprenant de l'information en connaissance suppose une adaptation des schémas mentaux préexistants ou la création de nouveaux schémas, mais toujours intégrés à l'ensemble du système mental de l'apprenant. La connaissance ainsi créée est intégrée à un usage, dans la mesure où elle est utilisée dans des processus qui permettent à l'apprenant de mener des actions dans son environnement.

Les informations, points de départ du processus d'apprentissage, sont mises à la disposition de l'apprenant par un autre acteur appelé « informateur ». L'informateur peut être une personne ou un groupe de personnes intervenant directement, ou médiatisées sous la forme d'un livre, d'un vidéo, d'un logiciel ou par tout autre instrument ou média permettant de rendre disponible une partie du savoir sous forme d'informations utilisables pour l'apprentissage.

Par « savoir », nous entendons un corps de connaissances reconnu socialement. Il peut faire partie d'un domaine général de connaissances comme la physique ou l'administration ; ce peut être aussi un savoir spécialisé concernant l'utilisation d'un logiciel ou une méthode de travail particulière à une organisation.

Dans un système d'apprentissage interviennent généralement d'autres acteurs. Ce sont des « facilitateurs » du processus d'élaboration des connaissances à partir de l'information. On distingue trois autres acteurs, à savoir le concepteur, le formateur et le gestionnaire. Le concepteur régit le processus d'ingénierie du système d'apprentissage ; il construit et maintient en état de fonctionnement un système d'apprentissage (SA) qui intègre des sources d'informations (informateurs humains ou instruments didactiques), des moyens de communication, d'interaction et de collaboration entre les acteurs, ainsi que des mécanismes d'assistance et de conseil (formateurs humains ou awareness). Le formateur régit le processus d'assistance et de conseil pédagogique ; il facilite l'acquisition des connaissances par l'apprenant en le conseillant dans sa démarche et sur les interactions qui peuvent lui être utiles en fonction des scénarios d'apprentissage définis par l'acteur concepteur.

#### *Définition 1.11 :*

Le gestionnaire régit le processus de gestion pédagogique ; il facilite l'apprentissage en gérant les acteurs et les événements pour assurer le bon déroulement des processus, en fonction des scénarios définis par l'acteur concepteur. Soulignons que les acteurs doivent être nettement distingués des types d'intervenants (ou personnages) tels qu'ils sont définis par une organisation donnée, et également des individus (ou participants) qui en assument la fonction et les rôles. Par exemple, à la Télé-université, le professeur est un personnage qui assume principalement les rôles de l'acteur concepteur (moins la réalisation et la production des documents non imprimés), de l'acteur informateur (lors des séances de formation des tuteurs ou en réponse directe à certains apprenants) et de l'acteur gestionnaire (dans la supervision des tuteurs). Le tuteur, quant à lui, est un autre personnage qui assume à la fois les rôles de l'acteur formateur et ceux de l'acteur gestionnaire, sans en épuiser la liste néanmoins.

Dans la formation en question interviennent des enseignants concepteurs, des tuteurs, des coordinateurs et des apprenants. Les tuteurs assurent le suivi et l'encadrement des travaux des étudiants tout au long de la formation. Une grande majorité d'entre eux sont aussi des concepteurs

de cours. Les coordinateurs (administratif, pédagogique et technique) garantissent le bon déroulement de la formation et assurent les liaisons entre l'administration, les tuteurs et les apprenants. Le coordinateur administratif s'occupe de l'inscription et de tout ce qui relève de l'administration. Le coordinateur pédagogique est l'interlocuteur principal des apprenants tout au long de leur parcours. Le coordinateur technique assure le bon déroulement de la formation sur la plateforme. Il est celui qui est contacté par tous les acteurs en cas de problème technique.

Les apprenants sont au centre de la formation et réalisent les activités aussi bien individuellement qu'en collaboration.

#### **1.4 Descriptif global de la formation**

La formation en licence, objet de la recherche, est une version plus approfondie de celle offerte pendant les deux premières années d'études universitaires. Elle vise principalement à former des techniciens supérieurs spécialisés, des professionnels capables de gérer et de mettre en valeur la spécialité ; à atteindre un nouveau public, plus particulièrement les cadres et les agents de maîtrise des entreprises de la spécialité.

Cette formation s'adresse aux titulaires d'un diplôme de technicien supérieur de la spécialité ou d'un diplôme de niveau Bac + 2 avec expérience professionnelle de deux ans minimums dans le domaine.

La formation a pour objectif de développer les compétences dans le domaine au niveau du savoir-faire et du savoir-être.

##### ***1.4.1. Déroulement de la formation***

Des tuteurs assurent le suivi effectif des activités d'apprentissage. Trois regroupements présentiels et obligatoires sont néanmoins organisés au début, en cours et en fin de formation dans les Campus Numériques Francophones (CNF) du pays du candidat ou du CNF le plus proche. Le premier regroupement a pour objet de maîtriser les techniques d'exploitation de la plateforme. Il dure une semaine avant la rentrée et regroupe tous les apprenants. Les deuxième et troisième regroupements



durent quelques jours et ont lieu à la fin du premier et du second semestre. Ils sont consacrés aux examens de ces deux semestres.

#### ***1.4.2. Suivi de la formation***

Chaque apprenant doit disposer d'un appareil (ordinateur, ou tablette ou téléphone) connecté au réseau Internet. Les candidats sélectionnés ont à leur disposition les moyens techniques et humains déployés par l'AUF dans ses CNF. Plusieurs modules de formation peuvent être étudiés pendant le semestre avec un accompagnement pédagogique concrétisé par le tutorat, la correction des travaux proposés, les réunions synchrones (ou chats) selon un calendrier prédéfini, les forums de discussions et la participation aux activités proposées.

#### ***1.4.3. Evaluation des apprenants***

L'évaluation est formative d'une part et sommative d'autre part. L'évaluation formative porte sur les travaux individuels et d'équipe résultant des activités d'apprentissage mises en ligne. Cette évaluation est effectuée par l'enseignant-tuteur. Elle prend en compte la participation effective de l'apprenant au travail d'équipe, son assiduité aux rencontres synchrones, le respect des délais de dépôt des travaux, le travail individuel et le travail d'équipe. L'apprenant dispose de la latitude de s'autoévaluer en effectuant les différents quiz ou QCM disséminés dans les cours. Cette autoévaluation a pour but de réguler l'activité d'apprentissage des apprenants par orientation ou remédiation. L'évaluation sommative quant à elle, est organisée sous forme d'examen en présentiel. Cet examen est individuel et porte sur les connaissances théoriques et pratiques acquises tout au long du semestre de la formation.

#### ***1.4.4. Descriptif détaillé d'un module de formation***

Chaque module de formation comporte un descriptif détaillé comme ci-dessous :

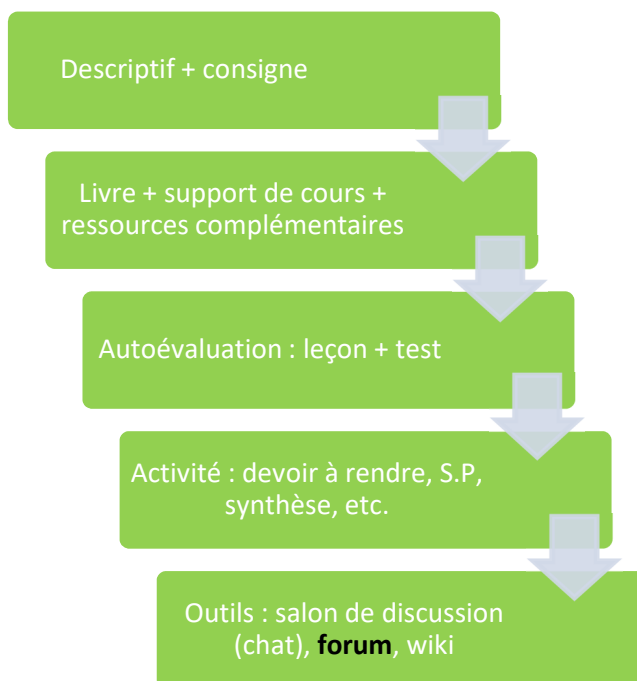
##### ***1. Titre du module : un titre concis, précis et révélateur du contenu du module***

2. **Durée du module :** *Dans cette section, les rédacteurs de modules doivent spécifier le nombre d'heures de travail que l'apprenant doit investir pour maîtriser les contenus de formation et réaliser les activités d'apprentissage et les auto-évaluations proposées dans le module.*
3. **Objectifs généraux du module :** *Des objectifs clairement écrits, instructifs, concis et compréhensibles sont extrêmement importants dans la réussite d'un module de formation ouverte et à distance. Les concepteurs de modules devront définir les objectifs généraux qu'ils assignent au module.*
4. **Objectifs spécifiques :** *Les rédacteurs de modules doivent identifier les objectifs spécifiques d'apprentissage pour le module. Chaque objectif spécifique sera au centre des activités d'enseignement/ apprentissage.*
5. **Glossaire des principaux concepts (à réaliser avec le module "Glossaire" de Moodle) :** *Cette section contient des définitions courtes et concises des termes utilisés dans la séquence. La section consacrée aux concepts clé est souvent appelée glossaire. Elle fournit aux apprenants des termes du module auxquels ils ne sont peut-être pas familiers.*
6. **Séquences :** *Donnez ici la liste des séquences pédagogiques prévues dans le module, en partant du principe que chaque objectif spécifique fera l'objet d'une séquence d'apprentissage indépendante. Chacune des séquences listées ici fera l'objet d'une description plus détaillée au moyen du canevas de présentation d'une séquence.*
7. **Activités d'apprentissage :** *Donnez ici la liste des activités d'apprentissage prévues dans le module. Chacune de ces activités fera l'objet d'une description plus détaillée au moyen du canevas de présentation d'une activité d'apprentissage.*
8. **Evaluation :** *Donnez ici un maximum d'indications sur la manière dont les apprenants seront évalués (type d'épreuves, durée, périodicité, pondération, etc.).*

Ce canevas est donné à tous les concepteurs de module lors de la formation sur la conception de contenus pédagogiques, dans la phase de projet. Il y a aussi d'autres canevas pour les activités (exercice ou étude de cas) ainsi que la SP ou situation problème.

### 1.4.5. Structure d'un module

Un module comporte deux ou plusieurs séquences. La figure ci-dessous illustre la composition d'une séquence.



**Figure 1.01 :** *Les constituants d'une séquence*

Dans une séquence, les apprenants doivent : connaître le descriptif et la consigne de travail de la séquence. Outre l'appropriation des contenus des cours et la consultation des ressources pédagogiques, les apprenants doivent s'autoévaluer pour vérifier leur compréhension en réalisant les activités sous forme de devoir à rendre. Ce devoir peut être une situation problème (SP), un exercice, une synthèse ou une étude de cas. Ils doivent exploiter de façon effective les outils de communication/collaboration intégrés dans la plateforme, le chat (synchrone), le forum de discussion et la messagerie (asynchrones).

#### 1.4.6. Scénario d'encadrement d'un module et scénario d'apprentissage

Le scénario pédagogique peut être défini selon l'université de Rouen, comme « le résultat du processus de conception d'une activité d'apprentissage, processus s'inscrivant dans un temps donné et aboutissant à la mise en œuvre du scénario. Dans un scénario, on trouve donc des objectifs, une planification des activités d'apprentissage, un horaire, une description des tâches des étudiants, des modalités d'évaluation qui sont définis, agencés et organisés au cours d'un processus de plan pédagogique. » [1.11].

Chaque enseignant a son propre scénario d'encadrement ou scénario pédagogique selon ses activités proposées. Il s'agit d'un travail préparatoire dans lequel le projet est décrit, raconté, découpé, structuré avant d'être réalisé. Pour le concepteur, le scénario pédagogique est l'occasion de formaliser dans un document les différents éléments et étapes nécessaires pour l'action de formation qu'il prépare. Le canevas standard qui a été donné lors d'un atelier de formation sur la conception de contenu pédagogique se présente comme suit :

**Module :**

**Nom du concepteur :**

Activités encadrées	Agent d'encadrement	Dimensions			Bénéficiaires
Activité « X »	Tuteur Co-apprenant ou Awareness	<b>Rôles Fonctions principaux</b>	et Pédagogique		Groupe plénier Equipe ou Individu
			Métacognitive		
			Technique		
			Organisationnelle		
			Evaluation		
			Motivationnelle		
			Socio-affective		

		<b>Contexte</b>	A l'entame de l'activité		
			En cours		
			A la fin		
		<b>Modalité</b>	Proactive		
			Réactive		
		<b>Temporalité</b>	Persistante		
			Ponctuelle		

Source : AUF

**Tableau 1.01:** *Scénario d'encadrement d'un module*

A partir de ce scénario pédagogique, le tuteur élabore un scénario d'apprentissage pour les étudiants. Le scénario d'apprentissage [1.12] représente la description, effectuée a priori (prévue) ou a posteriori (constatée), du déroulement d'une situation d'apprentissage ou unité d'apprentissage visant l'appropriation d'un ensemble précis de connaissances, en précisant les rôles, les activités ainsi que les ressources de manipulation de connaissances, outils et services nécessaires à la mise en œuvre des activités [1.11]. Le scénario d'apprentissage identifie, décrit et quantifie les différentes activités des apprenants tout en précisant leurs modalités d'intervention et les outils de communication qui les supportent.

#### ***1.4.7. Les dispositifs des trois formations ouvertes et à distance***

Les trois FOAD se développent depuis quelques années dans l'enseignement supérieur à Madagascar. Elles utilisent la plateforme Moodle comme support techno-pédagogique de la

formation, selon la 0Figure 1.03 : et la Figure 1.04 : . C'est un dispositif de formation ouverte et à distance, conçu avec un logiciel sous licence libre.

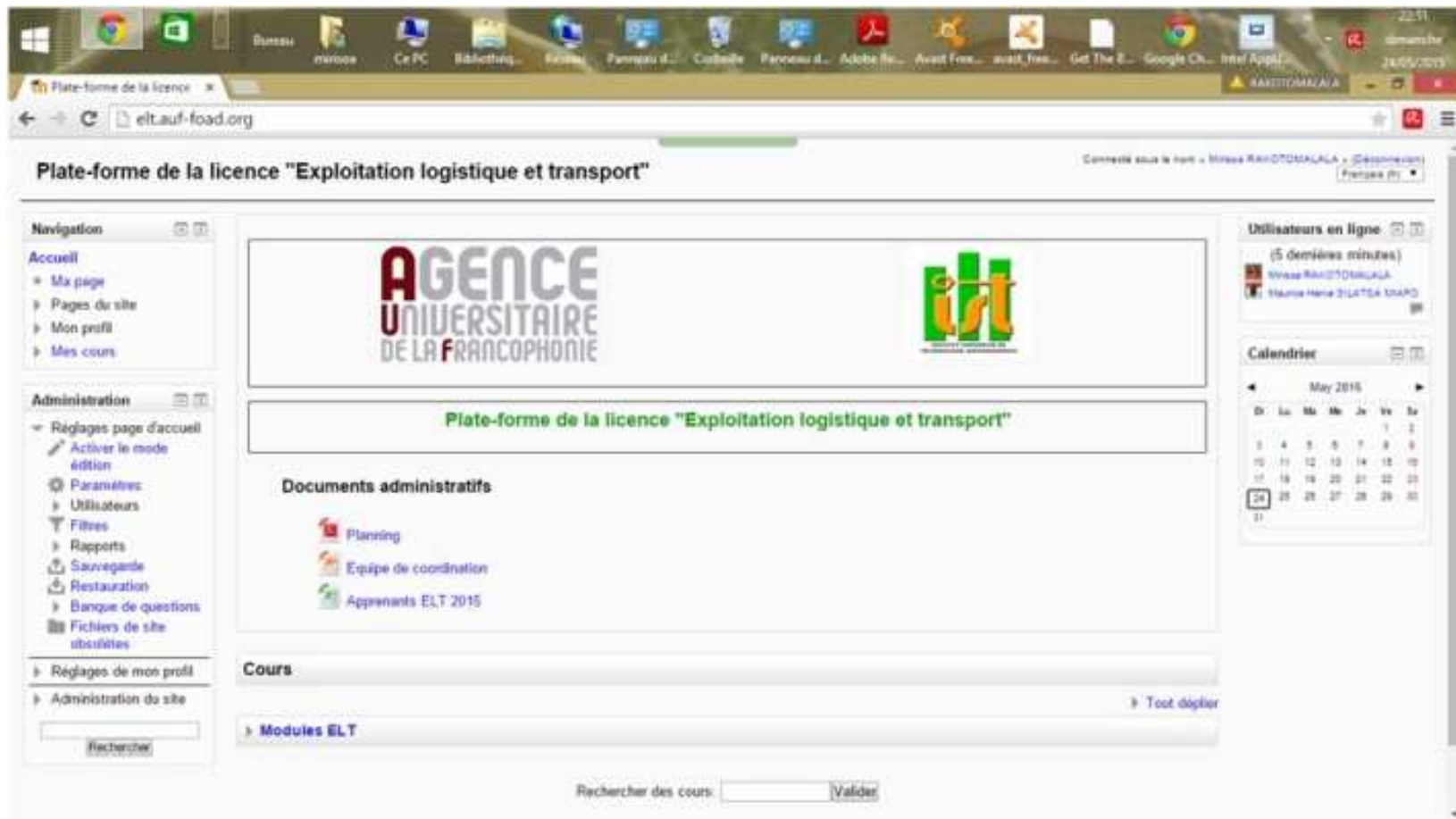


Figure 1.02 : Plateforme de la FOAD ELT de l'IST-T

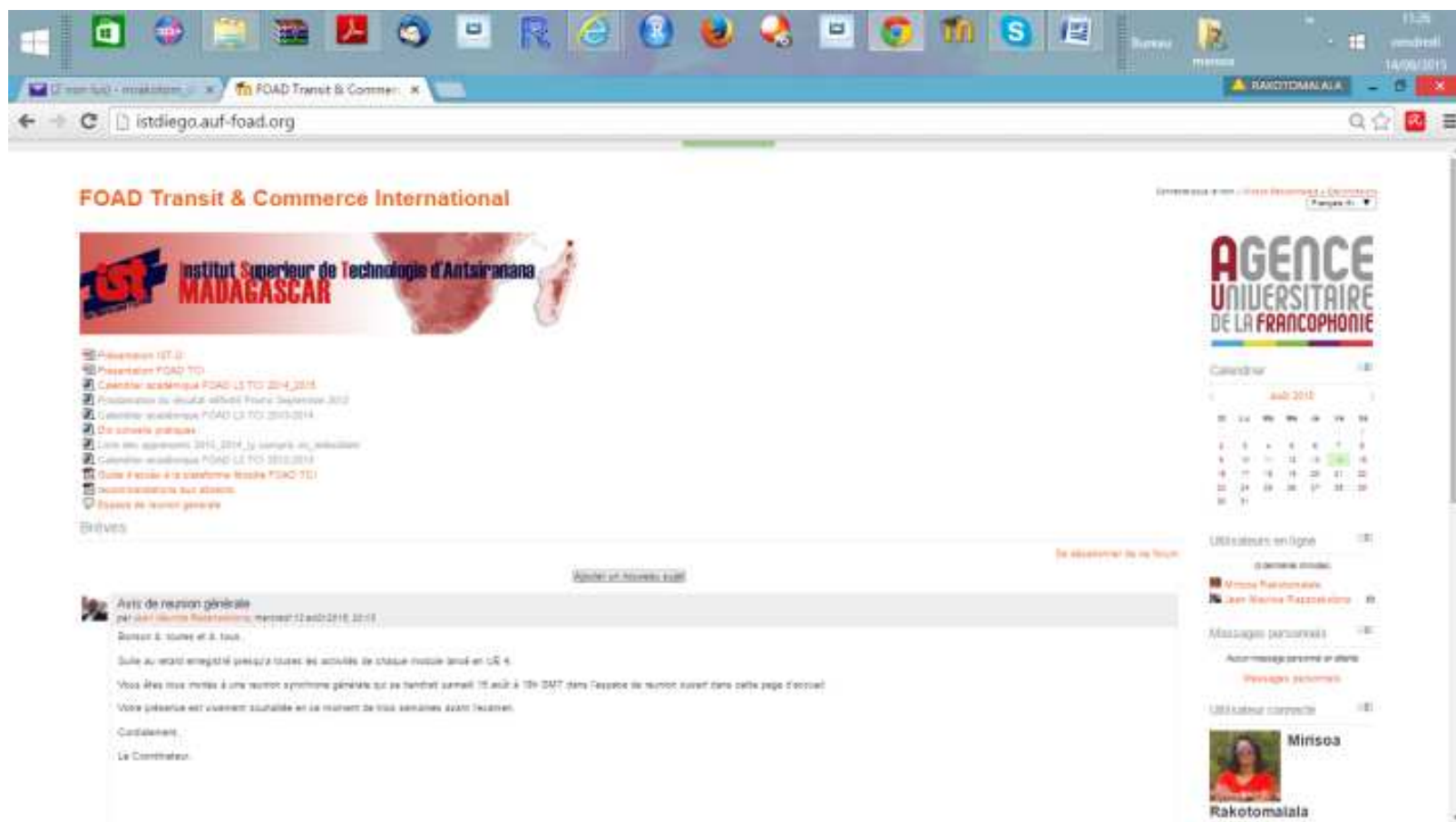


Figure 1.03 : Plateforme de la FOAD TCI de l'IST-D



PLATEFORME DE FORMATION - FOAD GSA/IST ANTANANARIVO

Français (fr) Mes cours 0 Mirisoa

Tableau de bord Cours Divers

Catégories de cours:

Divers

Rechercher des cours:

Valider

**Espace Tutorat**

**Dossiers administratifs**

**UE0-01 Guide de l'étudiant pour l'apprentissage en ligne**

Enseignant: Soloniaina RAKOTOARIJAONA

Enseignant: Mirisoa RAKOTOMALALA

Le module GEAL est destiné aux étudiants pour guider leurs premiers pas dans la FOAD. Ce module fournit les outils et méthodes nécessaires à la réussite de l'apprentissage à distance. Il se divise en 3 grandes parties : l'environnement d'apprentissage, les outils d'apprentissage à distance et le travail collaboratif.

**NAVIGATION**

- Tableau de bord
- Accueil du site
- Pages du site
- Mes cours
- Cours
  - Divers
    - ET
    - DA
    - GEAL
  - Semestre 5
  - Semestre 6
  - Ateliers de formation

Figure 1.04 : Plateforme de la FOAD GSA de l'IST-T

Les traces des activités sur cette plateforme constituent le cursus à étudier. Les dispositifs de formation donnent de nombreuses données dont le volume augmente quotidiennement. Comme le big data, le défi d'aujourd'hui n'est plus d'accéder à l'information mais de traiter cette masse d'information hétérogène qui arrive en flux continu durant une certaine période de la formation, en quantité énorme, et dans des formats divers. Cette recherche a alors pour objectif de traiter et d'exploiter les données des activités asynchrones afin de modéliser l'ingénierie d'apprentissage dans une formation ouverte et à distance.

### **1.5 Le module GEAL : Guide de l'Etudiant pour l'Apprentissage en Ligne**

En formation à distance, c'est-à-dire dans les cours où les étudiants sont séparés du professeur dans le temps et dans l'espace, les taux d'abandon des cours sont relativement élevés [1.13]. Ces taux d'abandon élevés seraient dus en partie aux difficultés particulières que les étudiants y rencontrent [1.14].

Lorsque le cours est offert en ligne et, selon la façon dont celui-ci a été conçu, d'autres difficultés peuvent s'ajouter. Selon Loisier [1.15], l'accès à des documents en ligne limite l'aptitude à interpréter les textes et à établir des connexions riches entre les notions. La non-linéarité peut aussi provoquer un sentiment d'égarement. Les étudiants peuvent se sentir perdus dans l'hyperespace et souffrir d'une surcharge cognitive lorsqu'ils passent du coq à l'âne. Selon Viau et Bouchard [1.16], l'étudiant doit avoir la perception de la pertinence de la tâche, de sa compétence à la réaliser et du contrôle nécessaire sur celle-ci pour pouvoir maintenir sa motivation à apprendre.

Viau [1.17] avance aussi que ce sont plutôt les considérations d'ordre pédagogique préalables à l'utilisation des TIC qui influencent surtout l'apprentissage. Par rapport à l'étudiant, ce sont surtout son engagement cognitif (approfondissement des activités) et sa persévérance (le temps consacré) qui lui permettent de bien réussir [1.15].

C'est dans ce sens que GEAL a été conçu. C'est un module sous forme de guide pour initier les connaissances des étudiants en matière d'e-learning. Il a la même structure que tous les modules de la formation. Il est destiné aux étudiants pour les accompagner lors de leurs premiers pas en formation à distance. Il peut également aider les étudiants confrontés à l'apprentissage en ligne sur une plateforme Moodle pendant une certaine période de leur cursus. Il a pour objectif de fournir les

outils, les conseils et méthodes nécessaires à la réussite de l'apprentissage en ligne. Il est composé de trois séquences.

GEAL fut lancé dans le cadre du premier regroupement de la formation. Ce regroupement s'est déroulé par visioconférence et les étudiants sont entrés par pays sur des plages d'une heure environ. Un mail de contact a été envoyé aux étudiants une semaine avant le regroupement. Il leur communiquait leur code d'accès à la plateforme et les instructions pour le remplissage du profil. Le jour du regroupement, les apprenants ont commencé à entrer dans GEAL, sous l'encadrement de deux ou trois tuteurs selon le nombre des apprenants.

The screenshot shows a web browser window with the URL `elt.auf-foad.org/course/view.php?id=86`. The page title is "UE0-00 : Guide de l'étudiant pour l'apprentissage en ligne". The user is logged in as "Mirisoa RAKOTOMALALA".

**Navigation Menu (Left):**

- Accueil
- Ma page
- Pages du site
- Mon profil
- Cours actuel
  - UE0-00 GEAL
    - Participants
    - Badges
    - Généralités
    - Séquence 1
    - Séquence 2
    - Séquence 3
- Mes cours

**Administration (Bottom Left):**

- Administration du cours
  - Activer le mode édition

**Main Content Area:**

## GUIDE DE L'ETUDIANT POUR L'APPRENTISSAGE EN LIGNE

**Responsable :** Mirisoa RAKOTOMALALA

**Concepteurs :** Mirisoa RAKOTOMALALA et Lova ZAKARIASY

L'objectif du module est de fournir les outils et méthodes nécessaires à la réussite de l'apprentissage en ligne.

- Forum des nouvelles
- Descriptif du module
- Introduction
- Organisation graphique du module
- Glossaire
- Présentez-vous
- Forum de discussions générales

**Right Sidebar:**

- Dernières nouvelles**
  - Ajouter un nouveau sujet...
  - 31 Jan, 09:40
  - Lova Zakariasy
  - Fin du module GEAL
  - Sujets antérieurs ...
- Événements à venir**
  - Aucun événement à venir
  - Aller au calendrier...
  - Nouvel événement...
- Activité récente**
  - Activité depuis le Wednesday 9 September 2015, 17:50
  - Rapport complet des activités récentes...
  - Rien de nouveau depuis votre dernière visite
- Utilisateurs en ligne**

Figure 1.05 : Le module GEAL

Le premier regroupement avait pour objectif de présenter le dispositif de formation (l'institut, la formation proprement dite, les différents responsables, le déroulement de la formation) et la prise en main de la plateforme (accès à la plateforme, modification du profil, les différents blocs, et le module GEAL). Après cette séance de regroupement, deux semaines environ ont été données aux apprenants pour prendre connaissance de la plateforme et maîtriser l'usage de l'environnement techno-pédagogique de la formation à travers le module GEAL : consultation des cours (sous forme de fichier PDF ou de page html) et des ressources complémentaires (sous forme de fichiers ou de liens), autoévaluations individuelles, dépôt de devoir individuel et travail collaboratif. GEAL restait ouvert par la suite. Les apprenants communiquaient entre pairs et avec les tuteurs en utilisant les outils de communication via la plateforme.

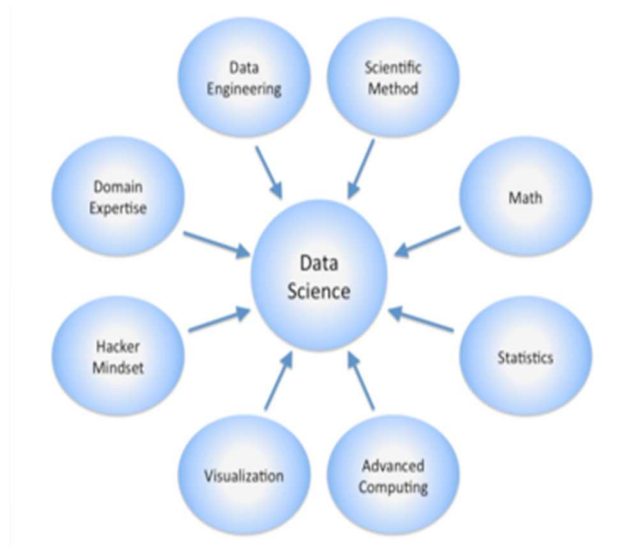
Puisque la recherche tend vers une modélisation cognitive de l'apprentissage dans une FOAD ; il s'avère nécessaire de procéder à l'analyse statistiques des données numériques du dispositif de la formation. Elle fait partie des sciences de données. Une présentation préalable du data science et de la formation à distance est une étape indispensable. Une approche pédagogique à distance a donc été développée.

## **1.6 Science de données et sciences cognitives**

### ***1.6.1.Science de données***

En général, la science de données (en anglais data science) est l'extraction de connaissance de données [1.18]. Elle est une discipline qui s'appuie sur des outils mathématiques, de statistiques, d'informatique et de visualisation des données. Elle est en plein développement, dans le monde universitaire, ainsi que dans le secteur privé et le secteur public. Moore [1.19] a défini la statistique comme la science des données.

La Figure 1.06 : montre les différentes disciplines de la science de données



Source : D. Calvin Andrus, Ph.D.

**Figure 1.06 :** *Les différentes disciplines de la science de données*

Ensuite, il a fallu explorer du côté de la neuroscience cognitive, car elle constitue la base de l'apprentissage des enfants.

### ***1.6.2. Sciences cognitives***

Les sciences cognitives regroupent des disciplines scientifiques dédiées à la description, l'explication, et le cas échéant à la simulation des mécanismes de la pensée humaine, animale ou artificielle, et plus généralement de tout système complexe de traitement de l'information capable d'acquérir, de conserver, d'utiliser et de transmettre des connaissances. Elles reposent donc sur l'étude et la modélisation de phénomènes aussi divers que la perception, l'intelligence, le langage, la mémoire, l'attention, le raisonnement, les émotions ou même la conscience. Les sciences cognitives utilisent conjointement des données issues d'une multitude de branches de la science et de l'ingénierie, comme la linguistique, l'anthropologie, la psychologie, les neurosciences, la philosophie et l'intelligence artificielle.

## 1.7 Neurosciences cognitives et facilité d'apprentissage

Contrairement à ce qu'envisageait Jean Piaget [1.20], l'enfant n'est pas dépourvu de compétences logiques abstraites. Bien au contraire, le cerveau de l'enfant est structuré dès la naissance, ce qui lui confère des intuitions profondes.

Il est doté de puissants et rigoureux algorithmes d'inférence statistique. En conséquence, l'école doit lui fournir un environnement enrichi : un enseignement structuré et exigeant, tout en étant accueillant et généreux [1.21]. Les neurosciences cognitives ont identifié quatre facteurs qui déterminent la facilité d'apprentissage.

En premier lieu, **l'attention** : elle fonctionne comme un projecteur, qui amplifie l'apprentissage, mais dont le rayon d'action est limité. Le plus grand talent d'un enseignant consiste donc à attirer, à chaque instant, l'attention de l'enfant sur le bon niveau d'analyse. *La présence de tuteur dans une FOAD pour assurer des fonctions correspondantes explique l'existence de ce facteur dans la formation.*

Deuxième facteur : **l'engagement actif**. Un organisme passif n'apprend pas. L'apprentissage est optimal lorsque l'enfant génère activement des réponses, et se teste régulièrement. L'auto-évaluation est donc une composante fondamentale de l'apprentissage, déjà identifié par Maria Montessori [1.22]. *Bien évidemment, chaque séquence comporte toujours un test d'auto-évaluation.*

Troisième facteur : **le retour d'information** (ou « feedback »). Le cerveau n'apprend que s'il reçoit des signaux d'erreur qui lui indiquent que son modèle interne doit être rectifié. L'erreur est donc non seulement normale, mais indispensable à l'apprentissage. *L'apprenant les reçoit, auprès des tuteurs ou de l'awareness.*

Quatrième pilier, enfin, **l'automatisation**. Au début de l'apprentissage, l'effort mobilise toutes les ressources du cortex frontal. Afin de libérer l'esprit pour d'autres tâches, il est indispensable que la connaissance devienne routinière. La répétition quotidienne va transférer l'apprentissage vers des circuits cérébraux automatiques et non conscients. Le sommeil fait partie intégrante de cet algorithme : dormir, c'est consolider les apprentissages de la journée.

Les deux facteurs, **l'engagement actif** et **l'automatisation**, seront pris en compte dans la recherche lors du traitement de données numériques recueillies sur la plateforme.

## **1.8 Ingénierie pédagogique**

L'ingénierie pédagogique s'appuie sur deux processus au cœur de la gestion des connaissances : d'abord l'extraction des connaissances que possèdent certaines personnes expertes dans leur domaine, ensuite l'acquisition des connaissances nouvelles grâce à l'apprentissage. C'est la transformation des informations en connaissances au moyen d'activités formelles ou informelles qui empruntent une variété de formés et de supports. Ce qui souligne évidemment l'importance de l'extraction des connaissances et leur modification.

Une définition de l'ingénierie pédagogique selon Gilbert Paquette est définie comme toute méthode de conception et de construction des systèmes permettant d'échanger, de partager et d'acquérir des informations dans le but de les transformer en connaissances, donc d'apprendre [1.23].

### ***1.8.1. Téléapprentissage***

Selon Gilbert Paquette, le téléapprentissage et son ingénierie sont l'avenir de la formation ; et la modélisation des connaissances est l'avenir de l'ingénierie du téléapprentissage. L'apprentissage par un individu consiste à transformer des informations en connaissances, à construire des connaissances à distance des sources d'information.

### ***1.8.2. L'ingénierie des systèmes d'apprentissage et les interactions***

Dans un système d'apprentissage, deux composantes jouent un rôle clé dans les interactions à savoir, le modèle des connaissances et le modèle pédagogique. L'analyse des acteurs d'un système d'apprentissage permet de préciser les différents types d'interaction.

### ***1.8.3. Le design pédagogique, une ingénierie des systèmes d'apprentissage***

Le design pédagogique est un processus complexe sur le plan des communications humaines, car il suppose l'apport de divers spécialistes : spécialistes du contenu faisant l'objet d'apprentissage, spécialistes du domaine de la pédagogie, spécialistes des médias et spécialistes de la gestion.



Mais, c'est d'abord et avant tout un processus complexe de résolution de problèmes tel qu'il est défini en sciences cognitives [1.24] et, parfois étudié comme tel en sciences de l'éducation [1.25] ; [1.26] ; [1.27].

On peut définir le **design pédagogique** comme une méthodologie visant à résoudre une classe particulière de problèmes : les problèmes de formation et d'apprentissage. Les problèmes de design pédagogique présentent des similitudes, qu'on travaille en architecture, en ingénierie d'un système physique, en génie logiciel ou en conception de système d'apprentissage. Dans tous les cas, la solution est un système à construire devant satisfaire certaines contraintes, très peu définies au départ, devant être spécifiées à la phase initiale d'analyse, puis précisées tout au long du processus. En observant les résolveurs à l'œuvre dans des problèmes de design, Goël et Pirolli [1.28] ont identifié un certain nombre de connaissances stratégiques invariantes, utilisées au cours du processus de résolution qui témoigne de la complexité de ce type de problème.

La méthode d'ingénierie des systèmes d'apprentissage (MISA) applique ces principes de résolution des problèmes de design aux processus d'ingénierie pédagogique. Elle innove par l'usage de techniques de modélisation cognitive à la fois pour la représentation des connaissances et celle des traitements pédagogiques et des traitements médiatiques. Ces trois dimensions du système d'apprentissage (SA) se distinguent nettement entre elles, mais sont aussi reliées par des associations précises. Par sa conception même, la méthode MISA fournit un cadre de référence bien défini mais suffisamment ouvert pour une étude de l'interactivité largement dégagée de son aspect médiatique.

#### ***1.8.4. Le concept de système d'apprentissage***

Le concept de système d'apprentissage produit par MISA est assez large. Un système d'apprentissage se compose des devis du système d'apprentissage, des matériels pédagogiques réalisés sur la base de ces devis et des infrastructures technologiques et organisationnelles qui le soutiennent. Le devis d'un système d'apprentissage se compose des éléments suivants :

- **un modèle des connaissances** représentant, sous forme graphique, le contenu du système d'apprentissage subdivisé en sous-modèles ; ceux-ci décrivent le contenu de chaque unité d'apprentissage et des instruments qui composent le matériel pédagogique ;

- **un modèle pédagogique** regroupant le réseau des événements d'apprentissage, notamment les unités d'apprentissage pour lesquelles on définit un ou plusieurs scénarios d'apprentissage qui regroupent les activités d'apprentissage et les ressources à la disposition des apprenants, ainsi qu'un ou plusieurs scénarios de formation qui réunissent les activités du formateur et les ressources à sa disposition ;

- **un modèle médiatique** définissant la manière selon laquelle les instruments sont regroupés en matériels pédagogiques sur différents types de supports, c'est-à-dire la structure de ces matériels : « storyboard », métaphore, unités et éléments médiatiques, règles d'organisation et de médiatisation, ainsi que les outils et les infrastructures technologiques qui permettent de les utiliser.

- **un plan de mise en place et d'entretien du système d'apprentissage** permettant de planifier la mise en place des infrastructures technologiques et organisationnelles requises pour diffuser le système d'apprentissage.

## 1.9 Approche d'apprentissage asynchrone dans une formation à distance

Apprendre consiste à acquérir de nouvelles connaissances, à les stocker en mémoire, à les organiser et à développer des automatismes. L'apprentissage est donc perçu comme une activité mentale réalisée individuellement à partir d'informations à traiter. En 2003, Entwistle [1.29] a identifié les différentes approches d'apprentissage chez les étudiants et a privilégié celles qui amélioreraient la qualité d'apprentissage. Il s'agit de l'approche d'apprentissage en profondeur, indiquant l'intention de comprendre par soi-même, l'approche métacognitive supposant pouvoir prendre du recul par rapport à ses manières d'étudier, et l'organisation des études pour gérer le temps efficacement et ménager ses efforts. Ces approches sont liées à **l'engagement de l'étudiant** dans la formation, ainsi qu'à sa **motivation** et ses **performances**, contrairement à une approche superficielle qui est plutôt liée à l'exploit à court terme.

Dans la modalité d'une formation à distance, l'apprentissage individuel nécessite aussi bien le soutien cognitif du tuteur, dans le sens où celui-ci aide l'apprenant à comprendre et l'oriente vers les informations pertinentes, que l'appui métacognitif et organisationnel ; afin de développer les approches d'apprentissages citées ci-dessus.

Dans une formation à distance, un tuteur devrait aider l'apprenant à développer ses compétences pour qu'il prenne conscience de la nécessité de planifier son apprentissage, d'organiser son travail, etc. [1.30] inciter les apprenants à auto-diriger leurs apprentissages, susciter des activités réflexives (méthodes de travail, évolution de l'apprentissage, acquisition de nouvelles compétences, etc.) ; et aider à identifier, à formuler et à surmonter les difficultés d'apprentissage. Ainsi, les tuteurs devraient assurer cette fonction métacognitive dans une FOAD [1.03].

Dans le type de formation asynchrone, l'échange avec les autres apprenants ou avec les tuteurs s'effectue via des modes de communication ne nécessitant pas une connexion simultanée. Il peut s'agir de forums de discussion ou bien encore d'échange d'e-mails. Par ailleurs, ce mode de formation repose souvent sur un apprentissage dit autodirigé, avec des cours, des exercices et des auto-évaluations impliquant une certaine autonomie de l'apprenant. La formation en ligne dans un environnement asynchrone implique que toute la responsabilité de l'apprentissage incombe aux apprenants. Ils prennent le contrôle de leur apprentissage et décident où, quand et comment apprendre. Mais cette flexibilité est-elle idéale pour réaliser les objectifs d'une telle formation ? Voyons les avantages et les inconvénients relatifs à l'apprentissage asynchrone.

### ***1.9.1. Contrôle absolu sur sa formation***

En tant que méthode centrée sur l'apprenant, l'apprentissage asynchrone donne aux étudiants l'entière responsabilité de leur expérience de formation en ligne. Ainsi, chacun peut décider du moment et de l'endroit où il veut suivre la formation, mais aussi de quelle manière. De plus, les problèmes relatifs à la distance et au temps sont éliminés. Puisque l'interaction apprenant/tuteur est effectuée selon des horaires individualisés, il est plus facile d'établir un suivi personnalisé de la part des facilitateurs ou des enseignants.

### ***1.9.2. L'apprentissage à son propre rythme***

L'apprentissage asynchrone offre la possibilité de réfléchir sur ce qu'ils ont appris avant de répondre aux questions ou de se joindre aux discussions sur forum. Parce que les apprenants assimilent de nouvelles informations de différentes façons, une solution de formation asynchrone peut être

bénéfique aux apprenants qui en ont le plus besoin. Il y a la possibilité de prendre le temps de répondre aux questions et de développer l'esprit critique.

### ***1.9.3. Très pratique***

L'apprentissage asynchrone est la solution idéale pour les apprenants avec un emploi du temps chargé. En effet, ils ne sont pas obligés d'être en ligne à un moment spécifique. Les apprenants peuvent communiquer avec le facilitateur en ligne ou les autres apprenants à leur propre convenance. Ils ont instantanément accès aux informations, aux travaux et aux ressources mis à disposition sur la plateforme.

### ***1.9.4. Moins d'obstacles sociaux***

Alors qu'il a été prouvé que l'interaction et la collaboration à distance contribuaient positivement à l'expérience de formation en ligne, il est aussi vrai que de nombreux apprenants ne cherchent pas forcément à sympathiser avec les autres. Certains ne sont pas toujours à l'aise à l'idée de participer à des discussions en ligne. Généralement, dans une discussion instantanée, les apprenants plus à l'aise dominent la conversation. Ainsi, l'approche de l'apprentissage asynchrone aide les apprenants introvertis à réduire leur anxiété sociale, puisqu'ils ne sont pas obligés d'être en groupe pour apprendre.

### ***1.9.5. Possibilité d'interaction malgré les barrières de lieu ou de temps***

Les méthodes d'apprentissage asynchrone permettent aux étudiants d'apprendre non seulement à leur rythme mais également d'interagir avec les autres apprenants et les tuteurs en ligne quel que soit leur fuseau horaire. Les forums de discussion et les e-mails via la plateforme servent à assurer une interaction efficace, la possibilité d'une collaboration en ligne pour les activités en groupe et le bon déroulement des conversations à distance.

#### ***1.9.6. Pas de feedback immédiat***

Le feedback dans le domaine du e-Learning est essentiel. En effet, il aide les apprenants à corriger les problèmes et les incompréhensions. Par conséquent, plus le feedback est rapide, plus vite les apprenants peuvent rectifier le tir et combler leurs lacunes. Dans un environnement d'apprentissage asynchrone, le feedback instantané est impossible et les apprenants risquent donc de perdre un temps précieux à attendre une réponse à leurs questions.

#### ***1.9.7. Sans interaction personnelle***

L'interaction personnelle entre les apprenants est éliminée dans le contexte de l'apprentissage asynchrone. L'absence d'une telle interaction peut non seulement entraîner une déconnexion entre les apprenants mais également avec le cours lui-même. Or, ce sentiment de déconnexion peut entraîner les apprenants à voir cette formation en ligne comme un fardeau. L'apprentissage individuel peut fonctionner pour certains, mais nombre d'apprenants ont besoin d'une interaction personnelle pour maintenir ou même renforcer leur motivation. De fait, l'impossibilité d'interagir avec d'autres apprenants peut entraver la réalisation des objectifs d'apprentissage et des acquis pédagogiques.

#### ***1.9.8. Pas de collaboration directe, ni d'activités en temps réel***

Apprendre à son propre rythme peut également signifier attendre que les autres répondent, souvent sur une période de temps relativement longue. L'apprentissage asynchrone n'offre pas la possibilité d'avoir une discussion ou des exercices de collaboration en direct. Ce sont pourtant deux éléments importants pour renforcer l'engagement et la motivation des apprenants. De plus, la communication générale entre les apprenants peut être entravée suite à ce sentiment d'isolement.

### ***1.9.9. Réduire la motivation***

Ce manque d'interaction peut désengager et démotiver les apprenants qui peuvent avoir besoin d'encouragement pour se connecter, lire les ressources et terminer le module en ligne. On remarque en effet que la procrastination est plus forte dans l'environnement d'apprentissage asynchrone que dans tout autre contexte pédagogique. L'interaction personnelle aide les apprenants à maintenir leur intérêt, alors que l'isolement renforce rarement la motivation.

### ***1.9.10. Nécessité d'un certain degré d'autodiscipline***

L'apprentissage asynchrone exige que l'apprenant soit concentré, déterminé et qu'il possède de bonnes compétences en gestion du temps. Réussir un apprentissage asynchrone nécessite de la part des apprenants un engagement et une discipline rigoureuse, ce qui peut être un réel inconvénient pour ceux qui ont du mal à se motiver.

Connaître les avantages et les inconvénients de l'apprentissage asynchrone peut aider à déterminer s'il s'agit de la meilleure méthode pour une formation à distance. Il ne fait aucun doute que l'apprentissage asynchrone est bénéfique, mais ses inconvénients doivent être pris en considération dans la recherche. Toutefois, il est tout à fait possible de réussir à développer un modèle pédagogique asynchrone.

## **1.10 Conclusion**

Les considérations qui précèdent ont leur importance pour identifier les types d'interaction qui peuvent être intégrés dans un système d'apprentissage. La recherche est basée sur les interactions dans lesquelles l'apprenant peut être impliqué au moment de sa formation. Ce sont les interactions entre pairs, entre les apprenants eux-mêmes, entre apprenants et tuteur et, entre apprenant et awareness. Ces interactions constituent une base de données importante sur la plateforme. Le traitement et l'analyse des données correspondantes permettent d'atteindre l'objectif de la recherche. En effet, l'analyse des traces d'activités de l'apprenant est l'une des méthodologies adoptées pour la recherche. Les théories de la science de données et de l'ingénierie de connaissance

ont été présentées, avant de passer à l'état de l'art sur les modèles d'apprentissage, à l'analyse des traces numériques d'activités d'apprentissage et à la modélisation dans la science cognitive.

## **CHAPITRE 2 – ETAT DE L’ART SUR LES THEORIES DE L’APPRENTISSAGE ET LA MODELISATION DANS UNE FORMATION A DISTANCE**

### **2.1 Introduction**

Un modèle prescriptif d'apprentissage à distance sera donné à l'issue de la recherche. Il représente " une solution " / adaptée au contexte malgache pour le suivi d'une formation en ligne.

Le présent chapitre expose l'état de l'art sur les modèles d'enseignement et d'apprentissage, l'analyse des traces numériques des activités d'apprentissage et la modélisation dans les sciences cognitives.

### **2.2 Les théories de l'apprentissage**

Les théories de l'apprentissage visent à expliquer le phénomène d'acquisition des connaissances. Elles servent à donner des explications de ce qui se passe lors du processus d'apprentissage. Les théories de l'apprentissage sont utiles pour deux raisons : elles fournissent un cadre conceptuel pour l'interprétation lors de l'observation et offrent également des orientations pour trouver des solutions aux problèmes rencontrés [2.01].

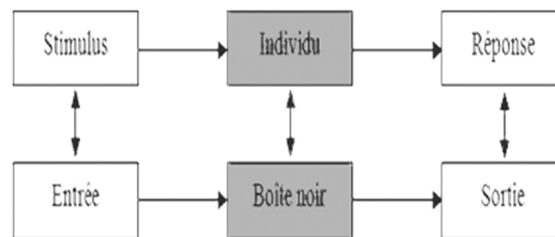
#### ***2.2.1. Le béhaviorisme***

Le béhaviorisme (ou comportementalisme) est une théorie de l'apprentissage qui s'intéresse à l'étude des comportements observables sans faire appel à des mécanismes internes au cerveau ou à des processus mentaux non directement observables [2.02]. Le terme béhaviorisme est apparu au début du XX<sup>e</sup> siècle en parallèle avec les travaux du psychologue américain John Watson. Ce dernier est considéré comme le pionnier du béhaviorisme, il proposait surtout de faire de la psychologie en général une discipline scientifique en utilisant seulement des procédures objectives, en vue d'établir des résultats exploitables statistiquement [2.03]. Watson a été influencé par les travaux du



physiologiste russe Ivan Pavlov sur le conditionnement des animaux. Cette conception l'entraîna à formuler la théorie psychologique du **stimulus-réponse** ou conditionnement classique.

Dans la même vision, Fechner compare l'individu à une boîte noire, dont on ne sait rien sur ce qui se passe à l'intérieur, mais dont on peut prévoir certains comportements, puisqu'en proposant des stimuli particuliers, on obtient toujours les mêmes résultats à la sortie [2.04], p. 55.



**Figure 2.01 :** *Modèle comportemental de Fechner*

Du point de vue de l'enseignement, le comportementisme considère l'apprentissage comme une modification durable du comportement résultant d'un entraînement particulier. Il part du principe que l'acquisition des connaissances s'effectue par paliers successifs. Le passage d'un niveau de connaissance à un autre s'opère par des renforcements positifs des réponses et comportements attendus.

De ce fait, l'enseignant répète une notion une ou plusieurs fois, lorsqu'il constate à travers le comportement observé que la notion en question n'est pas assimilée par les apprenants. De même, il a pour tâche de concevoir des exercices progressifs, de guider les élèves dans leurs réalisations et de leur communiquer les rétroactions nécessaires à la prochaine étape.

Néanmoins, les apprenants ne donnent souvent pas du sens aux connaissances qu'ils restituent et ils perdent le fil conducteur entre les différentes étapes de leur apprentissage. Dans cette théorie, l'apprenant est un élève qui écoute, regarde, réagit et tente de reproduire face à un enseignant qui est transmetteur d'information, de connaissances, qui présente, décrit, schématise, planifie et vérifie.

Le behaviorisme est confortable dans l'optique d'une introduction de la machine numérique dans le processus de l'enseignement-apprentissage. L'enseignant behavioriste sera porté à utiliser des exercices, des quizz, des jeux éducatifs et/ou des animations lors de la conception et la réalisation

d'une formation à distance. Si cette théorie est confortable dans l'optique d'une introduction de la machine numérique, elle semble trop pauvre pour être durable [2.05].

Contrairement aux behavioristes pour lesquels l'individu est modelé par son environnement, Piaget considère que l'apprentissage est le résultat d'une interaction entre le sujet et son environnement. Pour Piaget, l'apprentissage est le résultat d'un processus dynamique de recherche d'équilibre entre le sujet et son environnement. Il a notamment mis en évidence l'importance de l'action sur l'environnement et de la réflexion sur cette action pour construire des connaissances sur un objet [2.06].

### ***2.2.2. Le constructivisme***

Les constructivistes croient que chaque apprenant construit la réalité, ou du moins l'interprète, en se basant sur sa perception d'expériences passées. Selon le modèle constructiviste, l'acquisition de connaissance ne se réalise pas par simple empilement mais passe par une réorganisation de conceptions mentales précédentes, un travail de construction ou de reconstruction.

Selon Piaget, « La pensée humaine se construit progressivement lorsque l'individu, en particulier l'enfant, entre en contact avec le monde ». De plus, « Connaitre un objet, c'est agir sur lui, le transformer » [2.26] p. 45. Cette citation met en évidence l'importance de l'action dans le développement des connaissances. La conception piagétienne de l'apprentissage et du développement cognitif est habituellement qualifiée de constructiviste pour signifier que l'enfant se développe à travers l'interaction continue entre d'une part la structure cognitive qui le caractérise et son action sur le milieu et, d'autre part les informations qu'il reçoit en retour de ce milieu (difficulté, erreurs, succès, résistances, ...). Chaque action sur le milieu provoque ainsi une modification cognitive qui à son tour modifiera la prochaine action sur le milieu. C'est donc à partir des informations tirées de l'action sur le milieu que l'enfant construira ses connaissances nouvelles. Piaget a mis en évidence l'importance de l'action sur l'environnement et de la réflexion sur cette action pour construire des connaissances sur un objet [2.06]. Selon le modèle piagétien, l'enseignement a pour objet de confronter l'apprenant à des situations riches et diversifiées de manière à créer des interactions propices au développement cognitif.

Pour Piaget, l'assimilation et l'accommodation forment un couple indispensable à l'activité cognitive dont les différents processus d'équilibration seront développés dans l'équilibration des structures cognitives [2.07]. Selon le même auteur, l'assimilation désigne la réintégration d'éléments externes nouveaux dans une structure interne préexistante ; l'accommodation désigne l'adaptation de l'organisme aux variations externes qu'il ne réussit pas à assimiler.

Doolittle [2.08] insiste sur huit conditions nécessaires pour réussir une pédagogie constructiviste :

- a. Présenter aux apprenants des situations d'apprentissage complexes similaires à celles qu'ils rencontrent dans la vie courante.
- b. Favoriser l'interaction et la collaboration entre les apprenants.
- c. Donner le sens aux apprentissages des élèves.
- d. Tout apprentissage doit partir des acquis des élèves.
- e. Les élèves doivent bénéficier d'une évaluation formative continue.
- f. Les élèves doivent être responsables de leurs apprentissages.
- g. Les enseignants sont des guides et des agents qui facilitent l'apprentissage.
- h. Revoir des contenus et les présenter selon diverses perspectives.

A l'heure actuelle, le constructivisme apparaît toujours, il favorise des outils donnant une grande autonomie à l'élève lui permettant d'avancer à son rythme, comme dans une plateforme pédagogique ou avec l'utilisation de matériel didactiques, en recourant à des outils collaboratifs.

### ***2.2.3. Le cognitivisme***

Le cognitivisme (ou rationalisme) naît en même temps que l'Intelligence Artificielle, en 1956. Il est proposé par Miller et Bruner en réaction au béhaviorisme. Il est centré sur les manières de penser et de résoudre des problèmes. L'apprentissage ne peut être limité à un enregistrement conditionné, mais doit plutôt être envisagé comme nécessitant un traitement complexe de l'information reçue. La mémoire possède une structure propre qui implique l'organisation de l'information et le recours à des stratégies pour gérer cette organisation [2.09].

La psychologie cognitive considère qu'il y a fondamentalement trois grandes catégories de connaissances : les connaissances déclaratives, procédurales et conditionnelles. Elle invite

l'enseignant à développer des stratégies différentes pour faciliter l'intégration de chacune d'elles parce qu'elles sont représentées différemment dans la mémoire ; les connaissances déclaratives répondent en effet au QUOI, les connaissances procédurales au COMMENT et les connaissances conditionnelles au QUAND et au POURQUOI [2.10].

Pour les cognitivistes, l'apprenant est un système actif de traitement de l'information, semblable à un ordinateur : il perçoit des informations qui lui proviennent du monde extérieur, les reconnaît, les emmagasine en mémoire, puis les récupère de sa mémoire lorsqu'il en a besoin pour comprendre son environnement ou résoudre des problèmes [2.11]. L'enseignant est gestionnaire des apprentissages, il guide, anime, dirige, conseille, explique, régule, remédie. Les connaissances deviennent une réalité externe que l'apprenant doit intégrer à ses schémas mentaux et réutiliser plutôt qu'à acquérir des comportements observables [2.12]. De plus, la méthode d'enseignement favorisée laisse place à de multiples cheminements d'apprentissage afin de tenir compte des différentes variables individuelles pouvant influencer la manière dont les élèves traitent l'information. L'enseignant cognitiviste sera invité à utiliser des TIC qui favorisent une grande interactivité avec les élèves, telles que des simulateurs, des expériences et des tutoriels intelligents. Toutefois, le modèle cognitiviste a une limite importante, liée au fait qu'un matériel bien structuré n'est pas suffisant pour assurer un apprentissage. La motivation des élèves est un facteur déterminant puisqu'il fournit l'énergie nécessaire pour effectuer les apprentissages.

Le module GEAL, objet de la recherche, reprend l'approche de l'apprentissage par découverte [2.13] du modèle de cognitivisme pédagogique de Bruner. Ce dernier propose une approche basée sur la découverte active par l'élève des principes et des concepts à maîtriser. Pour lui, une telle approche conduit à un double bénéfice : d'une part, elle permet une maîtrise plus profonde des contenus qui font l'objet de l'apprentissage, d'autre part elle développe chez l'élève certaines démarches de pensée qui lui permettront par la suite d'être plus autonome dans son apprentissage (apprendre à apprendre). L'élève doit être préparé à ce type d'apprentissage à travers la capacité à mettre en œuvre certaines stratégies spécifiques telles que : récolter et sélectionner des informations, se poser des questions, ... Ensuite, il doit être guidé en cours d'apprentissage. Ce guidage peut bien entendu être assuré par l'enseignant mais aussi par d'autres élèves ou encore par un dispositif informatique. L'important est qu'un dialogue étroit s'installe grâce auquel l'élève sera épaulé dans les difficultés qu'il aura à dépasser pour résoudre le problème qui lui est posé. Bruner utilisera par la suite le terme « étayage » (scaffolding) pour désigner ce soutien apporté à l'élève en cours

d'apprentissage. Les idées exprimées par Bruner à propos de l'apprentissage par découverte s'inscrivent parfaitement dans le courant cognitiviste notamment par l'importance accordée au sujet qui est considéré comme l'acteur principal de son propre apprentissage [2.13]. Le processus d'apprentissage doit aider l'élève à saisir la structure des contenus qu'il aura à assimiler de manière à être capable de mettre en évidence les idées et les concepts essentiels et à établir des liens entre ceux-ci.

#### ***2.2.4. Le Socio-constructivisme***

Ce modèle proposé par Vygotsky reprend les idées principales du constructivisme de Piaget en y ajoutant le rôle social des apprentissages. L'apprentissage est vu comme l'acquisition de connaissances grâce aux échanges entre l'enseignant et les élèves ou entre élèves. Les élèves n'apprennent pas seulement grâce à la transmission de connaissances par l'enseignant mais aussi grâce aux interactions [2.14]. Selon ce modèle, les apprentissages doivent être compris dans leur zone proximale de développement : cette zone comprend les tâches que les élèves peuvent réussir à l'aide d'un adulte, elles ne sont ni trop difficiles, ni trop faciles. Cette zone augmente nettement le potentiel d'un élève à apprendre plus efficacement [2.15]. Le maître a pour rôle de définir précisément cette zone afin de donner des exercices appropriés. De plus, il va favoriser le débat entre les élèves (conflit socio-cognitif), en les faisant travailler en groupe. Dans ce modèle, les erreurs correspondent également à un point d'appui pour la construction de nouvelles connaissances.

Bruner [2.16] a également apporté sa contribution à la théorie socio-constructiviste en expliquant que le modèle transmissif place l'enseignant en situation de monopole ce qui empêche l'acquisition de l'autonomie des élèves. Pour lui, l'enseignant doit faire en sorte que la tâche soit plus agréable à réaliser avec son aide tout en évitant que l'élève devienne dépendant de lui. Il doit également mobiliser et motiver l'élève en maintenant l'intérêt de la tâche pour l'élève.

Dès les années 1930, le rôle des interactions sociales a commencé à être étudié. Les recherches portant sur l'importance des interactions entre apprenants pour construire de nouvelles connaissances dans le développement des technologies éducatives et les environnements d'apprentissage ont suscité les intérêts. Ces approches ont notamment fondé la conception des « Computer-supported collaborative learning » qui a conduit au développement de l'enseignement

à distance. Selon Lev Vygotski, les différentes fonctions cognitives (attention, mémoire, raisonnement, etc.) se développent d'abord dans des situations d'interactions sociales. « Dans le développement de l'enfant, toute fonction apparaît deux fois : dans un premier temps au niveau social, entre les individus et dans un second temps au niveau individuel, à l'intérieur de l'enfant » [2.17].

Le rôle des interactions sociales dans l'apprentissage a pris une place importante dans le modèle socioconstructiviste. Doise et Mugny ont mis en évidence l'importance des conflits socio-cognitifs dans l'apprentissage en situation de résolution de problèmes [2.18]. Ils ont montré que faire travailler ensemble des élèves ayant des positions contradictoires sur un même problème pouvait les conduire à comprendre qu'un autre point de vue que le leur était envisageable et à remettre en cause leur point de vue pour construire une connaissance plus élaborée. Le modèle artisanal [2.19] choisit de renforcer les interactions sociales entre tuteurs et apprenants. Il exploite pleinement les possibilités techniques des nouveaux médias numériques et permet une large diversification des échanges. Ce modèle pédagogique confère à l'apprenant un rôle actif dans la construction de ses connaissances. Il se base sur une approche socio-constructiviste des apprentissages en considérant que ces derniers résultent des interactions du sujet avec son environnement culturel et social.

Toutes personnes qui souhaitent analyser l'apprentissage et l'enseignement dans une formation en ligne sont intéressées par les traces des événements sur la plateforme. Ce sont des données extractibles automatiquement et très représentatives car très proches de l'interaction homme machine.

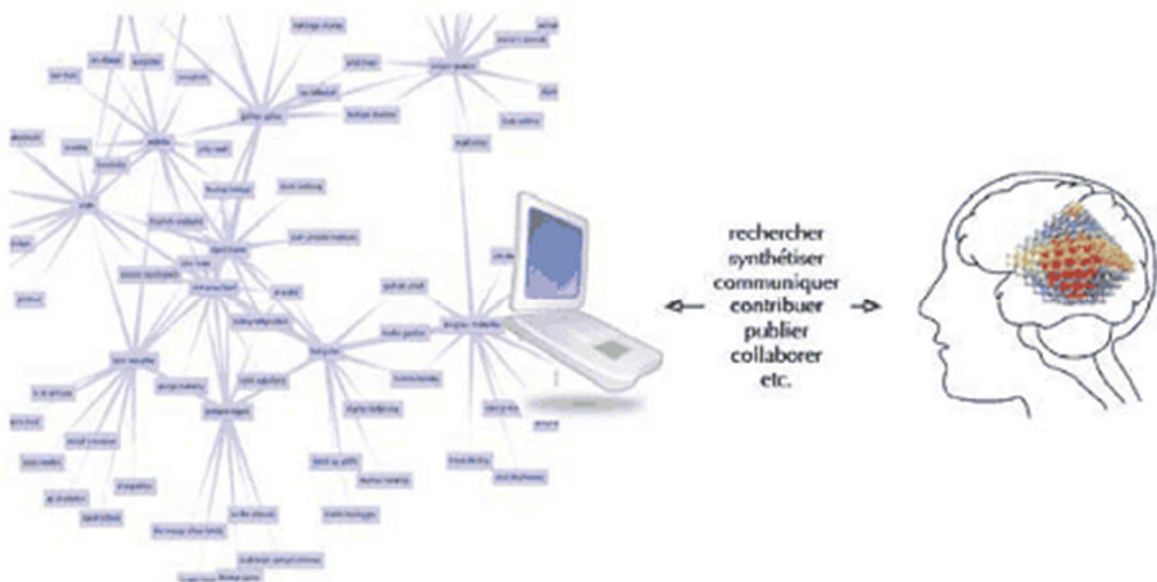
### ***2.2.5. Le connectivisme***

Face à l'évolution de l'usage des technologies dans l'enseignement, plusieurs chercheurs ont proposé le terme connectivisme pour désigner une nouvelle approche éducative qui s'adapterait à la formation en ligne. Développé par George Siemens et Stephen Downes, le connectivisme interroge le processus de l'apprentissage à l'ère du numérique et dans un monde connecté en réseaux en s'appuyant sur les limites du behaviorisme, du cognitivisme, du constructivisme et du socio-constructivisme [2.20]. Un aspect du connectivisme est l'utilisation d'un réseau composé de nœuds et de connexions comme métaphore centrale de l'apprentissage [2.21]. Dans cette métaphore, un

nœud peut être une information, des données, un sentiment, une image ou une simulation. L'apprentissage est le processus de connexions, englobant les connexions neuronales, les connexions entre les hommes, les ordinateurs et l'interconnexion entre les différents champs de savoirs [2.20].

*Définition 2.01 :*

Pour George Siemens : « *Le connectivisme est la somme de principes issus de la théorie du chaos, des réseaux, de l'auto-organisation et de la complexité. L'apprentissage est un processus qui se produit dans des environnements flous composés d'éléments de base changeants, et qui n'est pas entièrement sous le contrôle de l'individu. L'apprentissage peut résider en dehors de l'individu (au sein d'une organisation ou une base de données), et se concentre sur la connexion d'ensembles d'informations spécialisées. Les liens qui permettent d'apprendre davantage sont plus importants que l'état actuel de notre connaissance. Le connectivisme est motivé par la compréhension du fait que les prises de décision sont fondées sur des bases qui se modifient rapidement. De nouvelles informations sont constamment acquises. La capacité d'établir des distinctions entre l'information importante et sans importance est vitale. La capacité de reconnaître quand de nouvelles informations modifient le paysage en fonction des décisions prises hier est également critique.* » [2.20] p. 4.



**Figure 2.02 :** *Le connectivisme* [2.20]

Siemens [2.20] souligne les huit principes du connectivisme :

- a. L'apprentissage et la connaissance résident dans la diversité des opinions.
- b. L'apprentissage est un processus reliant des nœuds spécialisés ou des sources d'information.
- c. L'apprentissage peut résider dans des appareils non-humains.
- d. La capacité d'en savoir plus est plus critique que ce que l'on sait actuellement.
- e. Entretenir et maintenir des connexions est nécessaire pour faciliter l'apprentissage continu.
- f. La possibilité de voir les liens entre les domaines, les idées et les concepts est une compétence de base.
- g. Obtenir des connaissances précises avec la capacité de les mettre à jour est l'intention de l'apprentissage basé sur le connectivisme.
- h. La prise de décision est un processus d'apprentissage en soi. L'importance que l'on donne à une information est variable dans le temps, selon les modifications de l'environnement de cette information.

### **2.3 Analyse des traces numériques d'activités sur la plateforme**

En sciences de l'information et de la communication, l'étude des traces numériques, leur diffusion et leur centralisation constitue un risque pour la liberté et l'identité de l'individu, comme le montrent les travaux d'Arnaud et de Merzeau [2.22]. Cependant, ces traces ont un intérêt pédagogique important, comme le soulignent Larose et Jaillet [2.23].

En analysant les discours des acteurs à l'aide d'outils automatiques puis par une lecture critique, nous tentons de faire le bilan de leur apprentissage afin de mettre en évidence les actions pertinentes sur plateforme et de modéliser par la suite les activités d'apprentissage. Ces modèles doivent être compatibles avec la théorie des systèmes à base de traces pour l'apprentissage [2.24].

Réinterroger son activité peut en améliorer les résultats. Il est possible de renforcer cette réflexivité en favorisant la réappropriation et la restitution des traces numériques des apprenants sur la plateforme au lieu de se contenter d'une indexation « côté serveur ». Ainsi, l'étude des discours d'étudiants sur leur propre pratique permet d'avoir une meilleure compréhension des représentations de la recherche d'information et d'obtenir les modèles de leurs recherches [2.25].



### ***2.3.1. Notion de « trace d'activité »***

En pratique, l'apprenant permet de produire un ensemble de données qu'on peut interpréter après l'apprentissage pour comprendre certains aspects de l'activité. Le terme « données » désigne l'état d'une mémoire informatique à partir duquel peut être produite une interprétation (chiffres, textes et graphiques). Ces données peuvent être désignées sous le terme de « traces » au sens courant, dans la mesure où elles renseignent sur un phénomène passé.

#### *Définition 2.02 :*

La notion de trace d'activité est définie comme un ensemble de données interprétables pour comprendre le flux de discussions sur des aspects pertinents pour un apprenant, son interaction avec son environnement. Cette définition nous permet d'ores et déjà d'indiquer une caractéristique fondamentale des traces d'activité : elles doivent contenir des informations permettant de rendre compte de ce flux de discussions lorsqu'elles s'affichent sur l'écran. Ce sont donc des données « typées », chaque élément de donnée doit être associé à une fréquence d'intervention par type d'évènement que nous désignerons par un chiffre. Dans les chapitres précédents, nous avons présenté des méthodologies de recueil de ces traces : le recueil de données objectives à partir de l'observation des forums de discussion asynchrone, puis le recueil de données subjectives à partir d'une explicitation et d'une évaluation de sa propre discussion.

La question de l'analyse de ces traces pour produire des connaissances en ergonomie cognitive est maintenant abordée. Cette production de connaissances repose sur une transformation « intelligente » des traces pour les restituer de manière synthétique et expliquée. Pour cela, nous avons décrit successivement les techniques qui répondent à notre approche : la technique d'analyse de discours donnée dans 3.4- et la technique de l'analyse exploratoire de données.

Mais, avant de chercher des données à analyser, il est bon de se poser la question sur le modèle à explorer. Plusieurs acteurs interviennent dans un dispositif de formation à distance : les apprenants, les concepteurs, les tuteurs et les coordinateurs. Une attention particulière doit donc être portée sur ceux qui interviennent dans les dialogues. Plusieurs questions se posent : Quelles sont les entités qui interagissent ? Comment peut-on modéliser le comportement de l'apprenant ? Quelles techniques d'analyse peut-on adopter ? Produit-il des effets cognitifs ?

### ***2.3.2. Définition de traces numériques d'interaction***

Il convient tout d'abord de définir la notion d'interaction dans une formation ouverte et à distance.

#### *Définition 2.03 :*

Par définition, une interaction est une suite d'actions verbales ou non-verbales qui sont interdépendantes et qui s'influencent mutuellement. Une telle interaction présuppose des êtres capables d'exercer une volonté dans l'action ou d'être muni d'un pouvoir de délibération. Les êtres humains remplissent cette condition, mais non les ordinateurs.

Dans l'activité à analyser, il s'agit d'humains interagissant entre eux au travers une plateforme. Pour l'interaction entre êtres humains au travers d'un ordinateur, les expressions « d'interaction médiatisée » et « d'interaction médiée » par ordinateur se trouvent dans la littérature. Pour Peraya [2.26] et Meunier [2.27], les processus de médiatisation et de médiation portent sur des objets différents. La médiatisation désignerait le processus de mise en dispositif médiatique, incluant ainsi le choix de médias et d'une scénarisation, alors que la médiation serait le processus par lequel un artefact technique et symbolique peut modifier les processus de production et de communication des connaissances ou encore influencer sur l'apprentissage ou sur le développement des processus cognitifs des êtres humains. C'est la raison pour laquelle la définition d'une trace numérique d'interactions occupe le point central du présent chapitre.

#### *Définition 2.04 :*

Une trace numérique d'interactions est une suite de discussions située d'observés, qui relève soit d'une interaction entre humains, médiatisée et médiée de diverses façons par ordinateur soit d'une suite d'actions et de réactions entre un humain et une plateforme. Cette trace est éventuellement rejouable, auquel cas, elle devient dynamique. Elle est numérique puisqu'il s'agit d'enregistrements d'actions effectuées sur ordinateur.

À partir de là, nous pouvons distinguer plusieurs types de traces selon la façon dont les outils d'enregistrement rendent observable l'interaction humaine. En effet, les traces peuvent être constituées de sources multiples et multimodales (enregistrements informatiques y compris des vidéos d'écran, mais aussi des enregistrements audio et vidéo d'êtres humains en interaction entre eux, autour d'un ordinateur, ou à partir d'une analyse des traces de discussions qui s'influencent

mutuellement). Dans le cas où un humain interagit avec un ordinateur, il est préférable de parler d'une suite d'événements prenant la forme d'une action-réaction.

### ***2.3.3. Caractéristiques de l'interaction humaine***

Les études empiriques, qu'elles soient menées dans un paradigme de psychologie expérimentale ou qu'elles aient pour but d'observer des situations d'apprentissage, se focalisent sur une description des caractéristiques, voire de la qualité de l'interaction humaine, médiatisée et médiée par ordinateur. L'objectif de cette description est souvent de déterminer les interactions humaines produites dans le scénario pédagogique. Le scénario présente la succession des tâches demandées aux apprenants, dans les instructions précises qui leur sont données ou encore dans les outils mis à leur disposition. L'analyse des interactions humaines produites est donc utilisée pour évaluer l'apprentissage asynchrone du module d'initiation et permet de :

1. décrire les situations qui donnent lieu à des apprentissages spécifiques et donc de choisir les caractéristiques du scénario pédagogique feront atteindre les objectifs pédagogiques ;
2. proposer une réingénierie du système d'apprentissage en ligne et du contexte dans lequel il sera utilisé, basée sur la combinaison des caractéristiques de l'interaction humaine et des caractéristiques du scénario pédagogique et ce, selon les objectifs à atteindre ;
3. proposer des façons d'agir en asynchrone ; autrement dit, proposer des actions à intégrer dans le système d'apprentissage qui pourraient être déclenchées à la découverte d'un observable calculé de manière automatique dans 1.

La recherche a pour objectif d'étudier une centaine d'apprenants, regroupés dans trois formations ouvertes et à distance. Le scénario pédagogique est construit en trois phases. 1) production individuelle ; 2) collaboration en équipe où les apprenants produisent un document comme medium de débat ; 3) révision individuelle d'un document produit lors de la première phase. Quatre caractéristiques sont alors mesurées :

- La qualité d'intervention
- La quantité d'arguments et des relations exprimées
- La quantité et la nature des opinions exprimées

- La quantité de sujets traités dans l'espace de discussion.

#### ***2.3.4. Type et profil d'étudiants***

Comme la formation est destinée aux professionnels dans le domaine concerné, presque tous les apprenants sont des travailleurs en entreprise et des responsables issus de diverses sociétés. Ils n'ont pas le temps d'étudier dans la journée. La plateforme est conçue et configurée pour un dispositif correspondant à un tel type d'apprenant. Avec Moodle, l'apprenant est au centre de la formation, il suit un parcours pédagogique qui favorise l'apprentissage par l'activité. Les apprenants doivent alors entrer en contact avec l'environnement du dispositif de la formation, connaître un objet, agir sur lui et le transformer [2.26]. Ce processus met en évidence l'importance de l'action dans le développement des connaissances.

##### *Définition 2.05 :*

Un profil est un ensemble d'informations concernant un apprenant ou un groupe d'apprenants, saisies, collectées ou déduites à l'issue d'une ou de plusieurs activités pédagogiques. Ces informations peuvent concerner les connaissances, les compétences, les conceptions, les représentations ou encore les styles ergonomiques.

Les lurkers [2.28] par exemple sont des apprenants dans un séminaire synchrone, qui lisent les messages, ne perdent pas le fil des apprentissages mais n'interviennent pas. Leurs productions sont souvent de qualité car ils profitent des apports de leurs pairs. D'autres profils sont connus, souvent liés à des mécanismes sociaux : le social loafing, par exemple, se produit lorsqu'un étudiant travaille moins car il s'appuie sur le groupe. Le free rider effect voit quelques membres d'un groupe travailler indépendamment des autres et choisir d'avancer seuls pour économiser leur temps. La convergence précoce consiste à s'accorder sur une solution sous-optimale simplement parce qu'elle permet un consensus. La domination se produit lorsqu'un membre du groupe impose son point de vue [2.29]. On peut vérifier aussi les comportements des étudiants : proactif, réactif ou passif dans leurs activités.

## **2.4 Approche morphologique d'un corpus de message**

Pour analyser les échanges produits au sein d'une communauté en ligne, plusieurs aspects doivent être suivis :

- l'observation de la communauté, sans intervention aucune, pour recueillir les comportements effectifs. Elle peut se faire en différé, dès lors que la communication est asynchrone. La démarche est centrée sur les acteurs et leur contexte, avec une approche basée sur l'ethnographie.
- l'entretien : le choix des personnes doit être consécutif à un ensemble d'analyses préalables d'un corpus de messages, morphologique notamment.
- le questionnaire : il vient compléter l'entretien et/ou l'analyse du corpus de messages en les précédant (exploratoire) ou en les succédant (confirmatoire).
- la collecte et l'analyse de traces de l'activité principale de la communauté en ligne : les messages. Les extraits du forum sont traités puis analysés, selon des approches qui peuvent être très variables. Ces traces doivent toujours être contextualisées, pour donner sens aux synthèses conduites : chartes du forum, règlements, modalités d'organisation de la communauté en ligne, articulation de la liste ou du forum étudié avec d'autres instruments de communication ou d'information, finalités explicites de la communauté étudiée.

L'analyse des contenus d'échanges concerne les contenus manifestes ou latents, soit les deux. Nous allons nous attacher particulièrement dans cette section à l'analyse de ces derniers éléments, qui vont caractériser la morphologie [2.30] d'un corpus issu d'un forum.

### ***2.4.1. Premières analyses à mettre en œuvre***

Les forums sont des analyseurs de la dynamique d'un groupe. Les analyses permettent de révéler les valeurs, intérêts et règles partagés, ainsi que les points de vue et représentations des contextes.

D'un point de vue méthodologique, quels que soient les objets étudiés et les approches théoriques mobilisées en sociologie ou en psychologie sociale, les analyses échappent rarement à une approche

morphologique initiale du média considéré, caractérisant son activité générale. Une approche quantitative préliminaire à l'étude plus spécifique des contenus est nécessaire.

Des indications doivent être apportées sur l'adoption du médium, la participation et selon les axes de l'étude, la longévité des thèmes, les flux de contributeurs entrants et sortants, les variations d'effectifs, les fréquences de création de thèmes nouveaux sur des durées plus ou moins longues.

La répartition des propos ou leur concentration [2.31] constitue un indicateur pertinent permettant de repérer le leadership et de comprendre comment se forment et se disséminent les opinions dans ce type de collectif.

Les contributeurs principaux, des « sous communautés » de contributeurs prolixes et les contributeurs muets peuvent tous être repérés.

#### ***2.4.2. Caractéristiques quantitatives d'un forum***

##### **2.4.2.1. La participation**

Le simple fait de poster un message ou de déposer un document sur une plateforme de formation constitue un acte traduisant la participation. Le degré de participation, extrêmement variable d'un individu à l'autre, prête à confusion et ne traduit pas le degré d'investissement cognitif réel.

- La participation s'évalue en premier lieu selon des critères quantitatifs, par exemple : le nombre de messages et l'évolution de ce nombre.
- On observe la répartition dans le temps et la distribution des messages parmi les auteurs. On définit aussi le taux de participation en établissant le rapport entre le nombre de participants réels et le nombre d'inscrits.

##### **2.4.2.2. L'adoption**

Cette notion n'a pas le même sens selon qu'il s'agit d'un individu qui adhère volontairement à une communauté de pratiques/intérêt ou d'un formateur qui participe à des échanges d'une communauté d'apprentissage. En général, on évalue l'adoption d'un forum par le nombre d'inscrits.

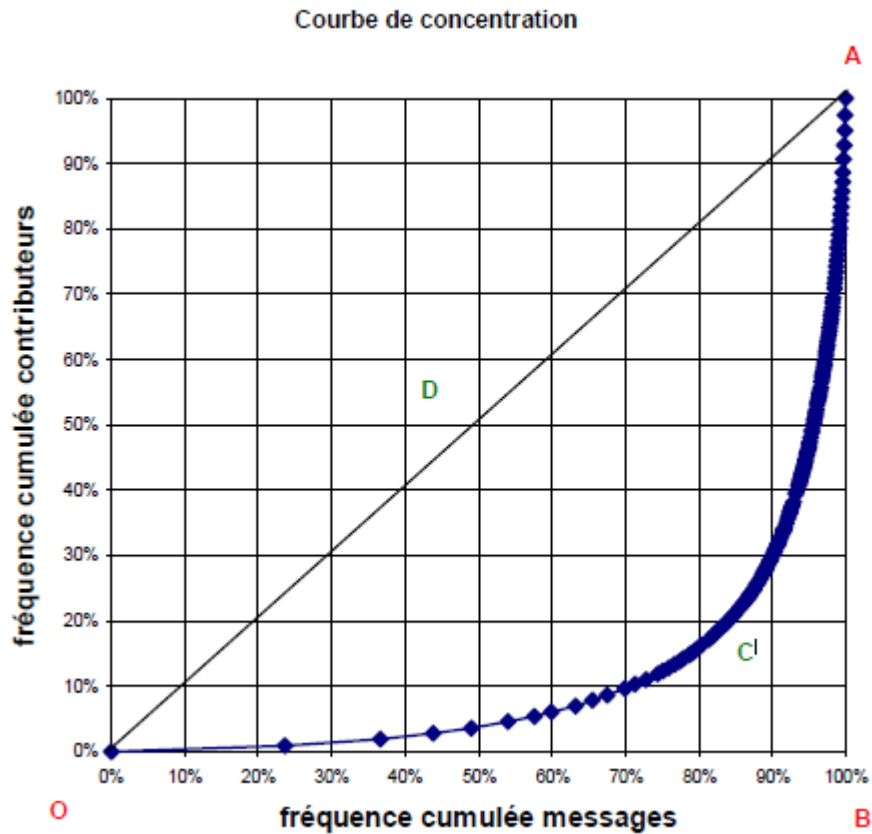
Le rapport entre le nombre d'individus potentiellement concernés et le nombre d'utilisateurs réels constitue le taux d'adoption d'une liste ou d'un forum. Le succès d'un forum dépend de son degré d'ouverture à d'autres professionnels non directement visés à l'origine [2.32].

#### 2.4.2.3. La concentration

L'étude de la concentration est un passage quasi obligé avant d'amorcer une étude qualitative des propos échangés.

Sur une liste de discussion ou sur un forum, la participation ne revêt pas un caractère obligatoire. Certains abonnés sont donc totalement invisibles, alors que d'autres sont sur-actifs. Plusieurs études ont démontré la présence d'un grand déséquilibre de la contribution des uns par rapport à celle des autres sur les listes de discussion [2.33], [2.34], [2.35] et de la forte concentration de l'essentiel des messages postés sur un nombre réduit de contributeurs. Sur la liste des ATICE, par exemple, et pour donner un ordre d'idée, près de 70% des messages sont postés par 10% des adhérents à la liste de discussion. Ces leaders d'opinion, une fois repérés peuvent constituer une sous-population intéressante à étudier : positionnement, thématiques privilégiées, etc.

La Figure 2.03 : ci-contre propose une allure de la courbe de Gini ou courbe de concentration [2.36].



**Figure 2.03 :** *Courbe de Gini ou courbe de concentration*

La droite D est la droite d'équirépartition de la variable sur la population. Ici, il s'agit de tous les contributeurs de la liste des ATICE (ici nombre de population  $N=1210$ ) [2.35].

Pour mesurer la concentration, on s'intéresse à l'aire comprise entre la droite D et la courbe C'.

On appelle *indice de GINI* ( $i$ ), le rapport de cette aire à celle du triangle OAB : La valeur O correspond à l'équirépartition. Les valeurs voisines de 1 révèlent une très forte concentration de la variable sur un faible nombre d'individus. Dans ce cas,  $i = 0,80$ .

Cette courbe s'obtient avec un tableur. La population des individus est ordonnée selon le nombre de messages qu'ils ont émis au cours de la période, par valeur croissante ( $N1=37\ 926$ ). On exprime la proportion que représente chaque individu dans sa population de référence et le pourcentage que représente le nombre de messages qu'il a émis par rapport au nombre total de messages postés durant



la période ciblée (5 ans). On obtient deux séries de pourcentages. En abscisse, on porte le pourcentage de messages cumulés et en ordonnée le pourcentage des individus cumulés.

## **2.5 La théorie sur l'analyse de discours : échanges écrits en ligne**

Cette section expose la démarche d'analyse de contenus de fils de discussion d'une liste ou d'interventions sur un forum. L'analyse de faits (ou actes) de langage exprimé sur une liste ou un forum contribue à l'analyse conversationnelle. L'objectif est d'entrer très modestement dans l'analyse de l'organisation d'une communication entre différents locuteurs. Les fils d'échanges sont des séquences conversationnelles, traduisant l'expression de pensées individuelles et des formes variées de rapports sociaux. Les approches et épistémologies sont multiples (linguistique, psychologie sociale, sociologie).

### ***2.5.1. Rigueur d'une analyse de contenu***

« L'analyse de discours s'intéresse à la forme du message médiatique, politique, public, gouvernemental, organisationnel, ... en rapport avec son lieu social de production. C'est donc une démarche fondée sur la linguistique, mais qui fait le lien entre le discours et le social, entre le verbal et l'institutionnel, entre les mots, les figures, les arguments et ceux qui les énoncent, dans leur dimension aussi bien individuelle que collective » [2.37].

L'analyse qualitative d'échanges sur des forums repose sur un ensemble d'enjeux méthodologiques, variant d'une étude à l'autre [2.38].

- Scientificité : objectivité (qualité des critères et grille d'analyse) ; fidélité (accord entre codeurs), répliquabilité, systématisme ;
- L'élaboration d'un protocole de recherche (descriptive ou expérimentale) ;
- La nature des contenus analysés : manifestes ou latents ;
- L'unité d'analyse : mots, phrases, paragraphes, messages, unité de sens, etc.
- L'utilisation de logiciels d'analyse de contenu ;
- Les aspects éthiques inhérents à la collecte et à l'analyse de communications privées.

Lorsqu'on étudie les échanges, il y a nécessité de coder selon des hypothèses, et d'agréger les données dans des catégories qui permettront une description précise des caractéristiques du contenu.

### ***2.5.2. Echanges, structures et transactions***

Dans une analyse de séquence conversationnelle ou une interlocution, le chercheur isole trois types d'objets : les échanges, les structures et les transactions [2.39].

L'échange est l'unité de base. Les échanges s'organisent en structure selon deux types d'enchaînements : soit les échanges sont d'un même niveau de discours, soit ils sont dépendants ou organisés hiérarchiquement.

Les structures sont fonctionnellement interprétables : communiquer une information, débattre, négocier, prendre une décision en groupe, etc.

Les transactions sont des échanges fortement déterminés par des normes et des conventions : entretien médical, procès, entretiens de conseil etc.

### ***2.5.3. L'interlocution : dimension cognitive et sociale***

Les échanges, les structures et les transactions véhiculent des événements cognitifs et expriment des rapports sociaux. Les échanges sont porteurs de cette double dimension.

Un échange est composé d'actes de langage qui définissent l'énonciation comme une action sociocognitive. Autrement dit, l'acte de langage est l'application d'une force sur le contenu propositionnel. Cette force définit le type d'action que vise à accomplir l'acte de langage (assertion, déclaration, direction, expression). Le contenu propositionnel, quant à lui définit la représentation (cognition).

#### ***2.5.4. Echanger par écrit pour apprendre***

Concernant les échanges écrits sur un forum, ils induiraient un effort cognitif jugé plus rigoureux que lors d'échanges oraux et pourraient aider les processus cognitifs en suscitant, une plus grande réflexivité [2.40]. Ces pratiques amèneraient les acteurs à injecter des significations, à clarifier leurs idées, à partager des points de vue, à rétroagir, à développer un langage commun et des solutions communes aux problèmes posés.

Dans le cadre de forums d'apprentissage, son usage contribuerait à « rendre les apprenants maîtres de leurs apprentissages et à favoriser ainsi le passage du modèle transmissif des connaissances vers le modèle collaboratif des communautés d'apprenants » [2.41].

#### ***2.5.5. Etapes du processus***

Le corpus de message est souvent particulièrement volumineux et peut atteindre parfois des quantités considérablement importantes, rendant illusoire le travail de fond, message par message. L'échantillonnage est donc de rigueur et différentes techniques sont mobilisables.

- Au hasard : l'avantage, c'est la représentativité, mais il y a un risque de perte du contexte et de conservation de l'intégralité des échanges.
- Par thème : les messages sont issus d'un même fil de discussion. Le risque est l'exclusion d'autres thématiques. L'avantage réside dans la conservation d'une certaine cohérence ;
- Par intervalle de temps : cette technique permet de préserver la richesse du contexte, mais le risque est de tronquer certaines interactions ou de générer de trop grands échantillons ;
- Par type d'acteur ou par groupe : permet de se focaliser sur un type d'acteur et de comparer (leaders/discret, étudiants/enseignants, etc.) mais le risque est là encore la perte du contexte.

Dans la recherche sur l'apprentissage dans une FOAD, nous sommes parties de l'hypothèse qu'il existe une relation étroite entre les thématiques traitées qui peut donc être perçue à travers le discours d'un type d'acteur donné. Nous avons repéré la relation entre le résultat de chacun et l'intervention et/ou la participation aux forums de discussion.

Pour cela, nous avons procédé au traitement d'une variable commune à tous les messages et à partir d'une première analyse, il en a résulté un échantillonnage thématique. Nous avons isolé les fils de

discussion portant sur les thématiques les plus fréquemment mobilisés par les tuteurs ; avant de saisir la fréquence d'intervention et le nombre de participation des apprenants dans le forum de discussion (affichage et consultation).

#### **2.5.6. *Le codage***

Les messages issus d'un forum doivent faire l'objet d'une catégorisation. La création des catégories d'analyse peut se faire :

- Selon un modèle ouvert : les catégories se créent au fur et à mesure, à la lecture des données elles-mêmes. Ce scénario vaut dans les recherches exploratoires. Laurence Bardin parle de catégorisation « par tas » [2.42].
- Selon un modèle plus fermé où les catégories sont délivrées dès le départ. Elles doivent en revanche être immuables.
- Un autre modèle hybride est encore possible : il s'agit de partir des catégories initiales qui serviront de guide, et la formulation, la modification, la suppression ainsi que l'ajout en fonction du sens de données sont possibles.

#### **2.5.7. *Propriétés des catégories***

Les catégories doivent obéir à certaines exigences principales [2.43]

- L'exhaustivité : les catégories couvrent l'ensemble du corpus ;
- La cohérence : les messages classés doivent bien aller dans le même sens. Les catégories doivent également être cohérentes entre elles du point de vue de leur niveau de catégorisation ;
- La simplicité : les dénominations des catégories ne doivent pas être équivoques ;
- L'objectivation : tous les codeurs doivent avoir les mêmes critères de codage pour chaque catégorie ;
- L'exclusivité (étanchéité) : un message ne peut être classé que dans une catégorie.

La catégorisation et le codage doivent être en ligne avec la réalisabilité et avec la fidélité entre les codeurs ; Autrement dit, une grille de codage doit être réutilisée dans un autre contexte, par d'autres codeurs. La robustesse de la grille d'analyse est un critère de validité externe des résultats.

## **2.6. Modélisation**

### *Définition 2.06 :*

Un modèle est une représentation simplifiée, où sont identifiés les différents éléments qui le constituent et dans laquelle les interactions entre ces éléments sont précisées et les résultats obtenus décrits.

### **2.6.1. Modélisation dans les sciences sociales**

Lévi-Strauss insiste sur la distinction des relations sociales avec des structures sociales (domaine du virtuel). L'analyse scientifique se propose de les mettre à jour. On ne peut y accéder que par la construction d'un modèle [2.44]. Il ne s'agit pas toutefois d'entendre l'activité de structuration que sous-tend le modèle au sens où les structuralistes entendaient rendre manifestes les structures du réel (tendance à ontologiser les structures, à y voir la « vraie » réalité...). Il s'agit de montrer en quoi la construction du modèle permet de passer d'une structuration faible, implicite, individuée, à une structuration forte, contrôlée, explicite des objets et, par-là, de gagner l'universel sans pour autant renvoyer le vécu à une simple illusion, ni tendre à ontologiser les structures. On ne se situe donc pas dans une opposition apparence/réalité, mais dans le passage de la perception d'un phénomène à la construction d'un objet. Il existe une pluralité de modèles scientifiques, ce qui rend difficile tout discours globalisant. On peut toutefois en soulever au moins trois traits distinctifs :

– le modèle est une représentation schématique qui s'efforce de reproduire les propriétés fondamentales d'un objet complexe. Le modèle est donc une simplification de la réalité. La difficulté d'une appréhension directe des phénomènes impose de passer par une « épure », un objet plus simple, plus facile à étudier ;

- le modèle ne se contente pas de fixer les connaissances relatives à un phénomène. Il a une fonction instrumentale et heuristique ; il devient l'objet même de l'analyse en tant qu'il a été réalisé sous une forme commode, apte à être manipulée : « Le modèle est un intermédiaire à qui nous déléguons la fonction de connaissance, plus précisément de réduction de l'encore énigmatique, en présence d'un champ d'étude dont l'accès, pour des raisons diverses, nous est difficile » [2.45], p. 3 ;
- le modèle s'exprime dans un langage aussi rigoureux que possible, qui exclut toute forme de métaphore. Il s'oppose clairement sur ce point au récit. Il est souvent le lieu d'une formalisation, à tel point que certains auteurs définissent tout modèle comme une représentation formelle. Ainsi, pour Malinvaud, « un modèle consiste en la représentation formelle d'idées ou de connaissances relatives au phénomène » [2.46], p. 45.

### ***2.6.2. Modèle psychocognitif d'encadrement selon Jacques Rodet***

Ce modèle psychocognitif [2.47] dont les éléments sont suggérés par les théories psycho cognitives a pour préoccupation majeure le développement chez l'apprenant des habiletés et savoir-faire cognitifs. En effet, l'acquisition de connaissances ne peut se réduire à l'enregistrement d'informations nouvelles par l'apprenant. La manière dont ce dernier va structurer ces nouvelles données est un facteur déterminant dans le processus d'apprentissage. Apprendre, c'est aussi savoir repérer ses erreurs conceptuelles et accepter de nouvelles définitions. Bachelard dit "... l'esprit scientifique ne peut se constituer qu'en détruisant l'esprit non scientifique... ". Il est important de noter que les connaissances préalables de l'apprenant ne constituent pas son point de départ mais sont les outils dont il dispose afin d'intégrer de nouvelles informations.

En formation à distance le questionnement de l'apprenant par l'encadrement peut se faire à travers différents médias. L'exemple de la Télé-université est instructif à cet égard. Au début du cours, le tuteur organise un dialogue téléphonique avec l'apprenant afin d'évaluer les stratégies d'apprentissage que ce dernier a choisies. Avant l'étude d'un nouveau concept l'apprenant est invité, par la réalisation d'une activité, à identifier ses connaissances préalables. L'utilisation d'un répertoire tout au long de sa formation permet à l'apprenant de préciser progressivement ses propres définitions et représentation des concepts abordés. Ces pratiques permettent à l'apprenant de prendre conscience de son bagage initial de connaissances, de prendre du recul et d'effectuer les ajustements nécessaires.

Ce dernier point peut se révéler conflictuel tant il est vrai que les obstacles de nature culturelle ou sociologique nécessitent parfois une véritable "conversion" de l'apprenant. Ainsi c'est tout au long du processus d'apprentissage que l'apprenant, guidé par le tuteur, doit progressivement modifier ses structures conceptuelles. En formation à distance l'utilisation de médias diversifiés et de manière synergique doit permettre à chaque apprenant de repérer son univers cognitif. L'apprentissage mise sur l'autonomie de l'apprenant. Le tuteur a le rôle de facilitateur.

L'interaction est la nature même de l'approche psycho-cognitive. Le suivi personnalisé de l'apprenant par l'encadrement est impératif. En formation à distance, où l'apprenant et l'enseignant ne sont pas en contact direct, c'est le tuteur qui assure cette interaction avec l'apprenant. Ainsi, dans une démarche psycho-cognitive, l'encadrement de l'apprenant est conçu de manière à permettre à l'apprenant d'exercer lui-même le contrôle de son apprentissage sur les plans cognitif, méthodologique, affectif et métacognitif.

### ***2.6.3. Modélisation dans les sciences cognitives***

Trois types de modélisations ont été présentés. Ces trois points de vue servent trois buts différents : l'analyse psychologique des apprenants, l'analyse des adaptations relatives à l'utilisation du module d'initiation à l'apprentissage en ligne et l'analyse de l'ergonomie de la plateforme déterminant la bonne utilisabilité du système. Chacune de ces analyses est perpétuellement soumise aux résultats des autres. Aussi, l'objet qui est au centre de l'analyse est l'étudiant mais la recherche est axée sur l'analyse des différentes activités de la plateforme.

La psychologie cognitive analyse le comportement des utilisateurs en identifiant les processus impliqués dans l'accomplissement d'actions servant la connaissance humaine. Aussi, chaque utilisateur est unique. Une fois un modèle de processus mental introduit, l'étude des psychologues se tourne vers la validation du modèle en y implémentant diverses catégories de comportements. Le modèle sera d'autant plus pertinent qu'il sera robuste aux différentes catégories d'utilisateurs. L'analyse permettra donc d'identifier des différences de comportements humains lors de la réalisation d'un même processus. Une définition généralement acceptée par la communauté scientifique décrit les processus cognitifs en ces termes :

*Définition 2.07 :*

*Les processus cognitifs sont les différents modes à travers lesquels les systèmes naturels ou artificiels traitent l'information en y répondant par une action. Ils sont les suivants :*

- Perception - Attention - Sensation
- Mémoire - Représentation - Langage
- Raisonnement - Catégorisation - Prise de décision - reconnaissance
- Apprentissage - Émotion - oubli
- Action - Comportement individuel et collectif et Phénomènes collectifs

2.6.3.1. Théorie de modélisation cognitive et analyse de tâches

*Définition 2.08 :*

Un modèle d'analyse cognitive décrit les processus mentaux par lesquels l'homme acquiert des informations sur son environnement et traite ces informations pour ajuster son comportement.

Afin de comprendre comment modéliser l'activité cognitive d'un utilisateur sur une plateforme, il importe de bien comprendre le mécanisme du processus cognitif associé. Les données récoltées lors de l'interaction ne concernent que les actions et les réactions des apprenants lors de leurs dialogues (entre pairs, avec le tuteur et avec l'awareness). Pour une analyse des tâches, il est intéressant de chercher le contexte de l'utilisateur. Kobsa et ses collègues [2.48] proposent plusieurs exemples de questions pour mettre à jour ce contexte :

- Quel est le type de l'utilisateur ?
- Quelles sont ses connaissances a priori ?
- Quelles sont ses capacités et ses compétences ?
- Quels sont ses intérêts et ses préférences ?
- Quels sont ses buts ?

La communauté scientifique cherchant des modèles pour l'analyse des tâches utilise également ces informations pour formaliser l'accomplissement des tâches par des utilisateurs. Ci-dessous quelques méthodes d'analyses de tâches sélectionnées dans la littérature :

- HTA : Hierarchical Task Analysis [2.49]



- GOMS : Goals, Operators, Methods and Selection rules [2.50]
- UAN : User Action Notation [2.51]
- MAD Méthode Analytique de Description de tâches [2.52]

#### 2.6.3.2. Quelques techniques d'analyse de tâches

a) L'évaluation d'une analyse HTA reste très subjective car les principaux critères de validité sont la consistance (le nombre de sous-tâches décrites pour chaque tâche doit être homogène) et la validité empirique (un novice doit être capable d'apprendre la tâche décrite). On trouve dans cette définition la notion de hiérarchie (une tâche se définit par un ensemble de sous-tâches) la notion d'ordonnancement, l'ordre des sous-tâches par rapport au même niveau hiérarchique et par rapport à d'autres niveaux.

b) La méthode GOMS a introduit la notion d'opérateur (outil pour faire la tâche), de méthode (comment utiliser l'outil), et l'idée d'une certaine optimisation en organisant des règles à appliquer sur les actions. Ces notions ont ensuite été largement reprises par la communauté et déclinées pour des applications plus spécifiques [2.53]. Par exemple, le modèle KLM-GOMS (Keystroke-Level Model) propose quelques règles d'optimisation et d'organisation dans l'agencement des actions à faire pour de l'édition de texte. L'automatisation d'un autre modèle de la famille GOMS, CPM-GOMS (Cognitive Perceptual Motor-GOMS), est décrite dans [2.54] ; où les auteurs comparent le temps d'exécution d'une tâche par des utilisateurs avec celui produit par un modèle fonctionnant selon les principes GOMS. Les résultats s'avèrent tout à fait satisfaisants dans les cas de tâches simples et les auteurs tentent alors de modéliser des tâches plus complexes.

c) La méthode UAN est un langage de description qui permet une évaluation analytique du scénario de la tâche. Les éléments qui composent ce langage sont :

- **Les symboles et les opérateurs.** Les opérateurs sont les différentes actions possibles sur une interface. Elles portent sur des symboles (généralement des menus, boutons, etc.).
- **Les conditions et les options.** Il s'agit des contraintes d'ordre logique permettant l'agencement de la tâche.
- **La table des actions utilisateurs, des réactions du système et des états du système.** Pour chaque action décrite par les opérateurs, une liste des réactions et des états du système doit être fournie.

- **Les relations et les contraintes temporelles.** Elles représentent les contraintes d'ordonnancement des actions.

d) La méthode MAD fournit un contexte d'application à la tâche et permet de décrire chaque tâche de manière similaire pour les comparer entre elles. Chaque intervenant dans une tâche collective à réaliser peut ainsi trouver son modèle d'activité. Une nouvelle version de ce modèle, appelée maintenant MDA (*Modèle de Description de l'Activité*) a donné lieu à une implémentation informatique [2.55]. Cet outil possède plusieurs options intéressantes pour décrire la réalisation d'une tâche. On peut ainsi éditer *l'arbre de tâches*, *les caractéristiques de tâches* et *les objets manipulés par l'utilisateur*. Un outil permet d'interroger la cohérence, l'analyse statistique et la recherche approfondie dans le modèle et un autre outil permet de simuler l'exécution d'une tâche pour vérifier notamment le bon ordonnancement des sous-tâches.

Une taxonomie des différents modèles de description de tâche est présentée dans [2.56]. La description cognitive la plus approfondie est obtenue par la méthode GOMS car elle permet de mettre en évidence chaque mouvement de l'utilisateur et ainsi, le processus cognitif associé. Comme pour tous les modèles énoncés ci-dessus et contrairement à ce que l'on souhaiterait pour une analyse automatique, les tâches ne sont pas découvertes sur les données, mais directement implantées dans le modèle.

#### 2.6.3.3. Analyse de l'engagement de l'apprenant dans son apprentissage

Il est essentiel de traiter la motivation en lien avec un outil d'apprentissage en ligne car "il n'y a pas d'apprentissage possible sans une mobilisation, une implication, un engagement plus ou moins important" dans une activité d'apprentissage sur le plan cognitif, émotionnel et comportemental [2.57].

L'engagement de l'apprenant sera analysé à partir de la motivation, de la collaboration et de la communication de l'apprenant via les forums de discussion. Un apprenant motivé, communique avec son entourage dans le dispositif de la formation, partage sa connaissance via la plateforme, développe son savoir-faire et son savoir être, et collabore entre les membres de son groupe pour les travaux demandés. L'évaluation de ces trois variables de mesure dans l'apprentissage du module d'initiation permet d'évaluer l'« engagement » de chaque apprenant.

La motivation est primordiale dans les apprentissages. Pour concevoir et développer un modèle d'apprentissage performant, nous nous sommes basés sur une combinaison d'éléments provenant de différents courants de pensée dans les sciences de l'éducation avec le modèle transmissif, behavioriste et constructiviste.

#### a) Motivation

Sur le plan motivationnel, l'autonomie de l'apprenant est souvent mise en avant dans l'apprentissage à distance. La motivation est liée aux objectifs que l'apprenant se donne en suivant une formation. Elle est très liée à la perception de l'utilité des tâches d'apprentissage, à la manière dont l'apprenant peut se faire une idée concrète du transfert dans son contexte des connaissances qu'il construit. Elle peut éliminer ou réduire l'abandon. Les motivations personnelles de l'apprenant sont agissantes pour le faire persévérer et atteindre ses objectifs. Quand un encouragement survient et parvient à atteindre l'apprenant, cela lui permet de trouver de nouvelles ressources pour continuer son action.

##### - *Quelques concepts sur la motivation*

Motivation et apprentissage sont intimement liés. De plus, s'il n'y pas d'apprentissage, il ne peut pas non plus y avoir de motivation. L'apprentissage et la motivation se renforcent donc mutuellement. Ce qui est vrai en situation scolaire ou d'études traditionnelles, mais à plus forte raison s'il s'agit de dispositifs d'apprentissage à distance de type e-learning.

Il existe deux positions extrêmement opposées qui prétendent que la motivation tire sa source soit de l'individu, c'est-à-dire sa personnalité, ses connaissances préalables, son histoire personnelle, le contexte familial et socioculturel, etc., soit de facteurs externes tels que le scénario d'apprentissage, les choix de l'enseignant.

Néanmoins, un autre point de vue, que l'on nomme couramment le "paradigme social-cognitif" a émergé dans le milieu des années 1960. Ce dernier argue que la motivation provient des représentations chez l'élève (étudiant ou apprenant) qui sont tout d'abord "la perception que l'effort investi dans la tâche conduira à sa réussite (expectancy)" p. 236, ensuite, "la perception de la réussite de la tâche permettra d'atteindre le but visé par l'élève (instrumentality)" et enfin, "la valeur et l'importance accordées au but visé en question (value)".

Cette voie se base sur trois postulats qui diffèrent des autres points de vue extrêmes évoqués plus haut.

Le premier postulat concerne les “représentations motivationnelles” soit des représentations mentales construites par l’élève dans la situation d’apprentissage (représentations de lui-même, de la tâche et de la situation).

En deuxième lieu, on considère “ces représentations motivationnelles” comme “situées”, c’est-à-dire dépendantes de la situation spécifique d’apprentissage dans laquelle se trouve l’apprenant et non pas “dispositionnelles”, c’est-à-dire dépendantes de la personnalité du sujet.

Finalement, pour l’élève ces “représentations motivationnelles” sont construites à partir d’interactions entre des facteurs internes (parcours de vie et scolaire, milieu familial et socioculturel, buts et projets personnels, conceptions de l’apprentissage, dispositions psychologiques, etc.) émanant de l’individu et des facteurs externes (le degré de difficulté de la tâche, son degré de contrôlabilité, le climat de la classe plus ou moins compétitif ou coopératif, l’autonomie ou le contrôle externe, etc.) provenant de la situation et du contexte.

- *La motivation intrinsèque et extrinsèque*

Plus récemment, à partir des années 2000, de nouvelles théories ont émergé. On pense alors que de la motivation est créée seulement si l’apprenant perçoit une valeur dans la tâche dans laquelle il ou elle s’engage et qu’il se sent capable de réussir ladite tâche. L’apprenant peut valoriser la tâche pour son intérêt intrinsèque ou extrinsèque.

*Définition 2.09 :*

La motivation est la source de ce qui fait qu’un apprenant se lance dans une tâche « non récompensée », par un système de note par exemple.

On peut toutefois scinder la motivation en motivation intrinsèque et motivation extrinsèque.

*Définition 2.10 :*

La motivation intrinsèque est le fait de vouloir exercer une activité par simple plaisir personnel sans en attendre un retour quelconque, sauf le plaisir d’exercer l’activité pour elle-même en tant que telle.

*Définition 2.11 :*

La motivation extrinsèque, quant à elle, est présente lorsqu'un sujet exerce une activité en vue d'obtenir quelque chose en échange, comme une note ou une autre forme de récompense (argent, félicitations d'une autorité parentale, familiale ou scolaire).

Par ailleurs, il est intéressant de souligner que la motivation s'en voit diminuée lorsque qu'une motivation intrinsèque se voit soumise à une contrainte ou un contrôle externe, comme par exemple le fait de noter des apprenants pour un travail.

La motivation intrinsèque s'avère plus efficace que la motivation extrinsèque. Des éléments qui peuvent influencer positivement la motivation intrinsèque d'un apprenant, comme l'autonomie (avec une suppression d'un contrôle externe) et où l'apprenant peut décider ce qu'il veut apprendre, quand il apprend, le temps qu'il veut passer à apprendre, etc.

Pour faire le lien avec la plateforme d'apprentissage, l'apprenant adulte est autonome et n'a aucun contrôle de la part de l'enseignant ou du tuteur, mais a des corrections automatiques liés aux exercices (feedbacks), à savoir une information sur sa compétence à solutionner un exercice, ce qui peut avoir un effet positif sur sa motivation intrinsèque.

Il est très probable aussi que les apprenants soient particulièrement sensibles à l'autonomie qui leur est laissée dans le développement de leurs compétences, soit avec une progression du programme de formation selon leurs choix et leurs possibilités en fonction de leur temps disponible et leur envie (ou motivation) à apprendre à un moment plutôt qu'à un moment imposé par une influence externe (des horaires stricts pour suivre un cours en présentiel par exemple). L'apprenant peut aussi en effet mesurer son degré de maîtrise d'une séquence par une évaluation de son plein gré sans jugement externe de l'enseignant ou du tuteur.

Il n'y aura pas de comparaison normative (avec un ou plusieurs membres d'un groupe ou globalement un groupe d'apprenants) qui pourrait faire baisser son estime de soi si sa performance est inférieure au benchmark (moyenne d'une classe ou un apprenant ayant plus de facilité ou ayant des prérequis ou tout simplement plus de temps à consacrer aux activités d'apprentissage).

Au contraire, l'apprenant évaluera par rapport à lui-même et décidera de travailler un sujet plutôt qu'un autre en fonction des indicateurs de performances et de progression présentés sur la plateforme.

- *La pédagogie de la maîtrise comme stratégie d'apprentissage*

Pour assurer un degré de motivation optimal, la tâche à accomplir ne doit ni être trop difficile (au risque de créer des blocages et de la frustration), ni trop facile (ce qui susciterait de l'ennui). Il s'agit de trouver un bon équilibre pour stimuler la motivation de l'apprenant pour l'aider à avancer dans son parcours d'apprentissage et à développer les compétences voulues.

C'est pourquoi, les activités à rendre sur la plateforme sont ventilées sur trois niveaux de difficulté, ce qui permet de trouver des exercices en phase avec les besoins d'apprentissage et participe à motiver l'apprenant tout au long de son programme de formation.

Notre plateforme pédagogique identifie les besoins d'apprentissage de ses apprenants avec des évaluations formatives et génère automatiquement un programme d'entraînement personnalisé, c'est-à-dire avec des exercices ciblés sur les besoins d'apprentissage, comme l'activité « Test » et l'activité « Leçon ». Ces exercices sont plus ou moins difficiles en fonction du degré de maîtrise de l'apprenant pour chaque chapitre évalué.

Nous avons emprunté cette stratégie pédagogique, à savoir la Pédagogie de la maîtrise, qui provient du courant de pensée des comportementalistes (béhavioristes) en psychologie de l'éducation. Nous avons mis un accent particulier sur cette stratégie pédagogique avérée car la pédagogie de la maîtrise avance que tout apprenant est capable de maîtriser une matière de 85 à 100%. Pour ce faire, il ou elle doit bénéficier de suffisamment de temps et d'une approche adéquate.

La pédagogie de la maîtrise comprend notamment les éléments suivants :

- des évaluations formatives (test en ligne) ;
- des procédures de régulation (programmes d'entraînement personnalisés) ;
- un séquençage des apprentissages par unités de cours ;
- des exercices de difficulté progressive ;
- des feedbacks immédiats (corrections et solutions automatiques).

## b) Interaction dans un forum de discussion

### *Définition 2.12 :*

Le forum de discussion est un outil simple d'utilisation, favorisant la collaboration. C'est un espace virtuel simple où le scripteur n'a pas besoin de manipuler des logiciels d'écriture complexes et de mise en forme spécifique afin de mettre en place des situations de communication interpersonnelles de différents types.

Hert [2.58] fera cependant, la différence entre les forums qualifiés de « libres » où les discussions se font spontanément et à bâtons rompus, sans objectif défini à l'avance et les forums qualifiés de « fermés » qui sont limités dans le temps et qui sont porteurs d'objectifs.

## c) Collaboration

### *Définition 2.13 :*

« Etymologiquement, collaborer signifie travailler ensemble, ce qui implique un concept de buts partagés et une intention explicite “d'ajouter de la valeur”, de créer quelque chose de nouveau ou de différent par la collaboration, par opposition à un simple échange d'informations ou à une transmission de consignes » [2.59].

En effet,

### *Définition 2.14 :*

« Est apprentissage collaboratif toute activité d'apprentissage réalisée par un groupe d'apprenants ayant un but commun, étant chacun source d'information, de motivation, d'interaction, d'entraide... et bénéficiant chacun des apports des autres, de la synergie du groupe et de l'aide d'un tuteur facilitant les apprentissages individuels et collectifs » [2.60]. Cette activité de communication entre personnes, en visualisant leurs traces scripturales, permet de développer des habilités intellectuelles et confronter ses propres stratégies de raisonnement et d'apprentissage par rapport à celles des autres.

### *Définition 2.15 :*

Le forum de discussion éducatif est un espace ouvert à un groupe restreint d'apprenants sur une plateforme d'enseignement en ligne et où ces derniers sont appelés à communiquer entre eux, à apporter leurs contributions à un sujet, à réagir aux messages postés par les autres participants et à partager des connaissances.

Le forum éducatif confère au temps une certaine souplesse [2.61] qui plaît aux apprenants dont le temps d'apprentissage est toujours en retard par rapport au temps didactique. Il permet à l'écrit d'être visible en permanence sur la plateforme, ce qui lui permet d'être visité à tout moment par les apprenants qui peuvent le partager et se l'approprier.

Les groupes peuvent se former spontanément pour mener un travail collectif, en ayant du plaisir à communiquer dans un climat convivial, où social, affectif et cognitif se conjuguent et où une information qui passe est une information répertoriée et affichée, mais surtout accessible à tous à tout moment. Le forum se substitue aux interactions en classe à travers des interactions à distance qui ont une dimension langagière essentiellement écrite et une flexibilité temporelle qui dépasse le cadre de la classe classique.

#### d) Communication

Dans son œuvre principale [2.62], Bavelas avait déjà montré que la structure de communication affecte la performance d'un groupe restreint (de 3 à 20 personnes) en mesurant la communication selon le concept de distance, et cela sur divers types de réseaux. Il montra que la configuration du réseau détermine l'organisation des échanges au sein du groupe. Il avait établi que les réseaux de transmission peuvent affecter l'accomplissement du groupe ainsi que la satisfaction des membres de groupe d'occuper certaines positions dans le réseau.

Les travaux de Bavelas sur le travail en réseau du type « all- channels » avaient déjà permis de comprendre que dans un travail de résolution de problème en réseau, les messages se classent en trois catégories : demandes d'information et réponses, propositions de méthode de travail et échanges d'informations relationnelles et affectives. Dans notre recherche, nous pouvons conserver ces catégories mais avec une touche de perfectionnement.



## **2.7. Conclusion**

Au cours de ces deux premiers chapitres, le dispositif de la formation à distance, le concept et le système d'apprentissage asynchrone dans une FOAD ont été exposés en premier lieu. Ensuite, les théories de l'apprentissage, la théorie de l'analyse des traces numériques, les modèles d'enseignement et d'apprentissage ont été présentés. Il contient également la modélisation dans les sciences sociales et dans les sciences cognitives.

Les traces des activités asynchrones, ont été récoltées : type d'acteur de la situation choisie, nombre de nouvelles discussions lancées et fréquence d'intervention (variable quantitative), en plus de la qualité de contenu posté (variable qualitative) au forum de discussion, avec le résultat (variable latente) de l'apprentissage. Les admis correspondent au code-résultat 1 tandis que les autres ont un code-résultat 0 (redoublants, renvoyés, abandons). L'étude descriptive documente comment le discours asynchrone élaboré collectivement est réinvesti par les apprenants. Les variables de l'intervention au forum de discussion favorisant l'apprentissage ont servi de cadre de référence. La méthodologie combine des analyses qualitatives et quantitatives. Le logiciel de traitement statistique R a été utilisé pour le traitement de données.

Ce deuxième chapitre permet d'avoir les techniques d'analyse des activités sur la plateforme, afin de modéliser l'apprentissage asynchrone dans une FOAD.

## **CHAPITRE 3 – THEORIES DE BASE SUR LE MODELE D'EQUATION STRUCTURELLE**

### **3.1 Introduction**

La culture numérique vient souvent au-devant de nos préoccupations, en particulier lorsqu'il s'agit de technologies liées aux apprentissages ou à l'enseignement à distance. Bien sûr, le numérique est quasiment toujours présent dans beaucoup de domaines de notre vie, de notre société et de notre culture. Cette recherche a été menée dans le cadre de cette culture numérique, plus précisément dans une FOAD après quelques années de fonctionnement. Une plateforme de formation est considérée comme terrain sur lequel l'apprentissage s'est déroulé et la trace des activités a été relevée.

Plusieurs communautés recourent à des environnements technologiques pour soutenir le travail, la collaboration et l'apprentissage, dont le forum de discussion qui procède par échange de messages textuels asynchrones [2.40]. Qu'elles soient constituées d'apprenants et de leurs formateurs ou de professionnels engagés dans une amélioration de leur pratique au quotidien, les communautés virtuelles ont un dénominateur commun : elles peuvent toutes être considérées comme des lieux d'apprentissage dans le plein sens du terme comme le proposent Henri et Pudelko [3.01] et, comme le démontre Wenger [3.02] en s'appuyant sur la théorie de l'apprentissage dont le concept central est la participation, synonyme d'apprentissage et de construction de connaissance. Cette recherche s'intéresse aux analyses de ces forums et tente de proposer un modèle d'équation conceptuelle structurelle, afin d'examiner la relation entre l'attitude face à l'apprentissage asynchrone et le résultat de la formation, pour ensuite démontrer les effets directs et indirects de l'engagement de l'apprenant dans une FOAD. Le présent chapitre présente alors dans un premier temps les outils et méthodes statistiques mobilisés pour pouvoir élaborer ce modèle d'équation structurelle. Et ce, en abordant la notion de base en statistique, l'analyse multivariée et l'équation structurelle. Dans un second temps, le modèle d'équation structurelle (SEM) proprement dite sera présenté à travers la modélisation, les fondements conceptuels et les différentes étapes d'élaboration du SEM.

### 3.2 Notion de base en statistique

La présente recherche utilise une approche empirique fondée sur la collecte et le traitement statistique de données par des méthodes quantitatives et qualitatives dans le domaine des sciences cognitives. Les avantages des méthodes quantitatives sont nombreux comparés aux méthodes qualitatives : l'observation, l'entretien de groupe de discussion et l'interview, en termes d'objectivité, grâce à l'emploi de méthodes statistiques et la possibilité de reproduction et de généralisation sur différents échantillons dans des contextes différents avec les mêmes outils de mesure. Une grande partie de la recherche quantitative a utilisé la modélisation d'équations structurelles. Avant d'entrer dans les équations structurelles, rappelons d'abord quelques définitions et termes clés utilisées dans les formules et les différentes fonctions de la recherche.

#### 3.2.1.Définitions

Commençons par la *statistique*,

*Définition 3.01 :*

La statistique est l'art et la science de collecter, analyser, présenter et interpréter des données.

*Définition 3.02 :*

Les *données* sont les faits et les chiffres qui sont collectés, analysés et résumés pour pouvoir ensuite être interprétés.

*Définition 3.03 :*

Toutes les données collectées dans la plateforme Moodle de formation à distance forment l'*ensemble de données* de la recherche.

*Définition 3.04 :*

Les *éléments* sont les entités auprès desquelles les données sont collectées.

*Définition 3.05 :*

Une *variable* est une caractéristique des éléments à laquelle on s'intéresse. Les données sont obtenues en collectant des informations sur chaque variable pour tous les éléments de l'étude.

*Définition 3.06 :*

L'ensemble des informations obtenues pour un élément particulier correspond à *une observation*.

De nombreuses situations requièrent des données relatives à un vaste ensemble d'éléments. Dans un souci de considérations de certains variables, les données ne peuvent être collectées qu'auprès d'une petite partie du groupe concerné. Le groupe considéré dans son ensemble est désigné par le terme *population* et la petite partie du groupe, par le terme *échantillon*. Formellement, les définitions suivantes sont utilisées.

*Définition 3.07 :*

Une *population* est l'ensemble des éléments considérés dans une étude particulière.

Un *échantillon* est un sous-ensemble de la population.

#### 3.2.1.1. Moyenne

*Définition 3.08 :*

La *moyenne*, ou *valeur moyenne* est la mesure de tendance centrale la plus importante pour une variable. Si les données sont issues d'un échantillon, la moyenne est notée  $\bar{x}$  ; si les données sont issues d'une population, la moyenne est notée  $\mu$ .

Moyenne d'échantillon :

$$\bar{x} = \frac{\sum x_i}{n} \quad (3.01)$$

Où

$n$  : nombre d'observation de l'échantillon et  $x_i$  : observation de rang  $i$

C'est une statistique d'échantillon.

Moyenne de la population :

$$\mu = \frac{\sum x_i}{N} \quad (3.02)$$

Avec

N : nombre d'observation dans une population et  $x_i$  : observation de rang i

La moyenne d'échantillon  $\bar{x}$  est un estimateur ponctuel de la moyenne de la population  $\mu$ .

#### 3.2.1.2. Variance

*Définition 3.09 :*

Il est souvent utile de considérer des mesures de variabilité ou de dispersion des données. La variance est basée sur la différence entre la valeur de chaque observation ( $x_i$ ) et la valeur moyenne ( $\bar{x}$  pour un échantillon,  $\mu$  pour la population). Cette différence entre chaque observation  $x_i$  et la moyenne est appelée *écart par rapport à la moyenne* ; et s'écrit  $(x_i - \bar{x})$  pour un échantillon et  $(x_i - \mu)$  pour une population. Pour calculer la variance, les écarts par rapport à la moyenne sont élevés au carré.

*Définition 3.10 :*

Si les données sont issues d'une population, la moyenne des écarts au carré est appelée *variance de la population*, et est notée par le symbole  $\sigma^2$ .

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N} \quad (3.03)$$

Dans la plupart des études statistiques, les données à analyser sont issues d'un échantillon. Le calcul de la variance d'un échantillon nous permet généralement ensuite d'estimer la variance de la population  $\sigma^2$ .

Si la somme des écarts par rapport à la moyenne au carré est divisée par  $n-1$  et non par  $n$ , la variance de l'échantillon fournira un estimateur sans biais de la variance de la population. C'est pour cette raison que la variance de l'échantillon, notées  $s^2$ , est définie de la façon suivante :

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} \quad (3.04)$$

La variance est utile pour comparer la dispersion de plusieurs variables.

### 3.2.1.3. Ecart-type

*Définition 3.11 :*

L'*écart-type* correspond à la racine carrée de la variance. En utilisant les notations adoptées pour définir la variance d'échantillon et la variance de la population, l'écart-type est déduit de la variance de la façon suivante :

$$\text{Ecart-type de l'échantillon} = s = \sqrt{s^2} \quad (3.05)$$

$$\text{Ecart-type de la population} = \sigma = \sqrt{\sigma^2} \quad (3.06)$$

L'écart-type de l'échantillon est l'estimateur de l'écart-type de la population  $\sigma$ .

### ***3.2.2. Mesure de la relation entre deux variables***

Jusqu'à présent, nous avons vu les méthodes numériques utilisées pour résumer les données d'une variable à un moment donné. Dans la recherche, nous nous intéressons à la relation entre deux variables. Dans cette section, nous rappelons la covariance et la corrélation, des mesures descriptives de la relation entre deux variables.

### 3.2.2.1. Covariance

*Définition 3.12 :*

Pour un échantillon de taille  $n$  composé des observations  $(x_1, y_1), (x_2, y_2), \text{etc.}$ , la *covariance de l'échantillon* est définie par :

$$s_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n - 1} \quad (3.07)$$

La formule de calcul de la **covariance pour une population** de taille  $N$  est similaire à la formule (3.07).

$$\sigma_{xy} = \frac{\sum (x_i - \mu_x)(y_i - \mu_y)}{N} \quad (3.08)$$

Avec :

$\mu_x$  : est la moyenne de la population de la variable  $x$  et

$\mu_y$  : est la moyenne de la population de la variable  $y$

La covariance de la population  $\sigma_{xy}$  est définie pour une population de taille  $N$ .

### 3.2.2.2. Interprétation de la covariance

La covariance est une mesure de la relation linéaire entre deux variables. Une valeur positive élevée de la covariance semble indiquer une forte relation positive et une valeur négative élevée de la covariance semble indiquer une forte relation négative. Cependant, l'utilisation de la covariance comme mesure de la robustesse de la relation linéaire présente un inconvénient : la valeur de la covariance dépend de l'unité de mesure des variables  $x$  et  $y$ . Le coefficient de corrélation est une mesure de la relation entre deux variables qui n'est pas exposée à ce type de problème.

### 3.2.2.3. Coefficient de corrélation

*Définition 3.13 :*

Pour un échantillon de données, le *coefficient de corrélation de Pearson* est défini par :

$$r_{xy} = \frac{S_{xy}}{S_x S_y} \quad (3.09)$$

Où

$r_{xy}$  correspond au coefficient de corrélation de l'échantillon

$S_{xy}$  correspond à la variance de l'échantillon

$S_x$  correspond à l'écart-type d'échantillon de  $x$

$S_y$  correspond à l'écart-type d'échantillon de  $y$

Pour des données issues d'une population, la formule de calcul du coefficient de corrélation est définie par :

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (3.10)$$

Où

$\rho_{xy}$  correspond au coefficient de corrélation de la population

$\sigma_{xy}$  correspond à la variance de la population

$\sigma_x$  correspond à l'écart-type d'échantillon de  $x$ , au niveau de la population

$\sigma_y$  correspond à l'écart-type de  $y$ , au niveau de la population

Le coefficient de corrélation de l'échantillon  $r_{xy}$  fournit une estimation du coefficient de corrélation de la population  $\rho_{xy}$ .

### 3.2.2.4. Interprétation du coefficient de corrélation

Le coefficient de corrélation varie entre -1 et +1. Des valeurs proches de -1 ou de +1 révèlent une forte relation linéaire. Plus le coefficient est proche de zéro, plus la relation est faible.



Un coefficient de corrélation +1 correspond à une relation parfaitement linéaire et positive entre  $x$  et  $y$ . A l'inverse, un coefficient de corrélation -1 correspond à une relation parfaitement linéaire et négative entre  $x$  et  $y$ . Une valeur de  $r_{xy}$  égale à zéro indique l'absence de relation linéaire entre  $x$  et  $y$ , et des valeurs de  $r_{xy}$  proches de zéro révèlent une faible relation linéaire.

Pour conclure, soulignons que la corrélation fournit une mesure de la relation linéaire mais pas nécessairement une relation de causalité. Une corrélation importante entre deux variables ne signifie pas nécessairement que des changements intervenant sur l'une des variables se traduiront par des changements sur l'autre variable [3.03].

### 3.2.3. Matrice de corrélation

Définition 3.14 :

La *matrice de corrélation* d'un vecteur de  $p$  variables aléatoires  $\vec{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}$  dont chacune possède une variance (finie) est la matrice carrée dont le terme générique est donné par :  $r_{i,j} = Cor(X_i, Y_j)$

Les termes diagonaux de cette matrice sont égaux à 1, elle est symétrique, semi-définie positive et ses valeurs propres sont positives ou nulles.

### 3.2.4. Estimation

Partant d'un échantillon  $(x_i, y_i), 1 \leq i \leq n$ , de réalisations indépendantes de deux variables  $X$  et  $Y$ , un *estimateur* (biaisé) *du coefficient de corrélation* est donné par

$$\hat{r}_p = \frac{\hat{\sigma}_{XY}}{\hat{\sigma}_X \hat{\sigma}_Y} \quad (3.11)$$

Avec :

$$- \hat{\sigma}_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{N}$$

$$- \hat{\sigma}_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \quad \text{et} \quad \hat{\sigma}_y = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2}$$

$$- \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad \text{et} \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$$

qui sont respectivement des estimateurs de la covariance, des écarts-types et des espérances des variables  $x$  et  $y$ .

-  $N$  : nombre d'observations de la population

Proposition 3.01.

Numériquement,  $\hat{r}_p$  est compris entre  $[-1,1]$

*Démonstration :*



On a par définition :

$$\hat{r}_p = \frac{\hat{\sigma}_{XY}}{\hat{\sigma}_X \hat{\sigma}_Y}$$

$$\text{Or } \hat{\sigma}_{xy} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{N}, \quad \hat{\sigma}_X = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \quad \text{et} \quad \hat{\sigma}_Y = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2}$$

Alors :

$$\hat{r}_p = \frac{\frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{N}}{\sqrt{\frac{1}{N^2} \sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}}$$

D'après l'inégalité triangulaire :

$$\sum_{i=1}^N (x_i - \bar{x})^2 \leq \left( \sum_{i=1}^N (x_i - \bar{x}) \right)^2$$

$$\sum_{i=1}^N (y_i - \bar{y})^2 \leq \left( \sum_{i=1}^N (y_i - \bar{y}) \right)^2$$

Il s'ensuit que :  $\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2 \leq (\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}))^2$

Et que :

$$\begin{aligned} -\sqrt{\left( \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \right)^2} &\leq \sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2} \leq \sqrt{\left( \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \right)^2} \\ -\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) &\leq \sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2} \leq \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \end{aligned}$$

Alors :

$$-\frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})} = -1 \leq \hat{r}_p \leq \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})} = 1$$

◆

Les deux séries ne sont pas linéairement corrélées si  $\hat{r}_p$  est nul. Les deux séries sont d'autant mieux corrélées que  $\hat{r}_p$  est proche de 1 ou de -1.

Appliqué à une matrice de corrélation, le test de sphéricité de Bartlett permet de juger si les coefficients extra-diagonaux sont globalement différents de zéro.

### 3.2.5. Matrice de covariance

Définition 3.15 :

Par définition, la *covariance de deux variables* aléatoires réelles  $X$  et  $Y$  ayant chacune une variance finie, notée  $Cov(X, Y)$  ou  $\sigma_{xy}$  comme indiquée dans la formule (3.08), est la valeur :

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y])] \quad (3.12)$$

Où

- $E$  désigne l'espérance mathématique
- La variance de  $X$  est donc la  $Cov(X, X)$ .

#### 3.2.5.1. Définition de la matrice de covariance

Définition 3.16 :

La *matrice de covariance d'un vecteur* de  $p$  variables aléatoires  $\vec{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}$  dont chacune possède une variance, est la matrice carrée dont le terme générique est donné par :  $a_{i,j} = Cov(X_i, X_j)$

La matrice de covariance, notée parfois  $\Sigma$ , est définie par :

$$Var(\vec{X}) = E \left[ \left( \vec{X} - E(\vec{X}) \right) \left( \vec{X} - E(\vec{X}) \right)^T \right] \quad (3.13)$$

En développant les termes :

$$Var(\vec{X}) = \begin{pmatrix} Var(X_1) & \cdots & Cov(X_1, X_p) \\ \vdots & \ddots & \vdots \\ Cov(X_p, X_1) & \cdots & Var(X_p) \end{pmatrix} = \begin{pmatrix} \sigma_{x_1}^2 & \cdots & \sigma_{x_1 x_p} \\ \vdots & \ddots & \vdots \\ \sigma_{x_p x_1} & \cdots & \sigma_{x_p}^2 \end{pmatrix} \quad (3.14)$$

### 3.2.5.2. Propriétés de la matrice de covariance

Proposition 3.02.

La matrice de covariance est symétrique.

*Démonstration :*



Par la définition de la matrice de covariance :

$$\text{Var}(\vec{X}) = \begin{pmatrix} \text{Var}(X_1) & \cdots & \text{Cov}(X_1, X_p) \\ \vdots & \ddots & \vdots \\ \text{Cov}(X_p, X_1) & \cdots & \text{Var}(X_p) \end{pmatrix} = \begin{pmatrix} \sigma_{x_1}^2 & \cdots & \sigma_{x_1 x_p} \\ \vdots & \ddots & \vdots \\ \sigma_{x_p x_1} & \cdots & \sigma_{x_p}^2 \end{pmatrix}$$

Or  $\text{Cov}(X_i, X_j) = \text{Cov}(X_j, X_i)$  par la commutativité du produit dans l'expression  $\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$ .

En effet :

$$\begin{aligned} \text{Cov}(X_i, X_j) &= E[(X_i - E[X_i])(X_j - E[X_j])] = E[X_i X_j - X_i E(X_j) - X_j E(X_i) + E(X_i)E(X_j)] \\ &= E(X_i X_j) - E(X_i)E(X_j) - E(X_j)E(X_i) + E(X_i)E(X_j) = E(X_i X_j) - E(X_i)E(X_j) \end{aligned}$$

Et  $E(X_i X_j) - E(X_i)E(X_j) = E(X_j)E(X_i) - E(X_j)E(X_i)$ , ou  $\text{Cov}(X_i, X_j) = \text{Cov}(X_j, X_i)$

Alors la matrice de covariance est symétrique.



Proposition 3.03.

Ses éléments diagonaux sont les variances et les éléments extra-diagonaux sont les covariances des couples de variables.

*Démonstration :*



Ceci découle de la définition même de la matrice de covariance. Les éléments diagonaux sont  $Var(X_i)$  et les éléments extra-diagonaux sont  $Cov(X_i, X_j)$ .

♦

Proposition 3.04.

Elle est semi-définie positive si ses valeurs propres sont positives ou nulles.

*Démonstration :*

♣

Notons  $M = Var(\vec{X})$  la matrice de covariance.

Supposons que  $M$  soit semi-définie positive, alors pour tout vecteur colonne  $Q$  non nul, on a :  $Q'MQ \geq 0$ . Soit  $\lambda \in \mathbb{R}$  une valeur propre de  $M$ . Puisque  $Q \neq 0$ , on a  $Q'Q = \|Q\|_2^2 > 0$  puis :

$$\lambda = \frac{Q'(\lambda Q)}{Q'Q} = \frac{Q'MQ}{Q'Q} \geq 0$$

Alors les valeurs propres de  $M$  sont positives ou nulles.

Réciproquement, on suppose que les valeurs propres de la matrice de covariance  $M$  sont positives ou nulles.

D'après le théorème spectral,  $M$  est orthogonalement semblable à une matrice diagonale. Donc,  $\exists P \in \mathcal{O}_n(\mathbb{R}), \exists D = diag(\lambda_i)_{1 \leq i \leq n} \in \mathcal{D}_n(\mathbb{R})$  tel que  $M = PDP^{-1} = PD'P$ .

Soit  $Q$  un vecteur colonne quelconque.

En posant :  $Q = (q_i)_{1 \leq i \leq n}$  puis  $Q_1 = P'Q = (q_{1i})_{1 \leq i \leq n}$ ,

$$Q'MQ = Q'(PD'P)Q = (P'Q)'D(P'Q) = Q_1'DQ_1 = \sum_{i=1}^n \lambda_i q_{1i}^2 \geq 0$$

Alors la matrice de covariance  $M$  est semi-positive.



*Définition 3.17 :*

L'inverse de la matrice de covariance est parfois désigné « matrice de précision ».

### 3.2.5.3. Interprétation de la matrice de covariance

La matrice de covariance est un outil essentiel pour l'analyse multivariée : l'analyse en composantes principales qui exploite la diagonalisation de cette matrice et l'analyse discriminante qui se fonde sur l'examen des coefficients de cette matrice. Dans la matrice de covariance, les éléments hors diagonale contiennent les covariances de chaque paire de variables. Les éléments sur la diagonale de la matrice de covariance contiennent les variances de chaque variable. La variance mesure le degré de dispersion des données autour de la moyenne. La variance est égale à la racine carrée de l'écart-type.

Concernant le test sur la matrice de covariance, le test de sphéricité de Bartlett permet de déterminer si les composantes hors de la diagonale de la matrice sont différentes de zéro, en d'autres termes s'il y a une relation entre les différentes variables prises en considération.

### **3.2.6. Test de sphéricité de Bartlett**

*Définition 3.18 :*

Le *test de sphéricité de Bartlett* est un test statistique relatif à l'indépendance globale des composantes d'un vecteur aléatoire. Il est basé sur le déterminant d'une estimation de la matrice de corrélation.

Partant d'un échantillon de  $n$  réalisations (indépendantes) d'un ensemble de  $p$  variables aléatoires réelles  $X_1, \dots, X_p$ , le test concerne la validité de :

- $H_0$  (Hypothèse nulle) : les variables sont globalement indépendantes.
- $H_1$  : les variables sont globalement dépendantes.

En se basant sur une estimation  $R$  de la matrice de corrélation, le test évalue

$$\chi^2 = -\left(n - 1 - \frac{2p + 5}{6}\right) \log(|\det R|) \quad (3.15)$$

qui, sous  $H_0$ , suit « approximativement » une loi du  $\chi^2$  disposant de  $\frac{p(p+1)}{2}$  degrés de liberté.

Remarque : Si les variables sont indépendantes, la matrice de corrélation est égale à la matrice identité, son estimation  $R$  devrait s'en approcher, son déterminant avoisine 1 et  $\log(|\det R|) \approx 0$ . Dans le cas contraire,  $R$  devient singulière, le déterminant s'approche de zéro et le  $\log(|\det R|)$  prend des valeurs négatives.

### 3.3 Analyse multivariée

En statistique, les analyses multivariées ont pour caractéristique de s'intéresser à la distribution conjointe de plusieurs variables. Les analyses bivariées sont des cas particuliers à deux variables. Les méthodes d'analyse des données en sont un sous-ensemble.

Les analyses multivariées sont très diverses selon l'objectif recherché, la nature des variables et la mise en œuvre formelle. Des méthodes descriptives visant à structurer et à résumer l'information ainsi que des méthodes explicatives visant à expliquer une ou des variables dites « dépendantes » (variables à expliquer) par un ensemble de variables dites « indépendantes » (variables explicatives), forment deux grandes familles de l'analyse multivariée.

#### 3.3.1. Régression linéaire

En statistiques, en économétrie et en apprentissage automatique, un modèle de régression linéaire est un modèle de régression qui cherche à établir une relation linéaire entre une variable, dite expliquée, et une ou plusieurs variables, dites explicatives. On parle aussi de modèle linéaire ou de modèle de régression linéaire.



Parmi les modèles de régression linéaire, le plus simple est l'ajustement affine. Celui-ci consiste à rechercher la droite permettant d'expliquer le comportement d'une variable statistique  $y$  comme étant une fonction affine d'une autre variable statistique  $x$ .

$$y = ax + b$$

En général, le modèle de régression linéaire désigne un modèle dans lequel l'espérance conditionnelle de  $y$  sachant  $x$  est une transformation affine en les paramètres. Cependant, on peut aussi considérer des modèles dans lesquels c'est la médiane conditionnelle de  $y$  sachant  $x$  ou n'importe quel quantile de la distribution de  $y$  sachant  $x$  qui est une transformation affine en les paramètres.

Le modèle de régression linéaire est souvent estimé par la méthode des moindres carrés mais il existe aussi de nombreuses autres méthodes pour estimer ce modèle. On peut par exemple estimer le modèle par maximum de vraisemblance ou encore par inférence bayésienne.

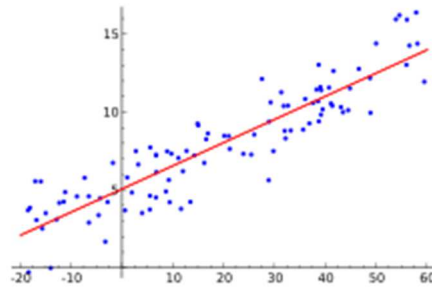
Bien qu'ils soient souvent présentés ensemble, le modèle linéaire et la méthode des moindres carrés ne désignent pas la même chose. Le modèle linéaire désigne une classe de modèles qui peuvent être estimés par un grand nombre de méthodes et, la méthode des moindres carrés désigne une méthode d'estimation. Elle peut être utilisée pour estimer différents types de modèles.

### ***3.3.2. Modèle linéaire simple***

On appelle généralement modèle linéaire simple un modèle de régression linéaire avec une seule variable explicative. On a donc deux variables aléatoires, une variable expliquée  $Y$ , qui est un scalaire, une variable explicative  $X$ , également scalaire. On dispose de  $n$  réalisations de ces variables  $(x_i) 1 \leq i \leq n$  et  $(y_i) 1 \leq i \leq n$ , soit :

$$y = \beta_0 + \beta_1 x_i + e_i \tag{3.16}$$

où  $e_i$  est le terme d'erreur ; chaque terme d'erreur lui-même est une réalisation d'une variable aléatoire  $E_i$ .



**Figure 3.01 : Droite de régression**

Dans le cadre d'un modèle linéaire simple, on peut représenter graphiquement la relation entre  $x$  et  $y$  à travers un nuage de points. L'estimation du modèle linéaire permet de tracer la droite de régression. Le paramètre  $\beta_0$  représente l'ordonnée à l'origine et  $\beta_1$  le coefficient directeur de la droite.

### 3.3.3. *Modèle linéaire multiple*

Par opposition au modèle de régression linéaire simple, le modèle de régression linéaire multiple est défini comme tout modèle de régression linéaire avec au moins deux variables explicatives. La régression linéaire multiple est une généralisation, à  $p$  variables explicatives, de la régression linéaire simple. Etant donné un échantillon  $(Y_i, X_{i1}, \dots, X_{ip})$ ,  $i = 1, \dots, n$  nous cherchons à expliquer, avec le plus de précision possible, les valeurs prises par  $Y_i$ , dite variable endogène, à partir d'une série de variables explicatives  $(X_{i1}, \dots, X_{ip})$ . Le modèle théorique, formulé en termes de variables aléatoires, prend la forme :

$$Y_i = a_0 + a_1X_{i1} + a_2X_{i2} + \dots + a_pX_{ip} + \epsilon_i \quad (3.17)$$

Avec  $i = 1, \dots, n$

où

- $\epsilon_i$  est l'erreur du modèle qui exprime, ou résume, l'information manquante dans l'explication linéaire des valeurs de  $Y_i$  à partir des  $X_{i1}, \dots, X_{ip}$  (problème de spécifications, variables non prises en compte, etc.)
- $a_0, a_1, \dots, a_p$  sont les paramètres à estimer.

### Notation matricielle

Nous pouvons adopter une écriture condensée qui rend la lecture et la manipulation de l'ensemble plus facile. Les équations suivantes

$$\begin{cases} y_1 = Y_i a_0 + a_1 x_{11} + a_2 x_{12} + \dots + a_p x_{1p} + \epsilon_1 \\ y_2 = Y_i a_0 + a_1 x_{21} + a_2 x_{22} + \dots + a_p x_{2p} + \epsilon_2 \\ \dots \\ y_n = Y_i a_0 + a_1 x_{n1} + a_2 x_{n2} + \dots + a_p x_{np} + \epsilon_n \end{cases} \quad (3.18)$$

peuvent être résumées avec la notation matricielle

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & \dots & x_{1p} \\ \vdots & \ddots & \vdots \\ 1 & \dots & x_{np} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix} \quad (3.19)$$

Soit de manière compacte :  $y = Xa + \epsilon$

avec

- $y$  est un vecteur colonne de dimension  $(n, 1)$
- $X$  est une matrice de dimension  $(n, p + 1)$
- $a$  est un vecteur de dimension  $(p + 1, 1)$
- $\epsilon$  est un vecteur colonne de dimension  $(n, 1)$
- la première colonne de la matrice  $X$  sert à indiquer que nous procédons à une régression avec une constante.

#### 3.3.4. Loi du Khi-deux

La **loi du  $\chi^2$**  (« khi-deux ») est une loi à densité de probabilité. Cette loi est caractérisée par un paramètre dit *degré de liberté* à valeur dans l'ensemble des entiers naturels non nuls.

Soient  $x_1, \dots, x_k$ ,  $k$  variables aléatoires indépendantes suivant des lois normales de moyennes respectives  $\mu_i$  et d'écart-type  $\sigma_i$  ;  $Y_i = \frac{x_i - \mu_i}{\sigma_i}$  leurs variables sont centrées et réduites.

*Définition 3.19 :*

La variable aléatoire  $X$ , telle que

$$X = \sum_{i=1}^k Y_i^2 = \sum_{i=1}^k \left( \frac{X_i - \mu_i}{\sigma_i} \right)^2 \quad (3.20)$$

suit une loi du  $\chi^2$  à  $k$  degrés de liberté a pour fonction densité  $f_X$

$$f_X(t) = \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} t^{\frac{k}{2}-1} e^{-\frac{t}{2}} \quad (3.21)$$

pour tout  $t$  positif où  $\Gamma$  est la fonction gamma.

### 3.3.5. Loi de Wishart

En théorie des probabilités et en statistique, la loi de Wishart est la généralisation multidimensionnelle de la loi du  $\chi^2$ . C'est une famille de lois de probabilité sur les matrices définies positives, symétriques. Une variable aléatoire de loi de Wishart est donc une matrice aléatoire. Trois lois sont d'une grande importance dans l'estimation des matrices de variance-covariance.

Si une variable aléatoire  $S$  suit une loi de Wishart, on notera

$$S \sim W_p(V, n) \text{ ou } W(V, p, n).$$

*Définition 3.20 :*

Supposons que  $X$  est une matrice  $n \times p$ , les lignes sont des vecteurs aléatoires indépendants et suivent une loi normale  $p$ -dimensionnelle centrée :  $X_{(i)} = (x_i^1, \dots, x_i^p) \sim N_p(0, V)$ .

Alors la loi de Wishart est la loi de probabilité de la matrice de dimension  $p \times p$

$$S = X^T X \quad (3.22)$$

connue sous le nom matrice de dispersion. L'entier naturel  $n$  est le nombre de degrés de liberté. Pour  $n > p$ , la matrice  $S$  est inversible avec probabilité 1 si  $V$  est inversible. Si  $p = 1$  et  $V = 1$ , alors la loi de Wishart est la loi du  $\chi^2$  à  $n$  degrés de liberté.

La loi de Wishart apparait comme la loi d'une matrice de variance-covariance d'un échantillon de valeurs suivant une loi normale multidimensionnelle. Elle peut être caractérisée par sa densité de probabilité de la manière suivante. On fixe  $V$  une matrice  $p \times p$  symétrique définie positive (paramètre d'échelle). Si  $n \geq p$ , alors la densité de probabilité de la loi de Wishart est donnée par :

$$f(X) = \frac{|X|^{\frac{1}{2}(n-p-1)} e^{\left(-\frac{1}{2} \text{tr}(XV^{-1})\right)}}{|V|^{\frac{n}{2}} \cdot 2^{\frac{np}{2}} \cdot \Gamma_p\left(\frac{n}{2}\right)} \quad (3.23)$$

pour toute  $n \times p$  -matrice  $X$  symétrique définie positive, et où  $\Gamma_p$  est la fonction gamma multidimensionnelle définie par :

$$\Gamma_p\left(\frac{n}{2}\right) = \pi^{\frac{p(p-1)}{4}} \prod_{j=1}^p \Gamma\left(\frac{n}{2} + \frac{1-j}{2}\right) \quad (3.24)$$

Comme la définition précédente peut être étendue à tout réel  $n \geq p$ . Si  $n < p$ , la loi de Wishart n'a plus de densité, mais devient une loi singulière.

### 3.3.6. Méthode des moindres carrés

La méthode des moindres carrés, indépendamment élaborée par Legendre et Gauss au début du XIX<sup>ème</sup> siècle, permet de comparer des données expérimentales, généralement entachées d'erreurs de mesure, à un modèle mathématique censé décrire ces données.

Ce modèle peut prendre diverses formes. Il peut s'agir de lois de conservation que les quantités mesurées doivent respecter. La méthode des moindres carrés permet alors de minimiser l'impact des erreurs expérimentales en « ajoutant de l'information » dans le processus de mesure.

Présentation de la méthode :

Soit  $f(x; \theta)$  le modèle théorique qui est une famille de fonction d'une ou plusieurs variables muettes  $x$  indexées par un ou plusieurs paramètres  $\theta$  inconnus. La méthode des moindres carrés permet de sélectionner parmi ces fonctions celle qui reproduit le mieux les données expérimentales. On parle dans ce cas d'ajustement par la méthode des moindres carrés. Si les paramètres  $\theta$  ont un sens physique, la procédure d'ajustement donne également une estimation indirecte de la valeur de ces paramètres.

La méthode consiste en une prescription (initialement empirique), qui est que la fonction  $f(x; \theta)$  qui décrit « le mieux » les données est celle qui minimise la somme quadratique des déviations des mesures aux prédictions de  $f(x; \theta)$ . Si, par exemple, nous disposons de  $N$  mesures  $(y_i)_{i=1, \dots, N}$ , les paramètres  $\theta$  « optimaux » au sens de la méthode des moindres carrés sont ceux qui minimisent la quantité :

$$S(\theta) = \sum_{i=1}^N (y_i - f(x_i; \theta))^2 = \sum_{i=1}^N r_i^2(\theta) \quad (3.25)$$

Où les  $r_i(\theta)$  sont les *résidus* du modèle, c'est l'écart entre la mesure  $y_i$  et la prédiction  $f(x_i; \theta)$  donnée par le modèle.  $S(\theta)$  peut être considéré comme une mesure de la distance entre les données expérimentales et le modèle théorique qui prédit ces données. La prescription des moindres carrés commande que cette distance soit minimale.

Si, comme c'est généralement le cas, on dispose d'une estimation de l'écart-type du bruit qui affecte chaque mesure, on l'utilise pour « peser » la contribution de la mesure au  $\chi^2$ . Une mesure aura d'autant plus de poids que son incertitude sera faible :

$$\chi^2_{(\theta)} = \sum_{i=1}^N \left( \frac{y_i - f(x_i; \theta)}{\sigma_i} \right)^2 = \sum_{i=1}^N w_i (y_i - f(x_i; \theta))^2 \quad (3.26)$$

Où

$\chi^2_{(\theta)}$  : khi-deux et la quantité  $w_i$ , inverse de la variance du bruit affectant la mesure, est appelé *poids* de la mesure. Si les erreurs de mesure qui entachent les  $y_i$  sont distribuées suivant une loi normale ; la méthode des moindres carrés permet de plus d'estimer quantitativement l'adéquation du modèle aux mesures, pour peu que l'on dispose d'une estimation fiable des erreurs  $\sigma_i$ . Si le modèle d'erreur est non gaussien, il faut généralement recourir à la méthode du maximum de vraisemblance, dont la méthode des moindres carrés est un cas particulier.

### 3.3.7. Test de Khi deux

Le test du  $\chi^2$ , prononcé « khi-deux » ou « khi carré », est un test statistique permettant de tester l'adéquation d'une série de données à une famille de lois de probabilités ou de tester l'indépendance entre deux variables aléatoires.

#### Principe :

À la base d'un test de statistique classique, il y a la formulation d'une hypothèse appelée *hypothèse nulle* (ou *hypothèse zéro*) [3.04], notée  $H_0$ . Elle suppose que les données considérées proviennent de variables aléatoires suivant une loi de probabilité donnée, et l'on souhaite tester la validité de cette hypothèse.

Ces données ayant été réparties en classes, il faut :

- Calculer algébriquement la distance entre les données observées et les données théoriques attendues ;
- Se donner *a priori* un risque d'erreur, celle consistant à rejeter l'hypothèse, alors qu'elle est vraie (la valeur 5 % est souvent choisie par défaut ; il s'agit plus souvent d'une coutume que du résultat d'une réflexion) ;
- Déterminer le nombre de degrés de liberté du problème à partir du nombre de classes et à l'aide d'une table de  $\chi^2$ , déduire, en tenant compte du nombre de degrés de liberté, la distance critique qui a une probabilité de dépassement égale à ce risque.

Si la distance calculée entre les données observées et théoriques est supérieure à la distance critique, on conclut que le résultat n'est pas dû seulement aux fluctuations d'échantillonnage, et que l'hypothèse nulle  $H_0$  doit être rejetée. Le risque choisi au départ est celui de donner une réponse fausse lorsque les fluctuations d'échantillonnage sont seules en cause. Le rejet est évidemment une réponse négative dans les tests d'adéquation et d'homogénéité mais il apporte une information positive dans les tests d'indépendance. Pour ceux-ci, il montre le caractère significatif de la différence.

#### 3.3.7.1. Test du $\chi^2$ d'adéquation

Le test du khi-carré d'ajustement ou d'adéquation qui compare globalement la distribution observée dans un échantillon statistique à une distribution théorique.

Ce test permet de vérifier si un échantillon d'une variable aléatoire  $Y$  donne des observations comparables à celles d'une loi de probabilité  $P$  définie a priori dont on pense, pour des raisons théoriques ou pratiques, qu'elle devrait être la loi de  $Y$ . L'hypothèse nulle ( $H_0$ ) d'un test du  $\chi^2$  d'adéquation (dénommé aussi test du  $\chi^2$  de conformité ou test du  $\chi^2$  d'ajustement) est donc la suivante : la variable aléatoire  $Y$  suit la loi de probabilité  $P$ .

En termes de valeur-p, l'hypothèse nulle (l'observation est suffisamment proche de la théorie) est généralement rejetée lorsque  $p \leq 0,05$



On observe un échantillon de données  $y_1, \dots, y_N$  d'une variable aléatoire  $Y$  qui prend un nombre fini  $J$  de valeurs. On veut tester l'hypothèse nulle selon laquelle les probabilités que  $Y$  prenne les valeurs 1 à  $J$  sont respectivement  $p_1, \dots, p_J$  avec  $\sum_{j=1}^J p_j = 1$ .

*Définition 2.16 :*

On appelle  $\hat{p}_j$  la probabilité empirique que  $Y$  prenne la valeur  $j$ , autrement dit, le nombre d'observations qui prennent la valeur  $j$  dans l'échantillon divisé par le nombre total d'observations  $N$  :

$$\hat{p}_j = \frac{1}{N} \sum_{i=1}^N [y_i = j] \quad (3.27)$$

$$\text{avec } [y_i = j] = \begin{cases} 1 & \text{si } y_i = j \\ 0 & \text{sinon} \end{cases}$$

On peut alors définir la statistique du  $\chi^2$  :  $T = \sum_j \frac{(\mathbf{1}_j(N\hat{p}_j - Np_j))^2}{Np_j}$  où  $\mathbf{1}_A(x)$  désigne la fonction indicatrice de  $A$  au point  $x$  :  $\mathbf{1}_A(x) = 1$  si  $x \in A$  et  $\mathbf{1}_A(x) = 0$  si  $x \notin A$ .

Sous l'hypothèse nulle, cette statistique suit asymptotiquement une loi du  $\chi^2$  à  $(J - 1)$  degrés de liberté.

### 3.3.7.2. Test du $\chi^2$ d'indépendance

Le test d'indépendance du khi-carré permet de contrôler l'indépendance de deux caractères dans une population donnée. Il sert à apprécier l'existence ou non d'une relation entre ces deux caractères.

Ce test permet de vérifier l'absence de lien statistique entre deux variables  $X$  et  $Y$ . Les variables  $X$  et  $Y$  sont dits indépendants lorsqu'il n'existe aucun lien statistique entre elles ; en d'autres termes, la connaissance de  $X$  ne permet en aucune manière de se prononcer sur  $Y$ . L'hypothèse nulle ( $H_0$ ) de ce test est la suivante : les deux variables  $X$  et  $Y$  sont indépendantes.

En termes de valeur-p, l'hypothèse nulle est généralement rejetée lorsque  $p \leq 0,05$

À noter que ce test permet de contrôler l'existence d'une dépendance mais en aucun cas le sens de cette dépendance (sauf dans certains cas particuliers où l'existence d'une relation implique une causalité univoque).

### 3.3.7.3. Etapes

Les principales étapes du test d'indépendance du Khi-carré sont :

1) Distribuer la population statistique dans un tableau à deux caractères où les modalités et/ou les valeurs sont regroupées par catégories.

2) Calculer le khi-carré dans l'hypothèse d'indépendance des deux caractères :

$$\chi^2_{calculé} = \sum_{i=1}^l \sum_{j=1}^c \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

Où :

-  $n$  : effectif de la population

-  $n_{ij}$  représente l'effectif qui appartient simultanément à la catégorie  $A_i$  de la dimension  $X$  et à la catégorie  $B_j$  de la dimension  $Y$ .

$$- e_{ij} = \frac{L_i \cdot C_j}{n} = \frac{C_j \cdot L_i}{n}$$

-  $L_i$  représente la somme des effectifs appartenant à la catégorie  $A_i$  de la dimension  $X$ .  $C_j$  représente la somme des effectifs de la catégorie  $B_j$ . Ce sont donc des distributions conditionnelles.

		Catégories du caractère Y						Total
		$B_1$	$B_2$	$\dots$	$B_j$	$\dots$	$B_c$	
Catégories du caractère X	$A_1$	$n_{11}$			$n_{1j}$		$n_{1c}$	
	$A_2$	$n_{21}$						
	$\vdots$							
	$A_i$	$n_{i1}$			$n_{ij}$		$n_{ic}$	$L_i$
	$\vdots$							
	$A_L$	$n_{L1}$			$n_{Lj}$			
	Total				$C_j$			$n$

**Figure 3.02 :** La population de la catégorie  $A_i$  et de la catégorie  $B_j$

- 3) Calculer le nombre de degrés de liberté par la formule :  $(\text{Nombre de lignes} - 1) \times (\text{Nombre de colonnes} - 1)$
- 4) Définir une probabilité d'erreur (en pratique 5%, 1% ou 1 pour mille)
- 5) Utiliser le nombre de degrés de liberté et la probabilité d'erreur pour déterminer le khi-carré théorique à partir de la table fournie.
- 6) Comparer la valeur khi-carré calculée avec la valeur khi-carrée théorique (issue de la table) et appliquer la règle suivante :
  - Si le khi-carré calculé est inférieur au khi-carré théorique : indépendance
  - Si le khi-carré calculé est supérieur au khi-carré théorique : dépendance

### 3.4 Les équations structurelles

La modélisation d'équations structurelles a été développée dans les années 1970 comme une approche analytique des travaux empiriques [3.05]. Les modèles d'équations structurelles qui s'intègrent dans le cadre des analyses multivariées de deuxième génération et constituent des méthodes avancées pour la recherche, ont été utiles pour appréhender de nombreux problèmes de fond et ont contribué à l'évolution de la théorie dans les sciences sociales et comportementales et plus tard dans la psychologie, l'éducation, etc. [3.06]. Ces méthodes permettent de tester de manière simultanée l'existence de relations causales entre plusieurs variables latentes explicatives et plusieurs variables à expliquer [3.07].

*Définition 3.21 :*

Une *variable latente* est une variable qui n'est pas directement observable et qui ne peut pas être mesurée directement. Elle nécessite par conséquent le passage par des indicateurs ou variables manifestes qui permettent de mesurer la variable.

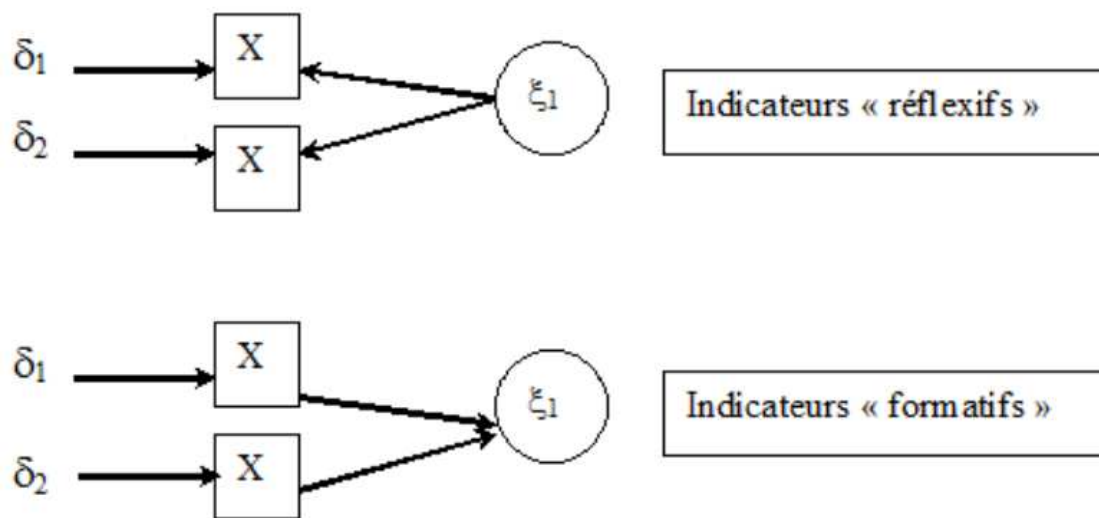
*Définition 3.22 :*

Une *variable manifeste* est une variable pour laquelle une mesure peut être directement recueillie (observée, mesurée, etc.).

Le paradigme dominant dans la théorie des tests fut celui représenté par Churchill [3.08] et qui considère que les variables manifestes sont toutes supposées représenter leur variable latente. Dans ce sens, il est postulé que tous les indicateurs concordent dans leur manière de mesurer le phénomène et permettent tous de refléter la même variable. En conséquence, le chercheur doit s'assurer de la significativité de la variable latente construite à la base de ces indicateurs, qui doivent être significativement corrélés. Dans ce sens, *une variable latente* bien construite est une variable dont la variation doit s'accompagner le plus fidèlement possible par la variation de tous les indicateurs qui la composent. Cette théorie classique des tests considérait alors toutes les variables comme réflexives. Le passage par une étape d'épuration des items par les différents chercheurs pour s'assurer de la cohérence interne et la validité des variables comme étape préalable au test attestent de cette considération de la réflexivité des variables. Néanmoins, dans certaines recherches, il est possible de se retrouver avec certaines variables latentes composées d'une combinaison d'indicateurs pas forcément corrélés, étant donné qu'ils mesurent des phénomènes théoriquement

différents. Ces indicateurs permettent alors de former la variable latente. On parle alors de variables formatives où les relations de causalité vont aller de ces indicateurs vers les construits.

La Figure 3.03 : selon [3.09] nous permet de distinguer entre les variables latentes réflexives et les variables latentes formatives.



**Figure 3.03 :** *Spécification des relations entre les variables latentes et leurs indicateurs*

Dans le premier type de spécification, les indicateurs constituent le reflet de la variable latente où elle demeure la cause des indicateurs, où chaque indicateur est lié à la variable latente par une équation de régression simple du type :

$$X_i = \lambda_i \xi_1 + \delta_i \quad (3.28)$$

Avec :

$\xi_1$  : construit latent

$X_i$  : indicateurs

$\lambda_i$  : coefficient représentant l'effet du construit sur l'indicateur (loading)

$\delta_i$  : indicateur de l'erreur de mesure

Dans le deuxième type de spécification, la variable latente constitue le reflet de tous les indicateurs où il demeure la cause de la variable latente. Le modèle se formalise alors par une équation de régression multiple :

$$\xi_1 = \sum_{i=1}^n \lambda_i X_i + \delta_i \quad (3.29)$$

Avec :

$\xi_1$  : construit latent

$X_i$  : indicateurs

$\lambda_i$  : coefficient représentant l'effet de l'indicateur sur le construit (coefficient de régression)

$\delta_i$  : indicateur de l'erreur de mesure

Il est important pour une recherche de bien distinguer entre les construits réflexifs et formatifs. En effet, une mauvaise spécification du modèle diminue la qualité du modèle de mesure et par conséquent, la qualité du modèle structurel [3.10]. Cette distinction aura pour conséquence un traitement différent au niveau de l'évaluation de la validité et de la fiabilité lorsqu'il s'agit d'une variable réflexive ou formative. Si les démarches classiques d'analyse de la validité et de la fiabilité reposent sur le postulat de la corrélation entre les indicateurs, et par conséquent la recherche de l'unidimensionnalité et de l'homogénéité des indicateurs dans le cas d'une variable latente réflexive, ceci peut, dans le cas d'une variable formative, éliminer un indicateur qui peut avoir du sens dans la recherche.

### ***3.4.1. Origine***

Les équations structurelles proviennent de trois “types” d'analyse : les analyses de cheminement (path analysis), les analyses de variables latentes (analyses factorielles) et les analyses d'équations simultanées. C'est la combinaison de ces trois chemins qui a mené aux équations structurelles. Il s'agit d'une analyse de cheminement, comprenant habituellement de multiples analyses simultanées et utilisant des variables latentes.

Selon Goodwin [3.11] l'analyse par équations structurelles est un outil heuristique pour organiser et exprimer la pensée du chercheur concernant la nature attendue des relations entre les variables ; un outil permettant de produire des estimations empiriques des relations théoriquement postulées entre des construits ; et une stratégie pour décider du modèle le plus approprié.

### **3.4.2. Causalité**

Quand on utilise des équations structurelles, on présume de l'existence de relations de causalité et donc répondant aux critères suivants : deux variables covariant ; il y a un ordre dans la relation, on a écarté les autres causes possibles de l'effet et l'explication de la chaîne de causalité est possible.

### **3.4.3. L'importance de la théorie**

“Rien de plus pratique qu'une bonne théorie”. Selon Kurt Lewin, La complexité doit être modélisée. C'est la base des équations structurelles : on cherche à valider un modèle théorique des relations empiriques. Il faut donc avoir élaboré le modèle théorique (model building) et être en mesure de le tester empiriquement (model testing).

- Dès le départ, on a utilisé les équations avec les lettres grecques,
- Deux séries d'équations :
  - a) Le modèle de mesure (équivalent à l'analyse factorielle) :

$$x = \lambda\xi + \delta \quad (3.30)$$

$$y = \lambda\eta + \varepsilon \quad (3.31)$$

- b) Le modèle de relation (équivalent à des régressions simultanées)

$$\eta = \beta\eta + \gamma\xi + \zeta \quad (3.32)$$

#### ***3.4.4. Quelques notions importantes***

Dans Lisrel ou tout autre logiciel, on fait la différence entre les variables mesurées ( $x_i, y_i$ ) et les variables latentes ( $\xi, \eta$ ). Et on fait la différence entre les variables exogènes, les  $\xi$ , soit les variables qui ne sont influencées par aucune autre variable ; et les variables endogènes, les  $\eta$ , qui sont influencées par au moins une autre variable, qu'elle soit exogène ou endogène.

#### ***3.4.5. Les étapes***

D'abord, il faut voir la théorie du modèle élaboré avant la cueillette de données (pas toujours possible mais c'est l'idéal). On recueille les données en fonction du modèle. Après la collecte de données (avant l'analyse d'équations structurelles), voir les statistiques descriptives univariées avant d'élaborer le modèle de mesure qui pourra effectuer les statistiques descriptives multivariées (analyses factorielles exploratoires). Réaliser les analyses bivariées et la vérification d'hypothèses "simples" ( $\chi^2$ , tests  $F$ , etc.). Utiliser les régressions avec entrée hiérarchique pour avoir une idée de "ce qui tient la route" et dresser le modèle à tester.

Une fois le modèle élaboré, on est prêt à le tester : Faire produire la matrice de covariance ou de corrélation (habituellement nécessaire) et donc déterminer le type de matrice à analyser. Ensuite, déterminer les paramètres fixes et ceux à estimer, les équations, le programme (langage Simplis dans Lisrel). Vérifier l'adéquation du modèle au moyen des divers indices d'ajustement et des indices de signification pour chaque paramètre, et réviser le modèle si nécessaire en fonction des indices fournis par le logiciel et de la théorie à l'origine du modèle.

En conclusion, les équations structurelles constituent un outil très intéressant puisqu'il permet de vérifier un ensemble de relations. Il permet entre autres de valider des modèles d'analyse factorielle en permettant la présence de variables complexes ou de facteurs de deuxième degré. Il permet de tester les effets de médiation de façon simultanée. On ne peut toutefois pas l'utiliser à toutes les sauces. On doit le garder pour une validation en fin de processus.



### 3.5 Modèles d'équations structurelles

Les modèles d'équations structurelles sont des modèles statistiques utilisés en sciences sociales et/ou comportementales. Ils sont assez méconnus des mathématiciens (statisticiens) mais pourtant, on peut les voir comme des généralisations des modèles et des méthodes statistiques comme l'analyse factorielle, la régression multiple et l'analyse de la variance (ANOVA).

#### 3.5.1. La modélisation par équations structurelles

La modélisation par équations structurelles est une méthode d'analyse de données utilisée pour représenter les hypothèses d'un modèle conceptuel complexe car composé de deux sous-modèles regroupant chacun une multitude de variables appelées manifestes (items prenant la forme carrée) ou latentes (format rond). Alors que les premières peuvent être mesurées directement, les secondes ne le sont qu'à travers les variables, dites manifestes.

Le premier sous-modèle appelé externe relie les variables latentes aux variables manifestes qui les mesurent ; le second, appelé interne, relie les variables latentes entre elles comme le montre la Figure 3.04 : .

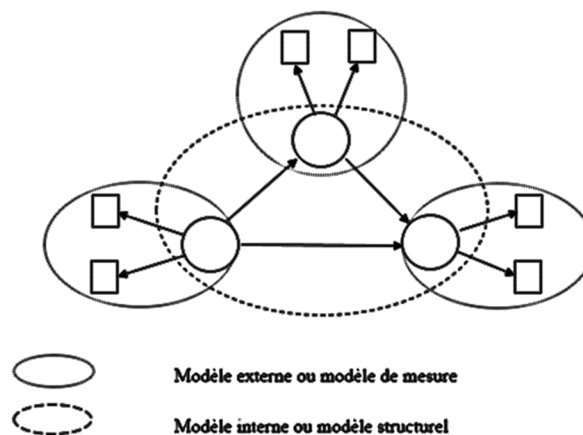


Figure 3.04 : Représentation du modèle

Les techniques de calcul des équations structurelles sont basées sur l'analyse de la covariance ou des moindres carrés partiels. Le choix de l'approche adaptée tient compte du type d'étude réalisée, de la taille de l'échantillon (plus ou moins de 100) et de la nature des modèles testés (formatifs ou

réflectifs). Selon Gefen et ses collègues [3.12], l'approche des moindres carrés, appelée également approche PLS est adaptée aux études prédictives ou visant à former des théories, menées sur des échantillons de petite taille et des modèles mixtes, à la fois formatifs et réflectifs.

Les résultats de la modélisation par équations structurelles sont évalués au regard de la variance expliquée, du score de concordance qui indique la qualité de l'ajustement entre les modèles interne et externe et du coefficient de détermination  $R^2$  rendant compte de l'exhaustivité des variables latentes étudiées. L'Alpha de Cronbach, quant à lui, mesure la fiabilité et la convergence des mesures ; il est à considérer en premier lieu.

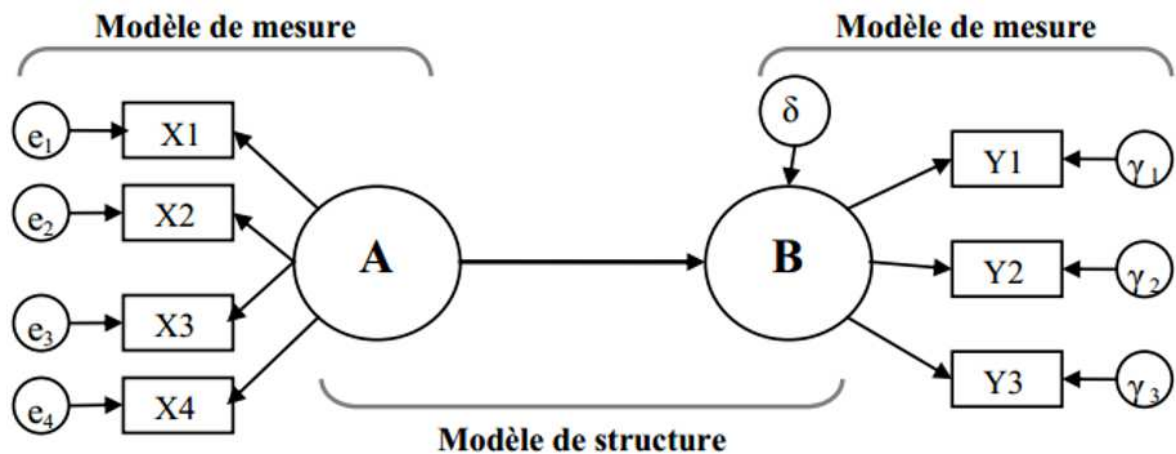
### ***3.5.2. Les fondements conceptuels de la modélisation par les équations structurelles (SEM)***

Selon Hoyle [3.13], la modélisation par les équations structurelles représente « *une approche statistique globale permettant de tester des hypothèses traitant des relations entre les variables observées et les variables latentes* ». De même, Kaplan [3.14] précise que cette méthode reflète « *une catégorie de méthodologies qui opte pour représenter des hypothèses sur les moyennes, les variances et les covariances des données observées en termes d'un nombre minimal de paramètres "structurels" définis par un modèle conceptuel sous-jacent* ». Ces techniques multivariées sont très adoptées dans le contexte des recherches en sciences sociales. Elles tiennent compte de manière claire et explicite des erreurs de mesures lors de l'étude de la relation entre les variables, comme elles ont l'exclusivité d'incorporer des variables latentes (ou encore des variables non directement observables) au niveau du modèle en question.

A ce titre, Hoyle [3.13] considère que les variables latentes forment un ensemble de construits ou encore de dimensions théoriques ou hypothétiques d'une importance majeure dans de nombreuses sciences. Elles représentent également des variables non observables au niveau des échantillons de la population étudiée.

Donc, un modèle d'équations structurelles se compose d'un ensemble d'indicateurs de mesure (appelés également variables manifestes), de variables latentes et d'erreurs. Les variables manifestes et latentes peuvent être indépendantes ou dépendantes, en fonction de leurs positions dans le modèle structurel. De même, la nature de la liaison entre les indicateurs de mesure et leurs variables latentes

permettent de déterminer si la nature des variables observées est réflexive ou formative (cf. figure 3.01). La figure 3.03 ci-après présente un exemple d'une structure réflexive d'un modèle causal.



**Figure 3.05 :** Exemple illustratif d'un modèle causal (indicateurs réfléchifs)

a) Equations du modèle de mesure

$$X_i = \lambda_i A + e_i \quad (3.33)$$

$$Y_i = \lambda_i B + \gamma_i \quad (3.34)$$

b) Equation du modèle de structure

$$B = rA + \delta \quad (3.35)$$

Comme le montre la Figure 3.05 : le modèle causal se compose d'un modèle de mesure et d'un modèle de structure (ou structurel). Il englobe également l'ensemble de variables observables indépendantes ( $X_i$ ), de variables observables dépendantes ( $Y_i$ ), d'une variable latente explicative ( $A$ ), d'une variable latente à expliquer ( $B$ ) et des termes d'erreurs ( $e_i$ ,  $\delta$  et  $\gamma_i$ ). Par conséquent, il est possible de distinguer deux types d'équations à savoir les équations du modèle de mesure et les équations du modèle de structure.

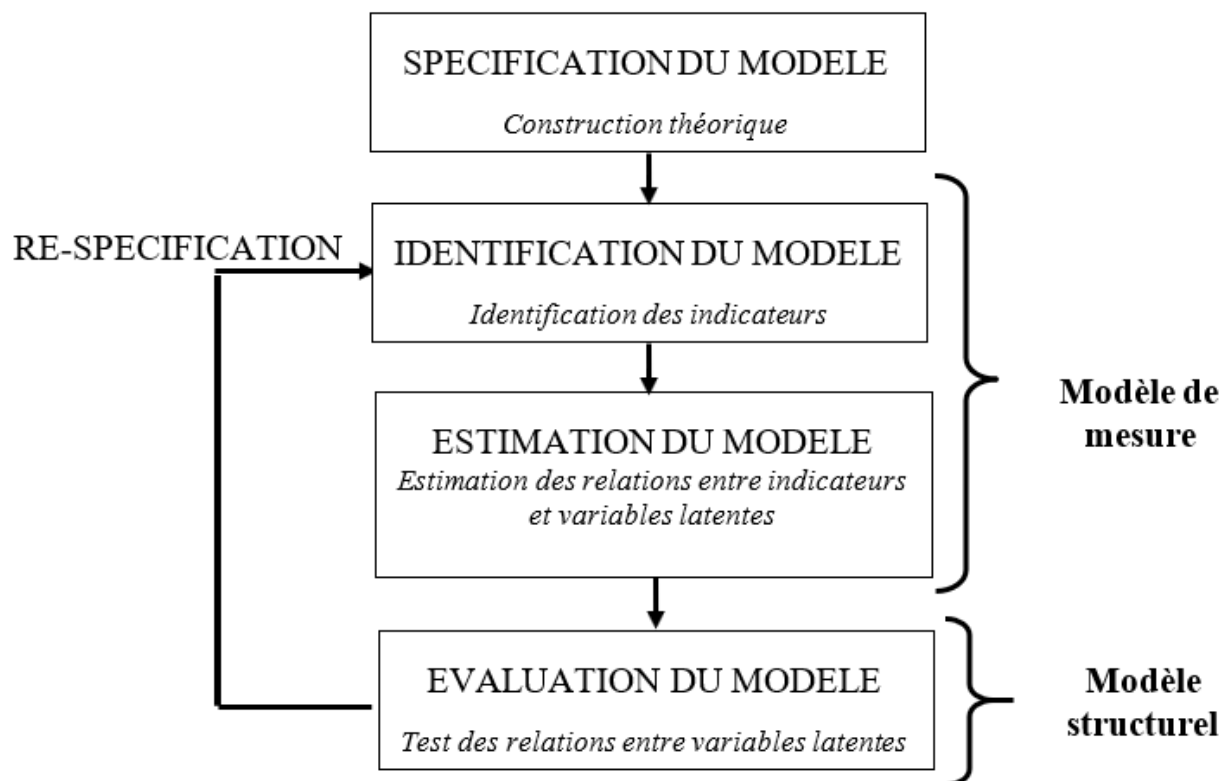
Le modèle de mesure spécifie la relation entre les variables observées et les variables latentes, alors que le modèle de structure permet d'examiner le lien entre les différentes variables latentes. Il en ressort que la méthode d'équations structurelles représente une technique multivariée qui combine

les modèles de mesure et les modèles de structures tout en examinant de manière simultanée une série de relations linéaires entre les variables observées et les variables latentes d'une part, et entre l'ensemble de variables latentes d'autre part [3.15].

### 3.5.3. Les étapes d'élaboration d'un modèle d'équations structurelles

L'élaboration d'un modèle d'équations structurelles rend nécessaire le suivi d'une démarche composée d'un ensemble d'étapes successives. Selon Bollen et Long [3.16] et Schumacker et Lomax [3.17], la spécification, l'identification, l'estimation, l'évaluation et la respécification du modèle représentent les étapes les plus communément admises au niveau de la conception d'un modèle d'équations structurelles.

La Figure 3.05 : basée sur [3.19], [3.16] et [3.17] nous permet de voir l'ensemble des différentes étapes pour l'élaboration d'un modèle d'équation structurelle.



**Figure 3.06 :** Les étapes de la construction du modèle des équations structurelles

### 3.5.3.1. La spécification du modèle

La spécification du modèle à étudier consiste à préciser ses principales caractéristiques en se basant sur la théorie. Elle est une étape importante dans la démarche d'utilisation des modèles d'équations structurelles. En effet, la présence d'un cadre conceptuel cohérent avec le modèle constitue un ingrédient préliminaire permettant de justifier théoriquement les relations linéaires à tester [3.18]. Par conséquent, cette étape contribue de manière effective à la formulation des hypothèses de recherche dans le but de tester la significativité des liens de causalité entre les différentes variables en question.

Les travaux de Roussel et al. [3.18] soulignent que la spécification du modèle conduit à la conception d'un « *schéma de relations linéaires* » qui se traduisent en modèles de mesure et en modèles de structure.

### 3.5.3.2. L'identification du modèle

L'identification est indispensable dans la mesure où elle permet d'attribuer une seule solution pour chacun des paramètres à estimer. Un modèle est identifié lorsqu'il converge vers une solution unique.

Un modèle conduit à exprimer la matrice de variance/covariance des variables manifestes  $\Sigma$  en fonction d'un ensemble de paramètres  $\theta$ . Ce modèle est identifié si  $\Sigma(\theta) = \Sigma(\theta')$  implique  $\theta = \theta'$ .

En pratique, l'identification d'un modèle suppose la satisfaction de deux conditions à savoir la condition d'ordre (condition nécessaire) et la condition de rang (condition nécessaire et suffisante).

La condition d'ordre dépend impérativement du degré de liberté *ddl* [3.18] et se rapporte à la correspondance entre les paramètres à estimer et le nombre de variances/covariances des variables observées [3.13]. En s'inspirant des travaux de Schumacker et Lomax [3.17], le degré de liberté s'exprime de la manière suivante :

$$ddl = \left( P \frac{(P + 1)}{2} \right) - N \quad (3.36)$$

Avec :

$P$  : le nombre d'indicateurs de mesure du modèle.

$N$  : le nombre de coefficients à estimer du modèle.

$ddl$  : le degré de liberté.

Par conséquent, trois situations sont possibles à savoir la sous identification, la juste identification et la sur identification du modèle.

a) La sous identification du modèle

$$\left( P \frac{(P+1)}{2} \right) < N \quad (3.37)$$

En d'autres termes, le degré de liberté est négatif ( $ddl < 0$ )

b) La juste identification du modèle

$$\left( P \frac{(P+1)}{2} \right) = N \quad (3.38)$$

En d'autres termes, le degré de liberté est égal à zéro ( $ddl = 0$ )

c) La sur identification du modèle

$$\left( P \frac{(P+1)}{2} \right) > N \quad (3.39)$$

En d'autres termes, le degré de liberté est positif ( $ddl > 0$ )

D'après ces trois situations, la condition d'ordre est vérifiée lorsque le modèle est juste identifié ou encore sur identifié, c'est-à-dire lorsque le degré de liberté est supérieur ou égal à zéro ( $ddl \geq 0$ ).

Par ailleurs, la condition de rang permet d'avoir une seule solution vis-à-vis de chacun des paramètres à estimer. Selon Schumaker et Lomax [3.17], la condition de rang est beaucoup plus complexe ; « *elle nécessite une détermination algébrique de manière que chaque paramètre du modèle peut être estimé de la matrice observée de variance/covariance* ». Ces auteurs précisent que

la condition de rang est atteinte lorsque la matrice de variances/covariances est définie positive. En d'autres termes, le déterminant de cette matrice est non nul.

### 3.5.3.3. L'estimation du modèle

L'estimation consiste à déterminer les valeurs relatives aux différents paramètres du modèle à tester. Au regard de la littérature, il existe plusieurs méthodes d'estimations. Le choix de l'une de ces méthodes dépend d'un certain nombre de conditions notamment la taille de l'échantillon et la distribution des données. Selon les travaux de Schumacker et Lomax [3.17] et de Chaput [3.19], la technique d'estimation du Maximum de Vraisemblance *ML* (Maximum Likelihood) et celle des Moindres Carrées Généralisées *GLS* (Generalized Least Squares) sont les plus utilisées étant donné qu'elles figurent dans la plupart des logiciels de traitement des données.

#### *Proposition 3.1.*

Selon Schumacker et Lomax [3.17], l'expression de la méthode de Maximum de vraisemblance se présente comme suit :

$$F_{ML} = \text{tr} (S . \hat{\Sigma}^{-1}) - p + \ln |\hat{\Sigma}| - \ln |S| \quad (3.40)$$

Si  $\theta$  soit un vecteur de paramètre libre estimé par le modèle, on peut écrire la relation 3.40 comme suit :

$$F_{ML} (\theta) = \text{tr} (S . \Sigma^{-1}(\theta)) - p + \ln |\Sigma(\theta)| - \ln |S| \quad (3.41)$$

Avec :

$S$  : la matrice de variance-covariance observée

$\ln$  : la fonction logarithmique naturelle

$\text{tr} (.)$  : la fonction trace d'une matrice

$\Sigma$  : la matrice de variances/covariances estimées et

$p$  : le nombre d'indicateurs de mesure ou variables observées.

$|S|$  : Déterminant d'une matrice  $S$

*Démonstration :*



Il provient de l'analyse de la matrice de covariance du modèle à partir du paramètre de l'estimation ( $\hat{\Sigma}$ ) et en le comparant à la valeur réelle de la matrice de covariance basée sur les données empiriques ( $S$ ).

On va d'abord évaluer la probabilité d'obtenir  $S$  étant donné  $\hat{\Sigma}$ . C'est possible en utilisant la distribution de Wishart.

La matrice de covariance  $S$ , de composante  $s_{ij}$ , de l'échantillon est :

$$S = \frac{1}{N-1} \sum_{i=1}^N (x_{ij} - \bar{x}_j)(x_{ij} - \bar{x}_j)' = \frac{1}{n} \sum_{i=1}^N (x_{ij} - \bar{x}_j)(x_{ij} - \bar{x}_j)'$$

Où  $n = N - 1$ , et  $\bar{x}_j = \frac{1}{N} \sum_{i=1}^N x_{ij}$ .

Ceci montre que  $nS$  suit la loi de Wishart et sa fonction densité est alors donnée par :

$$W(\hat{\Sigma}, n) = \frac{|nS|^{\frac{1}{2}(n-p-1)} e^{\left(-\frac{n}{2} \text{tr}(S\hat{\Sigma}^{-1})\right)}}{|\hat{\Sigma}|^{\frac{n}{2}} \cdot 2^{\frac{np}{2}} \cdot \Gamma_p\left(\frac{n}{2}\right)}$$

Où  $\Gamma_p\left(\frac{n}{2}\right) = \pi^{\frac{p(p-1)}{4}} \prod_{j=1}^p \Gamma\left(\frac{n}{2} + \frac{1-j}{2}\right)$ , la fonction gamma multidimensionnelle.

En notant  $T$  toutes les expressions constantes indépendantes de  $\hat{\Sigma}$  dans cette expression :

$$T = \frac{|nS|^{\frac{1}{2}(n-p-1)}}{2^{\frac{np}{2}} \cdot \Gamma_p\left(\frac{n}{2}\right)}, \text{ on aura :}$$

$$W(S, \hat{\Sigma}, n) = \frac{e^{\left(-\frac{n}{2} \text{tr}(S\hat{\Sigma}^{-1})\right)}}{|\hat{\Sigma}|^{\frac{n}{2}}} T$$



Le ratio de vraisemblance selon Neyman et Pearson compare le modèle théorique au modèle estimé à partir des données réelles en utilisant une sorte de distance. Ce ratio de vraisemblance est défini par :

$$LR = \frac{\text{Vraisemblance pour un modèle estimé sous } H_0}{\text{Vraisemblance avec un modèle des données réelles sous } H_1}$$

Ici, le ratio de vraisemblance est donc :

$$LR = \frac{e^{\left(-\frac{n}{2} \text{tr}(S\hat{\Sigma}^{-1})\right)} |\hat{\Sigma}|^{-\frac{n}{2}T}}{e^{\left(-\frac{n}{2} \text{tr}(SS^{-1})\right)} |S|^{-\frac{n}{2}T}}$$

On remarque que  $\hat{\Sigma}$  est remplacé par  $S$  dans le dénominateur car la matrice de covariance à partir des données réelles est égale à  $S$ .

D'où, le ratio de vraisemblance est :

$$LR = e^{\left(-\frac{n}{2} \text{tr}(S\hat{\Sigma}^{-1})\right)} |\hat{\Sigma}|^{-\frac{n}{2}} e^{\left(\frac{n}{2} \text{tr}(SS^{-1})\right)} |S|^{\frac{n}{2}}$$

En prenant le logarithme de chaque membre :

$$\ln(LR) = -\frac{n}{2} \text{tr}(S\hat{\Sigma}^{-1}) - \frac{n}{2} \ln|\hat{\Sigma}| + \frac{n}{2} \text{tr}(SS^{-1}) + \frac{n}{2} \ln|S|$$

Or  $SS^{-1}$  est une matrice identité de même taille que les variables d'observation, c'est-à-dire  $p$ . Alors, la trace de cette matrice est égale à  $p$  car les termes de diagonale sont tous égaux à 1.

La fonction de maximum de vraisemblance est alors donnée par, en simplifiant par  $-\frac{n}{2}$  :

$$F_{ML} = \text{tr}(S\hat{\Sigma}^{-1}) + \ln|\hat{\Sigma}| - p - \ln|S|$$

◆

*Proposition 3.2.*

La méthode des moindres carrées généralisées est exprimée par la relation ci-dessous qui est une forme particulière de la méthode de maximum de vraisemblance :

$$F_{GLS}(\theta) = \frac{1}{2} \text{tr} \left\{ [(S - \Sigma(\theta))S^{-1}]^2 \right\} \quad (3.42)$$

Avec :

- $S$  : la matrice de variance-covariance observée
- $\text{tr}(\cdot)$  : la fonction trace d'une matrice
- $\Sigma$  : la matrice de variances/covariances estimées

*Démonstration :*



$S$  est une matrice de dimension  $p \times p$ , où  $p$  est le nombre de variables observées. Et on a vu d'après la relation 3.41 que :

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

On déduit la méthode des moindres carrées généralisées  $F_{GLS}$  à partir de ce maximum de vraisemblance  $F_{ML}$ . On va noter  $s = \text{vech}(S)$  un vecteur de dimension  $p^*$  formé par les colonnes de  $S$  laissant les éléments au-dessus de sa diagonale, et soit  $\text{vec}(S)$  le vecteur formé par les colonnes de  $S$ . Alors selon Magnus et Neudecker [3.20], il existe une unique matrice  $D_p$  de dimension  $p^2 \times p^*$  tel que  $\text{vec}(S) = D_p s$ .

Soit  $\sigma = \text{vech}(\Sigma)$ . On va noter comme suit les dérivées matricielles :

$$\dot{F}(S, \Sigma) = \frac{\partial F(S, \Sigma)}{\partial \sigma'}$$

$$\ddot{F}(S, \Sigma) = \frac{\partial^2 F(S, \Sigma)}{\partial \sigma \partial \sigma'}$$

Comme  $F_{ML}(S, \Sigma)$  est une fonction de  $\sigma$ , on peut écrire le développement limité de Taylor à l'ordre de  $F_{ML}$  au voisinage de  $\sigma = s$  :

$$F_{ML}(S, \Sigma) = F_{ML}(S, S) + \dot{F}_{ML}(S, S)(\sigma - s) + \frac{1}{2}(\sigma - s)' \ddot{F}_{ML}(S, S)(\sigma - s) + r_{ML}(s, \sigma)$$

Où  $r_{ML}(s, \sigma)$  est le reste de développement limité, et est de l'ordre  $O(\|\sigma - s\|^3)$  au voisinage de  $\sigma = s$ .

Puisque  $F_{ML}(S, S) = 0$ ,  $\dot{F}_{ML}(S, S) = 0$ , et  $\ddot{F}_{ML}(S, S) = D_p'(S^{-1} \otimes S^{-1})D_p$ , on peut écrire que :

$$F_{ML}(S, \Sigma) = (\sigma - s)' W_1 (\sigma - s) + r_{ML}(s, \sigma)$$

Où  $W_1 = \frac{1}{2} D_p'(S^{-1} \otimes S^{-1})D_p$ . Le premier terme du second membre de cette égalité est égale à la méthode des moindres carrées  $F_{GLS}$ , et il s'ensuit alors que :

$$F_{GLS} = (\sigma - s)' W_1 (\sigma - s) = \frac{1}{2} tr((I - \Sigma S^{-1})^2) = \frac{1}{2} tr(((S - \Sigma)S^{-1})^2)$$



#### 3.5.3.4. L'évaluation du modèle

La qualité d'un modèle peut être appréciée à travers l'évaluation d'un ensemble d'indices d'ajustement. Chacun de ces indices dispose d'un ensemble de caractéristiques spécifiques. Plusieurs tests statistiques sont utilisés pour déterminer dans quelle mesure le modèle correspond aux données. On note qu'un bon ajustement entre le modèle et les données ne signifie pas que le modèle est « correct » ou même qu'il explique une grande proportion de la covariance. Un "bon ajustement de modèle" indique seulement que le modèle est plausible [3.21]. Lors de la présentation des résultats d'une analyse factorielle confirmatoire, il est recommandé de déclarer [3.22] :

- a) les modèles proposés ;
- b) les modifications apportées ;
- c) quelles mesures identifient chaque variable latente ;
- d) les corrélations entre variables latentes ;
- e) toute autre information pertinente, par exemple si des contraintes sont utilisées.

En ce qui concerne la sélection des statistiques d'ajustement des modèles à signaler, il ne faut pas simplement signaler les statistiques qui estiment le meilleur ajustement, bien que cela puisse être

tendant. Bien qu'il existe plusieurs opinions divergentes, Kline recommande de rapporter le test du chi carré, l'erreur quadratique moyenne de l'approximation (Root mean square error of approximation - RMSEA), l'indice comparatif d'ajustement (comparative fit index - CFI) et le résidu moyen carré normalisé (standardised root mean square residual SRMR).

### ***Indices d'ajustement absolu***

Les indices d'ajustement absolu déterminent l'adéquation du modèle a priori ou reproduisent les données [3.23]. Les indices d'ajustement absolu comprennent, mais sans s'y limiter, le test Chi-Squared, RMSEA, GFI, AGFI, RMR et SRMR.

#### **- Test du Chi-carré**

Le test du khi carré indique la différence entre les matrices de covariance observées et attendues. Les valeurs plus proches de zéro indiquent un meilleur ajustement ; différence plus faible entre les matrices de covariance attendues et observées. Les statistiques du Chi-carré peuvent également être utilisées pour comparer directement l'ajustement des modèles imbriqués aux données. Toutefois, une difficulté avec le test du khi-carré de l'ajustement du modèle est que les chercheurs peuvent ne pas rejeter un modèle inapproprié dans des échantillons de petite taille et rejeter un modèle approprié dans des échantillons de grande taille. En conséquence, d'autres mesures d'ajustement ont dû être développées [3.24].

#### **- Erreur quadratique moyenne d'approximation**

L'erreur quadratique moyenne de l'approximation (RMSEA) évite les problèmes de taille de l'échantillon en analysant l'écart entre le modèle hypothétique, avec des estimations de paramètres choisies de manière optimale et la matrice de covariance de la population. Le RMSEA va de 0 à 1, avec des valeurs plus petites indiquant un meilleur ajustement du modèle. Une valeur de 0,06 ou moins indique un ajustement acceptable du modèle [3.25].

#### **- Racine résiduelle moyenne carrée et racine normalisée moyenne résiduelle**

Le résidu quadratique moyen (root mean square residual - RMR) et le résidu moyen carré normalisé (standardised root mean square residual - SRMR) sont la racine carrée de la discordance entre la matrice de covariance de l'échantillon et la matrice de covariance du modèle. Cependant, le RMR peut être quelque peu difficile à interpréter, car sa portée est basée sur les échelles des indicateurs

du modèle (cela devient délicat lorsque vous avez plusieurs indicateurs à différentes échelles, par exemple, deux questionnaires, un sur une échelle de 0-10, l'autre sur une échelle 1-3) [3.21]. Le résidu du carré moyen normalisé élimine cette difficulté d'interprétation et varie de 0 à 1, une valeur de 0,08 ou moins étant indicative d'un modèle acceptable [3.25].

- Indice de qualité de l'ajustement et indice de qualité de l'ajustement ajusté

L'indice d'ajustement de la qualité (goodness of fit index - GFI) est une mesure d'ajustement entre le modèle hypothétique et la matrice de covariance observée. L'indice d'ajustement ajusté (adjusted goodness of fit index - AGFI) corrige le GFI, qui est affecté par le nombre d'indicateurs de chaque variable latente. Le GFI et l'AGFI se situent entre 0 et 1, avec une valeur de plus de 0.9 indiquant généralement un ajustement acceptable du modèle [3.26].

### ***Indices d'ajustement relatif***

Les indices d'ajustement relatif (également appelés « indices d'ajustement incrémental » et « indices d'ajustement comparatifs ») comparent le khi carré du modèle hypothétique à un modèle « nul » ou « de base ». Ce modèle nul contient presque toujours un modèle dans lequel toutes les variables ne sont pas corrélées et, par conséquent, a un chi-carré très grand (indiquant un mauvais ajustement). Les indices d'ajustement relatif comprennent l'indice d'ajustement normé et l'indice d'ajustement comparatif.

- Indice d'ajustement normé et indice d'ajustement non normé

L'indice d'ajustement normé (normed fit index - NFI) analyse l'écart entre la valeur du khi-deux du modèle hypothétique et la valeur du khi-deux du modèle nul. Cependant, le NFI tend à être biaisé négativement. L'indice d'ajustement non normé (non-normed fit index - NNFI, également connu sous le nom d'indice de Tucker-Lewis (Tucker-Lewis Index - TLI), fondé sur un indice formé par Tucker et Lewis en 1973 [3.27]) résout certains des problèmes de biais négatif, bien que les valeurs NNFI puissent parfois tomber au-delà de la gamme de 0 à 1. Les valeurs pour le NFI et le NNFI devraient être comprises entre 0 et 1, avec un seuil de 0,95 ou plus, indiquant un bon ajustement du modèle [3.25].

- Indice d'ajustement comparatif (Comparative fit index ou CFI)

L'indice d'ajustement comparatif (CFI) analyse l'ajustement du modèle en examinant l'écart entre les données et le modèle hypothétique, tout en ajustant les problèmes de taille d'échantillon inhérents au test du khi-carré de l'ajustement du modèle et l'indice d'ajustement normé. Les valeurs CFI vont de 0 à 1, avec des valeurs plus grandes indiquant un meilleur ajustement. Auparavant, une valeur CFI de 0,90 ou plus était considérée comme indiquant un ajustement acceptable du modèle. Cependant, des études récentes ont indiqué qu'une valeur supérieure à 0,90 est nécessaire pour s'assurer que des modèles mal spécifiés ne sont pas jugés acceptables. Ainsi, un indice CFI de 0,95 ou plus est actuellement accepté comme un indicateur de bon ajustement [3.25].

L'objectif étant d'aboutir à un meilleur dosage entre la maximisation de l'ajustement et la minimisation du nombre des coefficients estimés [3.18].

#### 3.5.3.5. La respécification du modèle

La respécification du modèle représente une phase de réflexion qui offre l'opportunité de reconsidérer la conception du modèle et d'en proposer des modifications éventuelles, en tenant compte du cadre théorique de la recherche en question. De même, Roussel et al. [3.18] soulignent que la phase de spécification du modèle doit être opérée avec précaution dans le but d'apporter des améliorations possibles au modèle d'origine. A titre d'exemple, il serait possible d'ajouter ou de supprimer des paramètres, de fixer ou de modifier des contraintes, etc. tout en prenant en considération les bases de la théorie qui fonde la conception du modèle de base.

#### 3.5.3.6. La présentation des résultats

A ce niveau, la sélection des résultats nécessaires permet de répondre aux objectifs de la recherche. En effet, la préparation des résultats de manière convenable constitue un préambule au développement des connaissances théoriques. De même, ces résultats doivent incorporer des informations pertinentes et précises, permettant d'aboutir facilement à des solutions aux problèmes susceptibles d'être rencontrés dans l'analyse.

### 3.6 Conclusion

Ce troisième chapitre permet d'avoir les statistiques de base nécessaires dans l'élaboration du modèle d'équation structurelle (SEM). Différentes étapes devraient être franchies dont la

spécification du modèle, l'identification, l'estimation et l'évaluation du modèle. Une respécification doit avoir lieu si nécessaire et doit revenir aux étapes précédentes avant de présenter le résultat. Toutes ces théories de base en statistiques et en SEM seront appliquées à l'apprentissage à distance dans le chapitre suivant.

## CHAPITRE 4 – MODELISATION DE L'APPRENTISSAGE ASYNCHRONE PAR EQUATIONS STRUCTURELLES LINEAIRES

### 4.1 Introduction

La modélisation par équation structurelle teste de façon empirique une série de relations linéaires hypothétiques entre variables. Elle permet de vérifier plusieurs relations linéaires simultanément par le truchement d'un enchaînement causal. L'application des théories de base présentées dans le chapitre précédent à l'apprentissage asynchrone dans une FOAD est l'objet du présent chapitre, afin d'obtenir les modèles théoriques de la recherche. Les données récoltées pour une étude empirique sont naturellement produites par les acteurs de la plateforme. Comme l'apprentissage est le centre de la recherche, les dialogues entre pairs et entre apprenant-tuteur sont les principales interactions étudiées. Ces deux sources produisent une très grande quantité de données à partir desquelles nous allons relier les engagements motivationnel et relationnel avec l'engagement comportemental de l'apprenant. Au niveau de l'ingénierie d'apprentissage, il s'agit alors d'exploiter un retour d'expérience pour concevoir un modèle d'équation structurelle en tenant compte des discussions observées lors de la mise en œuvre de l'apprentissage asynchrone. Ce retour d'expérience peut être organisé pour pouvoir mesurer les indicateurs selon la collaboration, la motivation et l'implication de l'étudiant lors de son apprentissage. Ces indicateurs faciliteront les adaptations à réaliser au moment de l'activité d'apprentissage [4.01] telle qu'elle se déroulera en réalité.

Ce chapitre comporte l'analyse d'une activité d'apprentissage asynchrone en le forum de discussion dans laquelle un module d'initiation a été mobilisé. Un modèle d'équations structurelles (SEM) sera proposé et les différentes étapes seront réalisées afin de sortir les deux modèles de mesures validées du SEM. R est utilisé pour le traitement statistique des données.

Ce chapitre est un développement de l'article de notoriété internationale et de celui de notoriété nationale publiés les mois d'avril et mai 2018 :

- M. Rakotomalala, T. B. Ravaliminoarimalalason, F. Randimbindrainibe, « *Models for Measuring Latent Variables in Asynchronous Learning* », ESRSA, International Journal of Engineering Research & Technology, Vol. 7 Issue 04, pp 146-150, April-2018.
- M. Rakotomalala, T. B. Ravaliminoarimalalason, F. Randimbindrainibe, « *Modèles de*



## **4.2 Analyse de l'activité**

L'analyse de l'apprentissage asynchrone d'un module d'initiation (GEAL) dans une formation à distance revêt une importance considérable. L'objectif est d'amener à s'interroger sur les dimensions à prendre en compte pour analyser l'acquisition de connaissance d'un apprenant en situation d'apprentissage.

Dans la recherche, l'analyse a été préférentiellement centrée sur le forum des apprenants qui se déroule sur la plateforme dans le module d'initiation au début de la formation. Ce module aide les apprenants dans la prise en main de la plateforme. Ils interagissent entre eux ou avec le tuteur.

Une formation ouverte et à distance en Génie du Système Automatique (GSA) est soulevée dans ce chapitre. L'utilisation d'une connexion internet est nécessaire pour pouvoir travailler sur la plateforme et accéder au GEAL. Les apprenants doivent avoir pris connaissance de son contenu et effectuer ensuite les activités individuelles et en groupe.

Le forum est un outil de communication et de collaboration asynchrone dans la formation. Cet outil a beaucoup d'avantages comme ceux cités dans le paragraphe 1.9 , et parmi lesquels on peut soulever l'intervention, laquelle peut se faire selon la disponibilité temporelle des acteurs. C'est un outil adapté au contexte malgache vu la qualité de la connexion qui est généralement lente dans presque toutes les régions de Madagascar et surtout le problème de délestage qui peut survenir à n'importe quel moment et pourrait empêcher les apprenants de joindre les réunions synchrones programmées dans une formation à distance réunions qui favorisent le travail collaboratif. L'enseignant est une personne ressource, il aide les apprenants à résoudre les problèmes techniques ou les oriente vers le responsable concerné. C'est également l'enseignant qui répond à leurs questions et joue le rôle d'animateur si cela s'avère nécessaire à un moment où les apprenants se lancent dans un travail collaboratif. Dans un forum, chaque apprenant peut poser n'importe quelle question pédagogique (cours, activité, évaluation) ou question d'organisation etc., et tous les participants peuvent réagir face aux préoccupations des autres sans attendre l'intervention du tuteur.

Dans les forums de discussion du module GEAL, les apprenants doivent répondre aux questions comme à celles posées dans « la présentation », trouver de manière collaborative les moyens de résoudre une situation problème, donner leur avis, commenter l'intervention des autres, interagir avec l'environnement, faire une proposition de synthèse pour le travail de groupe, contacter les tuteurs si nécessaire comme en cas d'absence des autres co-équipiers, s'exprimer sur leurs besoins en consignes supplémentaires et aviser le dépôt.

Le Tableau 4.01 : ci-dessous présente l'analyse d'un forum de discussion selon une grille d'analyse d'activité d'apprentissage. Le contexte de l'activité, l'établissement porteur de la formation, les matériels à disposition, la description et l'effet de cette activité sont les rubriques à analyser.

<b>Rubrique</b>		<b>Analyse de l'activité d'un ou de plusieurs apprenants</b>
<b>L'objet de la ou des activités réalisées</b>	<b>- Institutionnels (établissement, programmes, etc.),</b>	FOAD en L3 Exploitation logistique et transport (ELT) ; Transit et commerce internationale (TCI) ; Génie du Système Automatisé (GSA)
	<b>• Matériels (ex : organisation de l'espace, matériel à disposition),</b>	<ul style="list-style-type: none"> <li>- Dispositif de formation à distance centré sur l'apprenant contenant un module d'initiation.</li> <li>- Ordinateurs connectés à internet.</li> <li>- Activités basées sur l'apprentissage asynchrone dans des forums de discussion</li> </ul>
	<b>Les caractéristiques personnelles des apprenants (âge, niveau, etc.)</b>	<p>Les apprenants de la formation</p> <p>La plupart des apprenants</p>

L'objet de la ou des activités réalisées	<p><b>Description de l'activité :</b></p>	<p><b>Intervention :</b></p> <ul style="list-style-type: none"> <li>- Envoi de message</li> <li>- Création de discussion</li> <li>- Participation à un travail collaboratif</li> <li>- Modification ou suppression de message</li> <li>- Consultation des forums</li> </ul> <p><b>Communication :</b></p> <ul style="list-style-type: none"> <li>- Réponse à une question</li> <li>- Envoi d'une préoccupation</li> <li>- Appel des co-équipiers</li> <li>- Demande d'aide</li> </ul> <p><b>Consultation :</b></p> <ul style="list-style-type: none"> <li>- Prise de connaissance des forums sur la plateforme</li> <li>- Prise de connaissance d'une / des discussion(s)</li> </ul>
	<p><b>Description chronologique de l'activité :</b></p> <p>Temps 1 : intervention individuelle</p> <ul style="list-style-type: none"> <li>- Étape 2 : échange, confrontation, argumentation, justification et validation en petits groupes</li> <li>- Temps 3 : Rédaction du document de groupe</li> </ul>	<p>Réactions face au thème de discussion lancé par le tuteur</p> <p>Prise de connaissance et participation au fil de discussion en cours,</p> <p>Passage d'un travail individuel à des échanges collectifs sur le forum de discussion : prise de connaissance des travaux des autres, émission d'idées, de questions ou d'avis,</p> <p>Elaboration d'un document de synthèse : les apprenants comparent, discutent et se concertent pour avoir un document unique synthétique à déposer comme devoir de groupe</p>

<b>Artéfacts mobilisés dans cette activité</b>	<b>Fonction de chaque artéfact dans cette activité</b>	Les échanges et interactions s'effectuent dans l'espace de forum de discussion. Cet espace peut être utilisé comme un espace de travail collaboratif, là où tous les membres peuvent déposer leur contribution individuelle, leur avis par rapport aux travaux des autres.
	<b>Forme de médiation</b>	Les échanges entre apprenants permettent un apprentissage socio cognitif (échange sociocognitive au sein du groupe)
	<b>Relations entre les différents instruments impliqués dans cette activité (redondance complémentarité, existence d'un instrument pivot)</b>	La communication, l'interaction, la collaboration et l'entraide s'effectuent grâce à un seul outil : le forum de discussion, qui est un instrument pivot dans cette situation.
<b>Effets de l'activité</b>	<b>Sur les objets, l'environnement matériel et/ou social de cette activité</b>	Interagir et collaborer entre pairs pour atteindre l'objectif commun : pouvoir élaborer ensemble un devoir de groupe, travailler en équipe pour résoudre la situation problème posée.
	<b>Effets sur les compétences de l'acteur dont on analyse l'activité</b>	<p>L'interaction au sein du groupe facilite la compréhension du problème et sa résolution ; d'où la facilitation d'acquisition de connaissance dans un module de formation mobilisant cet outil.</p> <p>L'interaction favorise la sociabilité dans la communauté d'apprenant, d'où l'élimination du facteur d'isolement qui est une source d'abandon dans une FOAD. Les apprenants se sentent à l'aise dans l'apprentissage du module, ce qui améliore leur engagement cognitif.</p>

**Tableau 4.01 :** *Analyse d'une activité d'apprentissage : forum de discussion*

Ensuite, le terrain et le contexte à distance sont explicitement analysés. L'analyse doit tenir compte de l'approche thématique décrite dans le chapitre précédent.

Rappelons d'abord qu'une communauté désigne un groupe dont les membres exercent des fonctions similaires et se réunissent pour résoudre des problèmes et en même temps acquièrent des connaissances [3.07]. Dans la recherche, la communauté est formée par un groupe d'apprenants rassemblés pour une activité de formation afin de répondre à une préoccupation relative à une tâche dont le produit sera évalué.

D'un point de vue méthodologique, une approche quantitative préliminaire à l'étude plus spécifique des contenus est nécessaire. Tous les éléments qui constituent le contenu du forum sont traités et analysés. La collecte et l'analyse des traces de l'activité principale sont effectuées sur la base des messages émis. La participation, le nombre d'interventions, la consultation des forums et des discussions ainsi que la création de nouveaux thèmes sont des indicateurs à prendre en compte. Les contributeurs principaux peuvent être repérés, de même, les non-contributeurs ou contributeurs muets.

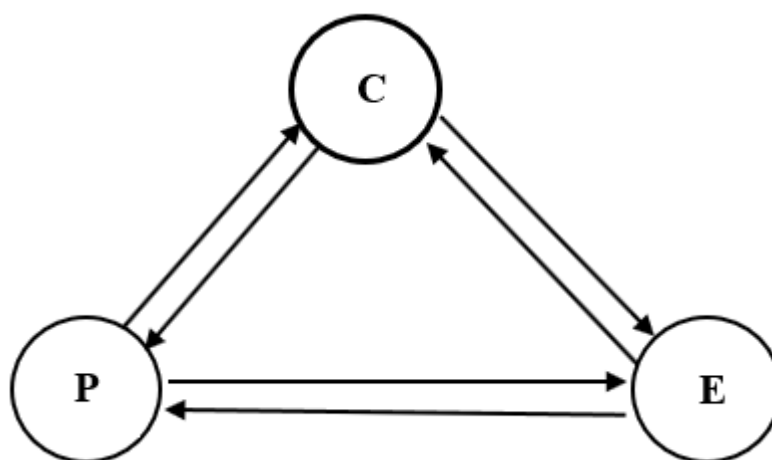
### **4.3 Approche théorique du SEM appliquée à la formation à distance**

Il existe peu de modèles qui explorent en profondeur les différentes dimensions cognitives, motivationnelles, et socio-affectives de l'engagement en contexte de formation à distance : ce sont ce que nous avons pris parmi les bases théoriques dans le présent chapitre. Nous présenterons d'abord l'engagement scolaire et ses dimensions avant d'exposer l'engagement en formation à distance.

#### ***4.3.1. L'engagement scolaire et l'approche sociocognitive de la motivation***

En matière d'éducation, l'engagement est un concept polysémique, faisant l'objet d'une grande variété de définitions et d'interprétations. L'approche sociocognitive de la motivation [4.02] , [4.03], est l'un des courants les plus utilisés pour étudier l'engagement des apprenants, et ce quel que soit le type d'enseignement visé. Elle postule une interaction entre les caractéristiques de l'environnement d'apprentissage, celles de l'apprenant et de son comportement [4.04] ; [4.05] ; [4.06]. Les caractéristiques de l'environnement interagissent avec les perceptions qu'un apprenant en a, au travers des processus cognitifs d'interprétation et d'anticipation. L'influence mutuelle de

l'environnement et de l'apprenant va alors impacter sur le comportement, lequel, selon les résultats obtenus, pourra lui aussi exercer un effet en retour. Voir



**Figure 4.01 :** *Le modèle de causalité triadique réciproque de Bandura [4.02]*

Dont « C » désigne les conduites, « P » les caractéristiques personnelles et « E » l'environnement. Ces trois dimensions sont en interaction continue et réciproque selon des importances variables et contingentes à des circonstances, des activités et des temporalités données.

#### ***4.3.2. Les dimensions de l'engagement scolaire***

Pour les auteurs nord-américains les plus influents dans le domaine de l'engagement scolaire [4.07], [4.08] ; ce dernier comprend plusieurs composantes en interaction dont une comportementale, une cognitive et une affective. Chacune de ces facettes peut être vue comme variant quantitativement et qualitativement. Selon la synthèse de la littérature effectuée par ces auteurs, la définition de l'engagement comportemental se fonde sur l'idée de participation et d'indicateurs observables de cette participation. La notion de participation renvoie elle-même à différents niveaux : le respect des règles de la classe et l'absence de comportements dérangeants ; l'effort, la concentration, l'attention et la participation aux activités en classe ; la participation à des activités extra-curriculaires. Certains auteurs abordent la question de l'engagement social en axant le concept sur l'idée de la participation des élèves aux activités du groupe et aux activités parascolaires [4.09]. L'engagement social est ainsi inclus dans la définition de l'engagement comportemental proposée par Fredricks et ses

collaborateurs [4.07] et, est aussi en lien avec l'engagement émotionnel car présuppose un sentiment d'appartenance à l'école.

#### 4.3.2.1. L'engagement cognitif

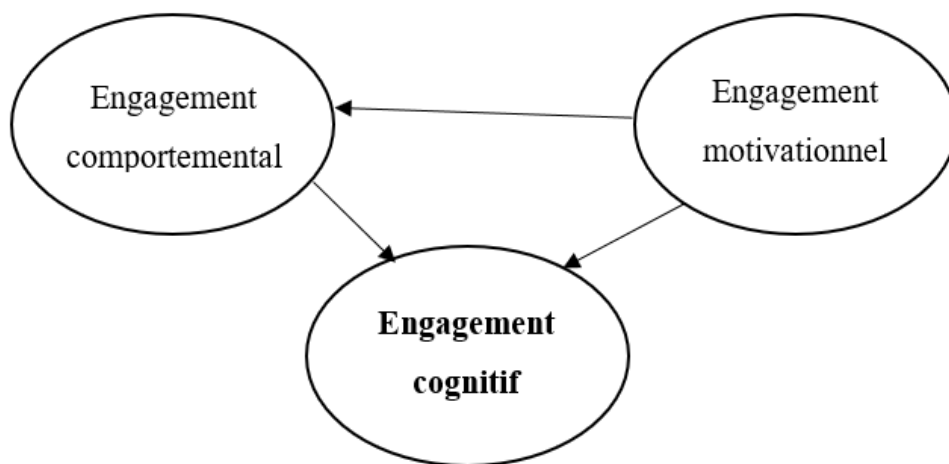
L'engagement cognitif se rapporte à l'investissement des ressources cognitives et à l'effort mental déployé lors de la réalisation d'une tâche. Il peut être envisagé sous un angle quantitatif (quantité de ressources allouées à la tâche, intensité des efforts) et qualitatif (degré de sophistication des stratégies d'apprentissage, adéquation des efforts). **L'engagement comportemental** est lié à l'engagement cognitif dont il est souvent un bon indicateur mais pas toujours. Ainsi, les manifestations observables de l'engagement nous renseignent peu sur les stratégies mobilisées durant l'effort. Par exemple, le fait de regarder le professeur en train de donner une explication peut traduire l'attention consacrée aux propos, mais n'exclut pas que l'apprenant soit en train de penser à tout autre chose. La notion d'effort mental est en lien avec à la fois l'engagement comportemental et l'engagement cognitif car l'effort comporte des signes visibles et donc observables.

#### 4.3.2.2. L'engagement émotionnel

L'engagement émotionnel se réfère aux émotions positives ou négatives (par exemple, l'intérêt, la tristesse ou l'anxiété) que les apprenants ressentent à l'égard de l'école, de leurs enseignants et leurs pairs, ou encore du contenu à apprendre voire de l'acte même d'apprendre. Dans une formation à distance **l'engagement émotionnel** (via les enseignants, les pairs et les caractéristiques instrumentales du dispositif) est lié au plaisir de se faire comprendre, au plaisir de partager et au plaisir d'apprendre, de découvrir et de créer.

#### *4.3.3. L'engagement en formation à distance dans une perspective socio-cognitive*

L'engagement, le comportement et leurs relations avec le résultat pédagogique sont les objets de préoccupation de cette recherche dans le domaine de la FOAD. Elle vise à comprendre la raison du résultat pédagogique insatisfaisant en s'intéressant aux conditions psychologiques qui permettent le maintien de l'engagement. L'angle adopté est celui du comportement et de l'engagement de l'apprenant.



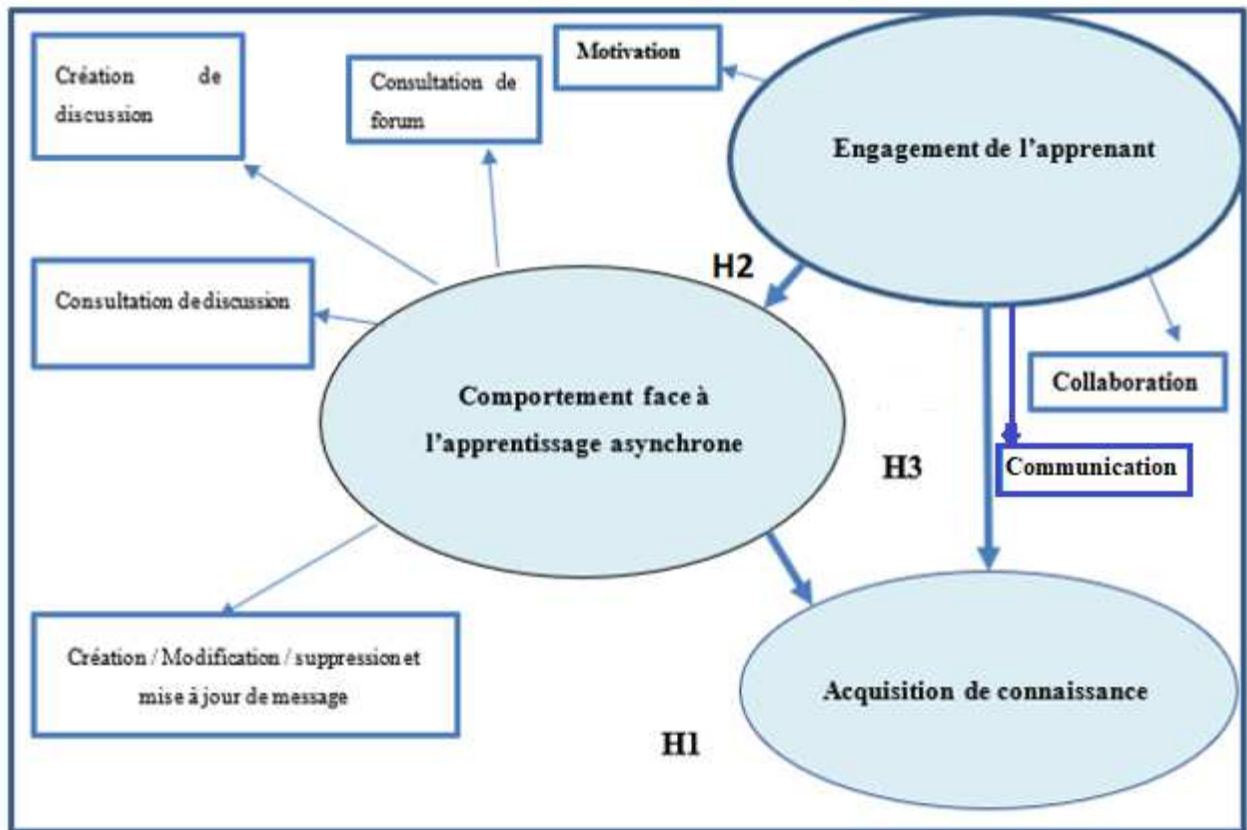
**Figure 4.02 :** *Modèle structurel du SEM*

#### ***4.3.4. Modèle d'équations structurelles en formation à distance***

En pratique, l'usage des techniques d'équations structurelles en apprentissage à distance est en évolution permanente comme elle reflète une opportunité pour le développement de la théorie et de la recherche. En effet, cette méthode est innovante dans la mesure où elle crée de la valeur et contribue à la proposition des travaux de plus en plus performants sur le plan méthodologique et empirique.

Nous avons choisi d'étudier l'impact de l'utilisation du forum de discussion sur l'acquisition de connaissance et la réussite de l'apprenant dans une formation à distance. Par conséquent, le modèle conceptuel de la recherche Figure 4.03 : présente les différents liens de causalité entre ces construits théoriques et montre que l'apprentissage asynchrone pourrait jouer un rôle médiateur au niveau de la relation entre l'exploitation du forum de discussion et la réussite de l'apprenant.





**Figure 4.03 : Le modèle conceptuel de la recherche**

Sur le plan théorique, les travaux de recherche en sciences de l'éducation ont souligné l'importance de l'intervention perçue dans la détermination du développement de connaissance de l'apprenant dans une formation en ligne. L'utilisation de forum de discussion comme outils de communication et de collaboration représente un déterminant incontournable de l'engagement de l'apprenant, de la motivation et du sentiment d'appartenance à un groupe pour rompre l'isolement dans la mesure où elle favorise le développement de connaissance. Ainsi, l'effet de développement de connaissance via le forum de discussion sur le résultat sera fortement démontré. En d'autres termes, les apprenants engagés, motivés et collaborant dans une formation à distance développent leur connaissance et leur savoir-faire de manière continue. L'impact de l'exploitation du forum sur l'acquisition de connaissance sera démontré empiriquement. Nous pouvons proposer l'ensemble des hypothèses suivantes :

*H1 : La forte utilisation du forum de discussion pourrait agir positivement sur le développement de connaissance de l'apprenant qui aboutit à la réussite dans une FOAD*

*H2 : L'engagement de l'apprenant dans un groupe pourrait soutenir son interaction asynchrone et sa motivation dans la formation*

*H3 : L'engagement de l'apprenant aurait un impact direct sur son résultat pédagogique dans une formation à distance.*

#### **4.4 Mesure de l'engagement**

L'engagement est une notion complexe, multidimensionnelle et multifactorielle, qui reste difficile à définir et à opérationnaliser [4.10]. Par ailleurs, il existe une variété de mesures de l'engagement, que ce soit des mesures subjectives comme des questionnaires auto-reportés ou des mesures plus « objectives » comme les indicateurs de participation. Les raisons qui président à l'engagement relèvent des modèles de motivation et sont tout à la fois internes et externes. S'engager dans une FOAD dépend de facteurs motivationnels comme les attentes par rapport à la formation (est-ce que cette formation fait vraiment sens pour moi ?) l'instrumentalité (est-ce que cette formation est le bon moyen pour moi d'atteindre mes objectifs personnels/professionnels ?). Pouvoir croire qu'on peut y arriver, seul et avec les autres, est une des raisons majeures de l'engagement [4.11]. Ces auteurs mettent également en lumière le rôle important que jouent les autres dans l'engagement en formation. (1) Le besoin d'être en relation avec autrui est une des sources d'engagement ; (2) le contexte climat social dans lequel a lieu l'apprentissage a une influence sur la motivation à s'engager et à apprendre ; (3) le besoin d'appartenance sociale a un impact positif sur l'engagement ; et (4) il existe une relation circulaire entre engagement cognitif et engagement relationnel dans l'apprentissage collaboratif.

Dans le modèle proposé sur la Figure 4.03 : l'engagement est le produit d'une motivation et d'un apprentissage social, c'est dans et par les interactions avec ses pairs (communication et collaboration) et ses enseignants que l'apprenant construit son engagement. Le dispositif de formation Moodle est conçu et est configuré de sorte à valoriser la dimension « personnes » et à soutenir la construction sociale de l'engagement.

#### ***4.4.1. Mesures de l'engagement scolaire en FAD***

La mesure des différentes formes d'engagement définies ci-avant consiste essentiellement en l'évaluation des perceptions que les apprenants ont d'eux-mêmes ou des perceptions que les enseignants ont de l'engagement des apprenants. Elle repose également sur des grilles d'observation souvent utilisées pour la mesure de l'engagement comportemental.

#### ***4.4.2. Mesure de l'engagement par les questionnaires***

Dans une recherche en formation à distance, on a traduit (en français), validé et utilisé le *Motivated Strategies Learning Questionnaire* (MSLQ) que Pintrich et ses collaborateurs [4.05] ont développé en lien avec leur modèle de l'apprentissage auto-régulé, en suivant le processus d'adaptation transculturelle suggéré par Vallerand et Thill [4.12]. Ce questionnaire comprend une section sur la motivation et une autre sur l'engagement cognitif envisagé sous l'angle des stratégies d'apprentissage. La section sur la motivation comprend deux sous-échelles liées aux attentes : une échelle d'auto-efficacité et une échelle des perceptions de contrôle. La composante cognitive de l'engagement est représentée par des échelles sur les stratégies cognitives (répétition, organisation, élaboration), métacognitives (auto-régulation, pensée critique, gestion du temps et de l'environnement) et socio-cognitives (apprentissage entre pairs, recherche d'aide).

L'échelle de motivation en éducation (EME) basée sur la TAD est aussi l'un des outils de mesure les plus utilisés [4.13]. Elle comprend 28 items mesurant 3 formes de motivation intrinsèque (connaissance, stimulation, accomplissement) et extrinsèque (externe, introjectée, identifiée).

Pour avoir des réponses plus objectives, nous n'avons pas procédé avec la mesure de l'engagement par questionnaire. Nous avons plutôt choisi d'aborder l'analyse des traces d'activités de l'apprenant comme mentionné dans le chapitre précédent pour la mesure de comportement.

#### ***4.4.3. L'analyse des traces pour mesurer l'engagement comportemental***

En contexte de classe, un chercheur peut utiliser les comportements observables des élèves pour évaluer la quantité d'efforts et mesurer l'engagement comportemental. Dans le contexte des FAD,

ce type de mesures est difficilement accessible. Cependant, dans les environnements d'apprentissage informatisés, les activités des apprenants peuvent être conservées sous la forme de *logs* permettant ainsi de garder la trace des différentes pages visitées et des différentes actions effectuées. Dans une plateforme en ligne, plusieurs chercheurs adoptent une approche fondée sur l'analyse des traces laissées par les apprenants dans leurs environnements (*learning analytics*) afin de mieux comprendre leurs comportements. Ces approches sont alternativement fondées sur une définition très globale de l'engagement (Downes, 2010, cité par [4.14]), ou sur la détection de *patterns* de comportement très spécifiques qu'on peut effectivement relier à des formes particulières d'engagement.

Les traces de l'activité des participants dans une formation en ligne peuvent être transformées pour représenter une mesure de l'engagement comportemental [4.15], ce qui a des incidences considérables sur la manière de les organiser et la manière de les simplifier afin qu'elles correspondent bien aux construits théoriques. Dans une initiative de recherche menée actuellement, l'engagement a été modélisé comme un continuum de participation aux diverses interventions accessibles sur des forums de discussion : consultation de forum, consultation d'une discussion, ajout d'un message à une discussion en cours, création d'une nouvelle discussion, mise à jour d'un message. Pour chacun de ces types d'interventions, des variables ordinales ont été créées à partir des traces. L'activité de chaque apprenant dans les forums a été recodée en une variable ordinale représentant un continuum : aucune activité dans les forums (niveau 0), activité de lecture des forums (niveau 1), activité de lecture des discussions (niveau 2), envoi d'un message dans un fil de discussion (niveau 3), mise à jour d'un message (niveau 4) et création d'une nouvelle discussion (niveau 5). Si les traces peuvent être considérées comme des mesures objectives de l'engagement comportemental, qu'en est-il de l'engagement cognitif ?

#### ***4.4.4. L'engagement dans l'apprentissage collaboratif à distance***

L'apprentissage collaboratif est une méthode pédagogique qui permet de renforcer les interactions entre les étudiants en formation à distance et qui vise un apprentissage profond. La participation aux échanges au sein de la plateforme, la richesse et la qualité des interactions entre apprenants sont reconnues comme des facteurs bénéfiques à l'engagement, à la satisfaction et à la réussite des apprenants en formation à distance (Lee, 2002 cité par [4.16]).

Toutefois, si l'apprentissage collaboratif est une méthode pédagogique de plus en plus encouragée dans les formations à distance du fait de ses effets positifs [4.17], travailler en groupe n'est pourtant pas systématiquement synonyme d'apprentissage. Barron [4.18], en référence aux travaux de Salomon et Globerson [4.19] parle du groupe comme d'une source potentielle d'aggravation pour les apprenants, pouvant générer un sentiment de perte de temps, de l'anxiété et du découragement.

Dans cette recherche, nous nous intéressons à l'engagement dans un apprentissage en ligne et plus particulièrement sur le fait de « s'engager dans une activité d'apprentissage asynchrone en le forum de discussion ». Les apprenants travaillent seuls sur une tâche et partagent leurs connaissances, problèmes et résultats aux autres, ou travaillent en collaboration sur la même tâche de façon asynchrone et synchrone pour produire quelque chose de commun.

Par définition, l'engagement dans l'apprentissage collaboratif est un processus mutuel. Il consiste en un effort conscient, volontaire et continu de la part de tous les participants, de mise en commun et de coordination en vue de résoudre ensemble le problème : « *collaboration as the mutual engagement of participants in a coordinated effort to solve the problem together* », [4.20], p. 70. L'engagement est ici cognitif. Les participants s'engagent à produire du sens non seulement à partir de ce qui est donné à apprendre (le contenu d'apprentissage) mais également à partir des idées de leurs partenaires ; il est question d'*intersubjective meaning-making* [4.21]. Dans la plateforme, l'engagement mutuel consiste également en la régulation par les apprenants non seulement de leurs propres processus épistémiques mais également de ceux de leurs coéquipiers.

Barron [4.18] envisage la collaboration comme deux espaces interdépendants dans lesquels les participants doivent s'impliquer conjointement, un espace de contenu (cognitif) et un espace relationnel. Dans ce dernier, l'engagement mutuel prend une dimension socio-affective, et renvoie à la façon dont les membres du groupe agissent les uns envers et avec les autres.

Dans la collaboration, l'engagement socio-affectif se traduit par des stratégies centrées sur la relation, comme des comportements d'entraide et d'encouragement mutuel [4.16] ; [4.22]. La qualité du climat socio-relationnel explique une part substantielle de la qualité des productions communes [4.23], et ce principalement pour des tâches où l'interdépendance entre les membres du groupe est importante.

L'engagement dans l'apprentissage collaboratif est un processus intentionnel. On retrouve cette idée dans le cadre théorique développé par [4.24] pour décrire les communautés de co-construction de connaissances (*knowledge building communities*). L'un des principes sous-jacents à ces communautés est celui de l'agentivité épistémique. D'après ce principe, les individus qui s'engagent dans une telle communauté « reconnaissent leurs responsabilités individuelle et collective dans le succès des efforts fournis pour construire des connaissances » (p. 10). Individuellement, l'apprenant engagé met en avant ses idées et négocie un ajustement entre ce qu'il propose et les idées des autres. Il utilise les divergences et les contrastes d'idées pour faire avancer les connaissances et « n'attend pas que ce soit les autres qui tracent cette voie pour lui » (p.10). Sur le plan collectif, l'apprenant engagé collabore avec ses pairs pour définir les objectifs communs, maintenir la motivation du groupe, planifier son travail et en évaluer les résultats.

#### ***4.4.5. Mesure de l'engagement dans l'apprentissage collaboratif à distance***

Il existe peu de recherches en matière d'apprentissage collaboratif médiatisé par l'ordinateur (CSCL : Computer-Supported Collaborative Learning) qui opérationnalisent le concept d'engagement mutuel. Dans un groupe où cet engagement est élevé, les membres contribuent à la tâche de façon plus ou moins égale : on parle de participation symétrique. Jermann [4.25] distingue la participation au dialogue (nombre de mots produits par participant) et la participation à l'activité (nombre d'actions réalisées). Dans une dyade avec C1 et C2 comme collaborateurs, l'asymétrie de participation, qui peut être vue comme un indicateur du niveau d'engagement mutuel, est calculée ainsi : il s'agit de la différence absolue du nombre de mots/actions produits par S1 et S2, divisée par la somme totale des mots/actions produits par S1 et S2 ([4.25], p. 142). La valeur de cet indicateur est comprise entre 0 et 1, les valeurs proches de 0 indiquant une participation symétrique, et donc un niveau élevé d'engagement.

L'asymétrie de participation est un indicateur de quantité ne rendant pas compte de la qualité de l'engagement mutuel tant sur le plan épistémique que relationnel. Pour mesurer la qualité de l'engagement dans la collaboration, il convient d'analyser plus finement les interactions entre participants. Les mesures de l'engagement épistémique sont des mesures de transactivité qui reflètent la façon dont les membres d'un groupe se réfèrent à, font quelque chose avec, construisent

sur la parole/les actions de l'autre. La transaction dans le dialogue est définie comme le « raisonnement qui opère sur le raisonnement de l'autre » ([4.26], p. 403).

Une contribution verbale est transactive lorsque le collaborateur qui la produit reprend les termes et les idées de son partenaire pour les transformer et les intégrer dans son propre discours. Une contribution transactive correspond à ce que Barron [4.18] identifie comme une réponse engagée. Pour Barron, les réponses engagées se répartissent en deux catégories : l'acceptation (approuver la proposition de l'autre, la soutenir, la documenter) et la discussion (questionner la proposition de l'autre, la contrer, demander des clarifications). Les réponses non engagées ne sont pas transactives ; le collaborateur ne tient pas compte de ce que son partenaire vient de dire. La transactivité est un phénomène multi-niveau. Il est ainsi possible d'observer de la transaction au niveau des actions et au niveau visuel. Au niveau des actions, Suthers [4.21] parle d'intersubjective uptake acts pour décrire les manipulations que les collaborateurs font des contributions de leur partenaire dans une carte conceptuelle collaborative. Sur le plan visuel, la transactivité concerne le degré d'attention visuelle accordée par les collaborateurs aux contributions de leur partenaire. Elle est aussi en lien avec la notion de couplage attentionnel : les partenaires regardent aux mêmes endroits aux mêmes moments [4.10].

Les mesures de l'engagement relationnel pourront être des mesures subjectives de perception recueillies par questionnaire à propos de la qualité de la relation de travail. Plusieurs questionnaires existent dont l'origine théorique est parfois inconnue. Certains questionnaires se recoupent au niveau de certaines dimensions mesurées. Toutefois, la façon dont ils définissent et évaluent ces dimensions communes est souvent très différente. Citons, à cet effet, quelques exemples de questionnaires sur la qualité socio-affective de la collaboration. Dans le questionnaire créé par Kreijns et ses collaborateurs [4.28] pour étudier la présence sociale dans les environnements d'apprentissage médiatisé, certaines échelles évaluent la dynamique socio-affective des groupes comme **l'échelle de sociabilité** (par exemple, « *L'environnement m'a permis de créer des liens étroits avec mes partenaires.* »), **l'échelle des comportements positifs** (par exemple, « *Les membres du groupe ont donné des informations sur eux-mêmes.* ») et **l'échelle des comportements négatifs** (par exemple, « *Les membres du groupe se sont méfiés les uns des autres.* »). L'échelle de communication relationnelle [4.29] évalue différentes dimensions relationnelles des interactions, dont l'intimité, la dominance, le calme/l'excitation émotionnelle, la formalité et l'orientation sur la tâche/la relation. L'intimité se décline en 7 sous-dimensions dont celle de l'engagement

(*involvement*) mesurée par 7 items comme « *Il/Elle a été très impliqué(e) dans la conversation.* » ou « *Il/Elle a créé de la distance entre nous* ». Cette échelle n'a jamais été utilisée dans le domaine CSCL mais pourrait être un outil intéressant pour étudier les processus relationnels dans l'apprentissage collaboratif à distance. Quintin et Masperi [4.23] ont quant à eux analysé les échanges verbaux asynchrones entre les membres d'un groupe restreint d'apprentissage en ligne pour rendre compte du climat socio-relationnel et sa relation aux résultats d'apprentissage. Ils ont analysé les marques linguistiques de liance qui témoignent de relations positives comme les expressions de satisfaction/d'empathie (« *Cool !* », « *Courage !* »), et les marques d'alliance qui montrent la volonté de s'engager dans le travail commun (actes promissifs ; « *Je compléterai notre document demain.* ») ou qui demandent la réalisation d'une action au profit du groupe (actes directifs ; « *Quand penses-tu pouvoir rédiger ta partie ?* »).

#### 4.5 La spécification du modèle

La spécification du modèle conduit à la conception d'un « *schéma de relations linéaires* » qui se traduisent en modèles de mesure et en modèles de structure selon Roussel et al. [3.18]. Trois à quatre variables observées pour chaque variable latente (ou dimension) ont été considérés pour un modèle de mesure. Ces modèles sont de nature réflexive, en d'autres termes, les indicateurs sont tous réflexifs. L'estimation, un peu plus bas, du modèle nécessite énormément de données ainsi que l'intégration de plusieurs techniques statistiques complexes. Plus précisément, au contraire des statistiques paramétriques traditionnelles, où l'on ne peut faire d'hypothèse sur la causalité, la modélisation par équation structurelle présente l'avantage relativement unique de permettre à la causalité d'être sous-entendue.

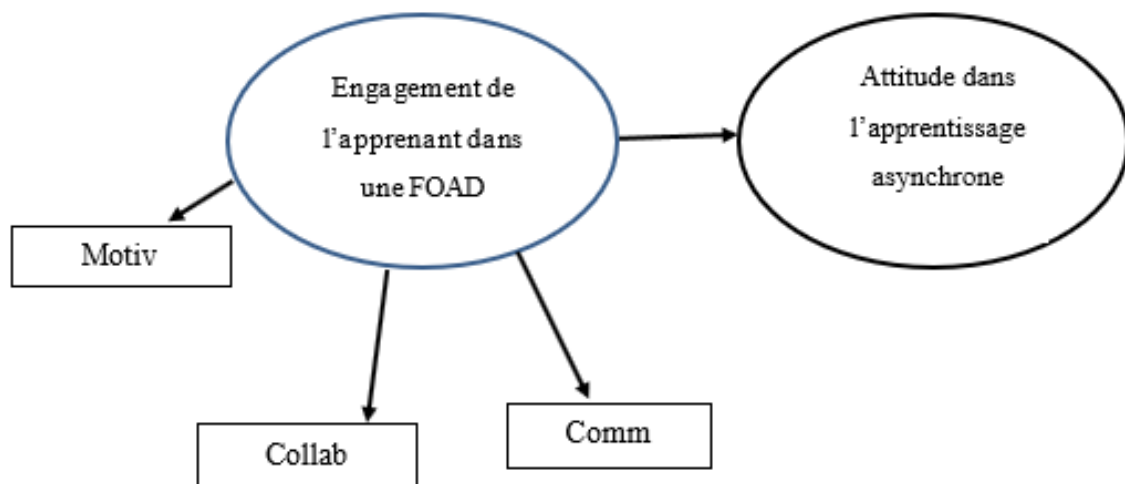
Les données utilisées ont été collectées auprès de trois plateformes de FOAD avec une centaine d'apprenants et analysées au moyen de méthodes de **modélisation par équations structurelles**. La relation entre les deux variables latentes a été développée pour décrire et quantifier l'influence des différentes causes.

Deux **modèles de mesure par équations structurelles** ont été élaborés selon la Figure 4.04 : et la Figure 4.05 :



#### 4.5.1. L'engagement de l'apprenant dans une formation à distance

Selon De Wulf et Odekerken-Schröder [4.30], l'engagement de l'apprenant est mesuré par la régularité au forum, la sensation d'être à l'aise par rapport au forum de discussion, la réalisation des activités asynchrones depuis la plateforme malgré les contraintes existantes et la consacrassion de plus d'effort afin de rester actif dans l'apprentissage asynchrone.



**Figure 4.04 :** *Modèle de mesure de l'engagement de l'apprenant*

$X_1$  = Collab : interaction collaborative dans le forum de discussion

$X_2$  = Comm : communication entre les protagonistes de la formation (apprenants, tuteur)

$X_3$  = Motiv : motivation dans une communauté d'apprentissage (groupe, sous-groupe, formation)

##### 4.5.1.1. Motivation

La motivation de l'apprenant est un des traits les plus importants lorsqu'on parle d'une formation ouverte et à distance. C'est un des plusieurs critères à étudier par l'administration pour la sélection des étudiants qui veulent s'inscrire dans cette formation. Le fait de donner du temps à un apprenant pour s'exprimer sur lui-même, sur son environnement d'apprentissage, et sa motivation constituent

la force motrice qui le pousse à être performant et couronné de succès dans son apprentissage à distance [4.30].

La motivation constitue le levier essentiel de tout apprentissage. A ce titre, les apprenants doivent conserver leur motivation, voire la renouveler.

**Motivation (Motiv) :** dans cette recherche, l'apprenant doit se présenter dans un forum spécifique « présentation » (forum fermé) afin que tout le monde puisse se connaître facilement. Les exemples ci-dessous présentent quelques postages exprimant la motivation de l'apprenant et copiés tels quels via le forum de la plateforme.

Exemple 1 : « *Bonsoir à tous et à toutes ;*

*je m'appelle NGAGOUE Epse NZADI DIANE MIREILLE. Je suis d'origine camerounaise résidant au Maroc.*

*Titulaire d'une licence en sciences de gestion obtenue à la faculté des sciences économiques et de gestion appliquée à l'Université de Douala (au Cameroun) et d'un master 1 en Management industriel et logistique à Moroccan Management Business Shool (au Maroc).*

*Je poursuis actuellement mes études en Master 2 en Management industriel et Logistique ainsi que la formation en ligne en Exploitation Logistique et Transport.*

*Passionnée de Logistique et transport, ainsi que de la supply chain management et désireuse de me lancer dans l'auto emploi, j'ai donc opté pour cette formation en ligne.*

*J'espère ainsi acquérir les armes nécessaires qui me permettront de réaliser mon projet professionnel. » ;*

Exemple 2 : « *bonjour,*

*je suis Franck-Michel LINGUISSY MAPAKOU, titulaire d'un brevet de technicien supérieur en Transit & commerce international. Précédemment superviseur Transit chez MTN CONGO et actuellement superviseur-Fleet toujours dans cette même société.*

*J'espère acquérir au cours de cette formation les outils nécessaires devant me permettre de mieux cerner les questions logistiques afin d'y apporter les solutions adéquates, aussi bien dans mon environnement de travail actuel et pour le reste de ma carrière professionnelle.*

*merci à tous pour la collaboration active ! »*

Exemple 3 : « *Bonsoir chers toutes et tous,*

*Je me nomme Yamtaryaoba Youssouf SAVADOGO, de nationalité bukinabè et résident à Ouagadougou. Je suis juriste de formation et je suis stagiaire depuis plus d'un an dans une société de transit (Emma Logistics Corporation) de la place. J'espère à travers cette formation acquérir des connaissances qui me seront utiles dans ma carrière et bénéficier des diverses expériences de tout un chacun. Ravi d'être parmi vous, je vous souhaite à tous bon courage et plein succès. »*

L'exemple ci-dessous est une intervention présentant une motivation dans un type de forum libre,

Exemple 1 : « *Bonjour Madame la Coordonnatrice Pédagogique,*

*Merci pour vos conseils et analyses. Pour ma part j'ai des préoccupations dont je voulais vous soumettre afin d'avoir des éléments de réponses.*

*Etant donné que nous sommes des professionnels , comment faire pour effectuer un stage dans une autre entreprise a partir du moment où cela entraverait notre fonction actuel?*

*Aussi concernant le thème de MEF , après choix pourrons nous faire suivre par un responsable de la structure ou nous travaillons déjà?*

*Dans l'attente de lire vos commentaires , merci. » ;* et la lecture des interventions des autres apprenants entraîne une sensation d'appartenance à un groupe. Aucun apprenant ne se sent isolé et chaque participant est à l'aise face à l'apprentissage en ligne.

Exemple 1 : « *Bjr ARCHANGE je partage ta préoccupation je suis la parceque on m'a parle d'une LICENCE LOGISTIQUE de surcroit sur cette plate forme JE NE VOIS NULLEMENT UN MASTER LOGISTIQUE . »* et

Exemple 2 : « *PATRICE* pourra nous expliquer cette situation. Si ce DTSS à grade de licence pourquoi ne pas l'appeler tout court *LICENCE* je crois qu'il existe bel et bien une différence entre ce DTSS et la licence. »

Il écrit son inquiétude afin d'avoir une réponse (de la part de ses pairs ou du tuteur)

Exemple : « *Bonsoir Madame Mirisoa,*

*A la lecture des généralités du module, 3 points ont attiré mon attention et je souhaite vous les soumettre sous forme de questions :*

*1 - si la Promotion en phase de fin d'étude est la 5ème promotion, pourquoi sommes-nous la promotion 7 au lieu de 6ème promotion ?*

*2 - il y a des erreurs sur la 8 page de la présentation du Foad\_ELT\_2016, au niveau de l'écriture en lettre des chiffres 33 et 53.*

*3 - je constate avec regret que depuis la première promotion, moins de la moitié des apprenants réussissent à décrocher leur diplôme,*

*3.1 - Avez vous fait une enquête pour comprendre le pourquoi ?*

*3.2 - Est ce la faute des apprenants ou bien la synchronisation de l'apprentissage qui fait problème ?*

*3.3 - Quels conseils vous nous donner pour nous permettre de maintenir le cap ?*

*Merci ».*

Il intervient pour débarrasser l'angoisse des autres.

Exemple : « *Bonjour à tous,*

*En attendant d'avoir la réponse du coordonnateur administratif merci a toi Archange pour l'inquiétude soulevée.*

*Il faut dire que nous partageons les mêmes inquiétudes car voulant à la fin un diplôme qui sied aux réalités du système universitaire qui a basculé sur le mode LMD. »*

#### 4.5.1.2. Communication

**Communication (comm)** : à part les communications proprement dites, du point de vue organisationnel, le forum est utilisé comme outil de communication afin de partager des informations entre membres et aussi pour s'organiser. Les exemples ci-dessous présentent des messages de communication et copiés tels quels via le forum de la plateforme :

Exemple 1 : *« Alerte, RDV ce soir 17H GMT pour échanger sur ce qui est déjà sur le wiki. » ;  
« Merci Mme pour la précision. » ;*

Exemple 2 : *« D'accord. A ce soir » et « Je pense que c'est le cas pour les apprenants du Cameroun, nous sommes toujours en attente. Merci de nous informer au cas ou le planning de regroupement soit modifié ou annulé. »*

#### 4.5.1.3. Collaboration

**Collaboration** : Le forum est utilisé principalement comme outil de collaboration (**Collab**). Les exemples ci-dessous présentent des messages de collaboration et copiés tels quels via le forum de la plateforme :

Exemple 1 : *« Bonjour a vous,*

*Je pense aussi qu'on pourrait travailler avant mercredi en soiree et demain de preference a partir de 19H GMT. Je serai connecte.*

*Vous trouverez egalement mon devoir ci-joint.*

*Bon weekend*

 [devoir3-christian.doc](#) » ; et

Exemple 2 : « *Bonjour à tous!*

*j ai remarqué l indisponibilité des uns et des autres sur le chat je pense que ceci est dû a vos divers obligations professionnelles mais n 'oubliez pas que nous sommes tenu par un délais de dépôt de nos devoirs et nous devons obligatoirement travailler en groupe*

*Svp faites un effort. je propose de reporter notre séance d' échange à ce soir 20 h GMT.*

*Merci d 'être ponctuel.*

*Genevieve. » ;*

D'après Walckiers [4.31], le travail collaboratif dans une formation à distance ou l'"apprentissage collaboratif" consiste en toute activité d'apprentissage réalisée par un groupe d'individus ayant un but commun, étant chacun une source d'information, de motivation, d'interaction, d'entraide ... et bénéficiant chacun des apports des autres, de la synergie du groupe et de l'aide d'un formateur facilitant les apprentissages individuels et collectifs. Chaque membre travaille sur les mêmes tâches. Il s'agit d'intégrer et de fusionner les contributions individuelles dans l'action commune afin de produire un résultat collectif.

D'après les travaux de Walckiers et De Praetere [4.31], on peut classifier les retombées de l'apprentissage collaboratif en trois catégories :

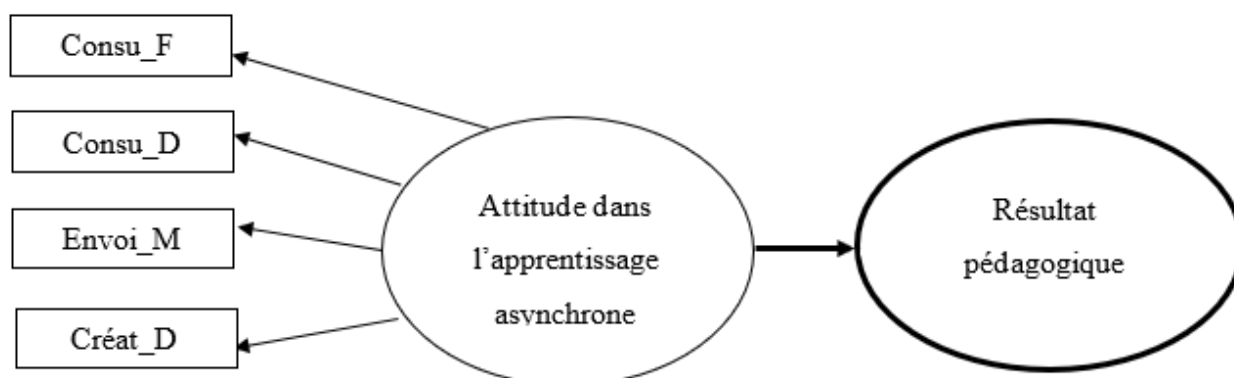
- **d'un point de vue pédagogique**, l'apprentissage collaboratif peut citer le développement de la pensée critique , le développement des compétences permettant la résolution de problèmes, la participation active à son apprentissage, le développement de son indépendance et de son autonomie, la concentration sur la tâche, la co-crédation de contenu, le développement de compétences professionnelles ; c'est l'occasion de dégager des significations communes à partir d'informations différentes, le développement d'aptitudes communicatives ; bref, un apprentissage plus en profondeur.

- **Dans une perspective plus sociale**, l'apprentissage collaboratif fournit un réseau de soutien aux apprenants ; il expose une gamme de perspectives différentes, crée un climat propice à la critique des idées et non des personnes. Il aide l'apprenant à développer des compétences sociétales, favorise l'apprentissage en tout temps et en tout lieu, constitue un lieu de pratique de l'exercice du leadership, contribue à l'établissement de communautés apprenantes et au renforcement des liens entre pairs et aide les professeurs à adopter un apprentissage centré sur l'apprenant.

- **D'un point de vue psychologique**, l'apprentissage collaboratif peut contribuer à augmenter la confiance en soi des apprenants, à réduire l'anxiété, à les aider à adopter des attitudes positives à l'égard de leurs professeurs et à vaincre le sentiment d'isolement de l'apprenant à distance.

#### ***4.5.2. Attitude dans l'apprentissage asynchrone***

Les traces de l'activité des apprenants dans une FOAD sont transformées pour représenter une mesure de l'engagement comportemental [4.15]. L'attitude dans l'apprentissage asynchrone a été modélisée comme un continuum de participation aux forums de discussion. Pour chacun de ces types d'actions, l'intervention de chaque apprenant a été comptée dans une variable. La même démarche a été utilisée pour les consultations. Si les traces peuvent être considérées comme des mesures objectives du comportement de l'apprenant, la recherche devrait montrer sa relation avec l'acquisition de connaissance présentée par le résultat pédagogique.



**Figure 4.05 :** *Modèle de mesure de l'attitude dans l'apprentissage asynchrone*

$Y_1 = \text{Consu\_F}$  : Nombre de forums consultés (forum\_view forum)

$Y_2 = \text{Consu\_D}$  : Nombre de discussions consultées (forum\_view discussion)

$Y_3 = \text{Envoi\_M}$  : Nombre de postages de contenus (forum\_add post)

$Y_4 = \text{Créat\_D}$  : Nombre de discussions créées (forum\_add discussion)

Les forums constituent les principaux outils de communication collective et de collaboration dans Moodle. Ils permettent de diffuser des informations, d'organiser une discussion sur un sujet donné, de dialoguer avec l'ensemble des apprenants, de leur proposer un espace pour discuter entre eux, et de répondre aux questions posées.

Le forum apporte beaucoup d'avantages par rapport à l'envoi de mail en nombre : c'est un espace de discussion structuré dont tout l'historique est conservé sur la plateforme et consultable à tout moment par chaque utilisateur. Les messages sont automatiquement envoyés à toutes les personnes inscrites, donc pas de risque d'oublier un destinataire. Quelles sont les actions des apprenants sur le forum dans leur apprentissage à distance ? Leurs gestes sont relevés ici dans quatre catégories d'actes qui sont : la consultation du forum, la consultation d'une discussion, l'envoi d'un contenu et la création d'une discussion.

#### 4.5.2.1. Consulter un forum / une discussion

Les forums sont structurés en **sujets de discussion** et en **messages**. Un forum contient un ou plusieurs **sujets** de discussion et chaque discussion comporte elle-même des **messages** que l'on peut consulter. L'apprenant peut prendre connaissance tout simplement de ces sujets de discussion ( $\text{Consu\_F}$ ) ou lire le contenu d'une discussion ( $\text{Consu\_D}$ ).

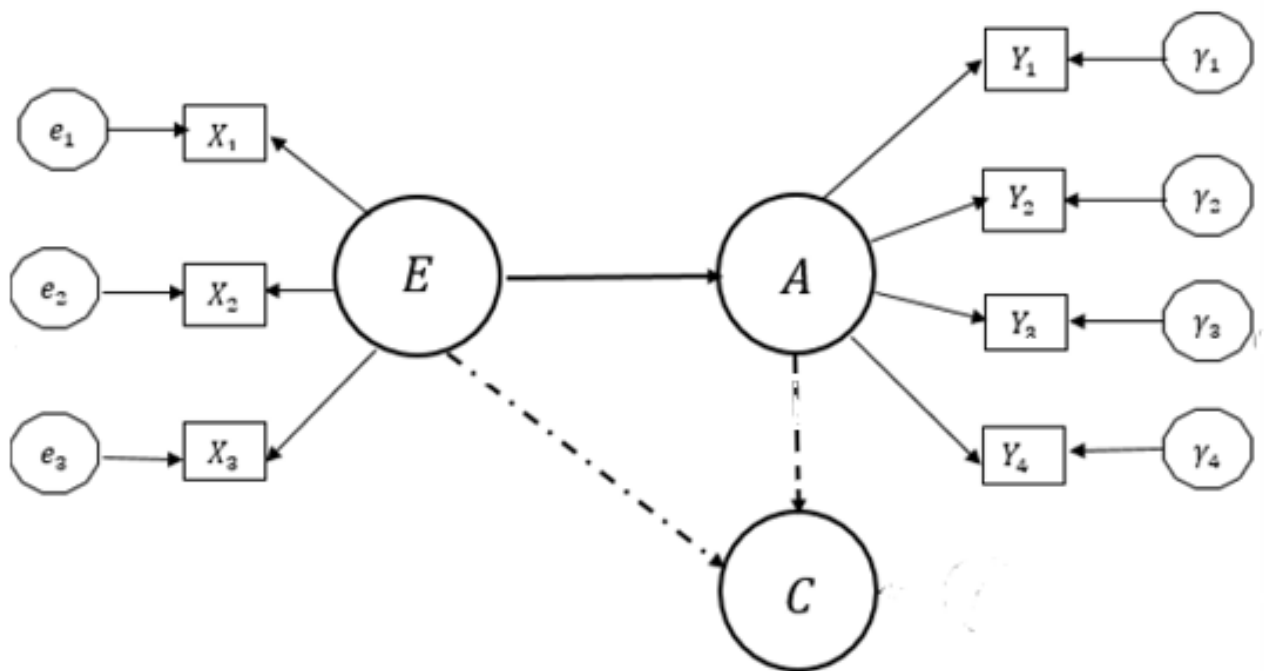
#### 4.5.2.2. Poster un message ou créer un nouveau sujet de discussion

Un apprenant a deux façons de publier dans un forum. Premièrement, en répondant à un message ( $\text{Envoi\_M}$ ), où il poste un message sur une discussion en cours ; et deuxièmement, en créant un nouveau sujet de discussion ( $\text{Créat\_D}$ ).



#### 4.5.3. Modélisation d'équations structurelles par le maximum de vraisemblance (LISREL)

Comme le montre la Figure 4.06 : , le modèle causal se compose d'un modèle de mesure et d'un modèle de structure. Il englobe l'ensemble de variables observables indépendantes ( $X_i$ ), de variables observables dépendantes ( $Y_i$ ), d'une variable latente explicative ( $E$ ), de deux variables latentes à expliquer ( $A$ ) et ( $C$ ) et des termes d'erreurs ( $e_i$ ,  $\delta$ ,  $\beta$  et  $\gamma_i$ ). Par conséquent, il est possible de distinguer entre deux types d'équations à savoir les équations du modèle de mesure et les équations du modèle de structure. Le modèle structurel de la covariance (LISREL) utilise une approche statistique pour tester des hypothèses sur les relations entre variables observées ou explicatives et latentes [3.13].



**Figure 4.06 :** *Modèle causal de la recherche*

#### 4.5.3.1. Equations du modèle de mesure

$$X_1 = \lambda_1 E + e_1$$

$$Y_1 = \alpha_1 A + \gamma_1$$

$$X_2 = \lambda_2 E + e_2$$

$$Y_2 = \alpha_2 A + \gamma_2$$

$$X_3 = \lambda_3 E + e_3$$

$$Y_3 = \alpha_3 A + \gamma_3$$

$$Y_4 = \alpha_4 A + \gamma_4$$

#### 4.5.3.2. Equations du modèle de structure

$$A = r_1 E$$

$$C = r_2 A + r_3 E$$

#### 4.5.3.3. Opérationnalisation des variables

L'engagement de l'apprenant est envisagé comme une variable réflexive dont les trois indicateurs sont la motivation, la communication et la collaboration.

Notation	Signification	Definition
$X_1 = Collab$	Interaction collaborative dans le forum de discussion	Indicateur de la variable latente $E$
$X_2 = Comm$	Communication avec les protagonistes de la formation (pairs, tuteur, ...)	Indicateur de la variable latente $E$
$X_3 = Motiv$	Motivation dans la communauté d'apprentissage (groupe, sous-groupe, formation)	Indicateur de la variable latente $E$
$Y_1 = Consu\_F$	forum consulté (forum_view forum)	Indicateur de la variable latente $A$
$Y_2 = Consu\_D$	discussion consultée (forum_view discussion)	Indicateur de la variable latente $A$
$Y_3 = Envoi\_M$	Postage de contenus (forum_add post)	Indicateur de la variable latente $A$
$Y_4 = Créat\_D$	Discussion créée (forum_add discussion)	Indicateur de la variable latente $A$
$E$	Engagement de l'apprenant dans une FOAD	Variable latente explicative
$A$	Attitude dans l'apprentissage asynchrone	Variable latente expliquée
$C$	Acquisition de Connaissances	Variable latente expliquée

**Tableau 4.02 :** *Les variables opérées dans la recherche*

#### 4.5.3.4. Les paramètres

Notation	Signification
$\lambda_1$	Coefficient représentant l'effet du construit E sur l'indicateur
$\lambda_2$	Coefficient représentant l'effet du construit E sur l'indicateur
$\lambda_3$	Coefficient représentant l'effet du construit E sur l'indicateur
$\alpha_1$	Coefficient représentant l'effet du construit A sur l'indicateur
$\alpha_2$	Coefficient représentant l'effet du construit A sur l'indicateur
$\alpha_3$	Coefficient représentant l'effet du construit A sur l'indicateur
$\alpha_4$	Coefficient représentant l'effet du construit A sur l'indicateur
$e_1$	Paramètre d'erreur
$e_2$	Paramètre d'erreur
$e_3$	Paramètre d'erreur
$\gamma_1$	Paramètre d'erreur
$\gamma_2$	Paramètre d'erreur
$\gamma_3$	Paramètre d'erreur
$\gamma_4$	Paramètre d'erreur
$r1$	Coefficient représentant l'effet du construit E sur la variable latente A
$r2$	Coefficient représentant l'effet du construit A sur la variable C
$r3$	Coefficient représentant l'effet du construit E sur la variable C

**Tableau 4.03 :** *Les paramètres à estimer dans la recherche*

## 4.6 Traitement de données sur R

Le logiciel R est un logiciel de statistique créé par Ross Ihaka & Robert Gentleman. Il est à la fois un langage informatique et un environnement de travail. Ce logiciel sert à manipuler des données,

à tracer des graphiques et à faire des analyses statistiques sur ces données. R est un logiciel libre et à code source ouvert. C'est aussi un outil très puissant et très complet, particulièrement bien adapté pour la mise en œuvre informatique de méthodes statistiques.

Les traces d'activités des apprenants dans la plateforme de la formation sont exploitées pour mesurer les effets directs et indirects des engagements des apprenants et de leurs comportements face au forum de discussion et les effets sur leurs cognitions qui se traduisent par leurs résultats pédagogiques.

#### ***4.6.1. Le modèle de mesure de la variable latente A***

Le modèle de mesure des attitudes de l'apprenant face au forum de discussion peut se traduire par les équations suivantes :

$$Y_1 = \alpha_1 A + \gamma_1$$

$$Y_2 = \alpha_2 A + \gamma_2$$

$$Y_3 = \alpha_3 A + \gamma_3$$

$$Y_4 = \alpha_4 A + \gamma_4$$

Qui peuvent s'écrire sous la forme matricielle :  $Y = \alpha A + \gamma$  et forment encore une régression linéaire

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} A + \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix}$$

En respectant les étapes décrites dans le chapitre précédent, nous pouvons commencer par

##### **4.6.1.1. La matrice de covariance entre les quatre variables**

```
> donnees.cov <- matrix(c(cov(donnees)),nrow=4,ncol=4)
```

```

> colnames(donnees.cov) <- rownames(donnees.cov) <-
c("Consu_F", "Consu_D", "Envoi_M", "Creat_D")
> donnees.cov

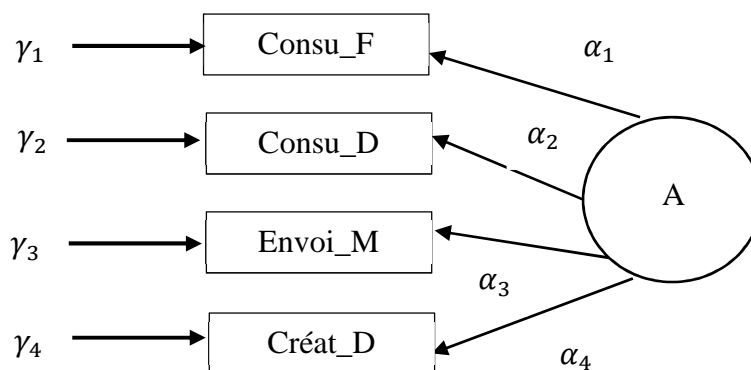
           Consu_F    Consu_D    Envoi_M    Creat_D
Consu_F 120.089744 156.134615  64.647436   3.7307692
Consu_D 156.134615  769.576923  92.929487  -1.5705128
Envoi_M  64.647436  92.929487  53.807692   0.4743590
Creat_D   3.730769  -1.570513   0.474359   0.6410256

```

Nous tenons à remarquer qu'il y a des covariances négatives entre les deux variables « Creat\_D » et « consu\_D » dans cette matrice. Ceci va engendrer des erreurs lors du calcul des paramètres du modèle comme le montre le résultat de la spécification ci-dessous.

#### 4.6.1.2. Spécification du modèle

On va spécifier le modèle de la Figure 4.07 : qui est un modèle de mesure extrait du modèle de la Figure 4.06 :



**Figure 4.07 :** *Modèle de mesure proposé pour l'attitude de l'apprenant*

Ce modèle induit le script suivant dans R :

```

> donnees.model <- 'A =~ l1*Consu_F + l2*Consu_D + l3*Envoi_M + l4*Creat_D'

```

#### 4.6.1.3. Estimation du modèle

La méthode LISREL (Linear Structural Relationship) est intéressante du point de vue de sa validité théorique et de l'optimalité des estimations des paramètres. Elle s'applique dans le cadre de la validation d'une théorie. Le fondement statistique de la méthode LISREL est **la covariance**. La méthode LISREL est une méthode a priori et nécessite que les chercheurs réfléchissent en termes de modèles et d'hypothèses.

L'estimation utilisée par la librairie lavaan est le maximum de vraisemblance.

```
> donnees.fit <- cfa(model=donnees.model, sample.cov=donnees.cov,  
sample.nobs=13, std.lv=FALSE)
```

```
> summary(donnees.fit, standardized= TRUE)
```

lavaan (0.5-20) converged normally after 7959 iterations

Number of observations	13
Estimator	ML
Minimum Function Test Statistic	17.612
Degrees of freedom	2
P-value (Chi-square)	0.000

Parameter Estimates:

Information	Expected
Standard Errors	Standard

Latent Variables:

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
A =~							
Consu_F	(11)	1.000				0.233	0.022
Consu_D	(12)	-0.763	NA			-0.178	-0.007
Envoi_M	(13)	-0.400	NA			-0.093	-0.013
Creat_D	(14)	174.569	NA			40.601	33.370

Variances:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
Consu_F	110.857	NA			110.857	1.000

Consu_D	710.347	NA	710.347	1.000
Envoi_M	49.660	NA	49.660	1.000
Creat_D	-1646.970	NA	-1646.970	-1112.532
A	0.054	NA	1.000	1.000

Le modèle ne converge qu'après 7959 itérations avec un test sur un échantillon de 13 observations. Le degré de liberté  $ddl = p(p + 1)/2 - N$  avec  $p = 4$  et  $N = 6$  est égal à 2 et, est vérifié par le « Degrees of freedom » retourné qui indique la sur-identification du modèle.

On note que l'erreur sur la variable Creat\_D est de -1112,532 (variance négative) qui définit une raison de respécification de notre modèle.

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

fmin	logl	chisq	df	npar
0.677	-164.987	17.612	2.000	8.000

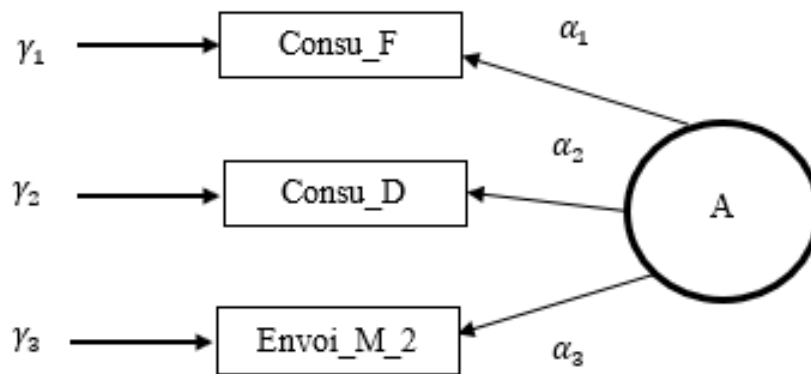
Le logarithme de la vraisemblance (logl) : -164,987 indique que le modèle n'a pas de vraisemblance par rapport aux données réelles.

#### 4.6.1.4. Respécification du modèle

Alors, il faut respécifier le modèle. La solution prise est de fusionner les deux variables « Envoi\_M » et « Creat\_D » car tous les deux expriment la participation de l'apprenant au forum de discussion .

Par conséquent, nous n'avons que trois variables observées au lieu de quatre comme auparavant.





**Figure 4.08 :** *Modèle théorique après respécification*

```

> #Chargement des données :
> gsa2 <- read.csv("H:/gsa2.csv", sep=";")
> donnees2 <- gsa2[6:8]

```

Nous avons repassé les étapes précédentes en traitant les nouvelles données. D'où, pour la matrice de covariance, nous avons :

```

> donnees2.cov <- matrix(c(cov(donnees2)),nrow=3,ncol=3)
> colnames(donnees2.cov) <- rownames(donnees2.cov) <-
c("Consu_F","Consu_D","Envoi_M_2")
> donnees2.cov

```

	Consu_F	Consu_D	Envoi_M_2
Consu_F	120.08974	156.13462	68.37821
Consu_D	156.13462	769.57692	91.35897
Envoi_M_2	68.37821	91.35897	55.39744

#### 4.6.1.5. Estimation du modèle 2

Le nouveau modèle a de la forme :  $A = l1 * Y_1 + l2 * Y_2 + l3 * Y_5$

```

> donnees2.model <- 'A =~ l1*Consu_F + l2*Consu_D + l3*Envoi_M_2'
> #Estimation du modèle

```

```
> donnees2.fit <- cfa(model=donnees2.model, sample.cov=donnees2.cov,
sample.nobs=13, std.lv=FALSE)
> #Description du modèle (coefficients, ...)
> summary(donnees2.fit, fit.measures = TRUE, standardized= TRUE)
```

lavaan (0.5-20) converged normally after 79 iterations

Number of observations	13
Estimator	ML
Minimum Function Test Statistic	0.000
Degrees of freedom	0

Parameter Estimates:

Information	Expected
Standard Errors	Standard

Model test baseline model:

Minimum Function Test Statistic	19.763
Degrees of freedom	3
P-value	0.000

User model versus baseline model:

Comparative Fit Index (CFI)	1.000
Tucker-Lewis Index (TLI)	1.000

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-144.312
Loglikelihood unrestricted model (H1)	-144.312
Number of free parameters	6
Akaike (AIC)	300.624
Bayesian (BIC)	304.014
Sample-size adjusted Bayesian (BIC)	285.804

Root Mean Square Error of Approximation:

RMSEA	0.000
-------	-------

90 Percent Confidence Interval	0.000	0.000
P-value RMSEA <= 0.05		1.000

Standardized Root Mean Square Residual:

SRMR	0.000
------	-------

Parameter Estimates:

Information	Expected
Standard Errors	Standard

Latent Variables:

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
A =~							
Consu_F	(11)	1.000				10.386	0.986
Consu_D	(12)	1.336	0.719	1.859	0.063	13.877	0.521
Envoi_M_2	(13)	0.585	0.200	2.924	0.003	6.077	0.850

Variances:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
Consu_F	2.981	31.672	0.094	0.925	2.981	0.027
Consu_D	517.817	210.816	2.456	0.014	517.817	0.729
Envoi_M_2	14.204	12.184	1.166	0.244	14.204	0.278
A	107.871	53.767	2.006	0.045	1.000	1.000

La dernière colonne « Std.all » présentent les paramètres estimés et les erreurs du modèle :

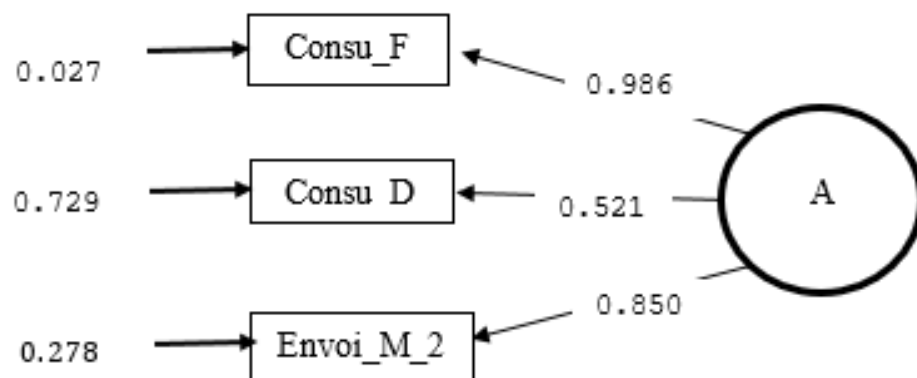
$$\left. \begin{array}{l} \alpha_1 = 0,986 \\ \alpha_2 = 0,521 \\ \alpha_3 = 0,850 \end{array} \right\} \Rightarrow \begin{pmatrix} \text{Consu}_F \\ \text{Consu}_D \\ \text{Envoi}_M \end{pmatrix} = \begin{pmatrix} 0,986 \\ 0,521 \\ 0,850 \end{pmatrix} * A + \begin{pmatrix} 0,027 \\ 0,729 \\ 0,278 \end{pmatrix}$$

$$\gamma_1 = 0,027$$

$$\gamma_2 = 0,729$$

$$\gamma_3 = 0,278$$

La variance de A est égale à 1 indique que A est une variable normalisée.



**Figure 4.09 :** *Modèle de mesure de la variable latente A avec les paramètres estimés*

#### 4.6.1.6. Identification du modèle

Le degré de liberté  $ddl = p(p + 1)/2 - N$  avec  $p=3$  et  $N=6$  devient nul. Le modèle devient juste identifié.

#### 4.6.1.7. Evaluation du modèle

Model test baseline model:

Minimum Function Test Statistic	19.763
Degrees of freedom	3
P-value	0.000

User model versus baseline model:

Comparative Fit Index (CFI)	1.000
Tucker-Lewis Index (TLI)	1.000

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-144.312
Loglikelihood unrestricted model (H1)	-144.312
Number of free parameters	6
Akaike (AIC)	300.624
Bayesian (BIC)	304.014
Sample-size adjusted Bayesian (BIC)	285.804

Root Mean Square Error of Approximation:

RMSEA	0.000
90 Percent Confidence Interval	0.000 0.000
P-value RMSEA <= 0.05	1.000

Standardized Root Mean Square Residual:

SRMR	0.000
------	-------

Le minimum de la fonction estimation (fmin) qui est égal à 0 nous indique que c'est un bon modèle par rapport au modèle précédent (fmin = 0,677)

Le logarithme de la vraisemblance (logl) est -144,312 et est plus proche de 0 par rapport à -164.987

Le degré de liberté du modèle (df) est 0, comme on l'a calculé manuellement.

Le nombre de paramètre estimé (npar) du modèle est 6.

### Indices d'ajustement absolu :

- Le  $\chi^2$  (chisq) est égal à 0 : test du chi-carré (Seuil : Valeurs proches de zero) : P-value = 0, alors le modèle est bien ajusté selon cet indice.
- Erreur quadratique moyenne d'approximation (seuil : Moins de 0.06) : RMSEA = 0.000, alors le modèle est bien ajusté selon cet indice.
- Résidu moyen carré normalisé (Seuil : Moins de 0.08) : SRMR = 0.000, le modèle est donc bien ajusté selon cet indice.

### Indices d'ajustement relatif :

- Indice d'ajustement non normé de Tucker-Lewis (Seuil : Plus de 0.90) : TLI = 1.000, alors le modèle est bien ajusté selon cet indice.
- Indice d'ajustement comparatif (Seuil : Plus de 0.95) : CFI = 1.000, alors le modèle est bien ajusté selon cet indice.

D'après ces indicateurs, nous pouvons conclure par la validation du second modèle.

#### ***4.6.2. Le modèle de mesure de la variable latente E***

Le modèle de mesure des engagements de l'apprenant peut se traduire par les équations suivantes :

$$X_1 = \lambda_1 E + e_1$$

$$X_2 = \lambda_2 E + e_2$$

$$X_3 = \lambda_3 E + e_3$$

Peuvent s'écrire sous la forme matricielle :  $\mathbf{X} = \boldsymbol{\lambda}E + \mathbf{e}$  qui forme une régression linéaire.

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} E + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix}$$

#### 4.6.2.1. La matrice de covariance entre les trois variables

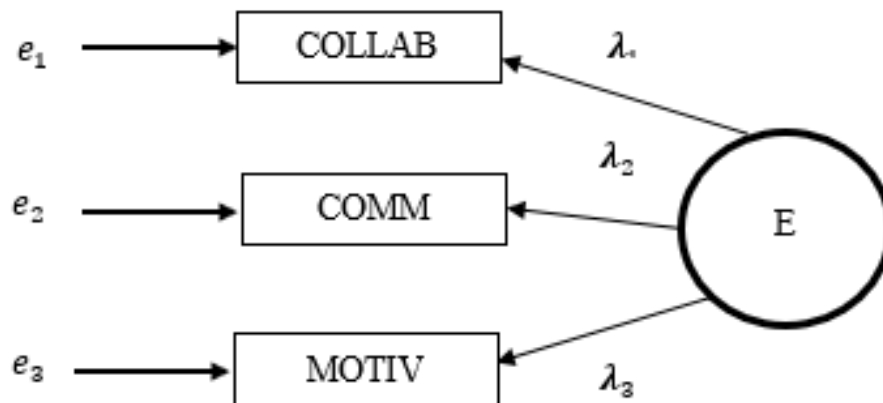
```
library(lavaan)

> gsa3 <- read.csv("H:/gsa2.csv", sep=";")
> donnees3 <- gsa3[2:4]
> donnees3.cov <- matrix(c(cov(donnees3)),nrow=3,ncol=3)
> colnames(donnees3.cov) <- rownames(donnees3.cov) <-
c("COLLAB", "COMM", "MOTIV")
> donnees3.cov
```

	COLLAB	COMM	MOTIV
COLLAB	2.230769	1.1602564	1.1858974
COMM	1.160256	2.1923077	0.7564103
MOTIV	1.185897	0.7564103	1.5641026

#### 4.6.2.2. Spécification du modèle

On va spécifier le modèle de la Figure 4.07 : qui est un modèle de mesure extrait du modèle de la Figure 4.06 :



**Figure 4.10 :** *Modèle de mesure à estimer pour l'engagement de l'apprenant*

#### 4.6.2.3. Estimation du modèle

Le script suivant est utilisé pour estimer le modèle :

```
> donnees3.model <- 'E =~ 11*COLLAB + 12*COMM + 13*MOTIV'
> donnees3.fit <- cfa(model=donnees3.model, sample.cov=donnees3.cov,
sample.nobs=13, std.lv=FALSE)
> summary(donnees3.fit, standardized= TRUE)
```

lavaan (0.5-20) converged normally after 22 iterations

Number of observations	13
Estimator	ML
Minimum Function Test Statistic	0.000
Degrees of freedom	0
Minimum Function Value	0.000000000000000

Parameter Estimates:

Information	Expected
Standard Errors	Standard

Model test baseline model:

Minimum Function Test Statistic	11.065
Degrees of freedom	3
P-value	0.011

User model versus baseline model:

Comparative Fit Index (CFI)	1.000
Tucker-Lewis Index (TLI)	1.000

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-61.470
Loglikelihood unrestricted model (H1)	-61.470
Number of free parameters	6
Akaike (AIC)	134.941
Bayesian (BIC)	138.331
Sample-size adjusted Bayesian (BIC)	120.121

Root Mean Square Error of Approximation:



RMSEA	0.000
90 Percent Confidence Interval	0.000 0.000
P-value RMSEA <= 0.05	1.000

Standardized Root Mean Square Residual:

SRMR	0.000
------	-------

Parameter Estimates:

Information	Expected
Standard Errors	Standard

Latent Variables:

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
A =~							
COLLAB	(11)	1.000				1.296	0.903
COMM	(12)	0.638	0.372	1.713	0.087	0.827	0.581
MOTIV	(13)	0.652	0.348	1.874	0.061	0.845	0.703

Variances:

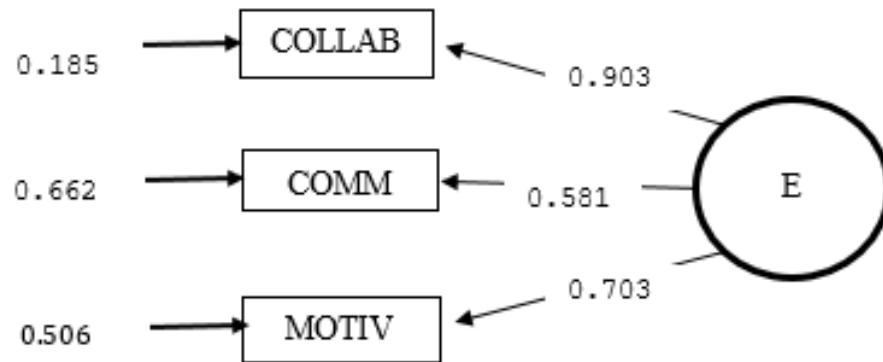
	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
COLLAB	0.380	0.777	0.489	0.625	0.380	0.185
COMM	1.341	0.611	2.196	0.028	1.341	0.662
MOTIV	0.730	0.433	1.688	0.091	0.730	0.506
A	1.679	1.101	1.525	0.127	1.000	1.000

$$\left. \begin{array}{l} \lambda_1 = 0,903 \\ \lambda_2 = 0,581 \\ \lambda_3 = 0,703 \end{array} \right\} \Rightarrow \begin{pmatrix} COLLAB \\ COMM \\ MOTIV \end{pmatrix} = \begin{pmatrix} 0,903 \\ 0,581 \\ 0,703 \end{pmatrix} * A + \begin{pmatrix} 0,185 \\ 0,662 \\ 0,506 \end{pmatrix}$$

$$e_1 = 0,185$$

$$e_2 = 0,662$$

$$e_3 = 0,506$$



**Figure 4.11 :** *Modèle de mesure de E avec les paramètres estimés*

#### 4.6.2.4. Identification du modèle

Manuellement  $Ddl = 3(3+1) / 2 - 6 = 0$  et d'après le logiciel R (cf. Annexe 5), « df » = 0, cela veut dire que le modèle est juste identifié.

#### 4.6.2.5. Evaluation du modèle

Model test baseline model:

Minimum Function Test Statistic	11.065
Degrees of freedom	3
P-value	0.011

User model versus baseline model:

Comparative Fit Index (CFI)	1.000
Tucker-Lewis Index (TLI)	1.000

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-61.470
Loglikelihood unrestricted model (H1)	-61.470
Number of free parameters	6
Akaike (AIC)	134.941
Bayesian (BIC)	138.331
Sample-size adjusted Bayesian (BIC)	120.121

Root Mean Square Error of Approximation:

RMSEA	0.000
90 Percent Confidence Interval	0.000 0.000
P-value RMSEA <= 0.05	1.000

Standardized Root Mean Square Residual:

SRMR	0.000
------	-------

Parameter Estimates:

Information	Expected
Standard Errors	Standard

Le minimum de la fonction estimation ( $f_{\min}$ ) est égal à 0.

Le logarithme de la vraisemblance ( $\log l$ ) est  $-61,47$

Le degré de liberté du modèle ( $df$ ) est 0, comme on l'a calculé manuellement.

Le nombre de paramètre estimé ( $n_{\text{par}}$ ) du modèle est 6.

### **Indices d'ajustement absolu :**

- Le  $\chi^2$  (chi sq) est égal à 0 : test du chi-carré (Seuil : Valeurs proches de zero) : P-value = 0, alors le modèle est bien ajusté selon cet indice.
- Erreur quadratique moyenne d'approximation (seuil : Moins de 0.06) : RMSEA = 0.000, alors le modèle est bien ajusté selon cet indice.
- Résidu moyen carré normalisé (Seuil : Moins de 0.08) : SRMR = 0.000, alors le modèle est bien ajusté selon cet indice.

### **Indices d'ajustement relatif :**

- Indice d'ajustement non normé de Tucker-Lewis (Seuil : Plus de 0.90) : TLI = 1.000, alors le modèle est bien ajusté selon cet indice.
- Indice d'ajustement comparatif (Seuil : Plus de 0.95) : CFI = 1.000, alors le modèle est bien ajusté selon cet indice.

## **4.7 Conclusion**

L'analyse du contenu est effectuée dans les forums de discussion du module d'initiation. Un modèle est proposé, et les différentes étapes du SEM sont réalisées. Ainsi, à l'issue du présent chapitre, nous avons obtenu les deux modèles de mesure validés. Ces modèles seront utilisés dans le chapitre suivant pour avoir un modèle d'équation structurelle de l'apprentissage asynchrone. Ils seront manipulés avec les données réelles des traces des activités de l'apprenant dans les trois formations ouvertes et à distance pour pouvoir valider un modèle structurel. La relation directe de l'engagement avec le résultat de la formation sera aussi à démontrer, autrement dit, les relations de causalité du résultat pédagogique, de l'engagement et de l'attitude de l'étudiant face à leur apprentissage asynchrone.

## CHAPITRE 5 – APPLICATION DU MODELE DE L'APPRENTISSAGE ASYNCHRONE

### 5.1 Introduction

L'engagement de l'apprenant dans l'apprentissage asynchrone est mesuré à partir de sa motivation, de sa collaboration et de sa communication via les forums de discussion. L'évaluation de ces trois variables de mesure dans l'apprentissage du module d'initiation permet de mesurer la variable latente « engagement » de chaque apprenant. De même, la consultation de forum, la consultation de discussion, la participation de l'apprenant par l'envoi et la mise à jour de messages sont toujours prises en compte et tous les messages sont énumérés, ceci permet de mesurer l'attitude de l'apprenant lors de son apprentissage à distance. Ainsi, les deux modèles validés dans le chapitre précédent, seront utilisés pour toutes les applications dans le présent chapitre, afin de sortir le modèle de structure de la recherche. Est-ce qu'il y a une relation de causalité entre l'engagement de l'apprenant et son attitude vis-à-vis du forum de discussion ? Est-ce que l'engagement permet le développement de connaissance de l'apprenant ? Est-ce que l'attitude de l'apprenant par rapport au forum de discussion a un impact positif sur son résultat pédagogique ? Voilà les différentes questions de recherche auxquelles nous allons répondre dans ce dernier chapitre du travail.

Ce chapitre est un développement de l'article de publication de notoriété internationale publié le 21 août 2018 :

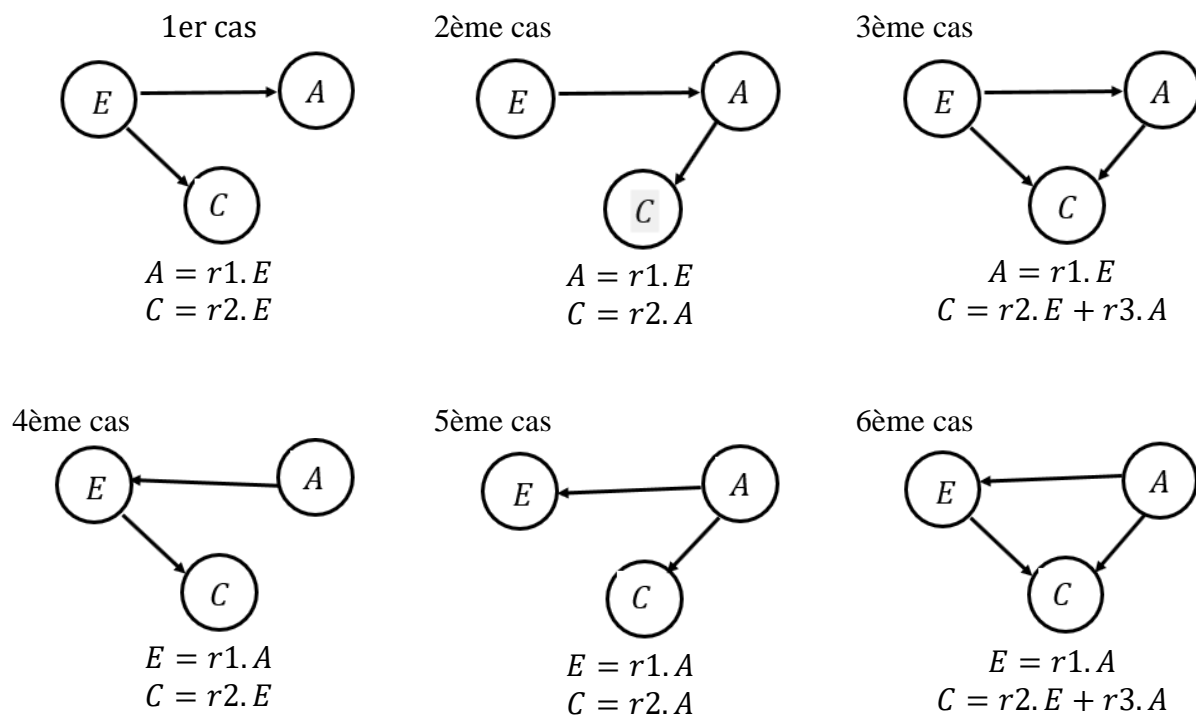
- M. Rakotomalala, T. B. Ravaliminoarimalalason, F. Randimbindrainibe, « *Asynchronous learning model in an ODL* », in International Journal of Engineering Research & Technology, Vol. 7 Issue 08, pp. 174-179, Août-2018.

### 5.2 Identification du modèle de structure de la recherche

Le modèle de structure permet d'examiner le lien entre les différentes variables latentes. Rappelons ce que nous avons vu dans le chapitre 3 : la méthode d'équations structurelles représente une technique multivariée qui combine les modèles de mesure et les modèles de structures tout en examinant de manière simultanée une série de relations linéaires entre les variables observées et les variables latentes d'une part, et entre l'ensemble des variables latentes d'autre part [3.15].

Nous avons travaillé avec les données obtenues sur GEAL à partir de l'analyse de traces numériques des activités des 96 apprenants des 5 promotions issues de 3 formations. Nous avons effectué l'analyse du contenu de tous les forums existants. 812 interventions ont été analysées. Lors de l'apprentissage, 9685 événements se sont passés. Les deux variables latentes ont été connues à partir de ces analyses ainsi que les deux modèles de mesure. La suite du travail permet de déterminer la troisième variable latente  $C$  qui reflète le résultat pédagogique de la formation.

Pour l'équation de structure du modèle, six cas sont probables. Prenons chaque cas un à un et effectuons la simulation pour identifier les modèles possibles.

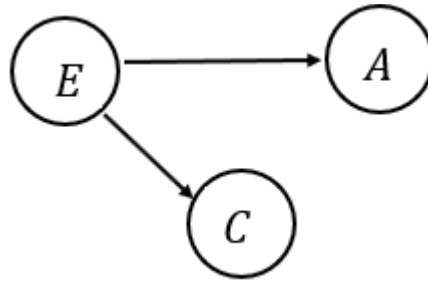


**Figure 5.01.** Les six cas possibles pour notre modèle

### 5.2.1. 1<sup>er</sup> cas de modèle

Pour le premier cas, la variable latente  $C$  est mesurée par la variable latente Engagement  $E$ , tel que l'attitude  $A$  de l'apprenant est une regression de cet engagement  $E$ .

Le modèle est représenté par la Figure 5.02



**Figure 5.02.** *Modèle de structure, proposition 1*

### Script et résultats pour l'estimation du modèle

```

> library(lavaan)

> #Chargement de données
> gsa <- read.csv("F:/Doctorat_2016/SCA/SCA/Thèses/D3/S5/sca/donnees.csv", sep="
;")

> donnee <- gsa[1:6]

> resultat <- gsa[7]

> colnames(donnee) <- c("COLLAB", "COMM", "MOTIV", "Consu_F", "Consu_
D", "Envoi_M")

> #Création du modele
> donnee.model <- '
+ A =~ y1*Consu_F + y2*Consu_D + y3*Envoi_M
+ E =~ l1*COLLAB + l2*COMM + l3*MOTIV
+ A =~ r1*E
+ C =~ r2*E
+ '

> #Estimation du modele
> donnee.fit <- cfa(model=donnee.model, data=donnee, se="robust")

> summary(donnee.fit, fit.measures= TRUE, standardized= TRUE)
lavaan (0.5-20) converged normally after 71 iterations

    Number of observations                    96

    Estimator                               ML
    Minimum Function Test Statistic          27.842
    Degrees of freedom                        8
    P-value (Chi-square)                     0.001

Model test baseline model:

    Minimum Function Test Statistic          413.264
    Degrees of freedom                       15
    P-value                                  0.000

User model versus baseline model:

```

Comparative Fit Index (CFI)	0.950
Tucker-Lewis Index (TLI)	0.907

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-2002.640
Loglikelihood unrestricted model (H1)	-1988.719
Number of free parameters	19
Akaike (AIC)	4043.280
Bayesian (BIC)	4092.003
Sample-size adjusted Bayesian (BIC)	4032.012

Root Mean Square Error of Approximation:

RMSEA	0.161
90 Percent Confidence Interval	0.098 0.228
P-value RMSEA <= 0.05	0.003

Standardized Root Mean Square Residual:

SRMR	0.046
------	-------

Parameter Estimates:

Information	Expected
Standard Errors	Robust.sem

Latent Variables:

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
A =~							
	Consu_F (y1)	1.000				29.102	0.870
	Consu_D (y2)	1.655	0.142	11.660	0.000	48.165	0.889
	Envoi_M (y3)	0.466	0.059	7.902	0.000	13.557	0.867
E =~							
	COLLAB (l1)	1.000				4.539	0.889
	COMM (l2)	0.717	0.245	2.928	0.003	3.255	0.627
	MOTIV (l3)	0.308	0.097	3.169	0.002	1.399	0.612
C =~							
	E (r2)	1.000				1.000	1.000

Regressions:

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
A ~							
	E (r1)	6.276	0.677	9.273	0.000	0.979	0.979

Intercepts:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
Consu_F	36.958	3.433	10.764	0.000	36.958	1.104
Consu_D	50.083	5.556	9.015	0.000	50.083	0.925
Envoi_M	13.844	1.604	8.631	0.000	13.844	0.886
COLLAB	3.750	0.524	7.162	0.000	3.750	0.735
COMM	2.510	0.532	4.714	0.000	2.510	0.484
MOTIV	2.198	0.234	9.375	0.000	2.198	0.962
A	0.000				0.000	0.000
E	0.000				0.000	0.000
C	0.000				0.000	0.000



Variances:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
Consu_F	273.009	110.834	2.463	0.014	273.009	0.244
Consu_D	612.393	324.960	1.885	0.059	612.393	0.209
Envoi_M	60.578	14.363	4.218	0.000	60.578	0.248
COLLAB	5.440	1.568	3.469	0.001	5.440	0.209
COMM	16.344	5.768	2.833	0.005	16.344	0.607
MOTIV	3.263	0.859	3.797	0.000	3.263	0.625
A	35.463	76.186	0.465	0.642	0.042	0.042
E	0.000				0.000	0.000
C	20.602	6.535	3.152	0.002	1.000	1.000

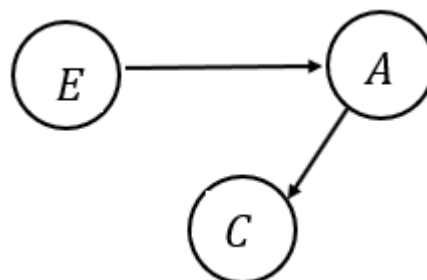
D'après ces résultats, on peut déduire que l'estimation du modèle a été effectuée avec succès.

Toutes les variables ont été correctement estimées selon les valeurs mises en exergue dans le résultat ci-dessus.

### 5.2.2. 2<sup>e</sup> cas de modèle

Pour ce deuxième cas, on va se pencher sur l'attitude qui décrit une regression de l'engagement, mais cette fois ci, le résultat *C* est mesurée par la variable latente attitude *A*.

Le modèle est représenté par la Figure 5.03.



**Figure 5.03.**Modèle de structure, proposition 2

Script et résultats pour l'estimation du modèle :

```

> library(lavaan)

> #Chargement de données
> gsa <- read.csv("F:/Doctorat_2016/SCA/SCA/Thèses/D3/S5/sca/donnees.csv", sep="
; ")

> donnee <- gsa[1:6]

> resultat <- gsa[7]

```

```

> colnames(donnee) <- c("COLLAB", "COMM", "MOTIV", "Consu_F", "Consu_
D", "Envoi_M")

> #Création du modele
> donnee.model <- '
+ A =~ y1*Consu_F + y2*Consu_D + y3*Envoi_M
+ E =~ l1*COLLAB + l2*COMM + l3*MOTIV
+ A =~ r1*E
+ C =~ r2*A
+ '

> #Estimation du modele
> donnee.fit <- cfa(model=donnee.model, data=donnee, se="robust")
Warning message:
In lav_model_vcov(lavmodel = lavmodel, lavsamplestats = lavsamplestats, :
lavaan WARNING: could not compute standard errors!
lavaan NOTE: this may be a symptom that the model is not identified.

> summary(donnee.fit, fit.measures= TRUE, standardized= TRUE)
lavaan (0.5-20) converged normally after 90 iterations

Number of observations                                96

Estimator                                             ML
Minimum Function Test Statistic                     27.842
Degrees of freedom                                   7
P-value (Chi-square)                                0.000

Model test baseline model:

Minimum Function Test Statistic                     413.264
Degrees of freedom                                   15
P-value                                              0.000

User model versus baseline model:

Comparative Fit Index (CFI)                         0.948
Tucker-Lewis Index (TLI)                           0.888

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)                       -2002.640
Loglikelihood unrestricted model (H1)               -1988.719

Number of free parameters                            20
Akaike (AIC)                                         4045.280
Bayesian (BIC)                                       4096.567
Sample-size adjusted Bayesian (BIC)                 4033.418

Root Mean Square Error of Approximation:

RMSEA                                                0.176
90 Percent Confidence Interval                     0.111 0.247
P-value RMSEA <= 0.05                               0.002

Standardized Root Mean Square Residual:

```

SRMR

0.046

## Parameter Estimates:

Information  
Standard ErrorsExpected  
Robust.sem

## Latent Variables:

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
A =~							
Consu_F	(y1)	1.000				29.102	0.870
Consu_D	(y2)	1.655	NA			48.165	0.889
Envoi_M	(y3)	0.466	NA			13.557	0.867
E =~							
COLLAB	(11)	1.000				4.539	0.889
COMM	(12)	0.717	NA			3.255	0.627
MOTIV	(13)	0.308	NA			1.399	0.612
C =~							
A	(r2)	1.000				0.242	0.242

## Regressions:

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
A ~							
E	(r1)	7.106	NA			1.108	1.108

## Covariances:

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
E ~~							
C		-17.097	NA			-0.535	-0.535

## Intercepts:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
Consu_F	36.958	NA			36.958	1.104
Consu_D	50.083	NA			50.083	0.925
Envoi_M	13.844	NA			13.844	0.886
COLLAB	3.750	NA			3.750	0.735
COMM	2.510	NA			2.510	0.484
MOTIV	2.198	NA			2.198	0.962
A	0.000				0.000	0.000
E	0.000				0.000	0.000
C	0.000				0.000	0.000

## Variances:

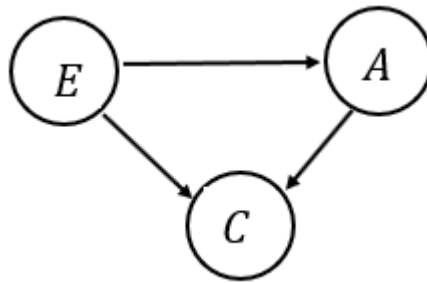
	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
Consu_F	273.009	NA			273.009	0.244
Consu_D	612.393	NA			612.393	0.209
Envoi_M	60.578	NA			60.578	0.248
COLLAB	5.440	NA			5.440	0.209
COMM	16.344	NA			16.344	0.607
MOTIV	3.263	NA			3.263	0.625
A	0.000				0.000	0.000
E	20.602	NA			1.000	1.000
C	49.652	NA			1.000	1.000

Des erreurs sur l'estimation du modèle ont été rencontrées. L'estimateur n'a pas pu évaluer l'écart-type (mis en exergue sur le modèle) suite à des variances et/ou covariances négatives.

### 5.2.3. 3<sup>e</sup> cas de modèle

Maintenant, on va fusionner les deux premiers cas. L'attitude est une régression de l'engagement, et la variable latente  $C$  du résultat est mesurée simultanément par les variables attitude et engagement.

Le modèle est représenté par la figure ci-dessous :



**Figure 5.04.**Modèle de structure, proposition 3

#### Script et résultats pour l'estimation du modèle

```
> library(lavaan)

> #Chargement de données
> gsa <- read.csv("F:/Doctorat_2016/SCA/SCA/Thèses/D3/S5/sca/donnees.csv", sep="
;")

> donnee <- gsa[1:6]

> resultat <- gsa[7]

> colnames(donnee) <- c("COLLAB", "COMM", "MOTIV", "Consu_F", "Consu_
D", "Envoi_M")

> #Création du modele
> donnee.model <- '
+ A =~ y1*Consu_F + y2*Consu_D + y3*Envoi_M
+ E =~ l1*COLLAB + l2*COMM + l3*MOTIV
+ A =~ r1*E
+ C =~ r2*E+r3*A

> #Estimation du modele
```

```
> donnee.fit <- cfa(model=donnee.model, data=donnee, se="robust")
Error in lav_data_full(data = data, group = group, group.label = group.label,
lavaan ERROR: missing observed variables in dataset: r3A
```

```
> summary(donnee.fit, fit.measures= TRUE, standardized= TRUE)
lavaan (0.5-20) converged normally after 59 iterations
```

Number of observations	96
Estimator	ML
Minimum Function Test Statistic	27.842
Degrees of freedom	6
P-value (Chi-square)	0.000

Model test baseline model:

Minimum Function Test Statistic	413.264
Degrees of freedom	15
P-value	0.000

User model versus baseline model:

Comparative Fit Index (CFI)	0.945
Tucker-Lewis Index (TLI)	0.863

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-2002.640
Loglikelihood unrestricted model (H1)	-1988.719
Number of free parameters	21
Akaike (AIC)	4047.280
Bayesian (BIC)	4101.131
Sample-size adjusted Bayesian (BIC)	4034.825

Root Mean Square Error of Approximation:

RMSEA	0.195
90 Percent Confidence Interval	0.125 0.271
P-value RMSEA <= 0.05	0.001

Standardized Root Mean Square Residual:

SRMR	0.046
------	-------

Parameter Estimates:

Information	Expected
Standard Errors	Robust.sem

Latent Variables:

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
A =~							
	Consu_F (y1)	1.000				29.102	0.870
	Consu_D (y2)	1.655	NA			48.165	0.889
	Envoi_M (y3)	0.466	NA			13.557	0.867
E =~							
	COLLAB (11)	1.000				4.539	0.889

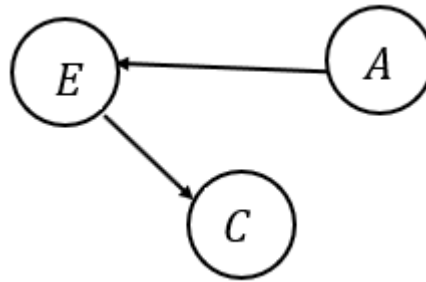
COMM	(12)	0.717	NA		3.255	0.627
MOTIV	(13)	0.308	NA		1.399	0.612
C =~						
E	(r2)	1.000			0.791	0.791
A	(r3)	2.704	NA		0.334	0.334
Regressions:						
		Estimate	Std.Err	Z-value	P(> z )	Std.lv Std.all
A ~						
E	(r1)	4.584	NA			0.715 0.715
Intercepts:						
		Estimate	Std.Err	Z-value	P(> z )	Std.lv Std.all
Consu_F		36.958	NA			36.958 1.104
Consu_D		50.083	NA			50.083 0.925
Envoi_M		13.844	NA			13.844 0.886
COLLAB		3.750	NA			3.750 0.735
COMM		2.510	NA			2.510 0.484
MOTIV		2.198	NA			2.198 0.962
A		0.000				0.000 0.000
E		0.000				0.000 0.000
C		0.000				0.000 0.000
Variances:						
		Estimate	Std.Err	Z-value	P(> z )	Std.lv Std.all
Consu_F		273.009	NA			273.009 0.244
Consu_D		612.393	NA			612.393 0.209
Envoi_M		60.578	NA			60.578 0.248
COLLAB		5.440	NA			5.440 0.209
COMM		16.344	NA			16.344 0.607
MOTIV		3.263	NA			3.263 0.625
A		0.195	NA			0.000 0.000
E		7.709	NA			0.374 0.374
C		12.893	NA			1.000 1.000

Des erreurs sur l'estimation du modèle ont été rencontrées. L'estimateur n'a pas pu évaluer l'écart-type (mis en exergue sur le modèle) suite à des variances et/ou covariances négatives.

#### 5.2.4. 4<sup>e</sup> cas de modèle

Quatrième cas : On va prendre l'engagement comme regression linéaire de l'attitude et la variable latente (C) du résultat est mesurée par l'engagement *E*.

Le modèle est représenté par la Figure 5.05



**Figure 5.05.** *Modèle de structure, proposition 4*

### Script et résultats pour l'estimation du modèle

```

> library(lavaan)

> #Chargement de données
> gsa <- read.csv("F:/Doctorat_2016/SCA/SCA/Thèses/D3/S5/sca/donnees.csv", sep="
;")

> donnee <- gsa[1:6]

> resultat <- gsa[7]

> colnames(donnee) <- c("COLLAB", "COMM", "MOTIV", "Consu_F", "Consu_
D", "Envoi_M")

> #Création du modele
> donnee.model <- '
+ A =~ y1*Consu_F + y2*Consu_D + y3*Envoi_M
+ E =~ l1*COLLAB + l2*COMM + l3*MOTIV
+ E =~ r1*A
+ C =~ r2*E
+ '

> #Estimation du modele
> donnee.fit <- cfa(model=donnee.model, data=donnee, se="robust")
Warning messages:
1: In lav_data_full(data = data, group = group, group.label = group.label, :
lavaan WARNING: some observed variances are (at least) a factor 1000 times la
rger than others; use varTable(fit) to investigate
2: In lavaan::lavaan(model = donnee.model, data = donnee, se = "robust", :
lavaan WARNING: model has NOT converged!

> summary(donnee.fit, fit.measures= TRUE, standardized= TRUE)
lavaan (0.5-20) converged normally after 104 iterations

Number of observations                96

Estimator                            ML
Minimum Function Test Statistic      27.842
Degrees of freedom                    7

```

P-value (Chi-square)	0.000
----------------------	-------

Model test baseline model:

Minimum Function Test Statistic	413.264
Degrees of freedom	15
P-value	0.000

User model versus baseline model:

Comparative Fit Index (CFI)	0.948
Tucker-Lewis Index (TLI)	0.888

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-2002.640
Loglikelihood unrestricted model (H1)	-1988.719
Number of free parameters	20
Akaike (AIC)	4045.280
Bayesian (BIC)	4096.567
Sample-size adjusted Bayesian (BIC)	4033.418

Root Mean Square Error of Approximation:

RMSEA	0.176
90 Percent Confidence Interval	0.111 0.247
P-value RMSEA <= 0.05	0.002

Standardized Root Mean Square Residual:

SRMR	0.046
------	-------

Parameter Estimates:

Information	Expected
Standard Errors	Robust.sem

Latent Variables:

	Estimate	Std.Err	Z-value	P(> z )
A =~				
Consu_F (y1)	1.000			
Consu_D (y2)	1.655	NA		
Envoi_M (y3)	0.466	NA		
E =~				
COLLAB (l1)	1.000			
COMM (l2)	0.717	NA		
MOTIV (l3)	0.308	NA		
C =~				
E (r2)	1.000			
Std.lv	Std.all			
	29.102	0.870		
	48.165	0.889		
	13.557	0.867		
	4.539	0.889		
	3.255	0.627		
	1.399	0.612		



0.205 0.205

Regressions:

		Estimate	Std.Err	Z-value	P(> z )
E ~					
A	(r1)	0.153	NA		
Std.lv	Std.all				
0.980	0.980				

Covariances:

		Estimate	Std.Err	Z-value	P(> z )
A ~~					
C		-0.144	NA		
Std.lv	Std.all				
-0.005	-0.005				

Intercepts:

		Estimate	Std.Err	Z-value	P(> z )
Consu_F		36.958	NA		
Consu_D		50.083	NA		
Envoi_M		13.844	NA		
COLLAB		3.750	NA		
COMM		2.510	NA		
MOTIV		2.198	NA		
A		0.000			
E		0.000			
C		0.000			
Std.lv	Std.all				
36.958	1.104				
50.083	0.925				
13.844	0.886				
3.750	0.735				
2.510	0.484				
2.198	0.962				
0.000	0.000				
0.000	0.000				
0.000	0.000				

Variances:

		Estimate	Std.Err	Z-value	P(> z )
Consu_F		273.009	NA		
Consu_D		612.393	NA		
Envoi_M		60.578	NA		
COLLAB		5.440	NA		
COMM		16.344	NA		
MOTIV		3.263	NA		
A		846.906	NA		
E		0.000			
C		0.863	NA		
Std.lv	Std.all				
273.009	0.244				
612.393	0.209				
60.578	0.248				
5.440	0.209				
16.344	0.607				
3.263	0.625				

```

1.000    1.000
0.000    0.000
1.000    1.000

```

```

Warning message:
In .local(object, ...) :
  lavaan WARNING: fit measures not available if model did not converge

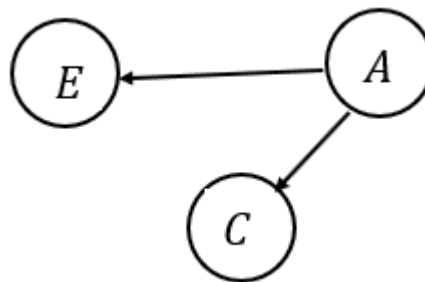
```

Des erreurs sur l'estimation du modèle ont été rencontrées. L'estimateur n'a pas pu évaluer l'écart-type (mis en exergue sur le modèle) suite à des variances et/ou covariances négatives.

### 5.2.5. 5<sup>e</sup> cas de modèle

Cette fois ci, on va prendre le résultat comme mesuré à partir de l'attitude de l'apprenant, tout en gardant l'engagement comme regression de l'attitude.

Ce modèle est représenté par la Figure 5.06



**Figure 5.06.**Modèle de structure, proposition 5

#### Script et résultats pour l'estimation du modèle :

```

> library(lavaan)

> #Chargement de données
> gsa <- read.csv("F:/Doctorat_2016/SCA/SCA/Thèses/D3/S5/sca/donnees.csv", sep="
;")

> donnee <- gsa[1:6]

> resultat <- gsa[7]

> colnames(donnee) <- c("COLLAB", "COMM", "MOTIV", "Consu_F", "Consu_
D", "Envoi_M")

> #Création du modele

```

```

> donnee.model <- '
+ A =~ y1*Consu_F + y2*Consu_D + y3*Envoi_M
+ E =~ l1*COLLAB + l2*COMM + l3*MOTIV
+ E =~ r1*A
+ C =~ r2*A
+ '

> #Estimation du modele
> donnee.fit <- cfa(model=donnee.model, data=donnee, se="robust")

> summary(donnee.fit, fit.measures= TRUE, standardized= TRUE)
lavaan (0.5-20) converged normally after 114 iterations

    Number of observations                    96

    Estimator                                ML
    Minimum Function Test Statistic          27.842
    Degrees of freedom                       8
    P-value (Chi-square)                     0.001

Model test baseline model:

    Minimum Function Test Statistic          413.264
    Degrees of freedom                       15
    P-value                                  0.000

User model versus baseline model:

    Comparative Fit Index (CFI)              0.950
    Tucker-Lewis Index (TLI)                0.907

Loglikelihood and Information Criteria:

    Loglikelihood user model (H0)            -2002.640
    Loglikelihood unrestricted model (H1)     -1988.719

    Number of free parameters                19
    Akaike (AIC)                             4043.280
    Bayesian (BIC)                           4092.003
    Sample-size adjusted Bayesian (BIC)      4032.012

Root Mean Square Error of Approximation:

    RMSEA                                    0.161
    90 Percent Confidence Interval           0.098 0.228
    P-value RMSEA <= 0.05                   0.003

Standardized Root Mean Square Residual:

    SRMR                                    0.046

Parameter Estimates:

    Information                                Expected
    Standard Errors                          Robust.sem

Latent Variables:
      Estimate Std.Err Z-value P(>|z|)
A =~

```

Consu_F	(y1)	1.000			
Consu_D	(y2)	1.655	0.142	11.660	0.000
Envoi_M	(y3)	0.466	0.059	7.902	0.000
E =~					
COLLAB	(11)	1.000			
COMM	(12)	0.717	0.245	2.928	0.003
MOTIV	(13)	0.308	0.097	3.169	0.002
C =~					
A	(r2)	1.000			
Std.lv	Std.all				
29.102	0.870				
48.165	0.889				
13.557	0.867				

4.539	0.889
3.255	0.627
1.399	0.612

1.000	1.000
-------	-------

#### Regressions:

		Estimate	Std.Err	Z-value	P(> z )
E ~					
A	(r1)	0.153	0.015	9.861	0.000
Std.lv	Std.all				
0.979	0.979				

#### Intercepts:

	Estimate	Std.Err	Z-value	P(> z )
Consu_F	36.958	3.433	10.764	0.000
Consu_D	50.083	5.556	9.015	0.000
Envoi_M	13.844	1.604	8.631	0.000
COLLAB	3.750	0.524	7.162	0.000
COMM	2.510	0.532	4.714	0.000
MOTIV	2.198	0.234	9.375	0.000
A	0.000			
E	0.000			
C	0.000			
Std.lv	Std.all			
36.958	1.104			
50.083	0.925			
13.844	0.886			
3.750	0.735			
2.510	0.484			
2.198	0.962			
0.000	0.000			
0.000	0.000			
0.000	0.000			

#### Variances:

	Estimate	Std.Err	Z-value	P(> z )
Consu_F	273.009	110.834	2.463	0.014
Consu_D	612.393	324.960	1.885	0.059
Envoi_M	60.577	14.363	4.218	0.000
COLLAB	5.440	1.568	3.469	0.001
COMM	16.344	5.768	2.833	0.005
MOTIV	3.263	0.859	3.797	0.000

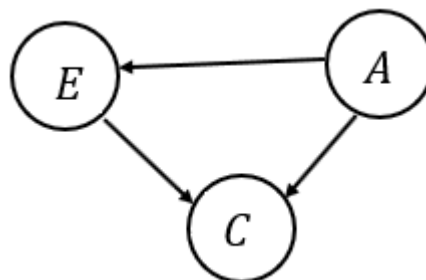
A		0.000			
E		0.863	1.874	0.460	0.645
C		846.905	201.982	4.193	0.000
Std.lv	Std.all				
273.009	0.244				
612.393	0.209				
60.577	0.248				
5.440	0.209				
16.344	0.607				
3.263	0.625				
0.000	0.000				
0.042	0.042				
1.000	1.000				

D'après ces résultats, on peut affirmer que l'estimation du modèle a été effectuée avec succès. Toutes les variables ont été bien estimées selon les valeurs mises en exergue dans le résultat ci-dessus.

### 5.2.6. 6<sup>e</sup> cas de modèle

Dans ce dernier cas, avec un engagement comme régression de l'attitude, on va mesurer la variable latente du résultat *C* à partir des deux variables de l'engagement *E* et de l'attitude *A*.

Ce modèle est représenté par la Figure 5.07.



**Figure 5.07.**Modèle de structure, proposition 6

#### Script et résultats pour l'estimation du modèle :

```

> library(lavaan)

> #Chargement de données
> gsa <- read.csv("F:/Doctorat_2016/SCA/SCA/Thèses/D3/S5/sca/donnees.csv", sep="
;")

> donnee <- gsa[1:6]

> resultat <- gsa[7]

```

```

> colnames(donnee) <- c("COLLAB", "COMM", "MOTIV", "Consu_F", "Consu_
D", "Envoi_M")

> #Création du modele
> donnee.model <- '
+ A =~ y1*Consu_F + y2*Consu_D + y3*Envoi_M
+ E =~ l1*COLLAB + l2*COMM + l3*MOTIV
+ E =~ r1*A
+ C =~ r2*E+r3*A

> #Estimation du modele
> donnee.fit <- cfa(model=donnee.model, data=donnee, se="robust")
Error in lav_data_full(data = data, group = group, group.label = group.label,
:
lavaan ERROR: missing observed variables in dataset: r3A

> summary(donnee.fit, fit.measures= TRUE, standardized= TRUE)
lavaan (0.5-20) converged normally after 58 iterations

Number of observations                                96

Estimator                                             ML
Minimum Function Test Statistic                     27.842
Degrees of freedom                                   6
P-value (Chi-square)                                0.000

Model test baseline model:

Minimum Function Test Statistic                     413.264
Degrees of freedom                                   15
P-value                                              0.000

User model versus baseline model:

Comparative Fit Index (CFI)                         0.945
Tucker-Lewis Index (TLI)                           0.863

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)                       -2002.640
Loglikelihood unrestricted model (H1)               -1988.719

Number of free parameters                            21
Akaike (AIC)                                         4047.280
Bayesian (BIC)                                       4101.131
Sample-size adjusted Bayesian (BIC)                 4034.825

Root Mean Square Error of Approximation:

RMSEA                                                0.195
90 Percent Confidence Interval                      0.125 0.271
P-value RMSEA <= 0.05                               0.001

Standardized Root Mean Square Residual:

SRMR                                                0.046

Parameter Estimates:

```

Information				Expected
Standard Errors				Robust.sem
Latent Variables:				
		Estimate	Std.Err	Z-value P(> z )
A =~				
Consu_F	(y1)	1.000		
Consu_D	(y2)	1.655	NA	
Envoi_M	(y3)	0.466	NA	
E =~				
COLLAB	(l1)	1.000		
COMM	(l2)	0.717	NA	
MOTIV	(l3)	0.308	NA	
C =~				
E	(r2)	1.000		
A	(r3)	10.340	NA	
Std.lv	Std.all			
29.102	0.870			
48.165	0.889			
13.557	0.867			
4.539	0.889			
3.255	0.627			
1.399	0.612			
0.620	0.620			
0.999	0.999			
Regressions:				
		Estimate	Std.Err	Z-value P(> z )
E ~				
A	(r1)	0.056	NA	
Std.lv	Std.all			
0.359	0.359			
Intercepts:				
		Estimate	Std.Err	Z-value P(> z )
Consu_F		36.958	NA	
Consu_D		50.083	NA	
Envoi_M		13.844	NA	
COLLAB		3.750	NA	
COMM		2.510	NA	
MOTIV		2.198	NA	
A		0.000		
E		0.000		
C		0.000		
Std.lv	Std.all			
36.958	1.104			
50.083	0.925			
13.844	0.886			
3.750	0.735			
2.510	0.484			
2.198	0.962			
0.000	0.000			
0.000	0.000			
0.000	0.000			

Variances:

	Estimate	Std.Err	Z-value	P(> z )
Consu_F	273.009	NA		
Consu_D	612.393	NA		
Envoi_M	60.578	NA		
COLLAB	5.440	NA		
COMM	16.344	NA		
MOTIV	3.263	NA		
A	0.986	NA		
E	0.853	NA		
C	7.912	NA		
Std.lv	Std.all			
273.009	0.244			
612.393	0.209			
60.578	0.248			
5.440	0.209			
16.344	0.607			
3.263	0.625			
0.001	0.001			
0.041	0.041			
1.000	1.000			

Des erreurs sur l'estimation du modèle ont été rencontrées. L'estimateur n'a pas pu évaluer l'écart-type (mis en exergue sur le modèle) suite à des variances et/ou covariances négatives.

### 5.2.7. Modèle identifié

D'après les six résultats obtenus, nous avons constaté que deux cas sont possibles, le 1<sup>er</sup> et le 5<sup>e</sup>. On n'a pas trouvé d'erreur lors de l'estimation de ces deux modèles.

On peut dire que ces deux cas sont similaires. En effet, si on dit que l'attitude est une régression linéaire de l'engagement (cas 1), réciproquement, l'engagement est aussi une régression linéaire de l'attitude (cas 5).

Selon la logique d'apprentissage, le comportement de l'apprenant est déterminé par rapport à son engagement motivationnel : initiative d'exprimer, volonté d'agir, question cognitive, etc. et sa façon de collaborer : organisation, partage, participation, animation, etc. dans un groupe. C'est la raison pour laquelle nous avons retenu le cas 1 comme modèle préconisé de la recherche.

## 5.3 Le modèle de structure avec les paramètres estimés

En prenant le modèle du cas 1, le résultat de l'estimation du modèle est donné comme suit :



```
> summary(donnee.fit, fit.measures= TRUE, standardized= TRUE)
lavaan (0.5-20) converged normally after 71 iterations
```

Number of observations	96
Estimator	ML
Minimum Function Test Statistic	27.842
Degrees of freedom	8
P-value (Chi-square)	0.001

Model test baseline model:

Minimum Function Test Statistic	413.264
Degrees of freedom	15
P-value	0.000

User model versus baseline model:

Comparative Fit Index (CFI)	0.950
Tucker-Lewis Index (TLI)	0.907

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-2002.640
Loglikelihood unrestricted model (H1)	-1988.719
Number of free parameters	19
Akaike (AIC)	4043.280
Bayesian (BIC)	4092.003
Sample-size adjusted Bayesian (BIC)	4032.012

Root Mean Square Error of Approximation:

RMSEA	0.161
90 Percent Confidence Interval	0.098 0.228
P-value RMSEA <= 0.05	0.003

Standardized Root Mean Square Residual:

SRMR	0.046
------	-------

Parameter Estimates:

Information	Expected
Standard Errors	Robust.sem

Latent Variables:

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
A =~							
	Consu_F (y1)	1.000				29.102	0.870
	Consu_D (y2)	1.655	0.142	11.660	0.000	48.165	0.889
	Envoi_M (y3)	0.466	0.059	7.902	0.000	13.557	0.867
E =~							
	COLLAB (l1)	1.000				4.539	0.889
	COMM (l2)	0.717	0.245	2.928	0.003	3.255	0.627
	MOTIV (l3)	0.308	0.097	3.169	0.002	1.399	0.612
C =~							
	E (r2)	1.000				1.000	1.000

Regressions:

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
A ~							
E	(r1)	6.276	0.677	9.273	0.000	0.979	0.979

Intercepts:

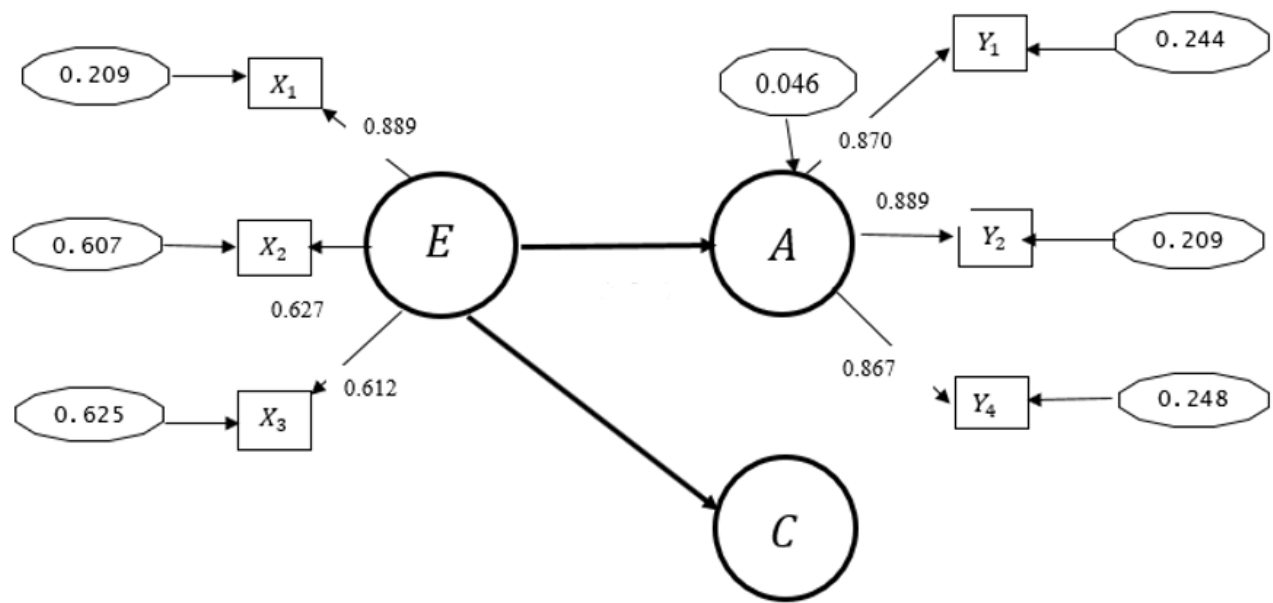
	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
Consu_F	36.958	3.433	10.764	0.000	36.958	1.104
Consu_D	50.083	5.556	9.015	0.000	50.083	0.925
Envoi_M	13.844	1.604	8.631	0.000	13.844	0.886
COLLAB	3.750	0.524	7.162	0.000	3.750	0.735
COMM	2.510	0.532	4.714	0.000	2.510	0.484
MOTIV	2.198	0.234	9.375	0.000	2.198	0.962
A	0.000				0.000	0.000
E	0.000				0.000	0.000
C	0.000				0.000	0.000

Variances:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
Consu_F	273.009	110.834	2.463	0.014	273.009	0.244
Consu_D	612.393	324.960	1.885	0.059	612.393	0.209
Envoi_M	60.578	14.363	4.218	0.000	60.578	0.248
COLLAB	5.440	1.568	3.469	0.001	5.440	0.209
COMM	16.344	5.768	2.833	0.005	16.344	0.607
MOTIV	3.263	0.859	3.797	0.000	3.263	0.625
A	35.463	76.186	0.465	0.642	0.042	0.042
E	0.000				0.000	0.000
C	20.602	6.535	3.152	0.002	1.000	1.000

Suite au résultat de traitement obtenu dans 5.2.1., le modèle de structure ci-dessous a été obtenu. On a remarqué que le modèle converge après 71 itérations. Le modèle a huit degrés de liberté, donc il est sur-identifié et la condition d'ordre est vérifiée.

Le modèle complet est représenté par la Figure 5.08 où on va afficher les variances des variables latentes mesurées.



**Figure 5.08.** *Modèle de structure de la recherche*

#### Indices d'ajustement absolu :

- Test du chi-carré (Seuil : Valeurs proches de zero) :  $P\text{-value} = 0$ , alors le modèle est bien ajusté selon cet indice.
- Erreur quadratique moyenne d'approximation (seuil : Moins de 0.06) :  $RMSEA = 0.0161$ , alors le modèle est moins ajusté selon cet indice.
- Résidu moyen carré normalisé (Seuil : Moins de 0.08) :  $SRMR = 0.046$ , alors le modèle est bien ajusté selon cet indice.

#### Indices d'ajustement relatif :

- Indice d'ajustement non normé de Tucker-Lewis (Seuil : Plus de 0.90) :  $TLI = 0.907$ , alors le modèle est bien ajusté selon cet indice.
- Indice d'ajustement comparatif (Seuil : Plus de 0.95) :  $CFI = 0.95$ , alors le modèle est bien ajusté selon cet indice.

D'après ces indicateurs, on peut retenir le modèle.

De tout ce qui précède on peut affirmer que l'engagement de l'apprenant a un impact direct sur l'acquisition de connaissance par l'apprenant.

## 5.4 Estimation du résultat à partir de la variable latente C

#### 5.4.1. Valeurs des variables latentes $A$ , $E$ et $C$

Trois variables latentes ont été mises en jeu dans ce modèle de structure. Le script ci-dessous permet de sortir les valeurs de A, E et C.

La variable latente C a une relation avec le résultat pédagogique de la formation. A Noter que « Rm » est le résultat à partir du modèle de la recherche et « Rr » le résultat pédagogique de la formation. Le résultat Rr vaut 1 lorsque l'apprenant est admis et est égal à 0 si l'apprenant redouble ou est renvoyé (non admis).

Le résultat montre les valeurs des variables latentes A, E et C en fonction des valeurs des variables observées CONSU\_F, CONSU\_D, ENVOI\_M, COLLAB, COMM et MOTIV (enregistrés dans chaque ligne de données).

Script et résultat sur les résultats du modèle :

```
> #2. Prediction
> predict(donnee.fit)
```

	A	E	C								
[1,]	-28.125	-4.264	-4.264	[18,]	16.731	2.347	2.347	[36,]	-15.279	-2.446	-2.446
[2,]	-19.769	-2.881	-2.881	[19,]	-24.002	-3.677	-3.677	[37,]	-17.272	-2.895	-2.895
[3,]	-13.783	-2.064	-2.064	[20,]	27.890	3.068	3.068	[38,]	-16.621	-2.484	-2.484
[4,]	-4.140	-0.487	-0.487	[21,]	39.883	6.194	6.194	[39,]	4.458	0.801	0.801
[5,]	-3.680	-0.727	-0.727	[22,]	-3.603	-0.775	-0.775	[40,]	2.251	0.812	0.812
[6,]	-17.787	-2.500	-2.500	[23,]	41.521	6.846	6.846	[41,]	-15.630	-2.391	-2.391
[7,]	0.210	-0.168	-0.168	[24,]	-0.226	0.238	0.238	[42,]	-1.948	-0.412	-0.412
[8,]	-18.600	-3.063	-3.063	[25,]	-13.416	-2.207	-2.207	[43,]	48.037	7.688	7.688
[9,]	4.553	0.779	0.779	[26,]	1.860	0.437	0.437	[44,]	-13.209	-2.127	-2.127
[10,]	-22.740	-3.386	-3.386	[27,]	0.633	-0.047	-0.047	[45,]	-11.093	-1.625	-1.625
[11,]	-23.937	-3.668	-3.668	[28,]	1.824	0.235	0.235	[46,]	15.830	2.422	2.422
[12,]	-26.275	-3.963	-3.963	[29,]	40.394	6.022	6.022	[47,]	-26.353	-3.973	-3.973
[13,]	-9.019	-1.495	-1.495	[30,]	-9.972	-1.908	-1.908	[48,]	-13.587	-1.899	-1.899
[14,]	-19.324	-3.055	-3.055	[31,]	108.158	16.410	16.410	[49,]	75.051	11.104	11.104
[15,]	-22.558	-3.427	-3.427	[32,]	-26.054	-4.003	-4.003	[50,]	22.352	3.611	3.611
[16,]	-12.169	-1.724	-1.724	[33,]	-9.779	-1.681	-1.681	[51,]	-16.368	-2.452	-2.452
[17,]	-27.917	-4.237	-4.237	[34,]	-25.251	-3.834	-3.834	[52,]	-13.096	-2.139	-2.139
				[35,]	-25.232	-3.899	-3.899	[53,]	-19.826	-2.920	-2.920

[54,]	2.712	0.670	0.670	[69,]	12.071	1.690	1.690	[84,]	-8.090	-1.021	-1.021
[55,]	-11.391	-1.432	-1.432	[70,]	36.412	4.998	4.998	[85,]	17.889	3.136	3.136
[56,]	46.751	7.675	7.675	[71,]	-27.600	-4.197	-4.197	[86,]	9.277	1.749	1.749
[57,]	-15.504	-2.272	-2.272	[72,]	-7.430	-1.362	-1.362	[87,]	26.585	5.238	5.238
[58,]	-5.861	-1.196	-1.196	[73,]	-7.699	-1.163	-1.163	[88,]	80.881	13.242	13.242
[59,]	13.287	1.619	1.619	[74,]	-10.619	-1.733	-1.733	[89,]	24.977	3.864	3.864
[60,]	-10.516	-1.354	-1.354	[75,]	-16.848	-2.644	-2.644	[90,]	-11.455	-1.639	-1.639
[61,]	0.931	-0.009	-0.009	[76,]	-2.493	-0.610	-0.610	[91,]	-20.802	-3.242	-3.242
[62,]	-20.218	-2.870	-2.870	[77,]	-25.021	-3.738	-3.738	[92,]	53.235	8.683	8.683
[63,]	9.303	1.406	1.406	[78,]	-25.071	-3.811	-3.811	[93,]	21.525	1.996	1.996
[64,]	-13.807	-2.165	-2.165	[79,]	12.862	1.951	1.951	[94,]	-7.963	-1.722	-1.722
[65,]	-26.824	-4.032	-4.032	[80,]	114.753	18.093	18.093	[95,]	-27.743	-4.215	-4.215
[66,]	-24.009	-3.677	-3.677	[81,]	-18.745	-2.851	-2.851	[96,]	-28.278	-4.283	-4.283
[67,]	8.767	0.980	0.980	[82,]	28.847	4.707	4.707				
[68,]	-18.171	-2.748	-2.748	[83,]	7.096	0.177	0.177				

**Tableau 5.01 :** *Les trois variables latentes de la recherche*

Nous avons exporté ces données sur Excel afin d'effectuer une analyse sur un tableur.

A titre de remarque, il est constaté que les variables latentes *E* et *C* sont identiques. D'après les valeurs des paramètres obtenues lors de l'estimation du modèle, le coefficient de régression entre *E* et *C* est égal à 1.

C	=~				
E	(r <sup>2</sup> )	1.000		1.000	1.000

A partir de cela, la proposition que nous avons avancée au paragraphe 5.3 selon laquelle *l'engagement a un impact direct sur l'acquisition de connaissance de l'apprenant* (son résultat) est corroborée. Donc, l'hypothèse *H3* de la recherche mentionnée dans le paragraphe 4.3 est confirmée.

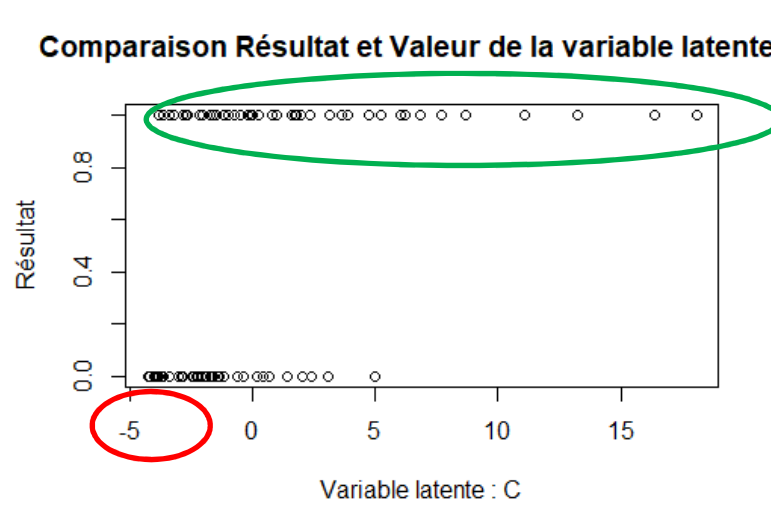
#### **5.4.2. Relation entre le résultat *Rr* et le résultat du modèle *Rm***

On va chercher une relation entre « *Rm* » le résultat à partir du modèle de la recherche (qui n'est autre que la variable latente *C* et « *Rr* » le résultat pédagogique de la formation.

Le graphe des points indiquant la valeur du résultat réel de l'apprenant (0 ou 1) en fonction de la variable latente *C* est représenté par la Figure 5.09.

### Script :

```
#Courbe résultat en fonction de la variable latente C  
> plot(xC,y,xlab="Variable latente : C",ylab="Résultat",main="Comparaison Résultat et Valeur de la variable latente C")
```



**Figure 5.09.** Comparaison du résultat par rapport à la variable latente  $C$

On constate une concentration de points à résultat égal à 0 autour de la zone encadrée en rouge sur la Figure 5.09 et une autre concentration de points à résultat égal à 1 autour de la zone encadrée en vert. On peut donc dire que les apprenants présentant une faible valeur de  $C$  ont un résultat 0 et ceux ayant une grande valeur de  $C$  ont un résultat 1. Mais une question se pose : quel est le seuil à partir duquel on peut décider que le résultat est 1 ou 0 ?

On pose une valeur seuil  $C_{seuil}$  de la variable latente  $C$  pour décider si un apprenant est admis ou non. En fonction de ce seuil quelconque, on va dresser le Tableau 5.02 : suivant pour déterminer une comparaison des résultats réels à ceux du modèle. On notera « VRAI » si le résultat réel de l'apprenant est identique au résultat prédit par notre modèle (que ce soit admis ou pas) et par « FAUX » dans le cas contraire.

Valeur seuil de C	-0,9	-1,0	-1,1	-1,2	-1,3	-1,4	-1,5
VRAI	72	72	73	73	73	71	71
FAUX	24	24	23	23	23	25	25
Nombre d'apprenants admis selon le "Modèle"	42	42	43	45	45	47	49
Nombre d'apprenants admis réellement	46	46	46	46	46	46	46

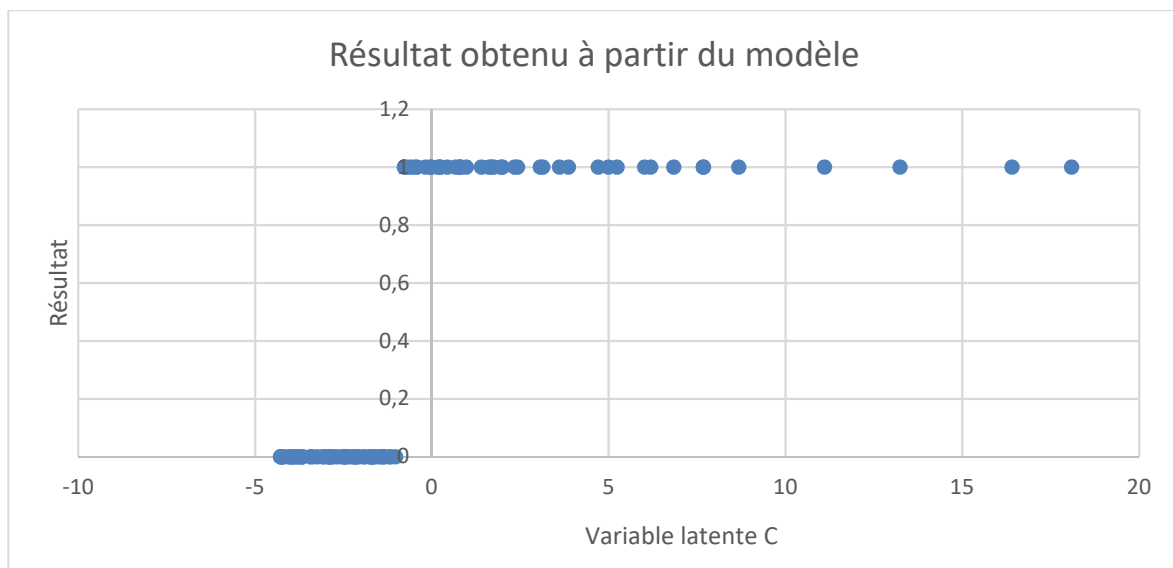
**Tableau 5.02 :** *Valeur seuil de C*

D'après le Tableau 5.02 : on peut déduire un maximum de vraisemblance entre les résultats du modèle et les résultats réels pour une valeur seuil allant de  $C_{seuil} = -1.1$  à  $-1.3$ . On peut déterminer la valeur exacte du seuil à adopter à partir de la distribution des valeurs de ces deux types de résultats dans le paragraphe 5.4.3.

Dans le cas  $C_{seuil} = -1.2$  ou  $C_{seuil} = -1.3$ , on observe aussi une similarité entre le nombre d'apprenants admis selon le modèle, et le nombre réel.

#### ***5.4.3 Résultat pédagogique par rapport à la Cognition selon le modèle***

En prenant  $C_{seuil} = -1.3$ , le graphe de la Figure 5.10 représente les décisions sur le résultat de l'apprenant en fonction de la « valeur de la variable latente C ».



**Figure 5.10.***Traitement sur Excel*

Nous avons constaté selon le graphe de la Figure 5.10 qu'à partir de la valeur «  $C_{seuil}$  » seuil de la cognition, le modèle donne le résultat 1, autrement dit, dès que cette valeur seuil de la cognition de l'apprenant est atteinte, on peut dire qu'il a réussi la formation.

N°	A	E	C	Rr	Rm	Comp
1	-28,125	-4,264	-4,264	0	0	VRAI
2	-19,769	-2,881	-2,881	0	0	VRAI
3	-13,783	-2,064	-2,064	1	0	FAUX
4	-4,14	-0,487	-0,487	1	1	VRAI
5	-3,68	-0,727	-0,727	1	1	VRAI
6	-17,787	-2,5	-2,5	0	0	VRAI
7	0,21	-0,168	-0,168	1	1	VRAI
8	-18,6	-3,063	-3,063	0	0	VRAI
9	4,553	0,779	0,779	1	1	VRAI
10	-22,74	-3,386	-3,386	0	0	VRAI
11	-23,937	-3,668	-3,668	1	0	FAUX
12	-26,275	-3,963	-3,963	0	0	VRAI
13	-9,019	-1,495	-1,495	1	0	FAUX
14	-19,324	-3,055	-3,055	0	0	VRAI
15	-22,558	-3,427	-3,427	1	0	FAUX
16	-12,169	-1,724	-1,724	0	0	VRAI
17	-27,917	-4,237	-4,237	0	0	VRAI
18	16,731	2,347	2,347	1	1	VRAI
19	-24,002	-3,677	-3,677	0	0	VRAI

20	27,89	3,068	3,068	0	1	FAUX
21	39,883	6,194	6,194	1	1	VRAI
22	-3,603	-0,775	-0,775	1	1	VRAI
23	41,521	6,846	6,846	1	1	VRAI
24	-0,226	0,238	0,238	1	1	VRAI
25	-13,416	-2,207	-2,207	0	0	VRAI
26	1,86	0,437	0,437	0	1	FAUX
27	0,633	-0,047	-0,047	1	1	VRAI
28	1,824	0,235	0,235	1	1	VRAI
29	40,394	6,022	6,022	1	1	VRAI
30	-9,972	-1,908	-1,908	0	0	VRAI
31	108,158	16,41	16,41	1	1	VRAI
32	-26,054	-4,003	-4,003	0	0	VRAI
33	-9,779	-1,681	-1,681	0	0	VRAI
34	-25,251	-3,834	-3,834	1	0	FAUX
35	-25,232	-3,899	-3,899	0	0	VRAI
36	-15,279	-2,446	-2,446	0	0	VRAI
37	-17,272	-2,895	-2,895	0	0	VRAI
38	-16,621	-2,484	-2,484	0	0	VRAI
39	4,458	0,801	0,801	1	1	VRAI



40	2,251	0,812	0,812	1	1	VRAI
41	-15,63	-2,391	-2,391	0	0	VRAI
42	-1,948	-0,412	-0,412	0	1	FAUX
43	48,037	7,688	7,688	1	1	VRAI
44	-13,209	-2,127	-2,127	0	0	VRAI
45	-11,093	-1,625	-1,625	1	0	FAUX
46	15,83	2,422	2,422	0	1	FAUX
47	-26,353	-3,973	-3,973	0	0	VRAI
48	-13,587	-1,899	-1,899	0	0	VRAI
49	75,051	11,104	11,104	1	1	VRAI
50	22,352	3,611	3,611	1	1	VRAI
51	-16,368	-2,452	-2,452	0	0	VRAI
52	-13,096	-2,139	-2,139	0	0	VRAI
53	-19,826	-2,92	-2,92	0	0	VRAI
54	2,712	0,67	0,67	0	1	FAUX
55	-11,391	-1,432	-1,432	0	0	VRAI
56	46,751	7,675	7,675	1	1	VRAI
57	-15,504	-2,272	-2,272	0	0	VRAI
58	-5,861	-1,196	-1,196	0	1	FAUX
59	13,287	1,619	1,619	1	1	VRAI
60	-10,516	-1,354	-1,354	0	0	VRAI
61	0,931	-0,009	-0,009	1	1	VRAI
62	-20,218	-2,87	-2,87	0	0	VRAI
63	9,303	1,406	1,406	0	1	FAUX
64	-13,807	-2,165	-2,165	1	0	FAUX
65	-26,824	-4,032	-4,032	0	0	VRAI
66	-24,009	-3,677	-3,677	0	0	VRAI
67	8,767	0,98	0,98	1	1	VRAI
68	-18,171	-2,748	-2,748	1	0	FAUX
69	12,071	1,69	1,69	1	1	VRAI
70	36,412	4,998	4,998	0	1	FAUX
71	-27,6	-4,197	-4,197	0	0	VRAI
72	-7,43	-1,362	-1,362	0	0	VRAI

73	-7,699	-1,163	-1,163	1	1	VRAI
74	-10,619	-1,733	-1,733	0	0	VRAI
75	-16,848	-2,644	-2,644	1	0	FAUX
76	-2,493	-0,61	-0,61	0	1	FAUX
77	-25,021	-3,738	-3,738	0	0	VRAI
78	-25,071	-3,811	-3,811	0	0	VRAI
79	12,862	1,951	1,951	1	1	VRAI
80	114,753	18,093	18,093	1	1	VRAI
81	-18,745	-2,851	-2,851	1	0	FAUX
82	28,847	4,707	4,707	1	1	VRAI
83	7,096	0,177	0,177	0	1	FAUX
84	-8,09	-1,021	-1,021	1	1	VRAI
85	17,889	3,136	3,136	1	1	VRAI
86	9,277	1,749	1,749	1	1	VRAI
87	26,585	5,238	5,238	1	1	VRAI
88	80,881	13,242	13,242	1	1	VRAI
89	24,977	3,864	3,864	1	1	VRAI
90	-11,455	-1,639	-1,639	0	0	VRAI
91	-20,802	-3,242	-3,242	1	0	FAUX
92	53,235	8,683	8,683	1	1	VRAI
93	21,525	1,996	1,996	0	1	FAUX
94	-7,963	-1,722	-1,722	1	0	FAUX
95	-27,743	-4,215	-4,215	0	0	VRAI
96	-28,278	-4,283	-4,283	0	0	VRAI

Légende :

A : variable latente « Attitude » de l'apprenant  
E : variable latente « Engagement » de l'apprenant  
C : variable latente « Cognition » de l'apprenant  
Comp : comparaison entre la prédiction du modèle et le résultat réel  
VRAI : la prédiction du modèle coïncide avec résultat réel  
FAUX : la prédiction du modèle ne coïncide pas avec résultat réel

**Tableau 5.03 : Traitement de données sur Excel**

A partir du Tableau 5.03 : on peut avancer que :

- 73 résultats sur 96 sont bien prédits par le modèle, et 23 ne le sont pas

- Il existe un taux d'erreur de l'ordre de 24% qu'on peut préciser à chaque prédiction de résultats. Pour des travaux ultérieurs, on peut creuser sur les faits ou les causes pouvant engendrer ces résultats sur ces étudiants ou d'autres paramètres que nous n'avons pas pu considérer pendant notre recherche.
- 3/4 des résultats sont bien prédits, et jugés acceptables avec les indices d'ajustement du modèle.

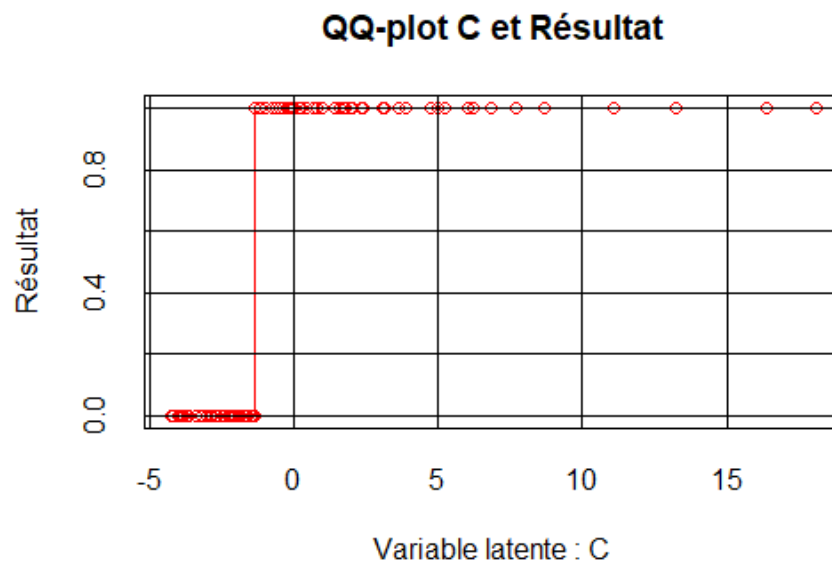
#### 5.4.3. Valeur seuil selon les distributions du résultat réel et celles du résultat du modèle

Outre la recherche manuelle effectuée dans le paragraphe précédent pour la détermination de la valeur du seuil  $C_{seuil}$ , celle-ci peut être aussi trouvée à partir des données statistiques que nous avons.

##### Script :

```
#Distribution des résultats en fonction de la variable C
> qqplot(xC,y,col=2,type="o",xlab="Variable latente : C",ylab="Résultat",axes=T
RUE,main="QQ-plot C et Résultat",tck=1)
```

##### Résultat :



**Figure 5.11.** Distribution du résultat pédagogique

Graphiquement la détermination du seuil est démontrée à partir de cette Figure 5.11

```
> #Identification de la valeur seuil
> locator()
$x
[1] -1.338206
```

Donc, la valeur du seuil  $C_{seuil}$  est égale à **-1.338**.

On va l'utiliser dans les paragraphes suivants pour la suite de notre analyse.

## 5.5 Prédiction de résultat pour l'année universitaire en cours

La formation ouverte et à distance GSA a recruté sa deuxième promotion pour l'année universitaire 2017/2018 et 15 nouveaux apprenants ont intégré le tutorat du module d'initiation GEAL au début de la formation.

### 5.5.1. Analyse des traces d'activités des apprenants 2<sup>e</sup> promotion FOAD GSA

Le module GEAL s'est déroulé normalement du 19 octobre au 05 novembre 2017, mais les traces ont été relevées jusqu'au 09 novembre 2017 car il y a encore des événements passés. Comme décrit dans le chapitre 1, GEAL est divisé en trois séquences, l'analyse a donc été effectuée par séquence afin de voir clairement la fiabilité du modèle avec les données par séquence.

#### 5.5.1.1. Comportement

Dans la première séquence, 510 événements se sont passés durant la période du 19 oct. 2017, 10:52 au 5 nov. 2017, 02:44 ; dont 29 pour le tuteur 1, 23 pour le tuteur 2 et 458 pour les apprenants. Ces 458 événements ont été classés dans les 3 variables suivantes : envoi de message Envoi\_M = 105, consultation de discussion Consu\_D = 273 et consultation de forum Consu\_F = 80.

Envoi\_M contient l'abonnement à une discussion, le contenu posté, le message créé, la discussion créée et le message modifié.

Apprenant	Consu_F	Consu_D	Envoi_M
1	10	14	6
2	8	9	
3	3	8	2
4	15	25	2

5	13	26	15
6	3	40	5
7	21	26	8
8	2	2	4
9	19	25	10
10	4	22	6
11	0	18	5
12	0	9	1
13	0	4	0
14	3	18	6
15	4	27	10

**Tableau 5.04 :** *Analyse de traces dans la première séquence*

#### 5.5.1.2. SEQ2\_2017\_2018

Période : 23 octobre 2017 – 05 novembre 2017 : 350 évènements

Enseignant 1 = 25 évènements

Enseignant 2 = 0 évènement

Apprenants = 325 évènements

Envoi\_M = 70

Consu\_D = 163

Consu\_F = 92

<b>Apprenant</b>	Consu_F	Consu_D	Envoi_M
1	6	16	3
2	1		
3	5	11	6
4	9	19	6
5	10	12	0
6	1	4	1
7	19	41	19
8	0	0	0
9	7	13	3
10	17	16	14
11	4	11	9
12	6	12	6

13	4	2	3
14	0	3	0
15	3	3	0

**Tableau 5.05 :** *Analyse de traces dans la deuxième séquence*

#### 5.5.1.3. SEQ3\_2017\_2018

Période : 20 octobre 2017 – 09 novembre 2017 : 2059 évènements

Enseignant 1 = 133 évènements

Enseignant 2 = 1 évènement

Apprenants = 1925 évènements

Envoi\_M = 512

Consu\_D = 886

Consu\_F = 527

<b>Apprenant</b>	Consu_F	Consu_D	Envoi_M
1	28	60	33
2	39	39	36
3	29	66	33
4	38	45	35
5	100	87	57
6	17	27	23
7	26	85	33
8	10	25	16
9	45	76	40
10	67	91	36
11	27	42	24
12	47	108	53
13	6	20	6
14	16	44	33
15	32	71	54

**Tableau 5.06 :** *Analyse de traces dans la troisième séquence*

### ***5.5.2. Analyse de contenu des forums de discussions de la 2<sup>e</sup> promotion FOAD GSA***

Un forum « Présentez-vous » contient la présentation de tous les apprenants. Pour chaque séquence, voici les nombres de fils de discussions trouvés :

- Première séquence : 4 fils de discussions
- Deuxième séquence : 6 fils de discussions
- Troisième séquence : 32 fils de discussions.

#### **5.5.2.1. Engagement**

Parmi les 276 messages postés dans les 4 forums du module GEAL, 260 messages ont été postés par les apprenants. Ils appartiennent à ces 43 fils de discussions et ont été groupés dans les trois affiliations considérées comme variables à étudier pour pouvoir mesurer la variable latente « engagement ».

#### **5.5.2.2. Communication**

**La variable COMM** représente toutes les simples communications effectuées par les apprenants. Les messages ci-dessous sont des exemples extraits des contenus des forums sur la plateforme et copiés tels quels :

- la communication entre les membres de l'équipe qui vont travailler ensemble

Exemple 1 : « *Ci dessous, vois trouverez la subdivision du Groupe 1\_Geal pour les travaux en équipe du devoir de la séquence 3...* »

Exemple 2 : « *Bonsoir mes pairs du groupe 2 GEAL,*

*Dans le but de réussir la séquence 3 du module GEAL, le tuteur du groupe 2 GEAL nous a demandé de constituer 2 sous-groupes dérivés du groupe 2.*

*Voici la répartition que nous avons déjà fait entre nous via la plateforme ».*

- Remerciement :

Exemple 1 : « *Bonjour Madame,*

*Message bien reçu.*

*Merci! »*

Exemple 2 : « *Bonsoir sidi, Merci de modifier l'orthographe de mon nom DJOGNOU WETIE »*

Exemple 3 : « *Merci beaucoup Patience pour votre esprit de collaboration. »*

- Excuse :

Exemple 1 : « *Bonjour Madame, moi je n'ai pas participer normalement à la discussion de ce matin, nous sommes confrontés à de sérieux problèmes de connexion Internet ces derniers jours »*

Exemple 2 : « *Slt les camarades,*

*dites moi, le temps de remise des devoirs indiqué. Est-il en GMT ou heure madagascar? »*

- Satisfaction :

Exemple : « *Bonjour Madame, Ce fut un plaisir de démarrer la formation avec vous.*

*Le module GEAL dans sa conception, son animation et sa coordination nous aidera beaucoup.*

*Meilleures salutations. »*

#### 5.5.2.3. Motivation

**La variable MOTIV** représente la motivation proprement dite. Les messages ci-dessous sont des exemples exprimant une motivation et copiés tels quels via le forum de la plateforme. La motivation est représentée par :

- Le dynamisme :

Exemple 1 :

« Question sur la dernière question du test d'autoévaluation de la séquence 2

Par [Apprenant1](#), mardi 24 octobre 2017, 16:32

« Bonjour Madame,

*Je n'ai pas compris pourquoi ma réponse à la dernière question du test d'autoévaluation de la séquence 2 est fausse:*

*Question: trouver la correspondance*

*Ne remettre que des travaux parfaits: c'est tout ou rien*

*j'ai répondu :facteur de réussite*

*le corrigé dit qu'il fallait répondre : facteur d'échec.*

*S'agit- il d'un bug ? »*

[Modifier](#) / [Supprimer](#) / [Répondre](#)

Re : Question sur la dernière question du test d'autoévaluation de la séquence 2

Par [Apprenant2](#), mardi 24 octobre 2017, 21:31

« Bonsoir Apprenant1 si tu remet uniquement les travaux parfait cela implique que tu ne fera jamais tous tes devoirs d'où un échec »

[Niveau supérieur](#) / [Modifier](#) / [Séparer](#) / [Supprimer](#) / [Répondre](#)

Re: Question sur la dernière question du test d'autoévaluation de la séquence 2

Par [Apprenant3](#), mardi 24 octobre 2017, 22:19

« Bonsoir Apprenant1,

*à mon avis c'est surtout l'expression "c'est tout ou rien" qui oriente la question.*



*Car si tu dis que tu ne rends que le travail parfait sinon tu ne rends pas tes devoirs, tu risques de ne pas rendre tous tes devoirs.*

*Et un travail sans une évaluation ne vaut rien. »*

[Niveau supérieur](#) / [Modifier](#) / [Séparer](#) / [Supprimer](#) / [Répondre](#)

Re: Question sur la dernière question du test d'autoévaluation de la séquence 2

Par [Apprenant1](#), mercredi 25 octobre 2017, 07:03

*« Merci pour vos réponses ».*

[Niveau supérieur](#) / [Modifier](#) / [Séparer](#) / [Supprimer](#) / [Répondre](#) »

Au moment de l'auto-évaluation, l'apprenant 1 n'a pas hésité à exposer tout de suite sa préoccupation dans le forum de discussion. Ses pairs Apprenant 2 et Apprenant 3 lui répondent immédiatement sans hésitation et sans attendre l'intervention du tuteur. Le dynamisme de tout un chacun comme décrit dans cet exemple a été classé comme une motivation dans cette recherche.

- La Question / Réponse pédagogique :

Exemple : *« Quel est le document qui nous donne les objectifs de l'apprentissage? »*

- Le partage de ressources :

Exemple : *« Bonjour chers collègues et tuteurs,*

*Le lien suivant <http://supportsfoad.com/articles-foad/34-fondamentaux/47-quest-ce-que-la-foad-> donné dans les **ressources complémentaires de la séquence 2** et ayant pour descriptif : [Qu'est-ce que la FOAD](#) ? par Anna Vetter*

*est introuvable. Sur la page d'accueil du site en question l'auteur informe que son site a été piraté.*

*Pour ma part j'ai fait des recherches sur internet pour me documenter sur la FOAD. Voici deux liens à propos:*

*1. La FOAD: Qu'est ce que c'est? <http://capfoad.chez.com/chapfoad.htm>*

*2. [Glossaire Formation Ouverte et à distance et e-learning](http://www.elearning-concepts.com/glossaire_foad_definition-en.htm) [http://www.elearning-concepts.com/glossaire\\_foad\\_definition-en.htm](http://www.elearning-concepts.com/glossaire_foad_definition-en.htm) »*

- Entraide technique

*Exemple : « il est possible de trouver les infos relatifs a la question sur le FOAD en écrivant dans l'espace dédié aux recherches sur le site fournis par le tuteur*

*a titre d'exemple veuillez entrée le questionnaire: qu'est ce que le foad? et cliquez sur l'icône recherche et les informations relatif au FOAD vous seras présentés »*

#### 5.5.2.4. Collaboration

**La variable COLLAB** représente la collaboration de l'apprenant.

Les messages postés dans la troisième séquence évoquent tous de la collaboration, car les étudiants ont effectué un travail collaboratif dans l'activité de la séquence 3.

#### 5.5.2.5. Les données issues de l'analyse de contenu

Apprenant	COLLAB	COMM	MOTIV
1	11	4	1
2	11	3	1
3	10	4	2
4	11	3	2
5	20	6	1
6	5	3	1
7	13	6	5
8	5	4	1
9	10	8	1
10	13	5	3
11	10	4	2
12	16	3	1

13	2	1	1
14	14	4	1
15	21	6	1

**Tableau 5.07 :** *Analyse de contenu des forums de discussions*

### 5.5.3. Les variables de mesures de A et de E

Après l'analyse des traces numériques d'activités sur GEAL, 2708 événements se sont passés. L'analyse du contenu des 260 messages enregistrés sur les forums a permis de mesurer la variable latente E, tandis que le traitement effectué à partir de ces événements a permis de mesurer la variable latente A.

Apprenant	COLLAB	COMM	MOTIV	Consu_F	Consu_D	Envoi_M
[1]	11	4	1	44	90	42
[2]	11	3	1	48	48	36
[3]	10	4	2	37	85	41
[4]	11	3	2	62	89	43
[5]	20	6	1	123	125	72
[6]	5	3	1	21	71	29
[7]	13	6	5	66	152	60
[8]	5	4	1	12	27	20
[9]	10	8	1	71	114	53
[10]	13	5	3	88	129	56
[11]	10	4	2	31	71	38
[12]	16	3	1	53	129	60
[13]	2	1	1	10	26	9
[14]	14	4	1	19	65	39
[15]	21	6	1	39	101	64

**Tableau 5.08 :** *Données après analyse de traces et analyse de contenus*

L'utilisation du modèle issu de la recherche permet de sortir la valeur de C pour la prévision du résultat pédagogique GSA 2017/2018. Le script ci-dessous doit être exécuté pour avoir ce résultat.

### Script et résultat :

```
#3. prévision
newDonnee <- read.csv("F:/Doctorat_2016/SCA/SCA/Thèses/D3/gsa_2017_2018.csv", sep=";")
ndonnee <- newDonnee[1:6]
colnames(ndonnee) <- c("COLLAB", "COMM", "MOTIV", "Consu_F", "Consu_D", "Envoi_M")
prediction = predict(donnee.fit, newdata=ndonnee)
write.csv(file = "F:/Doctorat_2016/SCA/SCA/Thèses/D3/S5/sca/newResult2-gsa.csv", x = prediction)
```

A l'issue de l'application du modèle de la recherche, nous avons obtenu la prévision du résultat dans la dernière colonne.

Apprenant	A	E	C	Rm
[1]	-23,907	-5,935	-5,935	0
[2]	15,167	3,908	3,908	1
[3]	15,763	4,086	4,086	1
[4]	-1,491	-0,255	-0,255	1
[5]	23,152	4,984	4,984	1
[6]	18,631	5,269	5,269	1
[7]	-1,963	-0,698	-0,698	1
[8]	-20,542	-4,598	-4,598	0
[9]	15,439	3,754	3,754	1
[10]	-18,005	-4,668	-4,668	0
[11]	12,096	3,168	3,168	1
[12]	59,066	14,486	14,486	1
[13]	-74,366	-17,643	-17,643	0
[14]	-53,902	-13,69	-13,69	0
[15]	34,86	7,832	7,832	1

**Tableau 5.09 :** *Prévision du résultat GSA 2<sup>e</sup> promotion selon le modèle*

Cette prévision nous donne un résultat de 10 admis parmi les 15 nouveaux apprenants de la deuxième promotion de GSA 2018.

## 5.6 Utilisation du modèle dans un cas général

On va utiliser notre modèle pour effectuer une prévision générale des résultats des apprenants dans une situation quelconque. Pour cela, on va générer des données avec des valeurs variants de 0 à 20 pour chacune des variables COLLAB, COMM, MOTIV, Consu\_F, Consu\_D, Envoi\_M.

	COLLAB	COMM	MOTIV	Consu_F	Consu_D	Envoi_M	A	E	C	R
2607	0	20	0	20	5	10	-13,06	-1,74	-1,74	0
2608	0	20	0	20	5	15	-10,87	-1,46	-1,46	0
2609	0	20	0	20	5	20	-8,69	-1,19	-1,19	1
2610	0	20	0	20	10	0	-16,66	-2,19	-2,19	0
2611	0	20	0	20	10	5	-14,47	-1,92	-1,92	0
2612	0	20	0	20	10	10	-12,29	-1,64	-1,64	0
2613	0	20	0	20	10	15	-10,11	-1,37	-1,37	0
2614	0	20	0	20	10	20	-7,92	-1,09	-1,09	1
2615	0	20	0	20	15	0	-15,89	-2,10	-2,10	0
2616	0	20	0	20	15	5	-13,71	-1,82	-1,82	0
2617	0	20	0	20	15	10	-11,52	-1,55	-1,55	0
2618	0	20	0	20	15	15	-9,34	-1,27	-1,27	1
2619	0	20	0	20	15	20	-7,15	-0,99	-0,99	1
2620	0	20	0	20	20	0	-15,12	-2,00	-2,00	0
2621	0	20	0	20	20	5	-12,94	-1,72	-1,72	0
2622	0	20	0	20	20	10	-10,75	-1,45	-1,45	0
2623	0	20	0	20	20	15	-8,57	-1,17	-1,17	1
2624	0	20	0	20	20	20	-6,39	-0,90	-0,90	1
2625	0	20	5	0	0	0	-18,97	-2,15	-2,15	0
2626	0	20	5	0	0	5	-16,79	-1,87	-1,87	0
2627	0	20	5	0	0	10	-14,60	-1,60	-1,60	0
2628	0	20	5	0	0	15	-12,42	-1,32	-1,32	1

**Figure 5.12.** Extraits de données générées avec les valeurs estimées des variables A, E, C et R

Les données générées seront utilisées dans la fonction « *predict* » avec notre modèle comme indiqué dans le script ci-dessous.

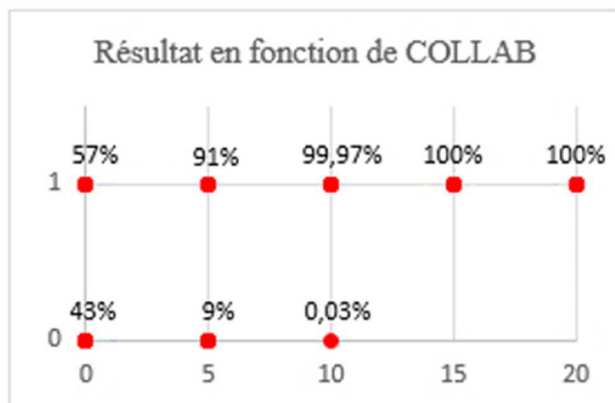
### Script :

```
> newDonnee <- read.csv("F:/Doctorat_2016/SCA/SCA/Thèses/D3/S5/sca/genereDonnee
.csv", sep=";")
> ndonnee <- newDonnee[1:6]
> colnames(ndonnee) <- c("COLLAB", "COMM", "MOTIV", "Consu_F", "Consu_
D", "Envoi_M")
> prediction = predict(donnee.fit, newdata=ndonnee)
```

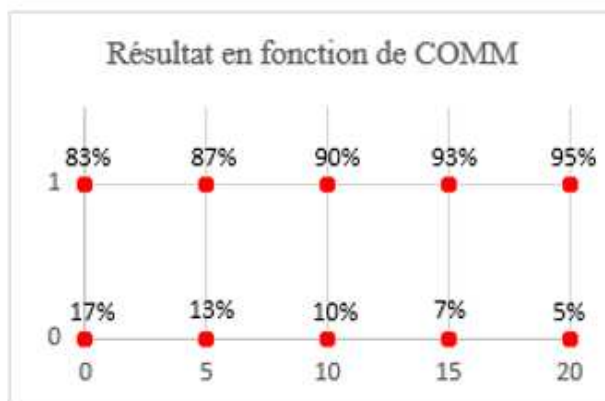
```
> write.csv(file = "F:/Doctorat_2016/SCA/SCA/Thèses/D3/S5/sca/newResult2_generD
onnee.csv",x =prediction)
```

### 5.6.1. Résultat en fonction de chaque variable

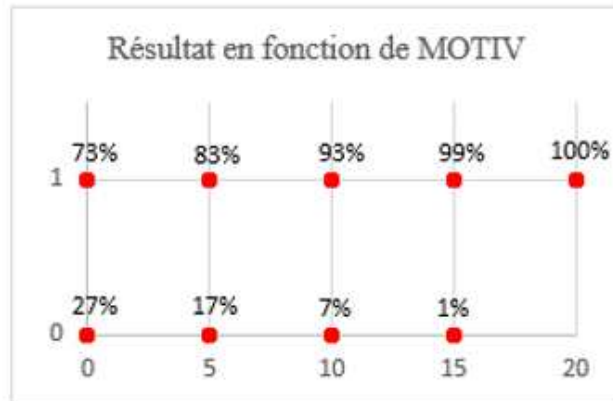
Les résultats sont résumés dans les graphes des figures suivantes.



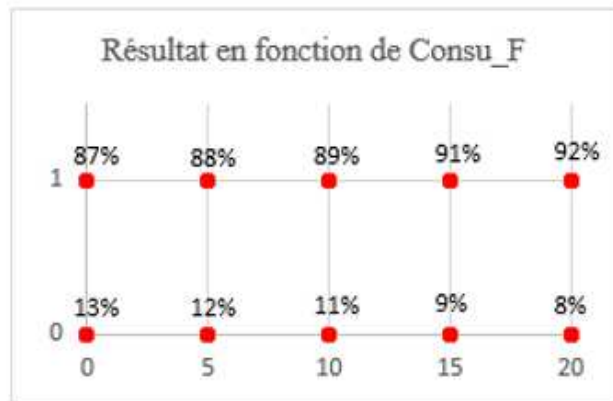
**Figure 5.13.** Résultat en fonction de la collaboration



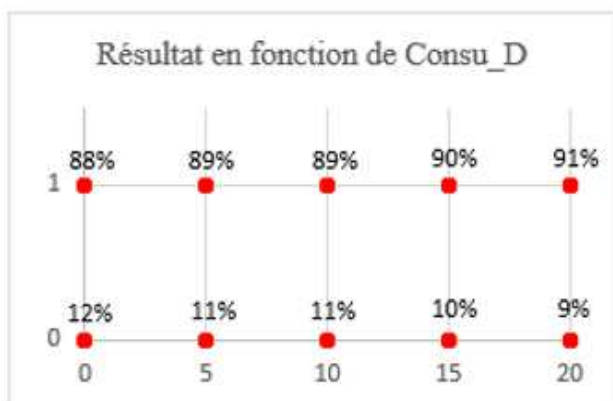
**Figure 5.14.** Résultat en fonction de la communication



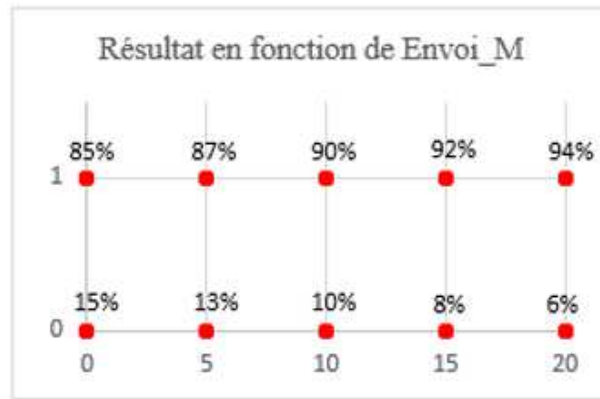
**Figure 5.15.** *Résultat en fonction de la motivation*



**Figure 5.16.** *Résultat en fonction de la consultation des forums*



**Figure 5.17.** *Résultat en fonction de la consultation des discussions*



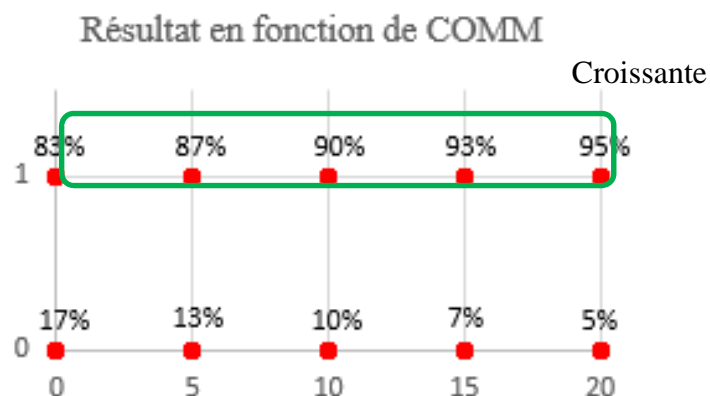
**Figure 5.18.** Résultat en fonction de l'envoi et de la mise à jour de message

Le résultat 1 indique un apprenant admis, tandis que 0 un apprenant non admis.

Dans la Figure 5.13 : Le marqueur rouge sur l'ordonnée 0 et l'abscisse 5 indique que des apprenants présentant un COLLAB = 5 ont échoué. Au-dessus, le marqueur rouge sur l'ordonnée 1 indique que d'autres apprenants ayant présenté la même valeur de COLLAB sont admis, mais évidemment avec d'autres valeurs pour les autres variables. La proportion entre les apprenants admis et ceux qui ont échoué est : 91% admis – 9% échoué.

### 5.6.2. Croissance d'admission en fonction des variables

A partir de la Figure 5.13, on note que la proportion d'apprenants admis augmente en fonction des valeurs des variables étudiées (exemple sur la Figure 5.19).



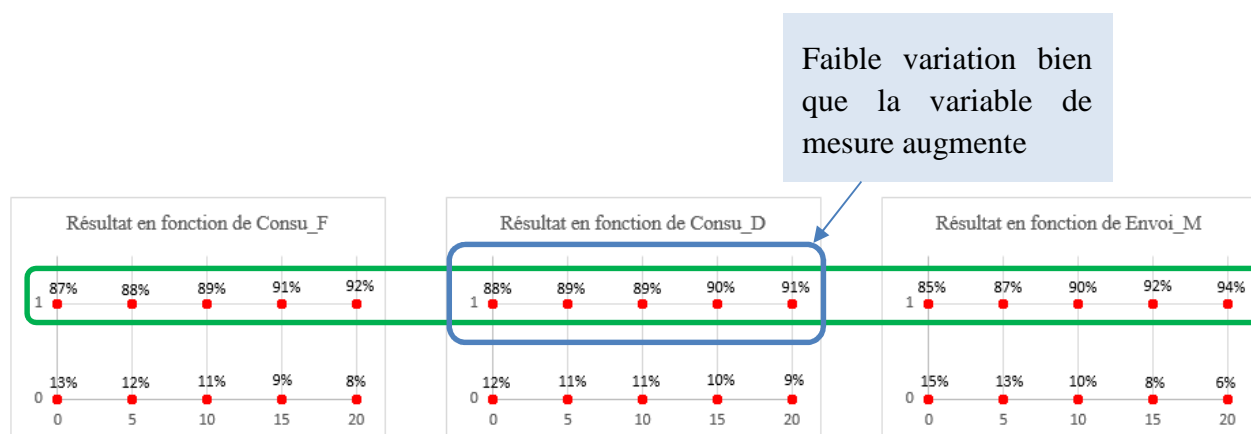
**Figure 5.19.** Taux croissant d'admission en fonction des variables



On peut déduire que plus les apprenants apportent leur collaboration (plus de communication, ou plus de motivation, ou plus de consultation dans les forums ou discussions, ou plus de postage dans les forums), plus ils ont des chances d’être admis. Un apprenant motivé collabore avec ses pairs et communique toujours pendant son apprentissage. Il consulte toujours les nouvelles, participe aux discussions en cours et crée des débats si nécessaire. Le dynamisme de l’apprenant dans la collaboration et dans la communication, le maintien de la motivation et la participation active dans un forum de discussion ont un impact positif sur sa cognition et le mènent vers un bon résultat pédagogique.

### 5.6.3. Faible variation d’admission même si la valeur des indicateurs de A augmente

Pour les variables observées engendrant la variable latente A (Consu\_F, Consud\_D, Envoi\_M), on note que le taux d’apprenants admis ne varie pas beaucoup, comme le montre la Figure 5.20.



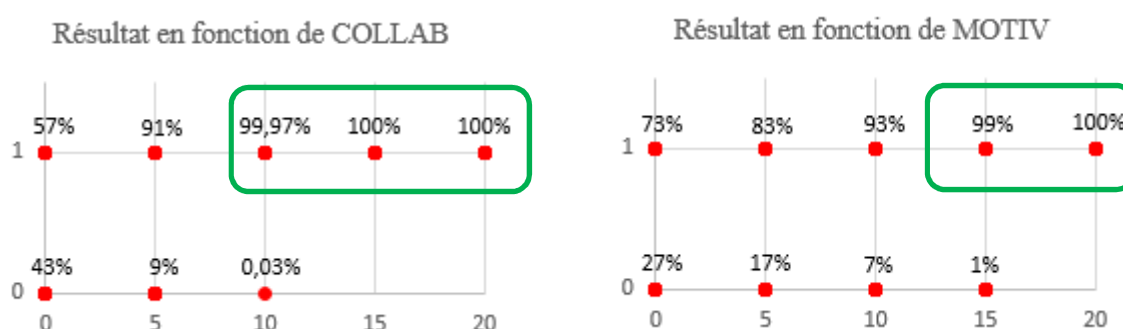
**Figure 5.20.** Taux d’admission avec la variable latente A

On déduit que les trois variables de mesure de la variable latente A n’ont pas d’impact direct sur la variable latente C ni sur le résultat pédagogique de l’apprenant. Autrement dit, le comportement de l’étudiant face au forum de discussion n’a pas d’influence directe sur sa capacité à acquérir des connaissances.

La participation à la discussion en cours et surtout la consultation des forums sur la plateforme sont importantes mais ne suffisent pas pour réussir un apprentissage dans une FOAD. Il faut réagir et toujours contribuer dans toutes les activités de groupe de la formation. Il est également nécessaire d'acquérir le réflexe de parler au responsable ou aux co-équipiers en cas de difficulté, afin qu'ils puissent aider dans la résolution du problème et prendre les mesures nécessaires afin de maintenir la persévérance et la motivation.

#### 5.6.4. Forts indicateurs pour la variable latente E

Un résultat intéressant peut aussi être déduit de la Figure 5.13 et la Figure 5.15 qu'on va mettre en exergue sur la Figure 5.21 ci-dessous.



**Figure 5.21.** Collaboration et motivation de l'apprenant

A partir d'une certaine valeur de collaboration apportée par un apprenant, il peut toujours être admis quelles que soient les valeurs des autres variables de mesure (0 échec). La collaboration est donc un indicateur fort pour la variable latente E et a un impact important sur C et sur le résultat pédagogique. Autrement dit, la collaboration privilégie une vision socioconstructiviste de l'apprentissage dans le cadre de la construction de connaissance [5.01], [5.02].

En outre, on peut citer aussi la motivation d'un apprenant. Elle crée un climat favorable à l'apprentissage. La motivation fait partie des compétences sociales nécessaires dans l'apprentissage en ligne [5.03]. Un apprenant motivé, communique avec son entourage dans un dispositif de formation, partage sa connaissance via la plateforme, développe son savoir-faire et son savoir être et collabore avec les membres de son groupe pour les travaux demandés.

L'hypothèse 2 de la recherche est donc vérifiée. *H2 : L'engagement de l'apprenant dans un groupe pourrait soutenir son interaction asynchrone et sa motivation dans la formation.*

On peut dire que pour réussir dans un apprentissage en ligne, l'étudiant doit être sérieusement engagé. L'étudiant doit être actif dans le travail collaboratif, interagir avec ses pairs et être motivé à suivre la formation et à réaliser toutes les activités d'apprentissage.

## **5.7 Discussion**

Les faits montrent d'une part, que la consultation de forum, la consultation de discussion, la participation aux fils de discussions en cours, la création de nouvelles discussions, la mise à jour de messages (création, modification, suppression) n'ont pas d'effet direct sur la cognition ; et d'autre part, que la collaboration et la motivation sont les causes dominantes de l'acquisition de connaissance dans une formation ouverte et à distance, suivie de la communication. Le dynamisme dans la collaboration implique une interaction forte au sein du groupe, d'où une meilleure efficacité pour l'atteinte d'un objectif commun et l'amélioration du résultat pédagogique qui s'en découle.

Le tutorat, s'il y en a, devrait alors favoriser la collaboration des apprenants pour avoir un bon résultat. Selon Dix et al. [5.04], le travail collaboratif s'élabore par paires, et mieux vaudrait présenter aux étudiants un forum organisé que leur proposer des pistes pour qu'ils l'organisent eux-mêmes.

### ***5.7.1. Collaboration à distance dans un forum structuré***

Dans le contexte de discussion collaborative à distance, les apprenants utilisent l'outil forum pour arriver à dégager une solution commune [5.05], [5.06] et doivent argumenter pour mieux se faire comprendre de leurs pairs [5.07].

Si le scénario a un rôle important à jouer pour stimuler les interactions entre apprenants [5.08], il importe également d'adapter les outils au contexte de l'apprentissage. C'est la raison pour laquelle le forum est structuré et oriente les apprenants vers des activités guidées comme la « présentation » et les scénarios de résolution d'une situation problème dans la séquence 3 de GEAL. Le résultat de

la recherche nous a prouvé l'importance de cette structuration et les impacts du travail collaboratif sur la réussite des apprenants actifs.

Les apprenants dans la formation à distance ont apporté les nuances à prendre en considération en termes de modalités de travail collaboratif. Bruno De Lièvre et son collègue [5.09] ont mis en évidence le fait que les apprenants apprécient et jugent bénéfique le travail collaboratif. Cette nuance a été constatée lors de l'analyse de contenu effectué dans quelques forums, et elle a un impact sur le comportement de l'apprenant.

La recherche a traité tous les forums, qu'ils soient structurés ou non. Les apprenants ont travaillé dans les deux premières séquences sans qu'aucune structure de forum ne leur ait été fournie. On a envisagé également la pertinence de leurs interactions dans des forums sans structure préalable et de faire confiance en leur capacité de construire eux-mêmes une organisation qui leur convienne. Nous avons constaté la perception par les apprenants de l'efficacité du travail collaboratif dans la séquence 3 du module GEAL, là où le forum est bien organisé selon le scénario d'apprentissage. Il est clair, que la structuration facilite l'usage du forum et son efficacité pour l'apprentissage à distance.

#### ***5.7.2. Intervention tutorale pour aider les apprenants à collaborer et à communiquer***

De même, l'intervention tutorale a une place importante dans la bonne réalisation d'une formation ouverte et à distance. Dans la résolution tutorale #8 de Jacques Rodet en avril 2018, il a soulevé la responsabilité du tuteur dans la collaboration, autrement dit aider les apprenants à collaborer. « *Le tuteur à distance doit intervenir auprès des groupes d'apprenants. Il ne lui appartient pas pour autant de jouer le rôle de chef d'équipe mais bien plus de venir en aide aux apprenants, tant de manière collective qu'individuelle, afin que les apprenants puissent atteindre leurs buts d'activités et de formation ainsi qu'une réelle autonomie de fonctionnement. Ses interventions ont pour cadre la collaboration des apprenants dont les principales phases sont les suivantes : l'engagement des apprenants envers l'équipe, la production en équipe et l'évaluation des activités de leur équipe* »

Dans ce cas, les apprenants doivent **s'engager dans l'équipe**. Se constituer en équipe suppose que ses différents membres partagent un but commun. De par les interventions tutorales, les apprenants doivent premièrement poser des actes fondateurs de leur équipe tels que le choix d'un nom, la mise

au point d'une identité visuelle. Deuxièmement, les apprenants doivent définir leur mode de fonctionnement. Il s'agit de définir à partir des compétences de chacun, les rôles que chaque membre va jouer au sein de l'équipe. Puis, les apprenants doivent définir les règles de fonctionnement de leur équipe. Ceci passe tant par l'identification des modalités de **communication entre les membres** : nature et fréquence des réunions, outils de communication utilisés..., que par la manière dont ils vont prendre des décisions. L'équipe doit toujours bien rester dans le cadre de l'activité **collaborative** et être ainsi dans le périmètre d'existence et d'autonomie.

Les apprenants doivent **produire en équipe**, de la planification à la réalisation des tâches. Tout travail collaboratif nécessite une projection dans le futur. Le planning est l'outil fédérateur des activités concrètes de l'équipe. Si le planning gagne à être co-construit et négocié par les apprenants, le tuteur peut avoir un certain nombre d'interventions à réaliser pour faire prendre conscience aux apprenants de son utilité, autrement dit aider les apprenants à planifier leur apprentissage. La détermination commune du planning n'élimine pas pour autant la nécessité d'un garant ou d'un responsable qui est chargé de vérifier et d'informer chacun des membres de l'avancée des tâches, des écarts constatés entre la planification et la réalité, des régulations et aménagements opérés. Le planning étant défini, chacun des apprenants se voit affecté un certain nombre de tâches à réaliser. Celles-ci sont directement liées au chemin choisi par l'équipe pour atteindre son but. La production en équipe est conditionnée par la capacité de production de chacun des membres. En fonction des besoins propres de chaque apprenant, le tuteur est amené à investir les différents plans de support à l'apprentissage : cognitif, socio-affectif, motivationnel et métacognitif. Les outils de communication doivent être accessibles à chaque apprenant.

Les apprenants doivent **évaluer les activités de leur équipe**. L'autoévaluation d'une équipe est non seulement souhaitable mais nécessaire. Cette autoévaluation devrait porter en particulier sur le fonctionnement de l'équipe et la cohésion entre ses membres ainsi que, de manière plus classique, sur les résultats qu'elle obtient : sa productivité tant qualitative que quantitative. La cohésion renvoie aux comportements positifs et négatifs des membres de l'équipe manifestés tout au long de la réalisation de l'activité collaborative. La productivité est liée à la réalisation des tâches et à l'atteinte des objectifs.

Les interventions tutorales aident les apprenants à collaborer. Elles sont ainsi fort diverses et marquées par les différentes phases de la collaboration. Les actions des tuteurs aident les apprenants à réaliser leurs tâches et à atteindre leurs objectifs en faisant vivre leur collaboration.

Dans l'apprentissage par problème et par projet, selon Caroline Verzat et ses collègues [5.10], l'efficacité de l'apprentissage est indissociable d'une réelle collaboration entre les étudiants.

Le rôle du tuteur est souvent de faciliter la collaboration, de créer un espace propice au conflit sociocognitif entre les étudiants et de savoir faciliter l'apprentissage dans le cadre d'une dynamique de groupe d'étudiants.

### ***5.7.3. Motivation, rôle de tous les protagonistes de la formation***

Du point de vue des neurosciences, la motivation est interdépendante de la sensation de plaisir et vice versa. D'où l'utilité d'un climat favorable à l'apprentissage, apaisant, clément, bienveillant, accueillant, d'entraide et de partage plutôt que de compétition ou de comparaison où l'apprenant a notamment le droit de commettre des erreurs sans être dévalorisé ni jugé par ses pairs et son professeur.

Ainsi, il incombe à l'institut de créer les conditions-cadre d'apprentissage qui favorisent l'échange et le développement des compétences professionnelles, sociales et personnelles de tous les apprenants.

Dans la même optique, c'est le rôle du tuteur d'assurer un lien émotionnel avec les apprenants et de leur garantir un accompagnement efficace dans la durée, et ce, dans un climat convivial et motivant.

## **5.8 Conclusion**

On peut conclure d'une part que le comportement de l'apprenant face au forum de discussion n'a pas d'effet direct sur sa cognition ; et d'autre part que le rôle crucial de la collaboration est mis en évidence sur l'acquisition de connaissance dans une formation ouverte et à distance. Par ailleurs, la motivation est également un facteur important, voire indispensable pour la persévérance dans la formation. Enfin, le rôle capital de communication a été également souligné car c'est une des conditions à remplir si on veut réussir dans ce type de formation.

Le dynamisme dans la collaboration implique une interaction forte au sein du groupe et c'est cette interaction qui aboutit à une production efficace qui mène vers l'atteinte de l'objectif commun et améliore par là même le résultat pédagogique.

## CONCLUSION GENERALE

Plusieurs universités publiques et privées créent des formations ouvertes et à distance actuellement à Madagascar pour différentes raisons et leur nombre ne cesse d'augmenter. Cependant, peu de recherches ont été conduites afin de vérifier comment les étudiants malgaches adoptent et utilisent les outils d'apprentissage. C'est ainsi que cette recherche a essayé de proposer un modèle d'apprentissage asynchrone afin d'examiner la relation entre l'engagement de l'apprenant, son attitude face au forum de discussion et son résultat pédagogique dans une FOAD.

Dans cet ouvrage, deux modèles de mesure d'équation structurelle de l'engagement et du comportement de l'apprenant ont été validés dans un premier temps et un modèle de structure a été validé dans un second temps. Un modèle complet d'équation structurelle approprié à l'apprentissage asynchrone a été obtenu à l'issue de la recherche.

L'approche par les relations structurelles linéaires dans la modélisation par équations structurelles sur variables latentes a été adoptée à titre de méthodologie. Les données ont été tirées à partir de l'analyse des traces numériques des activités et de l'analyse des contenus des forums de discussion. Le traitement statistique a été effectué au moyen du logiciel R.

Deux hypothèses se sont avérées. Premièrement, le comportement de l'étudiant face au forum de discussion dépend de son engagement communicationnel, son engagement motivationnel et surtout son interaction dans un groupe de travail pour la collaboration. Deuxièmement, l'engagement de l'apprenant dans son apprentissage a une relation directe avec son acquisition de connaissance. Le résultat pédagogique n'a rien à voir avec l'attitude de l'apprenant face au forum de discussion.

Une modélisation d'apprentissage asynchrone a été faite et le modèle obtenu pourra être appliqué à toutes formations ouvertes et à distance.

La recherche fait donc un lien entre le processus de l'engagement de l'apprenant et celui de la collaboration en équipe, la motivation et les motifs d'engagement en formation et, la communication entre tous les protagonistes de la formation. La recherche renseigne sur les taux de collaboration et de motivation nécessaires pour s'engager dans un processus d'apprentissage asynchrone.

La collaboration est la dimension la plus importante pour l'engagement de l'apprenant dans une formation à distance utilisant le forum comme outil principal de la formation. L'étudiant doit garder



sa motivation, son engagement cognitif et sa persévérance dans la formation. A tout cela s'ajoute la communication qui est indispensable pour maintenir la dimension relationnelle et sociocognitive.

### ***Applications***

L'application du modèle obtenu à la 2<sup>e</sup> promotion spécialisée en Génie du système automatique (GSA) de l'IST d'Antananarivo a révélé une prévision de résultat de 10 admis parmi les 15 nouveaux apprenants en cours d'études (année universitaire 2017/2018). Le conseil de classe correspondant va se tenir au mois d'octobre 2018. Les résultats réels pourront alors être confrontés aux prévisions.

On peut appliquer ce processus d'apprentissage et le modèle obtenu dans toutes les formations en ligne comme FOAD et MOOC (Massive Open Online Course) ou CLOM (Cours en Ligne Ouvert et Massif).

Enfin, le modèle obtenu a été testé et validé empiriquement et pourra être adapté facilement au contexte malgache étant donné qu'il fait recours à l'utilisation d'outils asynchrones.

### ***Perspectives***

- L'adaptation de ce modèle dans une formation hybride, qui est une autre forme de formation à distance, pourrait faire l'objet d'une autre recherche.
- Nous avons adopté dans cette recherche la méthode d'analyse des variances et covariances. On peut aussi étudier d'autres solutions algorithmiques si on ne cite que la méthode des moindres carrés ordinaires appliqués de façon indépendante pour chaque itération, ou même la régression simultanée développée par Tjalling Koopmans à la commission Cowles.
- L'exclusion de la variable Création de discussion (*Creat-Disc*) dans notre modèle (le fait de la mélanger avec la variable *Envoi\_M*) revient à dire que sa variation est légèrement transparente par rapport au résultat de l'apprenant quelle que soit la quantité apportée par un tel comportement, si le total *Envoi\_M* est maintenu constant. On peut alors étendre la recherche vers les SEM non paramétriques qui permettront l'estimation des effets totaux, directs et indirects, sans aucun engagement quant à la forme de nos équations aux distributions des erreurs. Ceci pourra enrichir notre analyse du système d'apprentissage asynchrone.

## Annexe 1 : LES ARTICLES DE NOTORIÉTÉ NATIONALE ET INTERNATIONALE

## A.1.1. Premier article de notoriété internationale

M. Rakotomalala, T. B. Ravaliminoarimalalason, F. Randimbindrainibe, «*Models for Measuring Latent Variables in Asynchronous Learning*», in International Journal of Engineering Research & Technology, Vol. 7 Issue 04, pp 146-150, April-2018.

**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:

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

With:  $\xi_1$ : latent construct,  $X_i$ : indicators,  $\lambda_i$ : coefficient representing the effect of the construct on the indicator (loading); and  $\delta_i$ : measurement error.

#### B. Model of structural equations

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

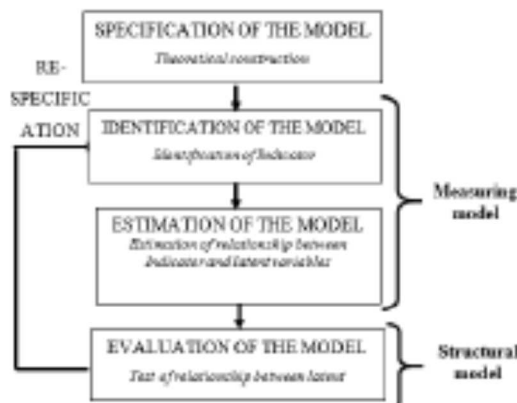


Fig. 1. Steps of structural equations modelling

The specification of the model leads to the design of a "linear relationship scheme" 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  $\Sigma$  as a function of a set of parameters  $\theta$ . This model is identified if  $\Sigma(\theta) = \Sigma(\theta')$  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]:

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

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 ( $ddl \geq 0$ ).

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:

$$F_{ML} = \text{tr}(S \cdot \hat{\Sigma}^{-1}) - p + \ln |\hat{\Sigma}| - \ln |S| \quad (3)$$

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

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

With:  $S$ : the observed variance-covariance matrix,  $\ln$ : the natural logarithmic function,  $\text{tr}()$ : The trace function of a matrix,  $\Sigma$ : 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 ( $\epsilon_i$ ,  $\delta$ ,  $\beta$  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.



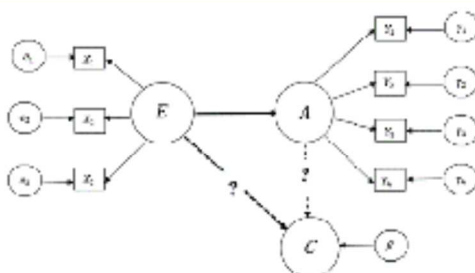


Fig. 2. The research conceptual model

### 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 20<sup>th</sup> and November 07<sup>th</sup>, 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:

$$\begin{aligned} V_1 &= \alpha_1 A + \gamma_1 \\ V_2 &= \alpha_2 A + \gamma_2 \\ V_3 &= \alpha_3 A + \gamma_3 \\ V_4 &= \alpha_4 A + \gamma_4 \end{aligned} \quad (5)$$

They can be written in the matrix form:  $Y = \alpha A + \gamma$  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 (logl): -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.

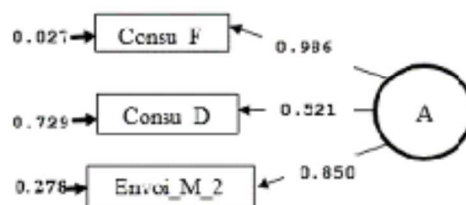


Fig. 3. Measurement model of latent variable A with estimated parameters

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  $\chi^2$  (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 (npars) 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:

$$\begin{aligned} X_1 &= \lambda_1 E + e_1 \\ X_2 &= \lambda_2 E + e_2 \\ X_3 &= \lambda_3 E + e_3 \end{aligned} \quad (6)$$

They can be written in the matrix form:  $X = \lambda E + e$  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:

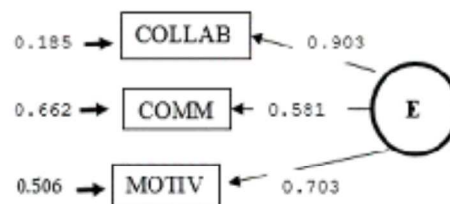


Fig. 4. Measurement model of latent variable E with estimated parameters

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 (logl) 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  $\chi^2$  (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. Paraya, "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. Geonac'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 modeling". *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 édition, London, 2004.
- [9] L. Chaput, « Modèles contemporains en gestion », Presses de l'Université du Québec, 2006.
- [10] P. Roussel, F. Durieu, 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.



*A.1.2. Lettre d'acceptation du premier article de notoriété internationale*



*A.1.3. Certificat de publication de l'article*





#### **A.1.4. Premier article de notoriété nationale**

M. Rakotomalala, T. B. Ravaliminoarimalalason, F. Randimbindrainibe, « *Analyse de l'apprentissage d'un module d'initiation dans une formation ouverte et à distance* », MADA-ETI, www.madarevues.gov.mg, Vol.1, pp. 18 - 31, mai 2018

##### **Résumé**

*Cet article vise à analyser l'apprentissage d'un module d'initiation dans une formation ouverte et à distance. Ce module est considéré non seulement comme un appui technique et un support cognitif pour les étudiants, mais également comme accompagnateur dans l'organisation de leur apprentissage. C'est ainsi qu'une enquête auprès des apprenants, une analyse de discours et une analyse de traces numériques d'activités sur la plate-forme ont été adoptées.*

**Mots clés :** *analyse, apprentissage, module d'initiation, formation ouverte et à distance*

##### **Abstract**

*This article aims to analyze the learning of an introductory module in open and distance learning. This module is considered not only as a technical support and a cognitive support for the students, but also as an accompanist in the organization of their learning. This is how a learner survey, a speech analysis and a digital traces analysis of activities on the platform were adopted.*

**Keywords :** *analysis, learning, introductory module, open and distance learning*

##### **1. Introduction**

La formation ouverte et à distance (FOAD) se développe actuellement dans l'enseignement supérieur. Les deux Instituts Supérieurs de Technologie (IST) de Madagascar possèdent ce type de formation en la FOAD licence en Exploitation logistique et transport (ELT) à l'IST d'Antananarivo (IST-T) et la FOAD licence en Transit et commerce international (TCI) à l'IST d'Antsiranana (IST-D). Ces deux formations utilisent la plate-forme Moodle comme support techno-pédagogique.

Certes, Moodle présente beaucoup d'avantages par rapport aux autres plateformes de formation en ligne, telles que la présence d'une panoplie complète d'outils, la disponibilité de suivi des apprenants et la possibilité de l'autonomie de l'apprenant. Cependant, il a quelques inconvénients, car à partir du moment où sa configuration a été adaptée à la formation proposée, trop de fonctionnalités peuvent rendre la prise en main délicate du point de vue des étudiants. L'idée de faciliter la prise en main de cette plate-forme a entraîné l'insertion d'un module Guide de

l'Étudiant pour l'Apprentissage en Ligne (GEAL) dans la formation.

Après une année de fonctionnement, GEAL devrait être évalué. La recherche penche sur l'analyse de traces d'apprentissage, et essaie de répondre aux questions ayant trait à l'intégration du module GEAL dans la formation à distance, de chercher à savoir comment les étudiants ont exploité ce module d'initiation dans leur formation. Elle va mettre en évidence les principaux outils adoptés par les étudiants et le temps nécessaire pour l'apprentissage de ce module.

La méthodologie mise en œuvre repose sur l'analyse de traces numériques d'activités des 38 nouveaux apprenants des deux FOAD dans le premier semestre de l'année académique 2014/2015. L'analyse de discours sur le forum de discussion a été réalisée, ainsi que l'analyse des fréquences d'intervention sur les événements des autres activités d'apprentissage. Afin de pouvoir confronter avec la trace, une enquête auprès des apprenants a été aussi effectuée, avec une dizaine de questions. Le questionnaire a pour objectif de savoir comment les étudiants ont exploité réellement GEAL.

L'article se divise en trois sections. La première concerne la formation, le module GEAL, et la méthodologie adoptée. Le cadre théorique forme la deuxième section ; et la dernière expose les résultats obtenus et la discussion. Il se termine par la conclusion et la perspective de la recherche.

## 2. La formation et le module d'initiation

Cette section présente les matériels et les méthodes utilisés dans la recherche.

### 2.1. La formation

La formation ouverte et à distance existe depuis plusieurs années aux Instituts Supérieurs de Technologie de Madagascar. Elle se fait entièrement à distance via une plate-forme Moodle. C'est une formation tutorée qui suit la théorie d'enseignement et d'apprentissage selon les approches constructiviste et socioconstructiviste. La formation s'organise autour de la résolution de situations à problèmes et s'appuie sur des démarches collaboratives entre pairs ainsi qu'une fréquence soutenue d'échanges avec le tuteur [1].

### 2.2. Les acteurs

Dans la formation interviennent des enseignants concepteurs, des tuteurs, des coordinateurs et des apprenants. Trente-trois tuteurs assurent le suivi et l'encadrement des travaux des étudiants tout au long de la formation. Une grande majorité d'entre eux sont aussi des concepteurs de cours.

Les coordinateurs (administratif, pédagogique et technique) garantissent le bon déroulement de la formation et établissent les liaisons entre l'administration, les tuteurs et les apprenants. Ils assurent le premier regroupement à la rentrée, durant lequel une présentation de la formation et une prise en main de la plate-forme sont effectuées.

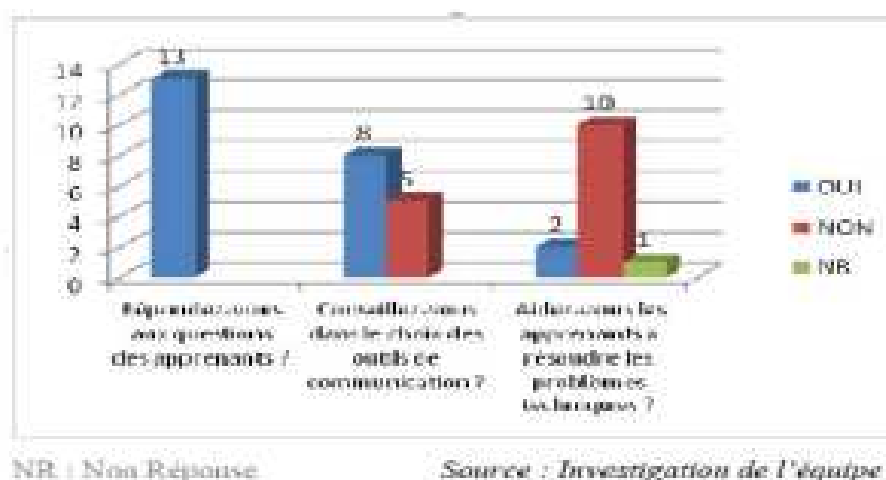


Figure 1. Fonction technique des tuteurs

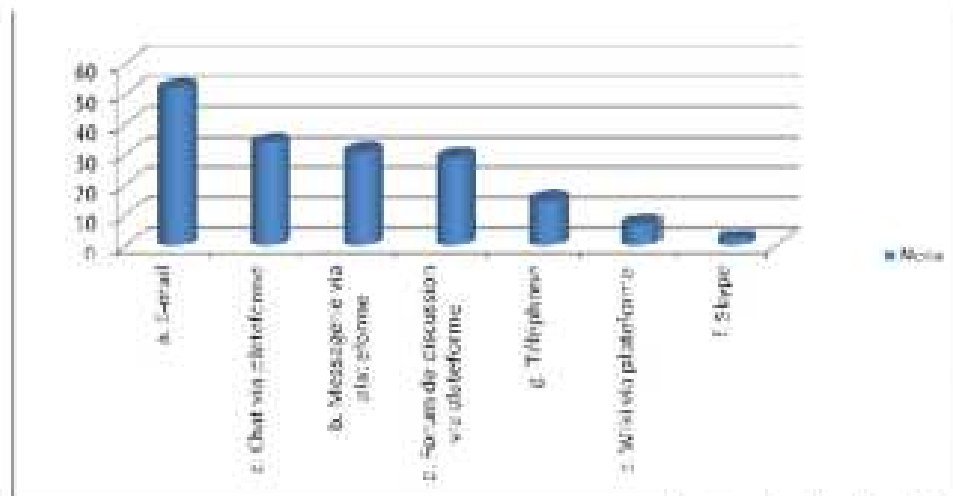
Les apprenants sont au centre de la formation et réalisent les activités aussi bien individuellement que collectivement. Ils furent encadrés pendant dix jours environ sur l'apprentissage du module GEAL. Par la suite, le module reste ouvert et les apprenants peuvent l'utiliser eux-mêmes, selon leur besoin.

Le dispositif a donné les moyens de s'approprier pleinement de la modalité à distance. Les tuteurs assurent un tutorat de ce module où les interactions ont lieu, et incitent fortement les étudiants à l'utilisation des outils via la plate-forme.

### 2.3. Études antérieures

Quelques données significatives lors de l'étude [2] effectuée auparavant ont été prises, afin d'effectuer la comparaison au niveau de l'utilisation des outils de communication et de collaboration (figure 2), et de l'assistance technique des apprenants par les tuteurs (figure 1).

La figure 1 indique que 15% seulement des tuteurs aident les apprenants à résoudre les problèmes d'ordre technique, et la figure 2 montre le niveau d'utilisation du forum de discussion auparavant.



Source : Investigation de l'équipe

Figure 2. Utilisation des outils de communication et de collaboration

#### 2.4. Le module d'initiation GEAL

GEAL est un module entier sous forme de guide pour initier les connaissances des étudiants en matière d'e-learning. Il est destiné aux étudiants pour les accompagner lors de leurs premiers pas en formation à distance. Il a pour objectif de fournir les outils et méthodes nécessaires à la réussite de l'apprentissage en ligne. Il est composé de trois séquences, en plus d'une séquence zéro (0) qui affiche les informations du module (figure 3).

La première séquence s'intitule « Environnement d'apprentissage ». Elle a pour objectif de connaître l'Institution, la formation et le dispositif de

formation. Dans cette séquence, les étudiants se familiarisent avec la plate-forme Moodle.

Pour chaque année universitaire, GEAL est lancé dans le cadre du premier regroupement qui se déroule par visioconférence. Les étudiants entrent par pays sur des plages d'une heure environ. Un mail de contact est envoyé aux étudiants une semaine avant le regroupement. Il contient leur code d'accès à la plate-forme et les instructions pour le remplissage du profil. Le jour du regroupement, les apprenants commencent à entrer dans GEAL, sous l'encadrement des tuteurs.



Figure 3. La séquence 0 du module GEAL.

### 2.5. Méthodologie adoptée

Les méthodologies mises en œuvre pour la réalisation de cette recherche reposent sur une technique d'enquête par questionnaire en ligne en direction des étudiants, une analyse de discours dans des forums de discussion, et une analyse de traces numériques des activités sur la plate-forme.

Afin de vérifier la cohérence, plusieurs instruments de collecte de données ont été pris pour croiser les informations par triangulation, dont un questionnaire pour la technique d'enquête, une grille d'analyse de discours et une grille d'analyse qui tient compte des variables à étudier pour les traces numériques d'activités.

Le questionnaire fut élaboré à l'aide d'un outil en ligne Google Forms. Il a été testé par deux experts en TICE avant l'envoi aux 38 nouveaux étudiants de la formation. Il a pour objectif de savoir comment les étudiants ont exploité le module

GEAL. Les questions concernent la contribution, la participation aux différentes activités de GEAL et les outils utilisés dans la formation. L'application des acquis aux autres modules et la raison de son retour vers GEAL s'il y en a, ont été aussi demandées. L'échelle de mesure de Wang [3] fut adaptée dans une question pour mesurer la satisfaction des étudiants [4]. Une échelle de Likert à 5 points allant de « tout à fait en désaccord » à « tout à fait d'accord » fut utilisée. Une dernière question est à propos du délai désiré pour l'apprentissage du module GEAL.

Afin de caractériser l'exploitation du module, une grille d'analyse de traces des activités a été établie de façon à analyser, d'une part, les interactions entre pairs, entre étudiants et tuteurs pendant le déroulement du module ; et d'autre part, les interactions entre GEAL et le reste de la formation.

### 3. Cadre théorique

Puisque cette recherche se penche sur l'exploitation du module d'introduction dans une formation à distance, il semble utile de présenter brièvement ce que sont à la fois une formation à distance, un dispositif d'apprentissage et une approche pédagogique à distance. Cette section contient également la partie théorique sur le type d'étudiant de la formation à distance et l'analyse de traces d'activités.

#### 3.1. Formation à distance

L'enseignement à distance a largement bénéficié des apports offerts par les technologies de l'information et de la communication : « Concevoir un dispositif efficace de formation à distance utilisant les TIC, c'est concevoir un artefact complexe qui favorise l'apprentissage et permet à l'apprenant de réaliser les tâches prescrites » [5].

La formation à distance est définie comme une formation médiatisée [6]. Celle-ci fait appel à des artefacts technologiques, et à des dispositifs médiatiques. « C'est ainsi que seront mis à disposition des apprenants : les ressources pédagogiques, les différents outils de communication et de collaboration, les activités d'apprentissage et les moyens de les réaliser. » C'est à travers de tels dispositifs que les apprenants communiqueront et collaboreront, en s'entraînant dans leurs travaux.

Dans cette recherche, ce dispositif s'appuie sur l'usage de la plate-forme Moodle. Il a été conçu pour favoriser un cadre de formation

socioconstructiviste et intègre des outils et des activités d'apprentissage axer sur les communications synchrones (personnes en ligne visibles, chat), sur les communications asynchrones (forum), sur l'apprentissage collaboratif (groupes, glossaire, wiki, atelier, base de données), sur la réflexion critique (sondage : vote, feedback : les attentes) et sur la personnalisation (page personnelle : profil). Il s'inspire aussi du constructivisme dans lequel l'apprentissage se construit à partir de l'interprétation et de la compréhension constamment renouvelées et élaborées à partir de représentations antérieures. Par ailleurs, avec Moodle, l'apprenant est au centre de la formation. Il suit un parcours pédagogique qui favorise l'apprentissage par l'activité. L'apprentissage est perçu comme une activité mentale réalisée individuellement à partir d'informations à traiter [7].

#### 3.2. Modèle d'enseignement et d'apprentissage

Contrairement aux behavioristes pour lesquels l'individu est modelé par son environnement, Piaget considère que l'apprentissage est le résultat d'une interaction entre le sujet et son environnement. Pour lui, l'apprentissage est le résultat d'un processus dynamique de recherche d'équilibre entre le sujet et son environnement. Il a notamment mis en évidence l'importance de l'action sur l'environnement et de la réflexion sur cette action pour construire des connaissances sur un objet [8].

### 3.3. Type et profil d'étudiants

Un profil est un ensemble d'informations concernant un apprenant ou un groupe d'apprenants, saisies, collectées ou déduites à l'issue d'une ou plusieurs activités pédagogiques. Ces informations peuvent concerner les connaissances, les compétences, les conceptions, les représentations ou encore les styles ergonomiques. Les lurkers [9] sont des apprenants dans un séminaire synchrone, qui lisent les messages, ne perdent pas le fil des apprentissages mais n'interviennent pas. Leurs productions sont souvent de qualité car ils profitent des apports de leurs pairs. D'autres profils sont connus, souvent liés à des mécanismes sociaux. Le social loafing, se produit lorsqu'un étudiant travaille moins car il s'appuie sur le groupe. Le free rider effect voit quelques membres d'un groupe travailler indépendamment des autres et choisir d'avancer seuls pour économiser leur temps. La convergence précoce consiste à s'accorder sur une solution sous-optimale simplement parce qu'elle permet un consensus. La domination se produit lorsqu'un membre du groupe impose son point de vue [10]. On peut vérifier aussi les intentions des étudiants : proactifs, réactifs ou passifs dans leurs activités.

### 3.4. Analyses de traces numériques d'activités

Cette méthodologie a été mise en œuvre pour récupérer les traces des apprenants sur la plate-forme, les transformer en données, puis traiter ces dernières. Sur le plan théorique, on considère que le log (connexion) informatique à partir duquel on va déduire une activité, se base sur un rapport

métonymique et synecdotique [11]. En effet, il est la conséquence d'un acte, et est en relation avec une activité. Le triplet d'activités [12], par exemple, articulant traces de connexions, durées de connexions, nombre de logs informe sur la manipulation de la plate-forme.

Une analyse des traces enregistrées dans un dispositif peut se concevoir selon une approche mixte qui combine méthode qualitative s'appuyant principalement sur des analyses de contenu des messages rédigés par les apprenants, et méthode quantitative se fondant sur un codage numérique suivi d'un traitement statistique d'une série d'informations parmi lesquelles figurent les actions effectuées par ces apprenants au sein de la plate-forme [13]. Ainsi, un travail d'analyse de participation aux forums de discussion du module GEAL fut effectué, en procédant l'analyse des traces d'interactions dans les forums pédagogiques [14].

Dans un premier temps, sont indiquées clairement les traces interceptant tous les messages et affichages (intervention et consultation des apprenants). Dans un second temps, la participation est évaluée selon le nombre de messages [11]. En effet, le simple fait de poster un message ou de déposer un document sur une plate-forme constitue un acte traduisant la participation. Une analyse thématique des échanges sur des forums apporte des compléments d'informations sur les interactions mises en œuvre. Les informations reçues furent ensuite entièrement transcrites sur



une feuille de calcul Excel. Ce dernier est le logiciel utilisé pour le traitement.

#### 4. Résultats et discussions

Les résultats sont classés en deux catégories : la première se réfère à l'enquête, la seconde à l'analyse de traces numériques des activités sur la plate-forme.

##### 4.1. Transfert des expériences vécues à travers GEAL aux autres modules

Comme la formation est destinée aux professionnels dans le domaine concerné, presque tous les apprenants sont des travailleurs en entreprise et des responsables dans plusieurs sociétés. Ils n'ont pas de temps pour étudier dans la journée. La plate-forme est conçue et configurée pour un dispositif correspondant à un tel type d'apprenant. Avec Moodle, les apprenants doivent entrer en contact avec l'environnement du dispositif de la formation, connaître un objet, agir sur lui et le transformer [15]. Ce processus met en évidence l'importance de l'action dans le développement des connaissances.

Lors de l'apprentissage de GEAL :

- Presque tous les apprenants (95%) ont rendu leur devoir dans la première séquence.
- Dans la deuxième séquence, 61% des apprenants ont effectué le test d'auto-évaluation avec de bons résultats ; la plupart des apprenants suivent les conseils et lisent le cours ;
- 56% seulement des apprenants utilisent un agenda pour noter les dates importantes.

- Tous les apprenants ont participé au forum de discussion.
- C'est à travers l'activité de la séquence 3 que les apprenants ont pratiqué pour la première fois le travail collaboratif dans cette formation. Les étudiants connaissent la nécessité, les avantages et les inconvénients du travail collaboratif. L'interaction lors de ce travail de groupe vise à soutenir les processus cognitifs d'une communauté virtuelle d'apprenants [8] et crée aussi le lien social entre pairs.

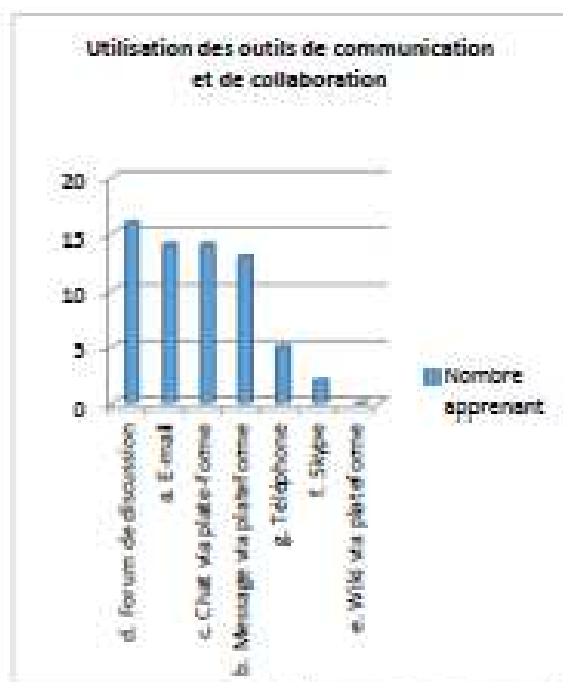
Beaucoup d'apprenants retiennent l'importance de l'esprit d'équipe et l'efficacité du travail collaboratif dans leur apprentissage. Ils sont proactifs dans leurs activités. En effet, ces expériences durant l'apprentissage de GEAL sont appliquées aux autres modules de la formation. Cependant, quelques profils sont connus. Un apprenant par exemple, a dit sans hésitation : « Je veux réussir ma formation et je n'attends pas les camarades, car la formation va vite et il faut faire le travail au jour le jour. Je n'aimerais pas être puni par la faute des autres ». Cet apprenant fait partie du profil free rider effect [10]. Le social loafing se produit, et la domination également se présente dans un petit groupe.

##### 4.2. Résultats selon les réponses au questionnaire

Après l'apprentissage du module GEAL, 81% des apprenants écrivent qu'ils ont maîtrisé les outils et procédures pour l'exploitation de la plate-forme Moodle. Un quart des répondants connaissent mal l'outil wiki, et un apprenant ne sait pas utiliser la messagerie via la plate-forme. Par contre, les 94%



des répondants ont affirmé que GEAL facilite l'usage des différents outils de communication via la plate-forme.



Source : investigation personnelle

Figure 4. Utilisation des outils de communication et de collaboration par les apprenants

Le forum de discussion est l'outil le plus utilisé dans les autres modules de la formation, selon la figure 4. Le chat et la messagerie électronique ont pris la seconde place car 88% des apprenants les utilisent. La troisième revient à la messagerie via plate-forme avec 81%. Presque tous les étudiant ont suivi la consigne d'utiliser les outils via la plate-forme.

La plupart des apprenants (94%) soulignent l'importance de GEAL. Par ailleurs, la navigation sur la plate-forme et la démarche pour le dépôt de devoir ont pris la première place dans la fonction d'aide du module, ensuite l'organisation

d'apprentissage, enfin l'usage des outils, et la gestion du temps.

A la question sur la durée du module, presque la moitié des répondants juge que deux semaines soient suffisantes pour assimiler le contenu de GEAL.

#### 4.3. Résultat selon la trace des activités numériques

L'analyse de traces des activités numérique présente, d'une part, la fréquentation des trente-huit apprenants dans GEAL pendant une durée de cinq mois et demi, soit le premier semestre de la formation, et d'autre part, la participation des apprenants aux activités du module.

Lors de l'apprentissage de GEAL, les étudiants ont beaucoup contribué aux activités proposées. Presque tous les apprenants, soient 95% ont rendu le Devoir 1. Le test d'auto-évaluation a été effectué par 61% des apprenants ; et trois quarts des apprenants ont participé au chat proposé. Tous les apprenants ont participé au forum de discussion avec divers thèmes ; ce qui confirme la réponse reçue lors du questionnaire.

##### 4.3.1 Forum de discussion, outil principal de la formation

Une première analyse de type quantitatif permet de voir que la participation au forum s'organise le plus souvent en « rafale » dans la semaine du premier regroupement. Les apprenants ont déposé en moyenne 6 messages (5,704), selon le tableau 1, avec un écart-type de 4,759 messages. Sur un total

de 38 apprenant, 37% des apprenants ont posté plus de 10 messages. Deux fils de discussions seulement furent initiés par les tuteurs : ce qui montre un rapport essentiellement proactif au forum.

Tableau 1 : Correspondance apprentissage - messages

Classe (Messages)	1	2 à 5	6 à 9	10 à 15	16 à 20	21 à 30	31 à 33
Efficacité	5	6	13	4	7	1	2
Efficacité Cumulée Croissante	5	11	24	28	35	36	38
Fréquence	0,11	0,29	0,63	0,74	0,92	0,99	1
Efficacité Cumulée Décroissante	38	31	27	14	10	5	2
Fréquence Cumulée Croissante	1,00	0,87	0,73	0,27	0,26	0,06	0,02

Source : Investigation personnelle

Tous les apprenants (100%) ont participé au forum de discussion avec divers thèmes. C'est à travers le forum « présentation » que les apprenants se connaissent mieux ; et à travers le forum « question / réponse », ils peuvent s'échanger. Ainsi, le forum est l'outil le plus utilisé dans l'activité de la séquence 3 relative au travail collaboratif. GEAL a participé, à la constitution pragmatique du lien social au sein de forums pédagogiques [16].

D'après la figure 2, l'e-mail est l'outil le plus utilisé par les apprenants auparavant. Viennent ensuite le chat et la messagerie via plate-forme. Le forum a pris la quatrième place. Après l'intégration du module GEAL, l'utilisation du forum de discussions prend la première place selon la figure 4. Il devient ainsi l'outil principal de la formation, suivi du mail, du chat et de la messagerie via la plate-forme.

Ce résultat peut s'expliquer en partie d'une part, par le fait qu'en tant qu'outil de communication et de

collaboration, l'asynchrone est très abordable pour les étudiants qui ont de difficultés sur la qualité non satisfaisante de leur connexion internet. L'asynchrone est également une solution pour les quelques étudiants qui se plaignent de la non-participation de leurs co-équipiers dans les rendez-vous synchrones.

L'utilisation de ces outils durant l'apprentissage du GEAL est par la suite appliquée dans toute la formation. Ceci est expliqué d'autre part, par le fait que les tuteurs du module incitent les apprenants à utiliser toujours les outils via la plate-forme. C'est la raison pour laquelle, l'utilisation d'e-mail est limitée. Contrairement à certaines formations à distance, cette recherche confirme celle de Roux et Mayen [17] selon lesquels le forum a une place déterminée, intégrée dans un scénario pédagogique explicite.

#### 4.4. Apprentissage de GEAL en trois semaines

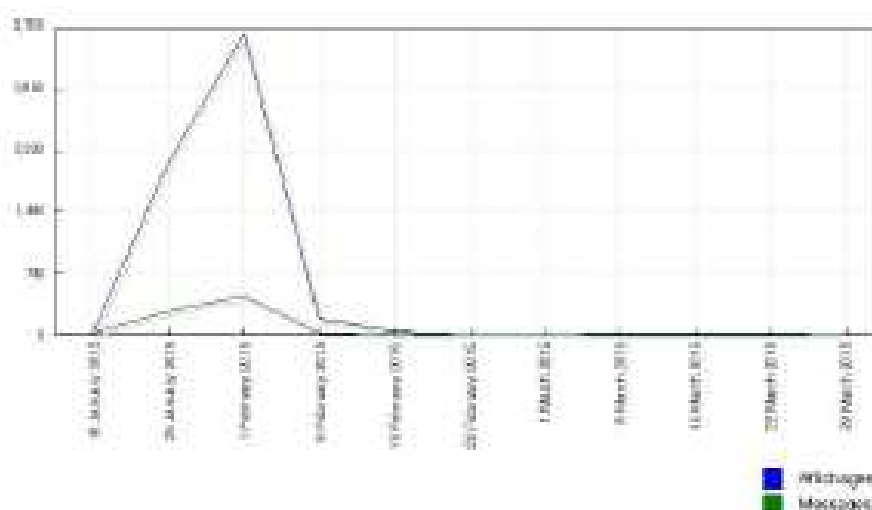
D'après la trace d'activité sur plate-forme, la durée moyenne de fréquentation dans GEAL est de 103,36 jours avec un écart-type de 45,88 jours. On observe une grande valeur de l'écart-type qui peut être interprétée comme une grande dispersion des valeurs par rapport à la moyenne. On constate que 7 apprenants seulement ont une durée de fréquentation supérieure à 100 jours. La durée moyenne de fréquentations pour les 7 étudiants est de 136,43 jours avec un écart-type de 20,13 jours, et celle pour les 31 apprenants est de 32,84 jours avec un écart-type de 22,63 jours. Pour éviter le trop grand écart entre les observations, seules les durées de fréquentations inférieures à 30 jours sont

prises en compte. Il est alors plus significatif de considérer que les apprenants ont fréquenté en moyenne le module pendant 33 jours.

Selon les réponses au questionnaire, 47% des répondants jugent que deux semaines suffisent pour assimiler le contenu de GEAL.

L'allure du rapport d'activité (Figure 5) sur la plate-forme nous donne un pic d'activités

(affichages et messages) au quatorzième jour. En effet, la participation est très active durant les deux premières semaines d'apprentissage, et cela diminue petit à petit lors de la troisième semaine. Une adaptation apparaît alors au bout de deux semaines. L'apprentissage se continue dans la troisième semaine et il y a encore quelques étudiants qui interviennent dans la quatrième semaine.



Source : Plate-forme de la formation

Figure 5. Allure d'intervention dans GEAL.

Vingt-et-un apprenant, soit 55% fréquentent le module GEAL en moins de 30 jours, selon la correspondance apprentissage-jour (tableau 2)

Tableau 2 : Correspondance apprentissage - jours

Classe (jour)	1- 15	16- 30	31- 45	46- 60	...	136- 150	151- 165
Effectif (apprenant)	7	14	3	1	...	2	2
Effectif Cumulé Cresissant	7	21	24	25	...	36	38
Fréquence Cumulée Cresissante	0,18	0,55	0,63	0,66	...	0,94	1

Source : Investigation personnelle

Alors, on peut en conclure que vingt et un jours soient trois semaines sont préférables pour l'apprentissage du module d'initiation GEAL.

#### 4.5. Analyse qualitative des discours

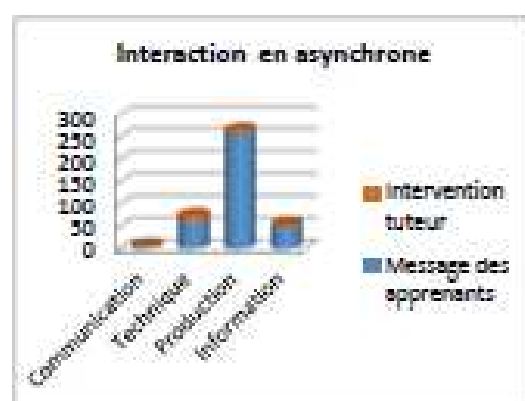
Le tableau ci-dessous montre les quatre thèmes principaux pour classer les 375 messages des apprenants et 33 messages des tuteurs sur les forums de discussion.

Tableau 3 : Analyse thématique des discussions asynchrones sur GEAL

Thème	Nombre d'intervention	
	Apprenant	Tuteur
Communication	6	3
Technique	60	16
Production	260	8
Information	49	6

Source : investigation personnelle

Les messages dans la production prennent la première place, soit 69%. Ils concernent la pratique du travail collaboratif et justifient ce que Henri et son collègue ont écrit : « L'apprentissage collaboratif est une démarche active par laquelle l'apprenant travaille à la construction de ses connaissances ... le groupe y participe comme source d'information, comme agent de motivation, comme moyen d'entraide et de soutien mutuel et comme lieu privilégié d'interaction pour la construction collective de connaissances » [17]. La représentation graphique ci-dessous met en évidence l'importance des messages de production par rapport aux autres formes de messages.



Source : investigation personnelle

Figure 6. Interaction dans les forums de discussion

L'utilisation d'un forum de discussion en contexte éducatif vise à susciter des interactions sociales, susceptibles de favoriser l'apprentissages [18].

## 5. Conclusion

Au terme de cet article, les participations aux activités, le délai d'apprentissage ainsi que l'usage des outils dans un module d'initiation d'une formation en ligne ont été analysées ; il convient de souligner les résultats suivants : tout d'abord, GEAL aide les étudiants sur le plan technique, notamment dans le choix et l'usage des outils de communication et de collaboration, et rend le forum de discussion comme outil primordial de la cognition.

Ensuite, tous les apprenants sont du même avis sur l'utilité, dans la formation, du module GEAL, qui donne les différents conseils à suivre pour faciliter l'apprentissage en ligne. Les étudiants n'ont plus de difficultés pour travailler sur la plate-forme. Ils sont aidés par GEAL dans l'organisation de leur apprentissage et la gestion de leur temps. Ce module est considéré comme accompagnateur et soutien dans l'organisation de leur parcours d'apprentissage. Il assure la fonction organisationnelle et la fonction métacognitive.

Aussi, cette investigation a permis de déterminer le délai préconisé de trois semaines pour l'apprentissage de GEAL.

Finalement, il serait intéressant de réfléchir sur la modélisation insistant sur l'utilisation de forum de

discussion comme outil cognitif d'une formation à distance.

## 6. Bibliographie

- [1] A. Jaillet, « *L'École à l'ère numérique* », Paris, L'Harmattan, Chapitre 3, p. 90-121, 2004.
- [2] M. Rakotomalala, L. Zakariasy, « *Pratiques tutorales et attentes des apprenants dans une FOAD* » - Tutoring practices and learner's expectations in an ODL, n°11 – décembre 2013, revue Franica, 2015.
- [3] Y. S. Wang, "Assessment of learner satisfaction with asynchronous electronic learning systems", in *Information and Management*, vol. 41, n° 1, p. 75-86, 2003.
- [4] E. Ben Romdhane, « *Étude des pratiques d'utilisation d'une plate-forme pédagogique et analyse de leurs relations avec la satisfaction des étudiants à l'égard du système* », Distances et médiations des savoirs n°2, 2013.
- [5] A. Tricot, F. Plégat-Soutjs, « *Pour une approche ergonomique du concept dispositif de formation à distance utilisant les TIC* ». IUFM de Midi-Pyrénées, 2003.
- [6] 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, 2003.
- [7] M. Fayol, et D. Gaonac'h, « *Le développement de la mémoire* ». Dans Blaya, A. et Lemaire, P. (Edit.) « *Psychologie du développement cognitif de l'enfant* » Paris-Bruxelles : De Boeck Université, p. 125-156, 2007.
- [8] S. Grosjean, « *L'apprentissage collaboratif à distance : Du scénario pédagogique à la dynamique interactionnelle* ». Technologies de l'Information et de la Connaissance dans l'Enseignement Supérieur et l'Industrie, Compiègne, France. Université de Technologie de Compiègne, pp.229-236, 2004.
- [9] C. Depover, et J.-J. Quintin « *Les modalités et les formes de l'enseignement à distance* », Le tutorat en formation à distance, Bruxelles, Éditions De Boeck, Coll. « Perspectives en éducation et formation », p. 29-38, 2011 b.
- [10] P. Dillenbourg, « *Pour une conception intégrée du tutorat de groupe* », Le tutorat en formation à distance, Bruxelles, Éditions De Boeck, Coll. « Perspectives en éducation et formation », p. 171-194, 2011.
- [11] J. Piaget, « *La Prise de conscience* », Paris, PUF, 1974a.
- [12] A. Jaillet, « *Y a-t-il un effet « instrument » sur l'activité des étudiants à Distance ?* », Distance et Savoirs, Vol. 3 – n°1/2003, Paris : CNED, Lavoisier, 49 - 66 (2003).
- [13] A. Strebelle et C. Depover, « *L'exploitation de la trace dans un dispositif d'apprentissage collaboratif à distance : Éléments méthodologiques et apports à l'étude du processus de modélisation* », dans F. Larose et A. Jaillet (dir.), *Le numérique dans l'enseignement et la formation. Analyses, traces et usages*. Paris, L'Harmattan, Collection « Sciences et société », 110-135, 2009.
- [14] F. Henri, D. Peraya, et B. Charlier, « *Méthodes d'analyse des forums de discussion dans leurs usages pédagogiques* ». In G. -L. Baron et E. Brullard, (Ed),

*Forums et communautés d'enseignants et d'apprentissage*, 2007.

[15] A. Jaillet, « *Traces et histoires de traces* ». Dans F. Larose et A. Jaillet (Dir.), *Traces numériques en enseignement et formation : analyses et usages*, Paris : L'Harmattan : coll. Sciences et Société, 15-36, 2009.

[16] C. Develotte, « *L'espace d'exposition discursive dans un campus numérique* ». Colloque « *Langage, objets enseignés et travail enseignant en didactique du français* », Université de Grenoble 3, 2005.

[17] F. Henri, et K. Lundgren-Cayrol, « *Apprentissage collaboratif à distance. Pour comprendre et concevoir les environnements d'apprentissage virtuels* ». Sainte-Foy, Québec : Presses Universitaires du Québec, (2001).

[18] C. Roux, et P. Mayen, « *Le forum de discussion en formation : un espace potentiel d'accès au rapport qu'entretient autrui avec les objets d'apprentissage* », Distances et médiations des savoirs N°3, 2013

### **A.1.5. Deuxième article de notoriété nationale**

M. Rakotomalala, T. B. Ravaliminoarimalalason, F. Randimbindrainibe, « *Modèles de mesure de variables latentes dans un apprentissage asynchrone* », MADA-ETI, [www.madarevues.gov.mg](http://www.madarevues.gov.mg), Vol. 1, pp 32-40, mai-2018.

#### **Résumé**

*La numérique touche à beaucoup de domaines de nos vies, de notre société et de notre culture. Cet article a été mené dans le cadre de cette culture numérique, plus précisément l'apprentissage asynchrone dans une formation ouverte et à distance. Les traces des activités de l'apprenant sur la plateforme ont été relevées et exploitées pour mesurer les effets directs et indirects de l'engagement des apprenants et de son comportement considéré comme variables latentes, sur sa cognition. Le résultat nous donne deux modèles de mesure de ces deux variables latentes selon le concept des équations structurelles de la recherche.*

**Mots-clés :** modèle de mesure, engagement, comportement, apprentissage, asynchrone,

#### **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

#### **1. Introduction**

Dans une formation ouverte et à distance (FOAD), la communauté est formée par un groupe d'apprenants, rassemblés pour une activité de formation, pour répondre à une préoccupation relative à une tâche dont le produit sera évalué.

Cet article s'intéresse aux analyses de ces forums dans cette FOAD et tente de proposer un modèle d'équation conceptuelle structurelle, afin d'examiner la relation

entre l'attitude de l'apprenant et son engagement face à son apprentissage asynchrone.

Le cadre théorique sur l'apprentissage asynchrone dans une formation à distance a été développé dans un premier temps, suivi de la méthodologie sur la modélisation d'équation structurelle. Dans un second temps, l'analyse de traces des forums de discussions et le modèle d'équation structurelle (SEM) proprement dit ont été développés. Enfin, une discussion et une perspective concluent l'article.

## 2. L'apprentissage asynchrone dans une FOAD

### 2.1. Définitions

Si la formation à distance existe depuis des décennies, la FOAD est un mode d'enseignement apparu dans les années 1990 qui rompt avec l'unité de lieux et de temps en s'appuyant sur les technologies numériques. Par définition, une FOAD est un dispositif :

- Organisé, finalisé, reconnu comme tel par les acteurs ;
- Qui prend en compte la singularité des personnes dans leurs dimensions individuelle et collective ;
- Et repose sur des situations d'apprentissage complémentaires et plurielles en termes de temps, de lieux, de médiations pédagogiques humaines et technologiques, et de ressources.

La *formation à distance* est définie comme une formation médiatisée [1]. Celle-ci fait appel à des artefacts technologiques, et à des dispositifs médiatiques.

Dans cet article, ces dispositifs s'appuient sur l'usage d'une plate-forme Moodle. Il a été conçu pour favoriser un cadre de formation socio-constructiviste et intègre des outils et des activités d'apprentissage orientés sur les communications synchrones et asynchrones. Par ailleurs, avec Moodle, l'apprenant est au centre de la formation.

Le forum est un outil de communication et outil de collaboration asynchrone dans la formation. Cet outil a beaucoup d'avantages. C'est un outil adapté aux contextes malgaches vu la qualité de connexion dans des différentes régions de Madagascar et surtout la présence de délestage à n'importe quel moment qui empêche de joindre des réunions synchrones programmées dans la

formation. Cet outil est aussi favorable pour le travail collaboratif ; il est accessible à tous, à tous les moments. L'enseignant tuteur est une personne ressource, il aide les apprenants à résoudre les problèmes techniques ou les oriente vers le responsable concerné, les réponds et prend le rôle d'animateur si nécessaire dans un travail collaboratif des apprenants. Dans un forum, chaque apprenant peut poser n'importe quelle question pédagogique (cours, activité, évaluation) ou question d'organisation (planning, groupe, emploi du temps, rendez-vous), etc., et tout le monde peut répondre aux préoccupations des autres sans attendre l'intervention du tuteur.

### 2.2. Approche d'apprentissage asynchrone dans une formation à distance

Apprendre consiste à acquérir de nouvelles connaissances, à les stocker en mémoire, à les organiser et à développer des automatismes. L'apprentissage est donc perçu comme une activité mentale réalisée individuellement à partir d'informations à traiter [02]. En 2003, Entwistle a identifié les différentes approches d'apprentissage chez les étudiants et a privilégié celles qui amélioreraient la qualité d'apprentissage [03]. Il s'agit de l'approche d'apprentissage en profondeur, indiquant l'intention de comprendre par soi-même, l'approche métacognitive supposant pouvoir prendre du recul par rapport à ses manières d'étudier, et l'organisation des études pour gérer le temps efficacement et manager ses efforts. Ces approches sont liées à l'engagement de l'étudiant dans la formation, ainsi qu'à sa motivation et ses performances.

Dans le type de formation asynchrone, l'échange avec les autres apprenants ou avec les tuteurs s'effectue via des modes de communication ne nécessitant pas une connexion simultanée. Il peut s'agir de forums de discussion ou bien encore de l'échange d'e-mails. Par



ailleurs, ce mode de formation repose souvent sur un apprentissage dit autodirigé, avec des cours, des exercices et des auto-évaluations impliquant une certaine autonomie de l'apprenant. La formation en ligne dans un environnement asynchrone implique que toute la responsabilité de l'apprentissage incombe aux apprenants.

En tant que méthode centrée sur l'apprenant, l'apprentissage asynchrone donne aux étudiants l'entière responsabilité de leur expérience de formation en ligne. L'apprentissage asynchrone exige que l'apprenant soit concentré, déterminé et qu'il possède de bonnes compétences en gestion du temps. Réussir un apprentissage asynchrone nécessite de la part des apprenants un engagement, une motivation et une discipline rigoureuse.

### 3. Methodologies

#### 3.1. Méthodologie utilisée

Les modèles d'équations structurelles ont été utiles pour attaquer de nombreux problèmes de fond et ont contribué à l'évolution de la théorie dans les sciences sociales et comportementales, dans la psychologie et l'éducation, etc. [04]. Ces méthodes permettant de tester de manière simultanée l'existence de relations causales entre plusieurs variables latentes explicatives et plusieurs variables à expliquer [05].

Une *variable manifeste* est une variable pour laquelle une mesure peut être directement recueillie (observée, mesurée, etc.). Le paradigme dominant dans la théorie des tests fut celui représenté par Churchill [06] et qui considère que les variables manifestes sont toutes supposées représenter leur variable latente. Dans ce sens, il est postulé que tous les indicateurs concordent dans leur manière de mesurer le phénomène, et permettent tous de refléter la même variable. En conséquence, on doit s'assurer de la significativité de la variable latente

construite à la base de ces indicateurs, qui doivent être significativement corrélés.

Une *variable latente* bien construite est une variable dont la variation doit s'accompagner le plus fidèlement possible par la variation de tous les indicateurs qui la composent.

Les traces d'activités des apprenants dans la plateforme de la formation sont exploitées pour mesurer les effets directs et indirects des engagements des apprenants et de leurs comportements face au forum de discussion, sur leurs cognitions qui se traduisent par leurs résultats pédagogiques.

Dans ce cas, les indicateurs constituent le reflet de la variable latente où elle demeure la cause des indicateurs, où chaque indicateur est lié à la variable latente par une équation de régression simple du type :

$$X_i = \lambda_i \xi_1 + \delta_i \quad (01)$$

Avec :  $\xi_1$  : construit latent,  $X_i$  : indicateurs,  $\lambda_i$  : coefficient représentant l'effet du construit sur l'indicateur (loading) ; et  $\delta_i$  : indicateur de l'erreur de mesure

#### 3.2. Modèle d'équations structurelles

Selon Bollen et Long [07] et Schumacker et Lomax [08], la spécification, l'identification, l'estimation, l'évaluation et la respecification du modèle représentent les étapes les plus communément admises au niveau de la conception d'un modèle d'équations structurelles.

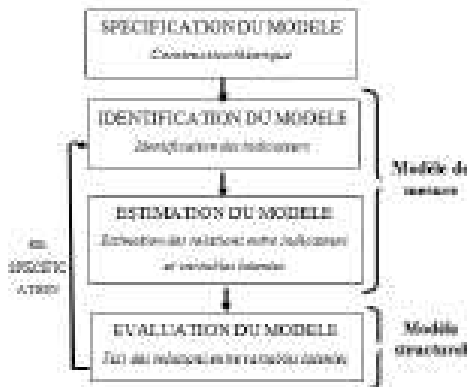


Figure 1. Les étapes de la construction du modèle des équations structurelles basée sur [09], [07] et [08]

La spécification du modèle conduit à la conception d'un « schéma de relations linéaires » qui se traduit en modèles de mesure et en modèles de structure. [10]

L'identification permet d'attribuer une seule solution pour chacun des paramètres à estimer. Un modèle conduit à exprimer la matrice de variance/covariance des variables manifestes  $\Sigma$  en fonction d'un ensemble de paramètres  $\theta$ . Ce modèle est identifié si  $\Sigma(\theta) = \Sigma(\theta')$  implique  $\theta = \theta'$ . En pratique, l'identification d'un modèle suppose la satisfaction de deux conditions : à savoir la condition d'ordre (condition nécessaire) et la condition de rang (condition nécessaire et suffisante).

La condition d'ordre dépend impérativement du degré de liberté  $ddl$  [10] et se rapporte à la correspondance entre les paramètres à estimer et le nombre de variances/covariances des variables observées [11]. Le degré de liberté s'exprime de la manière suivante [08] :

$$ddl = \left( P \frac{(P+1)}{2} \right) - N \quad (02)$$

Avec :  $P$  est le nombre d'indicateurs de mesure du modèle,  $N$  : le nombre de coefficients à estimer du modèle et  $ddl$  est le degré de liberté.

Par conséquent, trois situations sont possibles à savoir la sous identification ( $ddl < 0$ ), la juste identification

( $ddl = 0$ ) et la sur identification du modèle ( $ddl > 0$ ). D'après ces trois situations, la condition d'ordre est vérifiée lorsque le modèle est juste identifié ou encore sur identifié ( $ddl \geq 0$ ).

Par ailleurs, la condition de rang permet d'avoir une seule solution vis-à-vis de chacun des paramètres à estimer.

L'estimation consiste à déterminer les valeurs relatives aux différents paramètres du modèle à tester. La technique d'estimation du Maximum de Vraisemblance *ML* (Maximum Likelihood) et des Moindres Carrés Généralisés *GLS* (Generalized Least Squares) sont les méthodes les plus utilisées pour l'estimation d'un modèle étant donné qu'elles figurent dans la plupart des logiciels de traitement des données [08] [09]. L'expression de la méthode de Maximum de vraisemblance se présente comme suit :

$$F_{ML} = tr(S \cdot \Sigma^{-1}) - p + \ln |\hat{\Sigma}| - \ln |S| \quad (03)$$

Si  $\theta$  soit un vecteur de paramètre libre estimé par le modèle, on peut écrire la relation (03) comme suit :

$$F_{ML}(\theta) = tr(S \cdot \Sigma^{-1}(\theta)) - p + \ln |\Sigma(\theta)| - \ln |S| \quad (04)$$

Avec :  $S$  : la matrice de variance-covariance observée,  $\ln$  : la fonction logarithmique naturelle,  $tr(.)$  : la fonction trace d'une matrice,  $\Sigma$  : la matrice de variances/covariances estimées,  $p$  : le nombre d'indicateurs de mesure ou variables observées, et  $|S|$  : Déterminant d'une matrice  $S$ .

Pour l'évaluation du modèle, la qualité d'un modèle peut être appréciée à travers l'évaluation d'un ensemble d'indices d'ajustements. On dénombre plusieurs indices d'ajustements qui se regroupent en trois catégories : à savoir les indices absolus, les indices incrémentaux et les indices de parcimonie.

Les indices absolus sont qualifiés comme étant des indices classiques [09]. Ils examinent le niveau de correspondance entre le modèle proposé et les données observées [08]. Les indices incrémentaux permettent d'évaluer la contribution du modèle étudié par rapport à un modèle restrictif, c'est-à-dire une comparaison entre le modèle estimé et le modèle de référence ayant une corrélation nulle entre les données observées [10]. Les indices de parcimonie contrôlent la surestimation du modèle. Ils déterminent le nombre nécessaire de paramètres à estimer permettant d'atteindre le niveau d'ajustement spécifique [08]. L'objectif étant d'aboutir à un meilleur dosage entre la maximisation de l'ajustement et la minimisation du nombre des coefficients estimés [10].

La respecification du modèle représente une phase de réflexion qui offre l'opportunité de reconsidérer la conception du modèle et d'en proposer des modifications éventuelles, en tenant compte du cadre théorique de la recherche en question [10].

À un niveau de la présentation des résultats, ces derniers doivent incorporer des informations pertinentes et précises, permettant d'aboutir facilement à des solutions aux problèmes susceptibles d'être rencontrés dans l'analyse.

#### 4. Résultats

Le résultat se divise en deux parties, tout d'abord l'analyse des activités d'apprentissage asynchrone en le forum de discussion dans laquelle un module d'initiation a été mobilisé. Ensuite, un modèle conceptuel d'équations structurelles de la recherche sera proposé, et les différentes étapes seront réalisées afin de sortir les deux modèles de mesures validées du SEM. R est utilisé pour le traitement statistique des données.

#### 4.1. Analyse de l'activité

L'objectif est d'analyser l'apprentissage asynchrone dans une formation à distance. L'analyse se centre préférentiellement sur le forum des apprenants. La discussion se déroule sur la plate-forme dans un module d'initiation (GEAL) au début de la formation. Les apprenants interagissent entre eux ou avec le tuteur. Trois formations ouvertes et à distance sont concernées. L'apprenant a besoin de connexion pour pouvoir accéder à la plate-forme et y travailler.

##### 4.1.1. Les indicateurs du comportement de l'apprenant

Les traces de l'activité des participants dans une formation en ligne peuvent être transformées pour représenter une mesure de l'engagement comportemental [12], ce qui a des incidences considérables sur la manière de les organiser et les simplifier afin qu'elles correspondent bien aux construits théoriques. Dans une initiative de recherche menée actuellement, l'attitude de l'apprenant vis-à-vis d'un forum de discussion a été modélisée comme un continuum de participation aux diverses interventions accessibles sur des forums de discussion.

Du point de vue analyse de traces, une approche quantitative préliminaire à l'étude plus spécifique en statistiques est nécessaire. La participation en nombre d'intervention : la consultation des forums, la consultation des discussions, ajout ou mise à jour d'un message à une discussion en cours, et la création de thème nouveau sont des indicateurs pris en compte. L'ensemble de ces indicateurs caractérise le comportement de l'apprenant dans un forum de discussion.

##### 4.1.2. Les indicateurs de l'engagement de l'apprenant

L'engagement de l'apprenant est mesuré par la régularité au forum, la sensation à l'aise envers le forum de

discussion, la réalisation des activités depuis la plateforme malgré les contraintes existantes et la consécration plus d'effort afin de rester actif dans l'apprentissage asynchrone [13]. Dans la présente recherche, les modèles explorent les différentes dimensions communicationnelles, motivationnelles, et socio-cognitives de l'engagement en contexte de formation à distance. Ces trois dimensions sont en interaction continue et réciproque selon des importances variables et contingentes aux activités. L'analyse qualitative effectuée sur les contenus du forum de discussion nous donne les indicateurs à mesurer selon les trois dimensions ci-dessus pour l'engagement de l'apprenant. Ainsi, la collecte et l'analyse de traces de l'activité principale sont basées par les messages. Ces derniers sont classés selon le message de communication, l'indicateur de motivation et l'interaction pour la collaboration. Ces trois classes constituent l'engagement de l'apprenant dans un forum de discussion.

#### 4.2. Modélisation par des équations structurelles

Le modèle conceptuel de la recherche présente les différents liens de causalité entre les construits théoriques et montre que l'apprentissage asynchrone pourrait jouer un rôle médiateur au niveau de la relation entre l'engagement de l'apprenant et l'exploitation du forum de discussion. Le modèle causal se compose de deux modèles de mesure et des modèles de structure. Il englobe l'ensemble de variables observables indépendantes ( $X_i$ ), de variables observables dépendantes ( $Y_i$ ), d'une variable latente explicative ( $E$ ), de deux variables latentes à expliquer ( $A$ ) et ( $C$ ) et des termes d'erreurs ( $\epsilon_i$ ,  $\delta$ ,  $\beta$  et  $\gamma_i$ ). Par conséquent, il est possible de distinguer entre deux types d'équations à savoir les équations du modèle de mesure et les équations du modèle de structure.

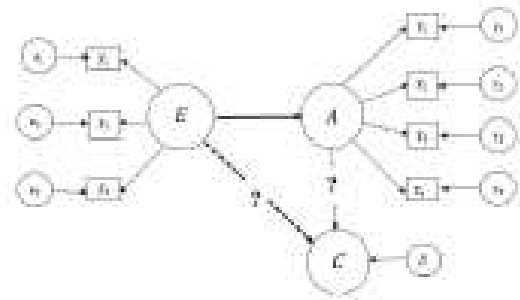


Figure 2 : Le modèle conceptuel de la recherche

Avec :

$E$  : Engagement de l'apprenant dans un forum de discussion

$A$  : Attitude de l'apprenant dans son apprentissage asynchrone

$C$  : Acquisition de Connaissance

$E$  et  $A$  sont envisagés comme des variables réflexives, dont les indicateurs sont :

$X_1 = Collab$  : Interaction collaborative dans le forum de discussion,

$X_2 = Comm$  : Communication avec les protagonistes de la formation (pairs, tuteur, ...), afin de partager des informations entre membres et aussi pour s'organiser.

$X_3 = Motiv$  : Motivation dans l'apprentissage, qui est la force même qui les pousse à être performants et couronnés de succès dans son apprentissage à distance [13]

$Y_1 = Consu_F$  : forum consulté (forum\_view forum)

$Y_2 = Consu_D$  : discussion consultée (forum\_view discussion)

$Y_3 = Envoi_M$  : Contenu posté (forum\_add post)

$Y_4 = Créat_D$  : discussion créée (forum\_add discussion)

Dans le modèle proposé dans la figure 2, l'engagement est le produit d'une motivation et d'un apprentissage social, c'est dans et par les interactions avec ses pairs que l'apprenant construit son engagement.

Par définition, l'engagement dans l'apprentissage collaboratif est un processus mutuel. Il consiste en un effort conscient, volontaire et continu de la part de tous

les participants, de mise en commun et de coordination en vue de résoudre ensemble le problème.

### 4.3. Les modèles de mesure

Deux modèles de mesure par équations structurelles ont été élaborés : L'attitude de l'apprenant dans l'apprentissage asynchrone, et son engagement dans une formation à distance.

On va appliquer ces modèles sur des données réelles d'une FOAD de l'IST d'Antananarivo. Il s'agit de traces d'activités asynchrones de 4804 événements de 13 apprenants à l'espace de 19 jours, du 20 octobre au 07 novembre 2016.

#### 4.3.1. Le modèle de mesure de l'attitude de l'apprenant

##### a. Première spécification du modèle

Le modèle de mesure de la variable latente A : attitude de l'apprenant face au forum de discussion, peut traduire par les équations suivantes :

$$\begin{aligned} Y_1 &= \alpha_1 A + \gamma_1 \\ Y_2 &= \alpha_2 A + \gamma_2 \\ Y_3 &= \alpha_3 A + \gamma_3 \\ Y_4 &= \alpha_4 A + \gamma_4 \end{aligned} \quad (05)$$

Il peut s'écrire sous la forme matricielle :  $Y = \alpha A + \gamma$  et forme une régression linéaire.

La matrice de covariance entre les quatre variables montre qu'il y a des covariances négatives entre les deux variables « Creat\_D » et « consu\_D ».

Pour l'estimation du modèle, nous avons utilisé l'estimation par le maximum de vraisemblance. Le modèle ne converge qu'après 7959 itérations, avec un test sur un échantillon de 13 observations.

Le degré de liberté  $ddl = p(p+1)/2 - N$  avec  $p = 4$  et  $N = 6$  est égal à 2, qui est vérifié par le « Degrees of freedom » retourné, indique la sur-identification du modèle.

On note que l'erreur sur la variable Creat\_D est de -1112,53 (variance négative) qui définit une raison de respecification de notre modèle.

Le logarithme de la vraisemblance (logl) : -164,987 indique que le modèle n'a pas de vraisemblance par rapport aux données réelles.

##### b. Respécification du modèle

Les deux variables « Envoi\_M » et « Creat\_D » expriment la participation de l'apprenant au forum de discussion ; on va les fusionner (Envoi\_M\_2). Par conséquent, nous n'avons que trois variables observées au lieu de quatre auparavant. Nous avons repassé les étapes précédentes en traitant les mêmes données avec le nouveau modèle.

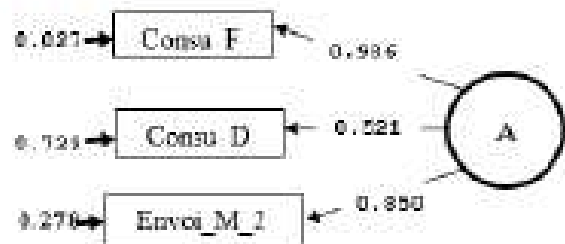


Figure 3 : Modèle de mesure de la variable latente A avec les paramètres estimés

Le degré de liberté  $ddl = p(p+1)/2 - N$  devient nul. Le modèle est juste identifié. En ce qui concerne l'évaluation du modèle, le minimum de la fonction estimation (fmin) est égal à 0, nous indique que c'est un bon modèle par rapport au modèle précédent qui a pour  $fmin = 0,677$ .

Le logarithme de la vraisemblance (logl) du modèle respécifié est -144,312 qui est meilleur que celui du modèle initial -164,987.

Le  $\chi^2$  (chisq) du modèle est égal à 0. Il n'y a aucune présomption contre l'hypothèse nulle. Les données réelles n'ont pas de grande conformité au modèle.

statistique sous-jacent, ceci peut être dû à leur répartition.

Le degré de liberté du modèle (df) est 0, comme on l'a calculé manuellement. On a estimé 6 paramètres (npar) du modèle.

On peut valider ce modèle respecifié par rapport au premier. Le khi deux ne nous empêche pas cette validation, la valeur de p ne devrait jamais être utilisée pour valider une hypothèse à partir de données puisque c'est l'inverse qui est calculé.

#### 4.3.2. Le modèle de mesure de la variable latente E

Le modèle de mesure des engagements de l'apprenant peut traduire par les équations suivantes :

$$\begin{aligned} X_1 &= \lambda_1 E + e_1 \\ X_2 &= \lambda_2 E + e_2 \\ X_3 &= \lambda_3 E + e_3 \end{aligned} \quad (06)$$

Ils peuvent s'écrire sous la forme matricielle :  $X = \lambda E + e$  qui forme une régression linéaire.

En repassant les différentes étapes pour le modèle de mesure de l'engagement de l'apprenant dont la spécification du modèle et l'estimation du modèle, nous avons obtenu le modèle :

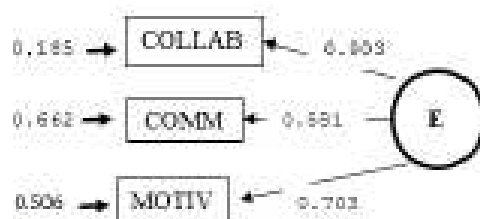


Figure 4 : Modèle de mesure de la variable latente E avec les paramètres estimés

Le degré de liberté du modèle est encore nul. Ceci nous donne une juste identification du modèle.

#### Validation du modèle

```
> 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
```

Le minimum de la fonction estimation (fmin) est égal à 0. Le logarithme de la vraisemblance (logl) est -61,47. On note une vraisemblance exacte pour une valeur nulle de ce logarithme de vraisemblance, et là, on approche petit à petit de cette valeur. On a vu logl = -144,312 pour le modèle des attitudes. Le  $\chi^2$  (chisq) est encore à 0. Le nombre de paramètre estimé (npar) du modèle est 6. Ces données nous donnent l'idée d'encore valider le modèle ainsi décrit.

#### 5. Conclusion

Les interactions entre les étudiants en formation à distance visent un apprentissage profond et ont été renforcées par une méthode pédagogique en l'apprentissage collaboratif. La participation aux échanges au sein de la plate-forme, la richesse et la qualité des interactions entre apprenants sont reconnues comme des facteurs bénéfiques à l'engagement, la satisfaction et la réussite des apprenants en formation à distance.

Quand même, la participation dans un forum de discussion, la lecture des interventions des autres apprenants et l'interaction entre pairs entraînent une sensation d'appartenance à un groupe. Cet engagement est considéré comme une variable latente et a été mesuré par un modèle de mesure dans une équation structurelle.

Pour conclure, nous avons obtenu les deux modèles de mesure validés. La modélisation structurelle de l'apprentissage asynchrone en manipulant les deux modèles avec les données réelles dans trois formations ouvertes et à distance nous permettront de valider les relations de causalité du résultat pédagogique, de l'engagement et de l'attitude de l'étudiant face au forum de discussion. L'effet de développement de connaissance via le forum de discussion sera fortement démontré. En d'autres termes, les apprenants engagés,

motivées et collaborées dans une formation à distance développent leur connaissance et leur savoir-faire de manière continue. L'impact de l'exploitation de forum sur l'acquisition de connaissance sera démontré empiriquement.

## 6. Bibliographie

- [01] D. Parrya, « *La formation à distance : un dispositif de formation et de communication médiatisée. Une approche des processus de médiation et de médiation. THÈSE et développement* », n°1, 2005.
- [02] M. Fayol, et D. Geonac'h, « *Le développement de la mémoire* ». Dans A. Blaya, et P. Lemaire, (Edit.) *Psychologie du développement cognitif de l'enfant*. Paris-Bruxelles : De Boeck Université, pp. 125-156, 2007.
- [03] N. Entwistle, "Concepts and conceptual frameworks underpinning the ETI Project", Edinburgh, School of Education, University of Edinburgh, 2003.
- [04] K. G. Jöreskog, D. Sörbom, « *Recent developments in structural equation modeling* ». *Journal of Marketing Research*, 19, pp 404-416, 1982.
- [05] Lacroix, « *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.
- [06] G. A. Churchill, « *A paradigm for developing better Measures of marketing constructs* », *Journal of Marketing Research*, vol. 16, pp 64-73, Feb 1979.
- [07] K.A. Bollen, J.S. Long, « *Testing structural equation models* ». Newbury Park, CA: Sage, 1993.
- [08] R.E. Schumaker, R.G. Lomax, « *A beginner's guide to structural equation modelling* », Lawrence Erlbaum Associates, 2ème édition, London, 2004.
- [09] L. Chaput, « *Modèles contemporains en gestion* », Presses de l'Université du Québec, 2006.
- [10] P. Roussel, F. Durrieu, E. Campoy, A. El Akraoui, « *Méthodes d'équations structurelles : Recherches et applications en gestion* », Édition Economica, Paris, 2002.
- [11] R.H. Hoyle, « *Structural equation modelling: Concepts issues and applications* », Edition Sage, London, 1995.
- [12] B. Poellnber, N. Ray 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.

## A.1.6. Deuxième article de notoriété internationale

M. Rakotomalala, T. B. Ravaliminoarimalalason, F. Randimbindrainibe, « *Asynchronous learning model in an ODL* », in International Journal of Engineering Research & Technology, Vol. 7 Issue 08, pp. 174-179, Août-2018.

**Abstract**— The analysis of the traces of digital activity and the content analysis of a discussion forum on an Open and Distance Learning (ODL) platform form a database to be processed on R, in order to release an asynchronous learning model, according to the approach by Linear Structural Relation (LISREL) of Structural Equation Modeling (SEM).

**Keywords**— Model, Asynchronous learning, involvement, collaboration, educational outcome, ODL

### I. INTRODUCTION

The appropriation of methods and techniques of distance learning is not obvious to students. Also, most tutors do not help learners solve technical problems [1]. To fill all these gaps, a "Student Guide for Online Learning" (SGOL) entry module was designed and subsequently introduced in three open and distance learning courses. It aims to support learners in taking charge of the platform and provides the tools, advice and methods necessary to succeed in open and distance learning.

The idea of exploiting this instrument mainly has been advanced. Reinforcement of learning in ODL devices was envisaged, through modeling of learning engineering. It was within this framework that research was conducted and was based on the analysis of ODL. A study that coming from cognitive science, in the field of engineering science.

Modeling focuses on learners' knowledge management and data analysis on the platform. In other words, how to establish a model of Structural Equation Modeling (SEM), which can apply to any kind of open and distance learning, according to the analysis of traces of learning activities on a platform? The problem was about the distance learning system. The aim was to study the behavior of the learners in front of the discussion forum and to know if in an online teaching, the engagement in an interaction between students can be valued. Which model of asynchronous learning could be a tool for improving teaching effectiveness?

The aim of the study was to present an online learning model. To achieve this goal, firstly, the trace of asynchronous activities was analyzed, then a structure model based on SEM was developed.

### II. MATERIALS

Data collection was conducted in 2016, 2017 and 2018. Three L3 ODLs of IST: (IST: High Institute of Technology, in Madagascar) were used as research ground: Automatic System Engineering (ASE), Transit and International Trade (TIT) and Logistic Operations and Transportation (LOT). The distance learning initiation module called SGOL or "Student's Guide to Online Learning" was put into play. It contains four forums for discussion. A digital activity trace analysis grid was used with six promotions, 109 new learners, 12393 past events and 1072 messages to analyze.

### III. METHODOLOGIES

A motivated learner communicates with his entourage in a training device, shares his knowledge via the platform, develops his know-how and his knowledge, and collaborates with the members of his group for a requested work. The evaluation of these three measurement variables in the learning of the initiation module made it possible to measure the latent variable "involvement" of the learner. Similarly, forum consultation, discussion consultation, learner participation by sending and updating messages, were all listed to be able to measure the learner's attitude when learning from a distance. Thus, the two models validated in [2] were used for all applications in the rest of the study, in order to leave the structure model of the research. Is there a causal relationship between the learner's involvement and his attitude in the discussion forum? Does involvement enable the development of learners' knowledge? Does the learner's attitude in the forum have a positive impact on their learning outcome? These are the different research questions that will be answered in this article.

### IV. RESULTS

The overall result of the research primarily covered the innovative strategy for distance learning. The expected result was the provision of an asynchronous learning model in ODL. This model can be adapted to the Malagasy context.

#### A. Asynchronous learning model

The structure model made it possible to examine the link between the different latent variables. The structural equation method represents a multivariate technique that combines measurement models and structural models while simultaneously examining a series of linear relationships



between observed variables and latent variables on the one hand, and between sets of latent variables on the other hand [3].

#### 1) Model identification

The data obtained from the analysis of digital traces of the activities and the content analysis of all existing forums were the basis of treatment on R. The two latent variables A and E were determined from these analyses according to the two previously validated measurement models [2]. The article presents the determination of the third latent variable C, which has a direct relationship with the educational outcome of the training. Six cases are likely to happen for the model structure equation. The figure below shows the identified model:

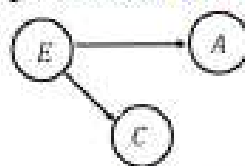


Fig. 1. Identified structure model

#### 2) Structured model with estimated parameters

The structure model below was obtained at the outcome of the processing of the first proposal. It was noticed that the model converges after 71 iterations.

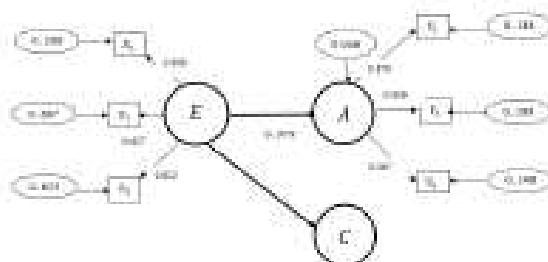




Fig. 5. Processing of the Threshold as Blood

It was found, according to the graph on Fig. 5, that starting from the value " $C_m$ " threshold of the cognition, the model gave the result 1, in other words, starting from the threshold value of the cognition, the learner has passed the training.

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	bb	bc	bd	be	bf	bg	bh	bi	bj	bk	bl	bm	bn	bo	bp	bq	br	bs	bt	bu	bv	bw	bx	by	bz	ca	cb	cc	cd	ce	cf	cg	ch	ci	cj	ck	cl	cm	cn	co	cp	cq	cr	cs	ct	cu	cv	cw	cx	cy	cz	da	db	dc	dd	de	df	dg	dh	di	dj	dk	dl	dm	dn	do	dp	dq	dr	ds	dt	du	dv	dw	dx	dy	dz	ea	eb	ec	ed	ee	ef	eg	eh	ei	ej	ek	el	em	en	eo	ep	eq	er	es	et	eu	ev	ew	ex	ey	ez	fa	fb	fc	fd	fe	ff	fg	fh	fi	fj	fk	fl	fm	fn	fo	fp	fq	fr	fs	ft	fu	fv	fw	fx	fy	fz	ga	gb	gc	gd	ge	gf	gg	gh	gi	gj	gk	gl	gm	gn	go	gp	gq	gr	gs	gt	gu	gv	gw	gx	gy	gz	ha	hb	hc	hd	he	hf	hg	hh	hi	hj	hk	hl	hm	hn	ho	hp	hq	hr	hs	ht	hu	hv	hw	hx	hy	hz	ia	ib	ic	id	ie	if	ig	ih	ii	ij	ik	il	im	in	io	ip	iq	ir	is	it	iu	iv	iw	ix	iy	iz	ja	jb	jc	jd	je	jf	jj	jh	ji	jk	jl	jm	jn	jo	jp	jq	jr	js	jt	ju	jv	jw	jx	ky	kz	la	lb	lc	ld	le	lf	lg	lh	li	lj	lk	ll	lm	ln	lo	lp	lq	lr	ls	lt	lu	lv	lw	lx	ly	lz	ma	mb	mc	md	me	mf	mg	mh	mi	mj	mk	ml	mm	mn	mo	mp	mq	mr	ms	mt	mu	mv	mw	mx	my	mz	na	nb	nc	nd	ne	nf	ng	nh	ni	nj	nk	nl	nm	nn	no	np	nq	nr	ns	nt	nu	nv	nw	nx	ny	nz	oa	ob	oc	od	oe	of	og	oh	oi	oj	ok	ol	om	on	oo	op	oq	or	os	ot	ou	ov	ow	ox	oy	oz	pa	pb	pc	pd	pe	pf	pg	ph	pi	pj	pk	pl	pm	pn	po	pp	pq	pr	ps	pt	pu	pv	pw	px	py	pz	qa	qb	qc	qd	qe	qf	qg	qh	qi	qj	qk	ql	qm	qn	qo	qp	qq	qr	qs	qt	qu	qv	qw	qx	qy	qz	ra	rb	rc	rd	re	rf	rg	rh	ri	rj	rk	rl	rm	rn	ro	rp	rq	rr	rs	rt	ru	rv	rw	rx	ry	rz	sa	sb	sc	sd	se	sf	sg	sh	si	sj	sk	sl	sm	sn	so	sp	sq	sr	ss	st	su	sv	sw	sx	sy	sz	ta	tb	tc	td	te	tf	tg	th	ti	tj	tk	tl	tm	tn	to	tp	tq	tr	ts	tt	tu	tv	tw	tx	ty	tz	ua	ub	uc	ud	ue	uf	ug	uh	ui	uj	uk	ul	um	un	uo	up	uq	ur	us	ut	uu	uv	uw	ux	uy	uz	va	vb	vc	vd	ve	vf	vg	vh	vi	vj	vk	vl	vm	vn	vo	vp	vq	vr	vs	vt	vu	vv	vw	vx	vy	vz	wa	wb	wc	wd	we	wf	wg	wh	wi	wj	wk	wl	wm	wn	wo	wp	wq	wr	ws	wt	wu	wv	ww	wx	wy	wz	xa	xb	xc	xd	xe	xf	xg	xh	xi	xj	xk	xl	xm	xn	xo	xp	xq	xr	xs	xt	xu	xv	xw	xx	xy	xz	ya	yb	yc	yd	ye	yf	yg	yh	yi	yj	yk	yl	ym	yn	yo	yp	yq	yr	ys	yt	yu	yv	yw	yx	yy	yz	za	zb	zc	zd	ze	zf	zg	zh	zi	zj	zk	zl	zm	zn	zo	zp	zq	zr	zs	zt	zu	zv	zw	zx	zy	zz
1	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000																																																																																																																																																																																																																																																																																																																																																																																			

Fig. 6. Data processing on board

With R, it is also possible to determine this threshold. According to Fig. 7, below, using locator () R-function,  $C_0$  was

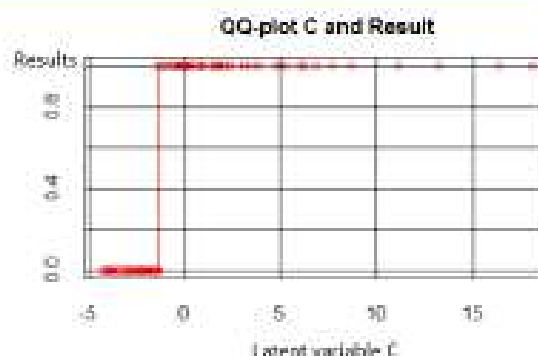


Fig. 7. Distribution of the educational results

This last value was used in the following analysis:

*A. Remita* *Forecast for the current academic year*

Open and distance learning ASE is recruiting its second class in the 2017/2018 academic year. Fifteen new learners followed the tutorial of the SGOL introductory module at the beginning of the training. After the digital trace analysis of activities on SGOL, 2708 events were parsed. The content analysis of 260 messages on forums made it possible to measure the latent variable E, whereas the processing carried out from these events made it possible to measure A.

Country	CORR	CORR2	MEAN	MEAN2	MEAN3	MEAN4	MEAN5
1	12	2	2	2	24	80	1
2	13	3	3	3	40	40	2
3	10	3	3	3	31	80	3
4	10	3	3	3	43	80	4
5	10	3	3	3	123	120	5
6	1	3	3	3	24	74	6
7	13	3	3	3	88	152	7
8	3	3	3	3	13	37	8
9	10	3	3	3	74	124	9
10	10	3	3	3	88	120	10
11	10	3	3	3	31	74	11
12	10	3	3	3	54	120	12
13	3	3	3	3	10	28	13
14	10	3	3	3	10	80	14
15	3	3	3	3	20	100	15

## Legend

- COLLAB: Collaboration of the learner
- COMM: Communication
- MOTIV: Motivation
- Conts\_E: Forum consultation
- Conts\_D: Discussion consultation
- Messs\_M: Message sending / forum participation

Fig. 8. Data after trace and content analysis.

As a result of the application of the research model, the result 10 admitted among the 15 new learners of the second class of ASE 2018 was obtained.

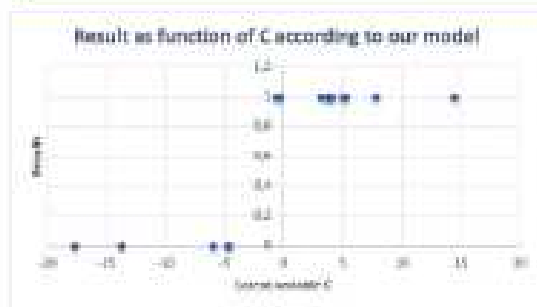


Fig. 9. Result forecast for ASE 2018

#### F. General production

Data with values varying from 0 to 20 for each variable COLLAB, COMM, MOTIV, Censu\_F, Censu\_D, Envoi\_M was created.

	COLLAB	COMM	MOTIV	Censu_F	Censu_D	Envoi_M	A	B	C	R
2007	0	0	0	0	0	0	1.00	1.00	1.00	1
2008	0	0	0	0	0	0	1.00	1.00	1.00	1
2009	0	0	0	0	0	0	1.00	1.00	1.00	1
2010	0	0	0	0	0	0	1.00	1.00	1.00	1
2011	0	0	0	0	0	0	1.00	1.00	1.00	1
2012	0	0	0	0	0	0	1.00	1.00	1.00	1
2013	0	0	0	0	0	0	1.00	1.00	1.00	1
2014	0	0	0	0	0	0	1.00	1.00	1.00	1
2015	0	0	0	0	0	0	1.00	1.00	1.00	1
2016	0	0	0	0	0	0	1.00	1.00	1.00	1
2017	0	0	0	0	0	0	1.00	1.00	1.00	1
2018	0	0	0	0	0	0	1.00	1.00	1.00	1
2019	0	0	0	0	0	0	1.00	1.00	1.00	1
2020	0	0	0	0	0	0	1.00	1.00	1.00	1
2021	0	0	0	0	0	0	1.00	1.00	1.00	1
2022	0	0	0	0	0	0	1.00	1.00	1.00	1
2023	0	0	0	0	0	0	1.00	1.00	1.00	1
2024	0	0	0	0	0	0	1.00	1.00	1.00	1
2025	0	0	0	0	0	0	1.00	1.00	1.00	1
2026	0	0	0	0	0	0	1.00	1.00	1.00	1
2027	0	0	0	0	0	0	1.00	1.00	1.00	1
2028	0	0	0	0	0	0	1.00	1.00	1.00	1
2029	0	0	0	0	0	0	1.00	1.00	1.00	1
2030	0	0	0	0	0	0	1.00	1.00	1.00	1
2031	0	0	0	0	0	0	1.00	1.00	1.00	1
2032	0	0	0	0	0	0	1.00	1.00	1.00	1
2033	0	0	0	0	0	0	1.00	1.00	1.00	1
2034	0	0	0	0	0	0	1.00	1.00	1.00	1
2035	0	0	0	0	0	0	1.00	1.00	1.00	1
2036	0	0	0	0	0	0	1.00	1.00	1.00	1
2037	0	0	0	0	0	0	1.00	1.00	1.00	1
2038	0	0	0	0	0	0	1.00	1.00	1.00	1
2039	0	0	0	0	0	0	1.00	1.00	1.00	1
2040	0	0	0	0	0	0	1.00	1.00	1.00	1
2041	0	0	0	0	0	0	1.00	1.00	1.00	1
2042	0	0	0	0	0	0	1.00	1.00	1.00	1
2043	0	0	0	0	0	0	1.00	1.00	1.00	1
2044	0	0	0	0	0	0	1.00	1.00	1.00	1
2045	0	0	0	0	0	0	1.00	1.00	1.00	1
2046	0	0	0	0	0	0	1.00	1.00	1.00	1
2047	0	0	0	0	0	0	1.00	1.00	1.00	1
2048	0	0	0	0	0	0	1.00	1.00	1.00	1
2049	0	0	0	0	0	0	1.00	1.00	1.00	1
2050	0	0	0	0	0	0	1.00	1.00	1.00	1
2051	0	0	0	0	0	0	1.00	1.00	1.00	1
2052	0	0	0	0	0	0	1.00	1.00	1.00	1
2053	0	0	0	0	0	0	1.00	1.00	1.00	1
2054	0	0	0	0	0	0	1.00	1.00	1.00	1
2055	0	0	0	0	0	0	1.00	1.00	1.00	1
2056	0	0	0	0	0	0	1.00	1.00	1.00	1
2057	0	0	0	0	0	0	1.00	1.00	1.00	1
2058	0	0	0	0	0	0	1.00	1.00	1.00	1
2059	0	0	0	0	0	0	1.00	1.00	1.00	1
2060	0	0	0	0	0	0	1.00	1.00	1.00	1
2061	0	0	0	0	0	0	1.00	1.00	1.00	1
2062	0	0	0	0	0	0	1.00	1.00	1.00	1
2063	0	0	0	0	0	0	1.00	1.00	1.00	1
2064	0	0	0	0	0	0	1.00	1.00	1.00	1
2065	0	0	0	0	0	0	1.00	1.00	1.00	1
2066	0	0	0	0	0	0	1.00	1.00	1.00	1
2067	0	0	0	0	0	0	1.00	1.00	1.00	1
2068	0	0	0	0	0	0	1.00	1.00	1.00	1
2069	0	0	0	0	0	0	1.00	1.00	1.00	1
2070	0	0	0	0	0	0	1.00	1.00	1.00	1
2071	0	0	0	0	0	0	1.00	1.00	1.00	1
2072	0	0	0	0	0	0	1.00	1.00	1.00	1
2073	0	0	0	0	0	0	1.00	1.00	1.00	1
2074	0	0	0	0	0	0	1.00	1.00	1.00	1
2075	0	0	0	0	0	0	1.00	1.00	1.00	1
2076	0	0	0	0	0	0	1.00	1.00	1.00	1
2077	0	0	0	0	0	0	1.00	1.00	1.00	1
2078	0	0	0	0	0	0	1.00	1.00	1.00	1
2079	0	0	0	0	0	0	1.00	1.00	1.00	1
2080	0	0	0	0	0	0	1.00	1.00	1.00	1
2081	0	0	0	0	0	0	1.00	1.00	1.00	1
2082	0	0	0	0	0	0	1.00	1.00	1.00	1
2083	0	0	0	0	0	0	1.00	1.00	1.00	1
2084	0	0	0	0	0	0	1.00	1.00	1.00	1
2085	0	0	0	0	0	0	1.00	1.00	1.00	1
2086	0	0	0	0	0	0	1.00	1.00	1.00	1
2087	0	0	0	0	0	0	1.00	1.00	1.00	1
2088	0	0	0	0	0	0	1.00	1.00	1.00	1
2089	0	0	0	0	0	0	1.00	1.00	1.00	1
2090	0	0	0	0	0	0	1.00	1.00	1.00	1
2091	0	0	0	0	0	0	1.00	1.00	1.00	1
2092	0	0	0	0	0	0	1.00	1.00	1.00	1
2093	0	0	0	0	0	0	1.00	1.00	1.00	1
2094	0	0	0	0	0	0	1.00	1.00	1.00	1
2095	0	0	0	0	0	0	1.00	1.00	1.00	1
2096	0	0	0	0	0	0	1.00	1.00	1.00	1
2097	0	0	0	0	0	0	1.00	1.00	1.00	1
2098	0	0	0	0	0	0	1.00	1.00	1.00	1
2099	0	0	0	0	0	0	1.00	1.00	1.00	1
2100	0	0	0	0	0	0	1.00	1.00	1.00	1

Fig. 10. Extracts of data generated with the estimated values of variables A, B, C and R

The results are summarized in the graphs in the following Figure.



Fig. 11. Result as function of each variable

A result 1 indicates an admitted learner, while 0 indicates as always, a non-admitted learner.

In Fig. 11, the red marker on the ordinate 0 and abscissa 5 indicates that learners having brought COLLAB = 5 failed. Above, the red marker on the ordinate 1 indicates that other learners having brought the same value of COLLAB are admitted, but obviously with other values of the other variables. The proportion between admitted and failed learners is: 91% admitted - 9% failed.

From Fig. 12, it is noted that the proportions of admitted learners increase according to the values of the variables studied (example in Figure 5.19).

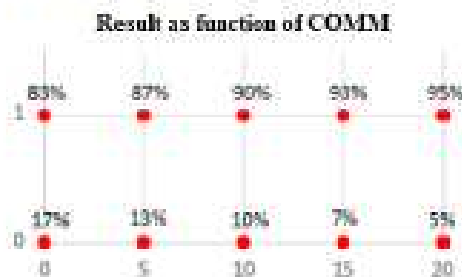


Fig. 12. Increasing rate of admission depending on the variable

It can be deduced that the more learners bring more collaboration, the more they can be admitted. A motivated learner collaborates with his peers and always communicates with them during his learning. He always reads news, participates in ongoing discussions and creates much needed debates. The dynamism of the learner in collaboration and communication, the maintenance of motivation and the active participation in a discussion forum has a positive impact on his cognition and brings him a good pedagogical result.

For the observed variables generating the latest variable A (Censu\_F, Censu\_D, Envoi\_M), it was noted that the rates of admitted learners do not vary much, as shown in Fig. 13.

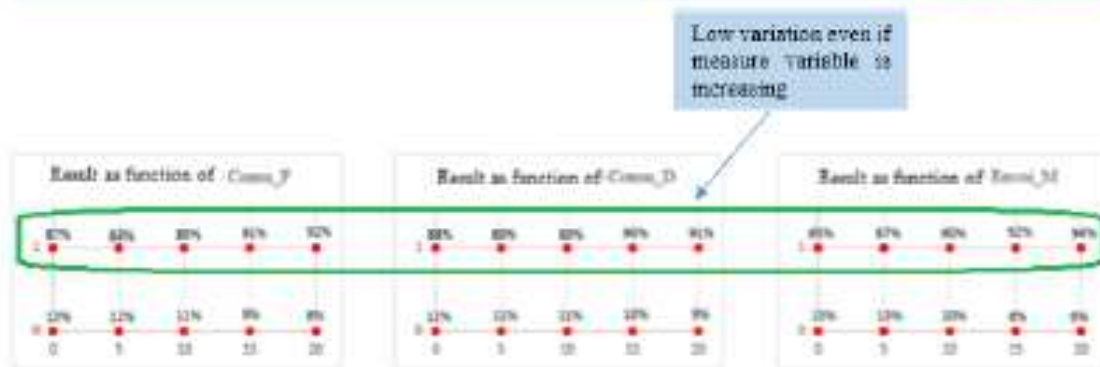


Fig. 13. Rate of admission with latent variable A

It was deduced that the three measurement variables of the latent variable A do not have a direct impact on the latent variable C, as well as on the pedagogical result of the learner. In other words, the student's behavior in front of the discussion forum does not have a direct influence on his / her acquisition of knowledge.

Participation in the ongoing discussion and any consultation of forums on the platform is not enough when learning in an

ODL. It is necessary to react, and to always contribute to the collaboration in all the group activities of the training. It is also necessary to talk to the manager or teammates about the difficulty if there is one, so that they can help solve it and take the necessary action to keep perseverance motivated.

An interesting result can also be deduced from Fig. 11. It is highlighted in the Fig. 14. below.



Fig. 14. Collaboration and motivation of learner

From a certain value of collaboration provided by a learner, that learner can always be admitted whatever the values of the other variables of measurement (0 failure). Collaboration is therefore a strong indicator for the latent variable E and has a significant impact on C and the educational outcome. In other words, collaboration favors a socio-constructivist vision of learning as part of the construction of knowledge [4], [5].

Afterwards, the motivation of a learner can also be mentioned. It creates a climate conducive for learning. Motivation is part of the social skills needed in e-learning. "The idea that we can build ourselves and rebuild ourselves, as we can partly build and rebuild the rest of society, permanently" [6]. A motivated learner communicates with his entourage in a training device, shares his knowledge via the platform, develops his know-how and his knowledge, and collaborates among the members of his group for the requested work.

The involvement of a learner in a group could support his asynchronous interaction and motivation in training.

It can be said that the involvement of the student must be serious to be successful in an ODL. The student should be active in collaborative work, interact with the peers, and he or she should be motivated by training and all learning activities.

## V. DISCUSSIONS

Tutoring, if there is one, should then favor the collaboration of the learners to have a good result. According to Dix et al. [7], collaborative work is developed in pairs, and it would be better to present students with an organized forum than to suggest ways for them to organize it themselves.

### A. Remote collaboration in a structured forum

In the context of collaborative remote discussion, learners use the forum tool to come up with a common solution [8] [9]; and they must argue to better understand their peers [10].

While the scenario has an important role to play in stimulating interaction between learners [11], it is also important to adapt the tools to the context of learning. The result of the research has approved the importance of this structuring, and the impact of collaborative work on the success of active learners.

Learners in distance education have brought nuances to collaborative working arrangements. Bruno De Lièvre and his colleague [12] have highlighted the fact that learners value and consider collaborative work to be beneficial. This nuance was found during the context analysis done in some forums, and has an impact on the behavior of the learner.

The study covered all structured forums. Learners work in the first two sequences, without any forum structure being provided to them. The relevance of their interactions in unstructured forums was also considered and their ability to build an organization that suited them was tested. The learners' perceptions of the effectiveness of collaborative work, in SGOL sequence 3, where the forum is well organized according to the learning scenario, was found. It is clear that the structuring facilitates the use of the forum and its effectiveness for distance learning.

#### B. Tutorial intervention to help learners collaborate

In the same way, the tutorial intervention has an important place in the good realization of an open and remote training. In tutorial resolution # 8 of Jacques Rodet in April 2018, he spoke about collaboration, in other words helping learners to collaborate. The remote tutor must intervene with groups of learners. It is not up to him to play the role of team leader but rather to help learners, both collectively and individually, so that learners can achieve their goals of activities and training as well as a real autonomy of operation. Its interventions are based on the collaboration of learners whose main phases are as follows: the learners' involvement to the team, the team production and the evaluation of their team's activities".

In this case, learners must get involve in the team, produce as a team, and evaluate the activities of their team.

Productivity is related to accomplishing tasks and achieving goals. Tutorial interventions help learners to collaborate. They are thus very diverse and marked by the different phases of the collaboration. Actions' learners help learners to accomplish their tasks and achieve their goals by keeping their collaboration alive.

In problem-based and project-based learning, according to Caroline Verzat and her colleagues [13], effective learning is inseparable from true collaboration among students. The tutor's role is often to facilitate collaboration, to create a space for socio-cognitive conflict between students and to be able to facilitate learning as part of a dynamic group of students.

#### VI. CONCLUSION

It can be concluded that the behavior of a learner in front of the discussion forum does not have a direct effect on his cognition; and then the domination of collaboration is highlighted on the acquisition of knowledge in open and distance learning. Motivation is also an important factor for perseverance in training, and in the end, communication is also necessary to

have success. The dynamism in the collaboration implies a strong interaction within the group, thus resulting in the effective production for the achievement of the common objective, and obviously improves the educational results.

#### REFERENCES

- [1] M. Rakotonirainy, L. Zakariasy, L. A. Radinomenantona, N. F. Limby, « Organisation du travail, une stratégie pour l'efficacité du dispositif de formation à distance », Colloque international sur les TIC, 23-29 juin 2012, Antananarivo-Madagascar, 2012.
- [2] M. Rakotonirainy, T. B. Ravallimimanantsoa, F. Randimandrainibe, « Models for Measuring Latent Variables in Asynchronous Learning », IJERTSA, International Journal of Engineering Research & Technology, Vol. 7 Issue 04, pp.146-150, April-2018.
- [3] J.F. Hair, B. Black, Babin B., R.E. Anderson, R.L. Tatham, « Méthodes d'analyse », Pearson Prentice-Hall, Nème ed, 928 pages, 2009.
- [4] B. Albert, « L'assimilation dans les dispositifs de PCAD : instrumenter le développement de l'autonomie dans les apprentissages », in Salah I., Lepage D. et Hocquart S. (dir.), Les TIC au cœur de l'enseignement supérieur, Laboratoire Pangraphie, Université Paris 8, novembre 2002.
- [5] M. Linaud, « Conception de dispositifs et changement de paradigme en formation », Éducation Permanente n° 152, 2002.
- [6] J. M. Destré, « La compétence sociale, diagnostic et Développement », L'Harmattan, Paris, 1997.
- [7] A. Dix, D. Ranchariy-Ulla, & J. Wilkinson, « Trigger analysis – understanding broken tasks ». In : D. Harper & N. Stanton (Eds.), « The Handbook of Task Analysis for Human-Computer Interaction ». Mahwah : Erlbaum 2004.
- [8] P. Dillenbourg, « What do you mean by collaborative learning? » In : P. Dillenbourg (Ed). Collaborative learning : Cognitive and Computational Approaches, pp. 1 - 19. Oxford, Elsevier, 1996.
- [9] K. Natchinas, D. Modeste, A. Oum, & J. Eari, « Web-supported emergent-collaboration in higher education courses », Educational Technology & Society, 14(3), pp. 94 -104, 2011.
- [10] M.J. Baker, « Forms of cooperation in dyadic problem-solving, *Revue d'Intelligence Artificielle* », 16, 4- 5, pp. 587 - 620, 2002.
- [11] B. De Lièvre, C. Depover, A. Strohle, « Les arts individuels sont-ils pris en compte lors d'une activité collaborative à distance ? », In : P. Tchoukline, M. Joub, L. Trouche (Eds.), « Environnements informatiques pour l'apprentissage humain », 2005, pp. 93-104. Montpellier, France, 2005.
- [12] B. De Lièvre, O. Tempierman, « Trois modalités de structuration d'un forum collaboratif : comment les étudiants les perçoivent-ils ? », Journées communication et apprentissage instrumentées en réseau 2008, 987-2-74622118-3. «halshs-01112579», 2008.
- [13] C. Verzat, N. O'Wan et B. Hancourt, « Réguler le leadership dans les groupes d'étudiants en APP », *Revue internationale de pédagogie de l'enseignement supérieur* pp. 31-1 | 2015



*A.1.7. Lettre d'acceptation du deuxième article de notoriété internationale*



### A.1.3. Certificat de publication de l'article



## Annexe 2 : ECHANTILLON DE DONNEES UTILISEES DANS LE CHAPITRE 4

N°	GSA_2016.2017	COLLAB	COMM	MOTIV	Consu_F.	Consu_D.	Envoi_M.
1	GSA1	0	0	0	1	2	0
	GSA2	2	0	1	8	14	4
3	GSA3	2	2	1	22	10	11
4	GSA4	4	3	2	33	3	22
5	GSA5	2	4	1	29	48	16
6	GSA6	3	0	1	7	14	6
7	GSA7	3	2	1	26	90	10
8	GSA8	0	0	0	23	20	5
9	GSA9	4	2	5	23	62	22
10	GSA10	1	0	1	11	2	3
11	GSA11	0	0	1	8	4	4
12	GSA12	0	0	1	0	11	0
13	GSA13	1	3	2	22	46	10



### **Annexe 3 : LES DONNEES UTILISEES DANS LE CHAPITRE 5**

N°	2016 & 2017	COLLAB	COMM	MOTIV	Consu_F	Consu_D	Envoi_M	Creat_D	Result
1	GSA1	0	0	0	1	2	0	0	0
2	GSA2	2	0	1	8	14	4	0	0
3	GSA3	2	2	1	22	10	8	3	1
4	GSA4	4	3	2	33	3	21	1	1
5	GSA5	2	4	1	29	48	15	1	1
6	GSA6	3	0	1	7	14	5	1	0
7	GSA7	3	2	1	26	90	9	1	1
8	GSA8	0	0	0	23	20	4	1	0
9	GSA9	4	2	5	23	62	22	0	1
10	GSA10	1	0	1	11	2	2	1	0
11	GSA11	0	0	1	8	4	3	1	1
12	GSA12	0	0	1	0	11	0	0	0
13	GSA13	1	3	2	22	46	9	1	1
14	ELT_1	0	1	1	13	11	6	3	0
15	ELT_2	0	0	2	12	6	2	1	1
16	ELT_3	3	0	2	19	10	9	4	0
17	ELT_4	0	0	0	2	2	0	0	0
18	ELT_5	5	1	4	50	77	21	10	1
19	ELT_6	0	0	1	12	1	2	1	0
20	ELT_7	0	0	4	145	100	14	5	0
21	ELT_8	9	1	10	74	89	27	20	1
22	ELT_9	2	0	2	51	31	9	4	1
23	ELT_10	16	1	3	52	106	26	19	1
24	ELT_11	7	0	1	25	27	13	8	1
25	ELT_12	0	0	3	20	19	9	5	0
26	ELT_13	6	0	2	28	35	14	9	0
27	ELT_14	4	0	1	48	30	14	6	1
28	ELT_15	5	0	1	42	43	11	7	1
29	ELT_16	10	0	5	43	184	21	14	1
30	ELT_17	0	0	1	17	60	10	2	0
31	ELT_18	21	0	11	147	210	53	36	1
32	ELT_19	0	0	0	8	6	0	0	0
33	ELT_20	0	0	4	34	25	8	4	0
34	ELT_21	0	0	1	6	1	2	1	1
35	ELT_22	0	0	0	9	10	0	0	0
36	ELT_23	1	0	1	27	9	8	2	0
37	ELT_24	0	0	0	27	9	8	2	0
38	ELT_25	2	0	1	17	28	2	0	0
39	ELT_26	2	5	8	33	57	14	4	1
40	ELT_27	9	0	1	39	44	4	4	1
41	ELT_28	1	1	2	29	26	0	0	0
42	ELT_29	1	7	3	43	51	9	1	0

43	ELT_30	15	8	2	120	109	21	5	1
44	ELT_31	1	4	0	32	29	4	0	0
45	ELT_32	3	1	1	25	34	3	2	1
46	ELT_33	6	5	3	66	94	11	0	0
47	ELT_34	0	0	1	7	1	0	0	0
48	ELT_35	2	0	4	17	26	4	1	0
49	ELT_36	15	4	4	106	309	21	3	1
50	ELT_37	9	2	4	86	74	9	6	1
51	ELT_38	2	0	1	16	31	2	0	0
52	ELT_39	1	1	1	24	48	2	0	0
53	ELT_40	1	1	2	14	19	0	0	0
54	ELT_41	6	4	2	29	68	10	0	0
55	ELT_42	5	0	1	16	32	4	0	0
56	ELT_43	15	2	7	114	99	19	7	1
57	ELT_44	1	0	4	14	29	2	2	0
58	ELT_45	1	2	1	58	47	1	1	0
59	ELT_46	2	2	5	77	97	9	1	1
60	ELT_47	4	1	2	22	26	6	0	0
61	ELT_48	4	0	1	61	57	4	1	1
62	ELT_49	2	0	2	9	11	2	0	0
63	ELT_50	5	3	3	66	67	5	5	0
64	ELT_51	2	1	0	25	35	1	2	1
65	TCI_1	0	0	1	4	2	0	0	0
66	TCI_2	0	0	1	11	8	0	1	0
67	TCI_3	1	4	5	62	44	23	4	1
68	TCI_4	1	2	1	9	19	7	0	1
69	TCI_5	3	5	5	51	47	26	6	1
70	TCI_6	4	2	8	91	93	40	7	0
71	TCI_7	0	0	0	5	0	0	0	0
72	TCI_8	1	3	1	25	51	10	2	0
73	TCI_9	1	4	4	23	31	12	2	1
74	TCI_10	1	4	1	24	21	12	3	0
75	TCI_11	1	0	1	19	21	5	1	1
76	TCI_12	1	5	2	26	42	18	5	0
77	TCI_13	0	0	2	3	5	2	0	0
78	TCI_14	0	0	1	8	8	0	0	0
79	TCI_15	7	2	1	51	53	25	3	1
80	TCI16	25	32	1	142	241	17	59	1
81	TCI17	1	1	1	17	15	3	1	1
82	TCI18	9	13	3	87	94	12	4	1
83	TCI19	0	0	0	36	110	3	23	0
84	TCI20	4	4	1	25	36	0	6	1
85	TCI21	8	9	4	30	98	10	11	1

86	TCI22	7	10	1	42	71	7	5	1
87	TCI23	15	14	3	32	69	9	18	1
88	TCI24	18	32	6	104	181	13	38	1
89	TCI25	8	8	2	71	108	5	13	1
90	TCI26	3	2	1	31	30	1	1	0
91	TCI27	0	1	1	19	16	1	0	1
92	TCI28	13	10	10	100	125	10	24	1
93	TCI29	0	0	0	60	143	1	35	0
94	TCI30	0	0	0	50	47	4	3	1
95	TCI31	0	0	0	0	3	1	0	0
96	TCI32	0	0	0	1	1	0	0	0

## Annexe 4 : ANALYSE DES TRACES D'INTERVENTION DANS LES FORUMS DE DISCUSSION DU GEAL - 2<sup>E</sup> PROMOTION FOAD GSA

*Période du 19 octobre au 09 novembre 2017*

La FOAD GSA a 15 nouveaux apprenants pour sa deuxième promotion (2017/2018).

N°	2017-2018	COLLAB	COMM	MOTIV	Consu_F	Consu_D	Envoi_M
97	GSA14	11	4	1	44	90	42
98	GSA15	11	3	1	48	48	36
99	GSA16	10	4	2	37	85	41
100	GSA17	11	3	2	62	89	43
101	GSA18	20	6	1	123	125	72
102	GSA19	5	3	1	21	71	29
103	GSA20	13	6	5	66	152	60
104	GSA21	5	4	1	12	27	20
105	GSA22	10	8	1	71	114	53
106	GSA23	13	5	3	88	129	56
107	GSA24	10	4	2	31	71	38
108	GSA25	16	3	1	53	129	60
109	GSA26	2	1	1	10	26	9
110	GSA27	14	4	1	19	65	39
111	GSA28	21	6	1	39	101	64

### Script & résultat

```
> source('F:/Doctorat_2016/SCA/SCA/Thèses/D3/S5/sca/gsa-script_2018.R', echo=TRUE)

> library(lavaan)

> #Chargement de données
> gsa <- read.csv("F:/Doctorat_2016/SCA/SCA/Thèses/D3/gsa_2017_2018.csv", sep=";")

> donnee <- gsa[1:6]

> colnames(donnee) <- c("COLLAB", "COMM", "MOTIV", "Consu_F", "Consu_D", "Envoi_M")

> #Création du modele
> donnee.model <- '
+ A =~ y1*Consu_F + y2*Consu_D + y3*Envoi_M
+ E =~ l1*COLLAB + l2*COMM + l3*MOTIV
+ A ~ r1*E
+ C =~ r2*E'
```

```

+ '

> #Estimation du modele
> donnee.fit <- cfa(model=donnee.model, data=donnee, se="robust")

> summary(donnee.fit, fit.measures= TRUE, standardized= TRUE)
lavaan (0.5-20) converged normally after 183 iterations

    Number of observations                  15

    Estimator                             ML
    Minimum Function Test Statistic        16.592
    Degrees of freedom                      8
    P-value (Chi-square)                   0.035

Model test baseline model:

    Minimum Function Test Statistic        116.505
    Degrees of freedom                     15
    P-value                                0.000

User model versus baseline model:

    Comparative Fit Index (CFI)            0.915
    Tucker-Lewis Index (TLI)              0.841

Loglikelihood and Information Criteria:

    Loglikelihood user model (H0)          -257.437
    Loglikelihood unrestricted model (H1)  -249.141

    Number of free parameters              19
    Akaike (AIC)                          552.874
    Bayesian (BIC)                        566.327
    Sample-size adjusted Bayesian (BIC)    508.322

Root Mean Square Error of Approximation:

    RMSEA                                0.268
    90 Percent Confidence Interval         0.069 0.450
    P-value RMSEA <= 0.05                 0.042

Standardized Root Mean Square Residual:

    SRMR                                0.174

Parameter Estimates:

    Information                           Expected
    Standard Errors                      Robust.sem

Latent Variables:

```

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
A =~							
	Consu_F (y1)	1.000				19.389	0.656
	Consu_D (y2)	1.622	0.521	3.116	0.002	31.451	0.862
	Envoi_M (y3)	0.900	0.278	3.236	0.001	17.453	1.058
E =~							
	COLLAB (11)	1.000				4.378	0.875

COMM	(12)	0.223	0.054	4.121	0.000	0.976	0.591
MOTIV	(13)	-0.018	0.024	-0.729	0.466	-0.078	-0.072

C =~

E	(r2)	1.000				1.000	1.000
---	------	-------	--	--	--	-------	-------

Regressions:

		Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
A ~							
E	(r1)	4.479	1.884	2.377	0.017	1.011	1.011

Intercepts:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
Consu_F	48.267	7.903	6.108	0.000	48.267	1.632
Consu_D	88.133	9.749	9.041	0.000	88.133	2.416
Envoi_M	44.133	4.411	10.006	0.000	44.133	2.674
COLLAB	11.467	1.338	8.572	0.000	11.467	2.291
COMM	4.267	0.441	9.664	0.000	4.267	2.583
MOTIV	1.600	0.289	5.527	0.000	1.600	1.477
A	0.000				0.000	0.000
E	0.000				0.000	0.000
C	0.000				0.000	0.000

Variances:

	Estimate	Std.Err	Z-value	P(> z )	Std.lv	Std.all
Consu_F	498.406	153.030	3.257	0.001	498.406	0.570
Consu_D	341.346	117.225	2.912	0.004	341.346	0.257
Envoi_M	-32.237	12.590	-2.560	0.010	-32.237	-0.118
COLLAB	5.878	3.057	1.923	0.054	5.878	0.235
COMM	1.776	0.937	1.895	0.058	1.776	0.651
MOTIV	1.167	0.743	1.572	0.116	1.167	0.995
A	-8.658	68.492	-0.126	0.899	-0.023	-0.023
E	0.000				0.000	0.000
C	19.170	9.835	1.949	0.051	1.000	1.000

```

> coef(donnee.fit)
              y2              y3              12              13
r1
1.622              0.900              0.223              -0.018
4.479
Consu_F~~Consu_F Consu_D~~Consu_D Envoi_M~~Envoi_M  COLLAB~~COLLAB      COMM~
~COMM
498.406              341.346              -32.237              5.878
1.776
MOTIV~~MOTIV              A~~A              C~~C              Consu_F~1      Cons
u_D~1
1.167              -8.658              19.170              48.267              8
8.133
Envoi_M~1              COLLAB~1              COMM~1              MOTIV~1
44.133              11.467              4.267              1.600

```

> #The values of the estimated parameters, the standard errors, the z-values, the standardized parameter values

```

> parameterEstimates(donnee.fit)
      lhs op      rhs label      est      se      z pvalue ci.lower ci.upper
1      A =~ Consu_F   y1    1.000    0.000    NA      NA    1.000    1.000
2      A =~ Consu_D   y2    1.622    0.521  3.116  0.002    0.602    2.642
3      A =~ Envoi_M   y3    0.900    0.278  3.236  0.001    0.355    1.445
4      E =~  COLLAB   l1    1.000    0.000    NA      NA    1.000    1.000

```

5	E	==	COMM	12	0.223	0.054	4.121	0.000	0.117	0.329
6	E	==	MOTIV	13	-0.018	0.024	-0.729	0.466	-0.066	0.030
7	A	~	E	r1	4.479	1.884	2.377	0.017	0.786	8.172
8	C	==	E	r2	1.000	0.000	NA	NA	1.000	1.000
9	Consu_F	~~	Consu_F		498.406	153.030	3.257	0.001	198.473	798.338
10	Consu_D	~~	Consu_D		341.346	117.225	2.912	0.004	111.589	571.103
11	Envoi_M	~~	Envoi_M		-32.237	12.590	-2.560	0.010	-56.914	-7.560
12	COLLAB	~~	COLLAB		5.878	3.057	1.923	0.054	-0.113	11.870
13	COMM	~~	COMM		1.776	0.937	1.895	0.058	-0.061	3.613
14	MOTIV	~~	MOTIV		1.167	0.743	1.572	0.116	-0.288	2.623
15	A	~~	A		-8.658	68.492	-0.126	0.899	-142.900	125.584
16	E	~~	E		0.000	0.000	NA	NA	0.000	0.000
17	C	~~	C		19.170	9.835	1.949	0.051	-0.106	38.447
18	Consu_F	~1			48.267	7.903	6.108	0.000	32.778	63.756
19	Consu_D	~1			88.133	9.749	9.041	0.000	69.026	107.240
20	Envoi_M	~1			44.133	4.411	10.006	0.000	35.488	52.779
21	COLLAB	~1			11.467	1.338	8.572	0.000	8.845	14.088
22	COMM	~1			4.267	0.441	9.664	0.000	3.401	5.132
23	MOTIV	~1			1.600	0.289	5.527	0.000	1.033	2.167
24	A	~1			0.000	0.000	NA	NA	0.000	0.000
25	E	~1			0.000	0.000	NA	NA	0.000	0.000
26	C	~1			0.000	0.000	NA	NA	0.000	0.000

```
> #The model-implied (fitted) covariance matrix (and mean vector) of a fitted model
```

```
> fitted(donnee.fit)
```

```
$cov
```

	Cons_F	Cons_D	Envo_M	COLLAB	COMM	MOTIV
Consu_F	874.329					
Consu_D	609.796	1330.515				
Envoi_M	338.398	548.926	272.382			
COLLAB	85.864	139.282	77.293	25.049		
COMM	19.143	31.052	17.232	4.274	2.729	
MOTIV	-1.531	-2.484	-1.378	-0.342	-0.076	1.173

```
$mean
```

	Consu_F	Consu_D	Envoi_M	COLLAB	COMM	MOTIV
	48.267	88.133	44.133	11.467	4.267	1.600

```
> inspect(donnee.fit)
```

```
$lambda
```

	A	E	C
Consu_F	0	0	0
Consu_D	1	0	0
Envoi_M	2	0	0
COLLAB	0	0	0
COMM	0	3	0
MOTIV	0	4	0

```
$theta
```

	Cons_F	Cons_D	Envo_M	COLLAB	COMM	MOTIV
Consu_F	6					
Consu_D	0	7				
Envoi_M	0	0	8			
COLLAB	0	0	0	9		
COMM	0	0	0	0	10	
MOTIV	0	0	0	0	0	11



```
$psi
  A  E  C
A 12
E  0  0
C  0  0 13
```

```
$beta
  A  E  C
A 0 5 0
E 0 0 0
C 0 0 0
```

```
$nu
      intrcp
Consu_F      14
Consu_D      15
Envoi_M      16
COLLAB       17
COMM         18
MOTIV        19
```

```
$alpha
      intrcp
A      0
E      0
C      0
```

## Annexe 5 : R ET LA BIBLIOTHEQUE LAVAAN

R stocke les variables, les données, les fonctions, les résultats, etc. dans la mémoire active de l'ordinateur sous la forme d'objets nommés. L'utilisateur peut alors faire des actions sur ces objets avec des opérateurs (arithmétique, logique, comparaison) et des fonctions (qui sont elles-mêmes des objets). Une grande partie de la fonctionnalité de R est de permettre d'appliquer des fonctions à des données ou d'autres objets.

Les fonctions R sont un ensemble d'instructions qui prennent l'entrée, calculent la (les) valeur (s) désirée (s) et les transforment en résultat. R pré-charge déjà un ensemble de fonctions couramment utilisées, mais il est possible d'en charger d'autres au moment où on charge des paquets ou simplement en écrivant une fonction.

Pour utiliser une fonction :

- (a) donner le nom de la fonction suivi de parenthèses ;
- (b) entre parenthèses, donner les valeurs nécessaires ou les arguments de la fonction.

### Les fonctions R utilisées dans la recherche :

- **Commentaire.** Ce n'est pas vraiment une fonction mais tout ce qui est écrit après le signe # est supposé être un commentaire et est ignoré par R. Les commentaires sont extrêmement utiles, comme annotation R. La syntaxe peut aider à économiser beaucoup de temps et d'efforts.

- **Attribuer.** Un autre symbole que la plupart des utilisateurs R rencontrent fréquemment est la flèche gauche, `<-`, qui est l'opérateur d'assignation standard de R (une autre option utilise `=`, mais c'est mieux de le réserver `=` pour définir des valeurs pour les arguments). Le signe `<-` est la manière d'assigner tout ce qui est à la droite de la flèche à l'objet à gauche de la flèche.

- **Concaténer.** La fonction de concaténation `c ()` concatène les arguments inclus dans la fonction. L'utilisation de `c ()` en conjonction avec `<-` affecte les objets concaténés dans un nouvel objet. Par

exemple, pour créer un jeu de données de 5 observations avec les valeurs 4, 5, 3, 6, 9, et nommer le newData, on utilise la syntaxe : `newData <- c (4, 5, 3, 6, 9)`

Avec R, si une fonction n'est pas disponible pour effectuer l'analyse ou la manipulation de données souhaitée, il existe une option pour écrire une nouvelle fonction en utilisant la fonction `function ()`.

Par exemple :

La covariance est calculée en utilisant la fonction `cov ()`

```
donnees.cov <- cov(donnees)
```

La matrice est calculée en utilisant la fonction `matrix ()`

```
donnees.cov <- matrix(c(cov(donnees)),nrow=4,ncol=4)
```

Pour entrer les noms de colonnes, on utilise la fonction `colnames ()`

```
colnames(donnees.cov) <- c("Consu_F", "Consu_D", "Envoi_M", "Creat_D")
```

Pour entrer les noms de lignes, on utilise la fonction `rownames ()`

```
rownames(donnees.cov) <- c("Consu_F", "Consu_D", "Envoi_M", "Creat_D")
```

## La bibliothèque Lavaan

Pour effectuer une analyse de chemin dans R, on utilise le paquet de lavaan.

Le paquet de lavaan (LATent VArable ANalysis) est un paquet de R conçu pour la modélisation générale de l'équation structurelle. L'information et la documentation à ce sujet peuvent être trouvées sur la page web du paquet : <http://www.lavaan.org>.

**colnames ()** est utilisé pour entrer les noms de colonnes, on fait usage du paquet de lavaan

```
# charger lavaan
```

```
library (lavaan)
```

```
# nommer les colonnes
```

```
Colnames (donnee) <- c("COLLAB", "COMM", "MOTIV", "Consu_F", "Consu_D",
"Envoi_M")
```

**sep = ";"** L'argument sep indique à la fonction comment les variables sont délimitées dans le fichier de données. `sep = ";"` est nécessaire pour indiquer une délimitation de point-virgule.

**read.csv ()**. La fonction `read.csv ()` fonctionne exactement comme la fonction `read.table ()`, mais elle suppose que les données sont délimitées par des virgules.

```
elt <- read.csv("F:/Doctorat_2016/SCA/SCA/Thèses/D3/ELT_donnees.csv", sep=";")
```

Cela ouvre une boîte de dialogue qui permet de choisir le fichier de manière interactive.

**read.table ()**. Par défaut, la fonction `read.table ()` renvoie une trame de données.

Une trame de données est un type de R objet qui stocke des variables en tant que colonnes. Les trames de données sont utiles car elles peuvent stocker différents types de variables, tels que les chaînes et les nombres.

**Donnees.model** est utilisé pour la spécification du modèle

```
donnees.model <- 'A =~ l1*Consu_F + l2*Consu_D + l3*Envoi_M + l4*Creat_D'
```

Une fois le modèle spécifié, on a utilisé `cfa ()` pour les modèles d'analyse factorielle confirmatoire, fonction pour s'adapter au modèle. Le `cfa ()` n'est en fait que l'appel à une fonction `lavaan ()`.

Estimation du modèle : Pour estimer les paramètres spécifiés dans `donnees.model`, on utilise **donnees.fit**

```
donnees.fit <- cfa(model=donnees.model):
```

```
donnee.fit <- cfa(model=donnee.model, data=donnee, se="robust")
```

Estimateur robuste : Estimateurs de paramètres dont les valeurs sont moins influencées par les violations d'hypothèses que l'estimateur traditionnel du maximum de vraisemblance.

La fonction **summary ()** est pour la description d'un modèle

```
summary(donnees2.fit, standardized= TRUE)
```

```
summary(donnee.fit, fit.measures= TRUE)
```

**coef()** : Discrimination et location

```
coef(donnee.fit)
```

Pour évaluer un modèle, on utilise la fonction **fitmeasures ()**

```
fitmeasures(donnees2.fit, fit.measures=c("fmin", "logl", "chisq", "df", "npar"))
```

**parameterEstimates(donnee.fit)**

Pour imprimer les estimations de paramètres, on utilise la fonction **summary ()** ou **parameterEstimates ()**.

La spécification de la valeur **standard = TRUE** dans l'une ou l'autre des fonctions produit à la fois les valeurs normalisées et non standardisées.

```
fitted(donnee.fit)
```

La fonction **fitted ()** dans le lavaan renvoie toutes les covariances impliquées par le modèle.

Pour m'assurer que mes valeurs ont été correctement spécifiées, j'ajuste le modèle en lavaan en utilisant la commande **do.fit = FALSE**.

De plus, l'utilisation de la fonction **fitted ()** renvoie les moyennes et covariances résultantes à partir du modèle en utilisant les valeurs de départ fixes ou par défaut pour les estimations de paramètres.

## Annexe 6 : CADRE LOGIQUE DU PROJET DE RECHERCHE

<b>Titre</b>	<b>Modélisation de l'apprentissage asynchrone dans une formation ouverte et à distance</b>	<b>Durée : 3 ans</b>
--------------	--	----------------------

<b>Objectif Général</b>	Modéliser l'interaction des étudiants dans des forums de discussion lors de leur apprentissage en ligne			
<b>Objectifs Spécifiques</b>	Analyser la trace des activités numériques de l'apprenant	Analyser le contenu des forums de discussions	Elaborer le modèle d'équation structurelle	Appliquer ce modèle dans les 3 FOAD

↑↓

↑↓

↑↓

↑↓

↑↓

<b>Résultats Attendus</b>	<ul style="list-style-type: none"> <li>- Les variables de mesure de l'attitude de l'apprenant face au forum de discussion sont identifiées</li> <li>- La base de données issue de l'analyse est constituée</li> </ul>	<ul style="list-style-type: none"> <li>- Les variables de mesures de l'engagement de l'apprenant sont identifiées</li> <li>- La base de données issue de l'analyse est constituée</li> </ul>	<ul style="list-style-type: none"> <li>- 2 modèles de mesures validés : relation entre la variable latente et ses variables de mesure</li> <li>- 1 modèle de structure identifié : relation entre les variables latentes</li> </ul>	1 modèle d'apprentissage asynchrone validé empiriquement
---------------------------	---	--	---	--

↑↓

↑↓

↑↓

↑↓

↑↓

<b>Activités</b>	Identification des 3 variables de mesure de la variable latente A (Attitude de l'apprenant face au forum de discussion)	Identification des 3 variables de mesure de la variable latente E (Engagement de l'apprenant)	Identification des 3 variables latentes mises en jeux, spécification, identification et estimation du modèle	<ul style="list-style-type: none"> <li>- Evaluation du modèle</li> <li>- Prévision de résultat : application dans une promotion en cours (2017/2018)</li> <li>- Prévision générale</li> </ul>
↕	↕	↕	↕	↕
<b>Moyens</b>	<ul style="list-style-type: none"> <li>- Grille d'analyse des traces d'activités numériques</li> <li>- 3 plateformes de FOAD, 4 forums, module GEAL, 6 promotions, 111 nouveaux apprenants, 12393 événements</li> </ul>	<ul style="list-style-type: none"> <li>- Grille d'analyse de contenus</li> <li>- 3 plateformes de FOAD, 4 forums, module GEAL, 6 promotions, 111 nouveaux apprenants, 1072 messages</li> </ul>	<ul style="list-style-type: none"> <li>- Théorie de statistique et analyse multivariée</li> <li>- Données relevées via la plate-forme</li> <li>- Approche LISREL</li> <li>- Logiciel R</li> <li>- SEM</li> </ul>	<ul style="list-style-type: none"> <li>- Données relevées via la plate-forme</li> <li>- Outils statistiques</li> <li>- Logiciel R</li> <li>- Logiciel Excel</li> </ul>

Source : conception personnelle

**Figure 4.01 :** *Cadre logique de la recherche*

## BIBLIOGRAPHIES

- [1.01] C. Develotte, « *Aspects interculturels de l'enseignement / apprentissage en ligne : le cas du programme franco-australien "le français en (première) ligne". Quelle didactique de l'interculturel dans les nouveaux contextes d'enseignement-apprentissage du FLE/S ?* », Louvain La Neuve, Belgique, 2005.
- [1.02] S. Berrouk & A. Jaillet « *Les fonctions tutorales : pour un déséquilibre dynamique* », Distances et médiations des savoirs, vol 1(2), 2013.
- [1.03] M. Rakotomalala, L. Zakariasy, L. A. Rafanomezantsoa, N. F. Leaby, « *Organisation du tutorat, une stratégie pour l'efficacité du dispositif de formation à distance* », Colloque international sur les TIC. 27-29 juin 2012, Antananarivo Madagascar, 2012.
- [1.04] J. Rodet, « *Proposition pour l'ingénierie tutorale* », la revue de t@d, la communauté de pratiques des tuteurs à distance, 2010.
- [1.05] M., Rakotomalala, L. Zakariasy, « *Pratiques tutorales et attentes des apprenants dans une FOAD - Tutoring practices and learner's expectations in an ODL* », n°11 – revue Frantice, déc. 2015.
- [1.06] 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.
- [1.07] M. Fayol, et D. Gaonac'h, « *Le développement de la mémoire* ». Dans A. Blaye, et P. Lemaire, (Edit.) Psychologie du développement cognitif de l'enfant. Paris-Bruxelles : De Boeck Université, pp. 125-156, 2007.
- [1.08] N. Guichon, « *Langues et Tice - Méthodologie de conception multimédia* », Paris : Ophrys, 2006.
- [1.09] D. Peraya, « *Médiation et médiatisation : le campus virtuel. Vers les campus virtuels* », *Hermès*, 25, Pp. 153-167, 1999.



- [1.10] B. Blandin, « *Usage des instruments de communication* », 2001.
- [1.11] G. Paquette, C. Ricciardi-Rigault, C. Paquin, S. Liégeois and E. Bleicher, « *Developing the Virtual Campus Environment* », ED-Media International Conference, Boston, Jun. 1996.
- [1.12] J. P. PERNIN, « *Quels modèles et quels outils pour la scénarisation d'activités dans les nouveaux dispositifs d'apprentissage ?* » INRP, Séminaire "TIC, nouveaux métiers et nouveaux dispositifs d'apprentissage", 2003.
- [1.13] L. Audet, « *Recherche sur les facteurs qui influencent la persévérance et la réussite scolaire en formation à distance. Recension des écrits* », 2008.
- [1.14] N. Racette, « *Augmenter la motivation dans un cours à distance* » : expérimentation d'un modèle. Sarrebruck : Éditions universitaires européennes, 2012.
- [1.15] J. Loisier, « *Les nouveaux outils d'apprentissage encouragent-ils réellement la performance et la réussite des étudiants en FAD ?* » 2011.
- [1.16] R. Viau & J. Bouchard, « *Validation d'un modèle de dynamique motivationnelle auprès d'élèves du secondaire* ». Revue canadienne de l'éducation, 25(1), pp. 16-26, 2000.
- [1.17] R. Viau, « *12 questions sur l'état de la recherche scientifique sur l'impact des TIC sur la motivation à apprendre* », 2005.
- [1.18] D. Vasant, « *Data Science and Prediction* », Communications of the ACM, no 12, pp. 64- 73, décembre 2013.
- [1.19] D. S. Moore, « *Teaching statistics as a respectable subject* ». In Gordon, F. and S. Gordon (eds.), Statistics for the twenty-first century, 14-25, Mathematical Association of America. 1991.
- [1.20] J. Piaget, « *Pionnier dans le domaine de la psychologie de l'enfant* », 1896-1980.
- [1.21] S. Dehaene, « *Le bébé statisticien : les théories bayésiennes de l'apprentissage* », 2013.

- [1.22] H. Röhrs, « Maria Montessori : 1870-1952 », *Perspectives : revue trimestrielle d'éducation comparée*, Paris, UNESCO : Bureau international d'éducation, vol. XXIV, n° 1-2, pp. 173-188, 1994.
- [1.23] G. Paquette, « *Modélisation des connaissances et des compétences, pour concevoir et apprendre* ». Sainte-Foy, Canada : Presses de l'Université du Québec, 2002 (b).
- [1.24] A. Newell, & H. Simon, "*Human problem solving. Englewood Cliffs*", NJ: Prentice-Hall, 1972.
- [1.25] D. Merrill, "*Principles of instructional design*". Englewood Cliffs, NJ: Educational Technology Publications, 1994.
- [1.26] C. Reigeluth, "*Instructional theories in action: Lessons illustrating selected theories and models. Hillsdale*", NJ: Lawrence Erlbaum, 1987.
- [1.27] A. J. Romiszowski, "*Designing instructional systems*". Londres/New York", Kogan Page/Nichols Publishing, 1981.
- [1.28] V. Goel, & P. Pirolli, "*Motivating the notion of generic design within information-processing theory: The design problem space*". *AI Magazine*, 10 (1), pp. 18-36, 1989.
- [1.29] N. Entwistle, "*Concepts and conceptual frameworks underpinning the ETL Project, Edinburgh, School of Education*", University of Edinburgh, 2003.
- [1.30] F. Henri, A. Kaye, « *Le savoir à domicile. Pédagogie et problématiques de la formation à distance* », Québec : Presses de l'Université du Québec/Télé-Université, 1985.
- [2.01] W. F. Hill, « *Learning: A survey of psychological interpretations* ». Thomas Y. Crowell, 1977.
- [2.02] T. Good, et J. Brophy, « *Educational Psychology: A Realistic Approach* », 4e éd., New York: Longman, 1995.

- [2.03] J. Watson, « *Le béhaviorisme* ». Paris. Editions Cepi, 1972.
- [2.04] F. Raynal, A. Rieunier et M. Postic, « *Apprentissages, formation, psychologie cognitive* », 5e édition augmentée d'une introduction / Issy-les-Moulineaux : ESF, 1997
- [2.05] J. El Bouhdid, « *Une Architecture Intelligente Orientée objectifs basée sur les Ontologies et les Systèmes Multi-agents pour la Génération des Parcours d'Apprentissage Personnalisés* », Doctorat, Université Abdelmalek Essaadi, 2013.
- [2.06] J. Piaget, « *Réussir et comprendre* ». Paris : PUF, p. 237, 1974.
- [2.07] J. Piaget, « *L'équilibration des structures cognitives* ». Paris, PUF, 1975.
- [2.08] P. E. Doolittle, « *Constructivism and online education. Virginia* »: Polytechnic Institute & State University, 1999.
- [2.09] S. Crozat, « *Éléments pour la conception industrialisée des supports pédagogiques numériques* », Doctoral dissertation, Université de Technologie de Compiègne, 2002.
- [2.10] B. Legault, « Le cognitivisme : théorie et pratique ». *Pédagogie collégiale*, 6(1), pp. 41-42, 1992.
- [2.11] R. Bibeau, « *École informatisée clés en main* ». Projet franco-qubécois de recherche-action. *Revue de l'EPI (Enseignement Public et Informatique)*, (82), pp. 137-147, 1996.
- [2.12] R. Bibeau, « *Les technologies de l'information et de la communication peuvent contribuer à améliorer les résultats scolaires des élèves* ». *Revue de l'EPI*, 2007.
- [2.13] J. S. Bruner, « *Le développement de l'enfant : savoir-faire, savoir dire* ». Paris, PUF, 1983.
- [2.14] W. Doise, & G. Mugny, « *Le développement social de l'intelligence* » (Vol. 1). Paris : Inter Editions, 1981.

- [2.15] L. S. Vygotsky, « *Mind in society: The development of higher psychological processes* ». Harvard university press, 1980.
- [2.16] J. S. Bruner, & Y. Bonin, « *L'éducation, entrée dans la culture : les problèmes de l'école à la lumière de la psychologie culturelle* », Retz, 1996.
- [2.17] L. S. Vygotsky, « *Pensée et langage* », trad. De F. Sève, Massidor, 1985.
- [2.18] W. Doise, & G. Mugny, " *sociale et développement cognitif* ». Paris : Armand Colin, 1997.
- [2.19] C. Depover, « *Modèles pédagogiques et tutorat dans la formation des maîtres à distance* », T. Karsenti, R.-P. Garry, A. Bensiane, B.-B. Ngoy-Fiama, F. Baudot, (Eds), *La formation de formateurs et d'enseignants à l'ère du numérique : stratégies politiques et accompagnement pédagogique, du présentiel à l'enseignement à distance*. Montréal : Réseau international francophone des établissements de formation de formateurs (RIFEFF) / Agence universitaire de la Francophonie (AUF), pp. 4-18, 2012.
- [2.20] G. Siemens, « *Connectivism: A learning theory for the digital age. International journal of instructional technology and distance learning* », 2(1), pp. 3-10, 2005.
- [2.21] E. Duplâa, & N. Talaat, « *Connectivisme et formation en ligne* ». Distances et savoirs, 9(4), pp. 541-564. (2012).
- [2.22] M. Arnaud, L. Merzeau, " *Traçabilité et réseaux* ", Hermès n° 53, Paris : CNRS éditions, 2009.
- [2.23] F. Larose, A. Jaillet, « *Le numérique dans l'enseignement et la formation, analyses, traces et usages* » Paris : L'Harmattan, 2009.
- [2.24] J. Laflaquière, « *Conception de système à base de traces numériques pour les environnements informatiques documentaires* », thèse de doctorat soutenue le 8 déc. 2009.

- [2.25] H. Thibaud, « *Représenter et modéliser l'activité de recherche d'information experte avec des traces d'activité pour l'apprentissage* », *Les Enjeux de l'information et de la communication* 1/2010, pp. 83-96 ; 2010.
- [2.26] D. Peraya, « *De la correspondance au campus virtuel : formation à distance et dispositifs médiatiques* ». In Charlier, B. & Peraya, D. (Éd.). « Technologie et innovation en pédagogie. Dispositifs innovants de formation pour l'enseignement supérieur » (pp. 79-92). Bruxelles : De Boeck, 2003.
- [2.27] J.-P. Meunier, D. Peraya, « *Introduction aux théories de la communication* », Bruxelles, De Boeck, 2004.
- [2.28] C. Depover B. De Lièvre, A. Jaillet, D. Peraya, et J.-J. Quintin, « *Le tutorat en formation à distance* ». Bruxelles : De Boeck-université, 2011.
- [2.29] P. Dillenbourg, « *Interactive tabletops in education* », *International Journal of Computer-Supported Collaborative Learning*, 2011.
- [2.30] Y. Jeanneret, « *Y a-t-il (vraiment) des technologies de l'information ?* » Presses du Septentrion, 2000.
- [2.31] Y. Aragon, C. Trinquier-Alcouffe, « *Introduction à la statistique en sciences sociales* ». Toulouse: Societas Privat, 1979.
- [2.32] J. Viens, « *L'évaluation des environnements, de l'usage aux impacts* ». In CHARLIER et DAELE (Eds.) *Comprendre les communautés virtuelles d'enseignants*. Paris : L'Harmattan, 2006.
- [2.33] E. Drot, « *Outils de communication et disciplines scolaires : quelle(s) rationalité(s) d'usage ?* » Thèse en sciences de l'éducation, ENS Cachan. 354 pages, 2001.
- [2.34] J. M. Turban, « *Listes de diffusion pour enseignants du premier degré : une expérience sociale formative, combinaison des logiques de l'action (intégration, stratégie, subjectivation)* ». Thèse de doctorat. UFR de sciences humaines. Rennes, Université de Rennes 2 - Haute Bretagne, 2004.

- [2.35] F. Villemonteix, « *Informatique scolaire à l'école primaire. Spécificités et devenir du groupe professionnel des animateurs TICE* ». Savoir et Formation. Paris : L'Harmattan, 2010.
- [2.36] Y. Aragon, C. Trinquier-alcouffe, « *Introduction à la statistique en sciences sociales* ». Toulouse : Societas Privat 1979.
- [2.37] Bonnafous, « *L'analyse du discours* », in Olivesi S., dir., Sciences de l'information et de la communication. Objet, savoirs, discipline, Grenoble, PUG, pp. 213-228, 2006.
- [2.38] L. Rourke, T. Anderson, R. Garrison, W. Archer, « *Methodological issues in the content Analysis of Computer Conference Transcripts* ». In International Journal of Artificial intelligence in Education, 2001.
- [2.39] A. Trognon, « *Eléments d'analyse interlocutoire. Apprendre dans l'interaction* », pp. 69-94, 1999.
- [2.40] F. Henri, K. Lundgren-Cayrol, « *Apprentissage collaboratif à distance. Pour comprendre et concevoir les environnements d'apprentissage virtuels* ». Sainte-Foy, Québec : Presses Universitaires du Québec, 2001.
- [2.41] M. Riel, « *Cross-classroom collaboration in global learning circles* ». Dans S. Star (éd.) The cultures of Computing, Oxford: Blackwell, 1995.
- [2.42] L. Bardin, « *L'analyse de contenu* ». Presses Universitaires de France – PUF, 2007.
- [2.43] B. Pudelko, A. Daele, F. Henri, « *Méthodes d'étude de communautés virtuelles* ». Comprendre les communautés virtuelles d'enseignants (daele, Charlier). Paris : L'Harmattan, 2006.
- [2.44] C. LEVI-STRAUSS – « *Anthropologie structurale* ». Paris : Plon, 1958.
- [2.45] S. BACHELARD, « *Quelques aspects historiques des notions de modèle et de justification des modèles* ». In P. Delattre & M. Thellier (Éds.), Élaboration et justification des modèles (p. 3-19). Paris : Maloine, 1979.

- [2.46] E. Malinvaud, « *Méthodes statistiques de l'économétrie* » (3e éd.). Paris : Bordas, 1978.
- [2.47] J. Rodet, « *L'encadrement à distance* », 2000.
- [2.48] A. Kobsa, J. Koenemann, and W. Pohl, “*Personnalised hypermedia presentation techniques for improving online customer relationships*”. The knowlegde engineering, 16(2), pp. 111-155, 2001.
- [2.49] J. Annett, and K. D. Duncan, “*Task analysis and training design*”. Journal of Occupational Psychology, 41, pp. 211-221, 1967.
- [2.50] S. K. Card, T. P. Moran, and A. Newell, “*The psychology of human computer interaction*”. Lawrence Erlbaum, 1983.
- [2.51] H. R. Hartson, A.C. Siochi and D. Hix, “*The UAN: a User Oriented Representation for Direct Manipulation Interface Design*”. ACM Transaction on Information Systems, 8(3), pp. 181-203, 1990.
- [2.52] D. L. Scapin, and C. Pierret-Golbreich, “*MAD : Une méthode analytique de description de tâches* ». Colloque sur l'ingénierie des Interfaces Homme-Machine, Sophia-Antipolis, France, pp. 131-148, 1989.
- [2.53] B.E. John & D.E. Kieras, “*The GOMS Family of User Interface Analysis Techniques: Comparison and Contrast*”. ACM Transaction on Computer-Human Interaction, Vol. 3, No. 4, December, pp. 320-351, 1996.
- [2.54] B. John, A. Vera, M. Matessa, M. Freed, R. Remington, “*Automating CPM-GOMS*”, SIGCHI Conference on Human factors in computing systems, Minneapolis, USA, 2002.
- [2.55] M. Baron, V. Lucquiaud, D. Autard, D.L. Scapin, « *Proceedings of the 18th international conference on Association Francophone d'Interaction Homme-Machine* ». Montreal, Canada, pp. 287 – 288, 2006.

- [2.56] S. Balbo, N. Ozkan, C. Paris, “*Choosing the right task-modeling notation: A taxonomy*”. The Handbook of Task Analysis for Human Computer Interaction, 2004.
- [2.57] E. Bourgeois & G. chapelle, “*Apprendre et faire apprendre*”, Presses Universitaires de France, 2011.
- [2.58] P. Hert, “*Quasi-oralité de l’écriture électronique et sentiment de communauté dans les débats scientifiques en ligne*”. Réseaux, 97, pp 211-259, 1999.
- [2.59] A. R. Kaye, “*Learning together apart*”. In A.R. Kaye (Ed.). Collaborative Learning through computer conferencing. The Najaden Papers. NATO ASI Series F. vol. 90. Berlin: Springer-Verlag, pp. 1-24, 1992.
- [2.60] M. Walckiers, & T. De Praetere, “*L’apprentissage collaboratif en ligne, huit avantages qui en font un must*”. Distances et savoirs, 2(1), pp. 53-75, 2004.
- [2.61] F. Mangenot, “*Analyse sémio-pragmatique des forums pédagogiques sur Internet*”. In J.-M. Salaün & C. Vandendorpe (Eds.) Les défis de la publication sur le Web : hyper lectures, cyber textes et méta-éditions (pp. 103-123). Villeurbanne : Presses de l’ENSSIB, 2004.
- [2.62] A. Bavelas, “*Communication Patterns in Task-oriented Groups*”. In Journal of the Acoustical Society of America, 22, pp. 271-282, 1950.
- [3.01] F. Henri, B. Pudelko, « *Understanding and analyzing activity and learning in virtual communities* ». Journal of Computer Assisted Learning. 19, pp. 474-487, 2003.
- [3.02] E. Wenger, « *Communities of practice. Learning, meaning and identity* ». Cambridge, UK: Cambridge University Press, 1998.
- [3.03] D. R. Anderson, D. J. Sweeney, T.A. Williams, “*Essentials of Statistics for Business and Economics*”, 5<sup>th</sup> edition, 2009.



- [3.04] Stephen Stigler, « *Karl Pearson's theoretical errors and the advances they inspired* », *Statistical Science*, no 23, 2008, p. 261–271
- [3.05] L. J. Williams, R. J. Vandenberg, J. R. Edwards, « *Structural equation modeling in management research: a guide for improved analysis* ». The Academy of Management Annals, 2009.
- [3.06] K. G. Jöreskog, D. Sörbom, « *Recent developments in structural equation modeling* ». *Journal of Marketing Research*, 19, pp 404-416, 1982.
- [3.07] A. 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.
- [3.08] G. A. Churchill, « *A paradigm for developing better Measures of marketing constructs* », *Journal of Marketing Research*, vol. 16, pp 64-73, Feb 1979.
- [3.09] L. Livolsi, P.-X. Meschi, « *Méthodologie quantitative de la recherche en gestion des ressources humaines* », in Allouche, J. (éditeur), *Encyclopédie des Ressources Humaines*, Vuibert, Paris, pp. 897-908, 2003.
- [3.10] S. B. MacKenzie, P. M. Podsakoff, C. B. Jarvis, « *The Problem of Measurement Model Misspecification in Behavioral and Organizational Research and Some Recommended Solutions* », *Journal of Applied Psychology*, Vol. 90, No. 4, pp 710–730, 2005.
- [3.11] C. Goodwin, "*Participation Frameworks in Children's Argument*", In *rowing Into A Modern World: Proceedings from An International Interdisciplinary Conference on the Life and Development of Children in Modern Society* (Karin Ekberg, Per Egil Mjaavatn, eds.), Trondheim, The Norwegian Centre for Child Research, pp. 1188–1195, 1988.
- [3.12] D. Gefen, D.W. Straub, M. Boudreau, « *Structural Equation Modeling Techniques and Regression: Guidelines for Research Practice* », *Communications of AIS* Volume 4, Article 7, 2000.

- [3.13] R.H. Hoyle, « *Structural equation modelling: Concepts issues and applications* », Edition Sage, London, 1995.
- [3.14] D. Kaplan, « *Structural equation modeling: Foundations and extension* », Sage, Thousand Oaks, CA., 2000.
- [3.15] J.F. Hair, B. Black, Babin B., R.E. Anderson, R.L. Tatham, « *Multivariate data analysis* », Pearson Prentice-Hall, 7ème ed, 928 pages, 2009.
- [3.16] K.A. Bollen, J.S. Long, « *Testing structural equation models* ». Newbury Park, CA : Sage, 1993.
- [3.17] R.E. Schumaker, R.G. Lomax, « *A beginner's guide to structural equation modelling* », Lawrence Erlbaum Associates, 2ème edition, London, 2004.
- [3.18] P. Roussel, F. Durrieu, E. Campoy, A. El Akremi, « *Méthodes d'équations structurelles : Recherches et applications en gestion* », Edition Economica, Paris, 2002.
- [3.19] L. Chaput, « *Modèles contemporains en gestion* », Presses de l'Université du Québec, 2006.
- [3.20] J.R. Magnus, H. Neudecker, "*Matrix differential calculus with applications in statistics and econometrics*", New York, Edition Wiley, 1999.
- [3.21] K. Schermelleh-Engel, H. Moosbrugger, & H. Müller, « *Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures, Methods of Psychological Research Online* », 8(2), 23-74, 2003.
- [3.22] D. L. Jackson, J. A. Gillaspay, & R. Purc-Stephenson, « *Reporting practices in confirmatory factor analysis: An overview and some recommendations* ». Psychological Methods, 14(1), 6-23, 2009.
- [3.23] R. P. McDonald, & M. H. R. Ho, « *Principles and practice in reporting statistical equation analyses* ». Psychological Methods, 7(1), 64-82, 2002.

- [3.24] H. Gatignon, « *Confirmatory Factor Analysis in Statistical analysis of management data* ». DOI : 10.1007/978-1-4419-1270-1\_4, 2010.
- [3.25] L. Hu, & P. M. Bentler, "*Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives*". *Structural Equation Modeling: A Multidisciplinary Journal*. 6 (1) : 1–55, 1999.
- [3.26] H. Baumgartner, & C. Hombur, « *Applications of structural equation modeling in marketing and consumer research: A review* ». *International Journal of Research in Marketing*, 13, 139-161, 1996.
- [3.27] L. R. Tucker, & C. Lewis, « *A reliability coefficient for maximum likelihood factor analysis* ». *Psychometrika*, 38, 1-10, 1973.
- [4.01] C. Choquet, S. Iksal, « *Modeling Tracks for the Model Driven Reengineering of a TEL System* », In: *Journal of Interactive Learning Research (JILR)*, Vol. 18 n°2, p. 161-184, 2007.
- [4.02] A. Bandura, "*Social foundations of thought and action: A social cognitive theory*". Prentice-Hall, 1986.
- [4.03] A. Bandura, « *Auto-efficacité : le sentiment d'efficacité personnelle* ». De Boeck Supérieur, 2003.
- [4.04] J. M. Keller, "*Motivational design of instruction*". Dans C.M. Reigeluth (dir.), *Instructional theories and models: An overview of their current status* pp. 383-434. New York: Lawrence Erlbaum Associates, 1983.
- [4.05] P. R. Pintrich, "*An achievement goal theory perspective on issues in motivation terminology: Theory and research*". *Contemporary Educational Psychology*, 25(1), pp. 92-104, 2000.

- [4.06] P. R. Pintrich, "*Motivation and classroom Learning*". Dans W. M. Reynolds et G. E. Miller (dir.), *Handbook of psychology*, vol 7: Educational psychology pp. 103-122, 2003.
- [4.07] J. A. Fredricks, P.C. Blumenfeld, et A.H. Paris, "*School engagement: Potential of the concept, state of évidence*". *Review of Educational Research*, 74(1), pp. 59-109, 2004.
- [4.08] E. A. Linnenbrink, et P. R. Pintrich, "*The role of self-efficacy beliefs in student engagement and learning in the classroom*". *Reading and Writing Quarterly: Overcoming Learning Difficulties*, 19(2), pp. 119-137, 2003.
- [4.09] J. D. Willms, S. Friesen, et P. Milton, « *What did you do in school today? Transforming classrooms through social, academic, and intellectual engagement* ». First National Report. Toronto, Canadian Education Association, 2009.
- [4.10] G. Molinari, B. Poellhuber, J. Heutte, E. Lavoué, D. S. Widmer et P. A. Caron, « *L'engagement et la persistance dans les dispositifs de formation en ligne : regards croisés* », *Distances et médiations des savoirs*, 13 | 2016, mars 2016,
- [4.11] J. A. Fredricks, W. McColskey, J. Meli, B. Montrosse, B. Mordica, et K. Mooney, "*Measuring student engagement in upper elementary through high school: a description of 21 instruments*". Regional Educational Laboratory, Greensboro, 2011.
- [4.12] R. J. Vallerand et E. E. Thill, « *Introduction au concept de motivation* ». Dans Vallerand, R. J. et Thill, E. E. (dir.), *Introduction à la psychologie de la motivation* pp.3-39. Éditions Études vivantes, 1993.
- [4.13] R. J. Vallerand, M. R. Blais, N. M. Briere, et L. G. Pelletier, "*Construction and validation of the Motivation toward Education Scale*". *Canadian Journal Of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, 21, pp. 323-349, 1989.

- [4.14] H. Fournier, R. Kop, et H. Sitlia, “*The value of learning analytics to networked learning on a personal learning environment*”. Proceedings of the 1st International Conference on Learning Analytics and Knowledge, pp. 104-109, 2011.
- [4.15] 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.
- [4.16] B. Poellhuber, M. Chomienne, et T. Karsenti, « *Quels sont les parcours menant à l’abandon en formation à distance au collégial ?* », Distance et Médiation de Savoir, 10(3), pp. 1-33, 2008.
- [4.17] D. W. Johnson, R. T. Johnson, et K. A. Smith, “*Cooperative learning: Improving university instruction by basing practice on validated theory*”. Journal on Excellence in University Teaching, 25(3-4), 85-118, 2013.
- [4.18] B. Barron, “*When smart groups fail*”. The journal of the learning sciences, 12(3), pp. 307-359, 2003.
- [4.19] G. Salomon, T. Globerson, & E. Guterman, « *The computer as a zone of proximal development: Internalizing reading-related metacognitions from a Reading Partner.* » Journal of Educational Psychology, 81(4), pp. 620-627, 1989.
- [4.20] J. Roschelle, et S. D. Teasley, “*The construction of shared knowledge in collaborative problem solving*”. Dans Computer supported collaborative Learning, pp. 69-97. New York, Springer, 1995.
- [4.21] D. D. Suthers, “*Technology affordances for intersubjective meaning making: A research agenda for CSCL*”. International Journal of Computer-Supported Collaborative Learning, 1(3), 315-337, 2006.
- [4.22] R. E. Slavin, “*Cooperative Learning*”. Dans V. G. Aukrust (dir.), Learning and Cognition in Education, pp. 160-66, Boston: Academic Press, 2011.

- [4.23] J. J. Quintin, et M. Masperi, « *Reliance, liance et alliance : opérationnalité des concepts dans l'analyse du climat socio-relationnel de groupes restreints d'apprentissage en ligne* », 2010.
- [4.24] M. Scardamalia, et C. Bereiter, "A *brief history of knowledge building*". Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie, 36(1), 2010.
- [4.25] P. Jermann, "Computer support for interaction regulation in collaborative problem-solving". Unpublished Ph. D. thesis, Genève, Université de Genève, 2004.
- [4.26] M. W. Berkowitz, et J. C. Gibbs, "Measuring the developmental features of moral discussion". Merrill-Palmer Quarterly, 29(4), pp. 399-410, 1983.
- [4.27] G. Molinari, M. Sangin, M. A. Nüssli, et P. Dillenbourg, "Effects of knowledge interdependence with the partner on visual and action transactivity in collaborative concept mapping". Dans Proceedings of the 8th international conference on International conference for the learning sciences, International Society of the Learning Sciences, 2, pp. 91-98, 2008.
- [4.28] K. Kreijns, P. A. Kirschner, W. Jochems, et H. Van Buuren, "Measuring perceived social presence in distributed learning groups". Education and Information Technologies, 16(4), pp. 365-381, 2011.
- [4.29] J. L. Hale, J. K. Burgoon, et B. Householder, "The relational communication scale". Dans V. Manusov (dir.), The sourcebook of nonverbal measures: Going beyond words pp. 127-139. Mahwah (N.J.): Erlbaum, 2005.
- [4.30] 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.
- [4.31] M. Walckiers, T. De Praetere, « *L'apprentissage collaboratif en ligne, huit avantages qui en font un must* », Lavoisier | « Distances et savoirs », page 60 à 62, 2004.

- [5.01] B. Albero, « *L'autoformation dans les dispositifs de FOAD : instrumenter le développement de l'autonomie dans les apprentissages* », in Saleh I., Lepage D. et Bouyahi S. (dir.), *Les TIC au coeur de l'enseignement supérieur*, Laboratoire Paragraphe, Université Paris 8, novembre 2002
- [5.02] M. Linard, « *Conception de dispositifs et changement de paradigme en formation* », *Education Permanente* n° 152, 2002
- [5.03] J. M. Dutrénit, « *La compétence sociale, diagnostic et Développement* », L'Harmattan, Paris, 1997
- [5.04] A. Dix, D. Ramduny-Ellis, & J. Wilkinson., « *Trigger analysis –understanding broken tasks* ». In : D. Diaper & N. Stanton (Eds.), « *The Handbook of Task Analysis for Human-Computer Interaction* ». Mahwah : Erlbaum 2004.
- [5.05] P. Dillenbourg, « *What do you mean by collaborative learning?* » In : P. Dillenbourg (Ed). *Collaborative –learning : Cognitive and Computational Approaches*. pp. 1 - 19. Oxford, Elsevier, 1996.
- [5.06] R. Nachmias, D. Mioduser, A. Oren, & J. Ram, « *Web-supported emergent-collaboration in higher education courses* ». *Educational Technology & Society*, 3(3), pp. 94 -104, 2000.
- [5.07] M.J. Baker, « *Forms of cooperation in dyadic problem-solving. Revue d'Intelligence Artificielle* », 16, 4 - 5, pp. 587 - 620. 2002.
- [5.08] B. De Lièvre, C. Depover, A. Strebelle, « *Les avis individuels sont-ils pris en compte lors d'une activité collaborative à distance ?* », In : P. Tchounikine, M. Joab, L. Trouche (Eds). « *Environnements informatiques pour l'apprentissage humain* », 2005, pp. 93-104. Montpellier, France, 2005.
- [5.09] B. De Lièvre, G. Temperman. « *Trois modalités de structuration d'un forum collaboratif : comment les étudiants les jugent-il ?* » *Journées communication et*

apprentissage instrumentées en réseau 2008, 987-2-74622138-3. <halshs-01112579>, 2008.

- [5.10] C. Verzat, N. O'Shea et B. Raucourt, « *Réguler le leadership dans les groupes d'étudiants en APP* », Revue internationale de pédagogie de l'enseignement supérieur pp. 31-1 | 2015.



## FICHE DE RENSEIGNEMENT

Nom : RAKOTOMALALA

Prénom : Mirisoa

Adresse de l'auteur : Lot II N 65 i Analamahitsy  
tanàna, 101 Antananarivo

Téléphone : +261 34 17 998 49

E-mail : mrakotom-diego@yahoo.fr



Titre de la thèse :

### **MODELISATION DE L'APPRENTISSAGE ASYNCHRONE DANS UNE FORMATION OUVERTE ET A DISTANCE**

Nombre de pages : 290

Nombre de tableaux : 13

Nombre de figures : 47

Directeur de thèse :

Monsieur RANDIMBINDRAINIBE Falimanana, *Professeur Titulaire*

Téléphone : +261 34 06 466 90

E-mail : falimanana@mail.ru

Co-directeur de thèse :

Monsieur RAVALIMINOARIMALALASON Toky Basilide, *Docteur*

Téléphone : +261 34 003 164 33

E-mail : tokybaz@gmail.com

## ***Résumé***

Cette thèse soutient la possibilité d'établir une modélisation mathématique qui s'applique à une formation ouverte et à distance. Un modèle d'apprentissage asynchrone a été rendu disponible via l'approche par la relation linéaire d'une modélisation d'équation structurelle à variables latentes. Ce modèle a été validé empiriquement et peut être utilisé dans toutes les formations en ligne. Son adaptabilité au contexte malagasy ne souffre d'aucune remise en question étant donné que l'outil principal de formation est un outil asynchrone : le forum. L'attitude (A) adoptée par l'apprenant face au forum de discussion va se définir en fonction de son engagement (E). Cette variable latente E a un impact important sur sa cognition (C) et une conséquence directe sur son résultat pédagogique. Par ailleurs, la collaboration de l'apprenant est un indicateur incontournable pour pouvoir mesurer la variable latente C. A cet indicateur s'ensuivent également sa motivation et sa relation communicationnelle avec tous les protagonistes de la formation.

***Mots-clés :*** *modélisation, apprentissage, asynchrone, collaboration, FOAD*

## ***Abstract:***

This thesis supports the possibility of establishing a mathematical modeling that applies to open and distance learning. An asynchronous learning model was made available through the linear relationship approach of a structural equation modeling with latent variables. This model was validated empirically and can be used in all online courses. Its adaptability to the Malagasy context does not suffer from any questioning since the main training tool is an asynchronous tool: the forum. The attitude (A) adopted by the learners in the discussion forum is defined according to their commitment (E). This latent variable E has a significant impact on the learners' cognition (C) and a direct consequence on their educational outcome. In addition, the learners' collaboration is an essential indicator for measuring the latent variable C. Their motivation and their communication relationship with all the protagonists of the training are also taken into account.

***Keywords :*** *modeling, learning, asynchronous, collaboration, ODL*