

Les aspects méthodologiques d'une recherche portant sur la persévérance et l'abandon des études en formation à distance

Présentation des paramètres
d'une recherche au Québec

Guillaume **DESJARDINS**, Louise **SAUVÉ**,
Cathia **PAPI** et Serge **GÉRIN-LAJOIE**

Étudier les facteurs d'abandon et de persévérance dans les études en formation à distance (FAD) nécessite d'examiner à la fois le cadre théorique des recherches et les analyses statistiques qui y sont utilisés. Dans le chapitre précédent, les modèles théoriques sur l'abandon des études ont été présentés, et les facteurs qui semblent avoir une influence sur la persévérance et l'abandon des études en FAD ont été identifiés. Dans ce chapitre, nous décrivons en détail les instruments de mesure et les analyses statistiques qui ont fait l'objet de l'étude.

Nous procédons d'abord à une brève revue de la littérature portant sur les instruments de mesure et les analyses statistiques utilisés dans les recherches sur l'abandon et la persévérance dans les études avant d'en venir à préciser ces notions. Ensuite, nous décrivons la méthodologie employée dans le cadre d'une étude financée par le Fonds de recherche du Québec – Société

et culture (FRQSC) intitulée *Regard des apprenant.es universitaires sur les modes d'organisation et d'encadrement pédagogique en formation à distance et en ligne* en précisant la méthode d'échantillonnage et les instruments de mesure de la recherche¹. Finalement, les trois catégories d'analyses statistiques – tris à plat, analyses croisées, analyses de *clusters* – mises en œuvre sont explicitées. Ce chapitre détaille ainsi les fondements méthodologiques de la recherche dont les résultats sont documentés et discutés dans les chapitres 4 à 6.

1 QUE DIT LA LITTÉRATURE SUR L'ASPECT MÉTHODOLOGIQUE DE LA RECHERCHE ?

Les recherches qui traitent de l'abandon des études proviennent d'un large spectre tant sur le plan de la population visée (enseignement primaire, secondaire, supérieur) que du format de formation offerte (en classe ou sur le campus, hybride ou complètement en ligne). Néanmoins, les méthodologies habituellement utilisées dans ces recherches empiriques (Tan et Shao, 2015) sont : 1) la collecte de données grâce à des questionnaires et l'analyse des corrélations entre les variables ; 2) le relevé des tendances de la distribution des variables démographiques en relation avec des variables d'abandon, en utilisant les données acquises au moyen de questionnaires ou d'un système d'information ; 3) l'identification des raisons de l'abandon par l'analyse de transcriptions d'entrevues réalisée avec les parties concernées ; ou 4) une combinaison de ces trois méthodes. Examinons-les plus en détail.

Tout d'abord, certaines recherches utilisent des statistiques descriptives et des propos recueillis lors d'entrevues pour étudier les facteurs associés à la persévérance ou à l'abandon des études. Ces recherches peuvent utiliser uniquement des statistiques descriptives résultant d'enquêtes par questionnaire (Rekkedal, 1985 ; Garrison, 1987), des transcriptions d'entrevues (Li, Chen et Han, 2010 ; Ramsdal, Bergvik et Wynn, 2018 ; Bakker *et al.*, 2019), ou employer une méthodologie mixte (Li, Niu et Ding, 2012 ; Arifin, 2018). Par exemple, l'étude de Bakker et ses collaborateurs (2019) s'appuie sur une méthodologie comprenant des entrevues

1. Dans ce chapitre, nous ne traitons pas des instruments de mesure et des analyses statistiques que nous avons réalisées avec les intervenants (professeurs et tuteurs). Pour en savoir davantage sur cet aspect, vous pouvez consulter le rapport de recherche (Sauvé *et al.*, 2020).

semi-structurées afin de comprendre pourquoi certains étudiants abandonnent leur baccalauréat en sciences infirmières lors de la dernière année du programme. Le recours à une approche qualitative, telle que l'utilisation d'entrevues ou de groupes de discussion, est recommandé pour un protocole de recherche exploratoire (Frey et Fontana, 1991; Rubin et Rubin, 2011) qui vise la découverte de nouveaux facteurs associés à l'abandon (Andrà *et al.*, 2011; Gupta *et al.*, 2020).

D'autres recherches utilisent la régression logistique pour analyser la relation entre les différentes variables en lien avec l'abandon des étudiants (Zhang *et al.*, 2004; Simpson, 2006; Doherty, 2006; Roblyer *et al.*, 2008; Nichols et Levy, 2009). Cette méthodologie quantitative est appropriée lorsque le chercheur tente de mesurer le degré d'association entre deux ou plusieurs variables (No et Hirakawa, 2012; Yasmin, 2013). Par exemple, l'étude de Nichols et Levy (2009) croise diverses variables telles que l'attitude des étudiants face à l'informatique, leur motivation intrinsèque et extrinsèque ainsi que leurs performances scolaires antérieures afin d'élaborer un modèle prédicteur de l'abandon des « étudiants-athlètes » suivant des cours en ligne.

Plus récemment, des chercheurs ont tenté d'utiliser l'apprentissage automatique (*machine-learning*) (Kotsiantis *et al.* 2003) pour prédire la possibilité que des étudiants abandonnent leurs études (Dalipi, Imran et Kastrati, 2018; Chung et Lee, 2019; Figueroa-Cañas et Sancho-Vinuesa, 2019; Del Bonifro *et al.*, 2020; Kemper, Vorhoff et Wigger, 2020; Muthukumar et Bhalaji, 2020). L'apprentissage automatique est un champ d'études de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune d'elles (Kotsiantis *et al.*, 2003). Cependant, les travaux fondés sur cette méthodologie ont une visée pragmatique et ne cherchent pas à comprendre le phénomène d'abandon. Alors que les modèles d'apprentissage automatique sont conçus pour faire des prévisions les plus précises possibles, les modèles statistiques, eux, sont conçus pour faire des inférences sur les relations entre les variables. En ce sens, leur finalité diffère.

Enfin, une dernière catégorie d'études tente de regrouper les différents facteurs relevés dans la littérature en un tout cohérent (López-Goñi, Fernández-Montalvo et Arteaga, 2012; Crain-Dorough, 2003; Iam-On et Boongoen, 2017; Rahmat, 2017; Hsu et Yeh, 2020). Les auteurs de ces recherches emploient l'analyse de *cluster* pour effectuer ces regroupements. Les deux modèles statistiques les plus fréquemment utilisés sont le modèle hiérarchique (Bowers, 2010; Thanassoulis *et al.*, 2017; McDermott, Anderson et Zaff, 2018) et le modèle en deux étapes (*two-step*) (Krstić *et al.*, 2019; Hsu et Yeh, 2020; Gilar-Corbi *et al.*, 2020; Lottering, Hans et Lall, 2020). Malgré des résultats prometteurs, Battaglia et ses collaborateurs (2016) indiquent que l'analyse de *cluster* est encore peu exploitée dans le domaine de l'éducation.

2 L'OPÉRATIONNALISATION DES NOTIONS DE PERSÉVÉRANCE ET D'ABANDON DES ÉTUDES

Tinto (2005) relève qu'il est difficile de définir ce qu'est la persévérance dans les études en raison de la complexité du phénomène auquel elle se rapporte. La persévérance dans les études prend une grande variété de formes selon l'angle d'approche, que ce soit la notion de rétention (Yorke et Thomas, 2003; Schertzer et Schertzer, 2004; Phillion *et al.*, 2010), de maintien d'effectifs (Grayson et Grayson, 2003; Millet, 2012), de diplomation ou d'inscription continue de l'étudiant à son programme jusqu'à l'obtention de son diplôme (Conseil supérieur de l'éducation [CSE], 2000).

De ces différents points de vue, il ressort que la constance de l'étudiant dans son programme universitaire, et ce, jusqu'à l'obtention de son diplôme ou de son certificat, constitue l'élément central sur lequel l'accent est généralement mis (Romainville et Michaut, 2012). Cependant, les conditions qui favorisent cela demeurent assez diverses et varient selon les étudiants, les programmes d'études ou même les institutions d'enseignement et leurs contextes (Barr-Telford *et al.*, 2003).

Dans le contexte particulier de notre recherche portant sur les étudiants inscrits dans un programme universitaire de 1^{er} cycle (certificat ou baccalauréat) en FAD qui sont invités à remplir un questionnaire sur l'engagement pendant trois sessions d'études successives, la persévérance est définie comme la réinscription de

l'étudiant à son institution à chaque session d'étude ou jusqu'à l'obtention d'un diplôme (certificat, baccalauréat). Les étudiants dont l'absence est justifiée et qui se réinscrivent dans les délais prescrits sont considérés comme persévérants.

Tout comme la notion de persévérance, celle d'abandon recouvre de nombreuses définitions selon les angles d'analyse du problème (Grayson et Grayson, 2003; King, 2005): absence, interruption, retour, transfert, départ de l'institution, sortie du système, etc. Pour certains chercheurs, cette notion est assimilée à celle de décrochage (Fortin *et al.*, 2006; Blaya, 2010; Bernard, 2011) ou d'attrition (Tinto, 1987; Grayson et Grayson, 2003). Spady (1970) se réfère au changement d'institution ou au départ du système universitaire et Sauvé et Viau (2003) précisent que l'abandon est un départ volontaire ou non des étudiants de leur programme universitaire sans avoir obtenu leur diplôme. Comme précisé dans le chapitre 1, au Québec, l'abandon lors des études universitaires correspond à une non-réinscription depuis au moins cinq trimestres consécutifs. Dans cette optique, l'abandon se constate par une sortie du système scolaire par l'étudiant avant l'obtention du diplôme.

En définitive, le départ de l'étudiant de son institution universitaire avant l'obtention de son diplôme apparaît comme l'élément central des différentes définitions de l'abandon recensées. En regard du contexte de notre étude et du temps alloué, nous distinguerons l'étudiant ayant abandonné de l'étudiant qui a persévéré par sa réinscription ou non à au moins deux sessions consécutives pendant la période de notre recherche. Nous prendrons aussi bien en compte l'abandon volontaire signalé ou non à l'établissement d'attache que l'abandon involontaire causé par une contrainte particulière de quitter l'établissement d'attache.

Enfin, de la même manière, nous considérerons l'abandon d'un cours comme l'arrêt du travail dans un cours sans obtention d'une note finale (Stiller et Bachmaier, 2017).

3 UNE MÉTHODOLOGIE EN QUATRE ÉTAPES

Nous proposons de présenter les quatre étapes mises en œuvre pour atteindre l'objectif de cette recherche. La première étape se concentre sur la méthode d'échantillonnage des participants. La

deuxième étape porte sur l'élaboration et la distribution d'un questionnaire en vue d'une collecte de données auprès des étudiants. La troisième étape traite de la création d'une grille d'analyse de design de cours. Cette grille permet de recueillir les variables qui sont présentes (ou non) dans les cours en ligne. Enfin, la quatrième étape élabore une série de questions pour soutenir les groupes de discussion réalisés avec les étudiants afin d'approfondir certains éléments relevés lors de la collecte de données quantitatives.

3.1 La méthode d'échantillonnage

Un échantillonnage stratifié a été utilisé dans le cadre de ce projet pour sonder la population étudiante ainsi que ses cours. Trois éléments ont été pris en compte pour choisir les cours à analyser. Premièrement, le département : les cours sélectionnés proviennent d'au moins trois départements différents (sciences humaines, lettres et communication ; sciences de l'éducation ; sciences de l'administration)². Deuxièmement, le nombre d'étudiants par cours : prenant en considération que les cours ont en moyenne 312 inscriptions par an (selon les données obtenues par le registraire pour l'année 2016), nous avons décidé de ne retenir que des cours ayant plus de 300 inscriptions annuelles. Troisièmement, la variabilité des taux d'échec et d'abandon des cours : afin de refléter la diversité de la réalité, nous avons sélectionné 19 cours ayant des taux d'échec allant de 4,35 % à 26,51 % et des taux d'abandon allant de 4,3 % à 26,35 %.

3.2 Les instruments de mesure de la recherche

Afin de rejoindre la population ciblée, des instruments de mesure pour collecter des données quantitatives et qualitatives ont été créés. Ces outils comprennent un questionnaire autoadministré, une grille visant la description du design pédagogique des cours ainsi que l'élaboration d'un protocole pour la réalisation de groupes de discussion avec les étudiants ayant accepté d'y participer.

2. Cette première stratégie évite les biais que provoquerait l'utilisation du nombre d'inscriptions comme unique critère de sélection. En effet, les données institutionnelles où la recherche a eu lieu démontrent que les cours à forte inscription proviennent tous du même département, soit celui d'administration.

3.2.1 Un questionnaire en ligne pour les étudiants

Le questionnaire destiné aux étudiants comprend huit sections et traite de différents sujets incluant les données sociodémographiques, l'utilisation de stratégies d'apprentissage, la satisfaction de l'encadrement reçu et des outils numériques proposés dans le cours. D'autres variables incluent le degré de satisfaction des interactions des étudiants par rapport à leurs pairs ainsi qu'à leurs proches.

La section des données sociodémographiques et scolaires comprend 23 questions (tableau 3.1). Excluant le matricule, l'établissement et l'adresse courriel, chaque variable établie dans cette section du questionnaire provient d'écrits scientifiques ayant indiqué une relation entre cette dernière et la persévérance dans les études (Sauvé *et al.*, 2020).

TABLEAU 3.1
Variables sociodémographiques et scolaires dans le questionnaire

L'âge	La scolarité du père
Le genre	La scolarité de la mère
L'état civil	La distance (en temps) de l'établissement d'études
La citoyenneté/le statut légal canadien	La session d'études de complétude du questionnaire
La situation familiale	Le nombre de crédits suivis pendant la session
Le diplôme le plus élevé atteint	Le financement des études
La présence de diplôme à l'étranger	Les heures de travail par semaine
Le nombre de cours crédités	Le jugement de la situation financière
Le nombre de crédits complétés	L'établissement d'attache
Le type de programme	L'adresse courriel
La présence d'un handicap	Le matricule scolaire
La langue maternelle	

La construction du questionnaire a été réalisée en respectant les trois étapes proposées dans la littérature à ce sujet (Franklin *et al.*, 2010; Cooper et Schindler, 2011), soit le prétest, le pilote et l'administration.

Le prétest comprend deux étapes, à savoir la création du questionnaire ainsi que sa validation. Concernant la création, il a été convenu que le questionnaire serait divisé en trois sections, chacune faisant l'objet d'une sollicitation distincte

des participants. Cette stratégie a été employée afin d'éviter un effet d'essoufflement (Burchell et Marsh, 1992; Krosnick, 2018) des étudiants.

La première sollicitation comprend des questions qui portent sur les variables sociodémographiques que les étudiants remplissent en sélectionnant des items parmi de multiples propositions (p. ex. «Célibataire, Marié, Conjoint de fait, Divorcée, Autre») ou en écrivant une courte réponse (p. ex. le participant doit inscrire le nombre de crédits validés dans son programme). Cette section comprend aussi des énoncés portant sur l'utilisation de stratégies d'apprentissage par les étudiants (p. ex. «Je suis en mesure de me fixer un horaire d'étude et de le respecter»). Les participants répondent à ces questions en se positionnant sur des échelles de Likert à cinq points.

La majorité du contenu des questions de la deuxième sollicitation se concentre sur le cours dans lequel l'étudiant a été sollicité. Le questionnaire traite d'abord du mode d'organisation pédagogique qui regroupe les stratégies déployées dans les cours et dans le programme (19 items) ainsi que les technologies présentes, utilisées et souhaitées par les étudiants dans leur cours en FAD (14 items). Il collecte également des informations sur le type d'encadrement, les outils de communication utilisés par les intervenants et les étudiants, les interactions avec la personne tutrice, les autres étudiants et les proches (35 items). Le questionnaire inclut enfin les énoncés sur la persévérance dans les études qui touchent l'intégration sociale (4 items) et l'intégration scolaire (11 items). Les étudiants répondent à ces questions majoritairement à l'aide d'une échelle de Likert.

Lors de la troisième sollicitation est proposé un questionnaire sur les styles d'apprentissage (ISALEM-97)³ développés par une équipe interdisciplinaire du Laboratoire d'Enseignement Multimédia de l'Université de Liège, sous la direction scientifique de Brigitte Monfort (Honorez *et al.*, 2000). Ce questionnaire est inspiré des travaux de Kolb (1984). Il comprend 12 situations tirées de la vie quotidienne et universitaire. Chaque question

3. N'ayant pas obtenu de différence statistiquement significative ($p = .05$) entre les styles d'apprentissage et l'abandon des cours ou la non-réinscription, les résultats ne seront pas présentés dans cet ouvrage.

inclut quatre énoncés que les étudiants doivent coter différemment selon l'échelle à quatre items: «Tout à fait moi, Souvent moi, Parfois moi, Rarement moi».

Ainsi, une première itération en format papier a été distribuée à quatre chercheurs du projet de recherche afin d'assurer la validité apparente des construits (Leiva *et al.*, 2006). Aucune modification majeure n'a été notée à cette étape. Par la suite, l'équipe de recherche a réalisé un groupe de discussion avec cinq étudiants universitaires de premier cycle. Le groupe de discussion avait pour objectifs de s'assurer que les questions étaient uniformément comprises par les participants (c'est-à-dire que la question est compréhensible) et sans ambiguïté (c'est-à-dire que la question est comprise uniformément chez les participants). Il visait également à vérifier que les énoncés portaient sur des phénomènes qui sont analytiquement conséquents (c'est-à-dire que les énoncés ont une validité conceptuelle pour les chercheurs). En outre, les participants ont été invités à faire part de leur réaction générale au questionnaire et à indiquer les améliorations qu'ils estimaient devoir y apporter. Hormis des modifications mineures, aucune modification sur le fond des énoncés n'a été proposée par le groupe de discussion.

Selon Teijlingen et Hundley (2001), les tests pilotes sont utilisés pour évaluer chaque phase de la mise en œuvre du questionnaire, y compris les étapes administratives. Les études pilotes reproduisent l'enquête finale à une échelle plus petite. Pour ce faire, les chercheurs ont sollicité une cohorte d'étudiants de la session d'été 2018 dans trois cours de l'université. Afin d'encourager la participation, une entente établie avec les trois professeurs responsables des cours sélectionnés offrait de bonifier de 2% la note finale de l'étudiant au cours s'il remplissait le questionnaire dans son entièreté. Sur 201 étudiants sollicités, 53 ont rempli le questionnaire, soit un taux de réponse de 26,36%. Ce test pilote a notamment permis l'établissement des règles et procédures de l'équipe de recherche lors des sollicitations étudiantes (format, invitation, paramètres). Puisqu'aucune modification n'a été apportée lors du test pilote, les données étudiantes de l'été 2018 ont été conservées dans l'échantillon final (Holloway, 1997; Xuan *et al.*, 2020).

Le déploiement du questionnaire s'est réalisé en ligne à l'aide de la plateforme LimeSurvey (<limesurvey.org>) en suivant les règles et restrictions indiquées par le comité d'éthique. Les questionnaires étaient disponibles entre l'été 2018 et l'été 2019. Les participants ayant reçu une invitation de la part des chercheurs devaient cliquer sur un lien hypertexte qui les dirigeait directement sur la page Web du questionnaire. Le questionnaire était disponible uniquement en français et pouvait être rempli à l'aide d'un ordinateur, d'une tablette ou d'un téléphone intelligent.

À la fin de la session, le registraire de l'établissement a fourni à l'équipe de recherche la note finale du cours ainsi que son statut d'inscription (inscrit ou abandon du cours). Les étudiants ayant répondu à la sollicitation ont indiqué leur accord concernant cette transmission d'informations comme condition préalable à l'accès au questionnaire.

3.2.2 Une grille d'analyse des cours universitaires en ligne

La grille d'analyse des cours contient 6 sections et comprend un total de 623 variables par cours analysées. La première section contient les informations générales du cours comme son titre, son programme d'attache, son adresse Web (URL) ainsi que le nombre d'étudiants qui y sont inscrits. La deuxième section concerne le mode d'organisation pédagogique du cours, plus précisément la présence (ou non) d'outils technologiques offerts. La disponibilité des outils pouvait être considérée au sein même du cours, du département ou de l'université. Un total de 14 outils a été répertorié dans la grille d'analyse (annexe 3.1).

La troisième section comprend une analyse séance par séance du cours sur 67 variables. Ces variables incluent les modalités d'apprentissage, la finalité des activités, les formules pédagogiques employées, les modalités d'apprentissage, les formules d'évaluation, le type de correction, etc. À titre d'exemple, sur le plan de la finalité des activités, nous retrouvons les items suivants: découverte (sensibilisation, exploration), appropriation de nouvelles connaissances, entraînement (exercices), évaluation diagnostique (vérification des connaissances avant de faire une activité), évaluation formative (vérification des apprentissages sans notation), évaluation sommative (vérification notée des apprentissages) et encadrement (intervention).

La quatrième section traite des éléments relatifs au guide d'étude (aussi appelé syllabus, plan de cours): les objectifs généraux et spécifiques du cours, le cheminement dans le cours (liste des activités ou des actions à réaliser par les étudiants sous forme de feuille de route ou de calendrier détaillé), la description ou la présentation du contenu du cours, des activités d'apprentissage, des activités d'évaluation (diagnostique, formative et sommative), des aides à la navigation dans l'environnement Web (consignes, plan du site, guide) et des ressources pour faciliter le travail collaboratif.

La cinquième section traite du mode d'encadrement pédagogique des étudiants dans le cours. Plus précisément, quatre éléments sont compilés. Tout d'abord, le type d'encadrement est noté et peut être répertorié dans quatre catégories :

1. En présence physique (période en classe, période de disponibilité au bureau, etc.)
2. En ligne synchrone (téléphone, clavardage [*chat*], audio ou visioconférence de type Skype, Adobe Connect, etc.)
3. En ligne asynchrone (courriel, liste de distribution, forum, etc.)
4. Aucun encadrement prévu

Par la suite, une sous-section est prévue afin d'indiquer si (et le cas échéant, comment) l'encadrement est présenté aux étudiants. Par exemple, les modalités d'encadrement pourraient être présentées dans le plan de cours, par rencontre présentielle ou par une vidéo d'introduction au cours. Les autres variables de la section du mode d'encadrement comprennent les moyens offerts aux étudiants pour communiquer entre eux (de façon collaborative ou non) ainsi que les interactions possibles entre l'intervenant et les étudiants comme l'illustre le tableau 3.2.

Enfin, la sixième section tente d'illustrer la typologie du cours. Ainsi, les éléments du cours tels que le démarrage du cours, le déroulement des séances et les activités d'évaluation sont pondérés selon leur modalité (en présentiel, en ligne ou à distance). Le tableau 3.3 est un exemple de recueil de données que l'on peut trouver dans cette section.

TABLEAU 3.2

Interactions entre l'intervenant et les étudiants

Interactions entre l'intervenant et les étudiants	Présentées	Description	Commentaire
a) Indication que l'enseignant prend contact avec les étudiants dès la première semaine et se présente.			
b) Le rôle de l'enseignant et, le cas échéant, des autres intervenants impliqués dans le cours est expliqué aux étudiants, notamment :			
• Réponses aux questions sur le contenu du cours			
• Rétroactions sur des activités non notées			
• Rétroactions sur des activités notées			
• Réponses à des questions administratives			
• Réponses à des questions d'ordre technologique			
• Autre. Précisez.			
c) Les moyens utilisés pour répondre aux questions des étudiants sont clairement indiqués.			
d) Le délai de réponse aux questions des étudiants est clairement indiqué dans le plan de cours.			
e) Le délai de correction des évaluations est indiqué.			
g) Autre. Précisez.			

Source : Sauv   *et al.*, 2020, p. 241.

Deux strat  gies pour s'assurer de la validit   et de la fiabilit   de l'instrument de mesure ont   t   employ  es. Tout d'abord, concernant la validation de la grille du design des cours, une m  thode de groupe de discussion a   t   employ  e (Leiva *et al.*, 2006). En effet, lors de l'analyse des deux premiers cours par les auxiliaires de recherche, un groupe de discussion a   t     labor   conjointement avec les chercheurs et les auxiliaires afin que ces derniers puissent indiquer les difficult  s rencontr  es lors de l'utilisation de la grille et y apporter des modifications si n  cessaires. Les chercheurs   taient l   pour   clairer les auxiliaires si les d  finitions des variables n'  taient pas claires et, le cas   ch  ant, reformuler certaines sections. Aucune nouvelle difficult   n'a   t   relev  e par les auxiliaires de recherche apr  s le deuxi  me cours. Ensuite, concernant la fiabilit   des r  sultats, une m  thode interjuges a   t   utilis  e (Leiva *et al.*, 2006). Une fois que l'auxiliaire de recherche avait fini de remplir la grille d'analyse du design de cours, celle-ci

était soumise à une seconde évaluation par le deuxième auxiliaire qui vérifiait aléatoirement certaines données inscrites en consultant le site du cours.

TABLEAU 3.3

Types de cours

	En présence physique (face à face) (%)	En ligne ou à distance		À distance (%)				Sous-total
		Synchrone (%)	Asynchrone (%)	Papier	DVD-CD	TV	Autre. Précisez.	
Déroulement des modules, séances, unités, etc.								
Présentation du cours ou son démarrage								
Activités d'évaluation sous surveillance								
Autres événements :								
							Total (100%)	0%

Source : Sauv  et al., 2020, p. 243.

Le remplissage de la grille de design des cours a  t  r alis  pour 19 cours d'une universit  canadienne sp cialis e dans la formation en ligne, entre l'automne 2018 et l' t  2019. L'auxiliaire de recherche charg  de l'analyse des cours a navigu  sur le site Web du cours et a inscrit de fa on syst matique les variables recens es dans la grille.

La navigation en ligne s'est effectu e en trois  tapes :

1. Visite de la page d'accueil
 - a) Prendre connaissance du plan de cours et de ses modalit s.
 - b) Prendre connaissance des modalit s d'encadrement et des personnes-ressources.

- c) Prendre connaissance des outils offerts lors de la formation.
 - d) Prendre connaissance des modalités des examens.
- 2. Visite des séances/modules**
- a) Prendre connaissance des différentes composantes du cours : modalités présence/distance, finalité de l'activité, formules pédagogiques, modalités d'apprentissage, formule d'évaluation et sa valeur, type de correction, matériel didactique, outils de communication, encadrement.
- 3. Visite des autres pages**
- a) Prendre connaissance de la présence d'un forum ou d'autres outils de communication.
 - b) Prendre connaissance du matériel complémentaire.
 - c) Prendre connaissance des exercices formatifs.

Les informations ont été systématiquement colligées par l'auxiliaire dans la grille du design des cours.

3.2.3 Les groupes de discussion

À la suite de la première période de sollicitation des étudiants les invitant à répondre aux questionnaires en ligne, l'équipe de recherche s'est rencontrée afin d'établir un canevas d'entretien. Un protocole d'entretien semi-structuré a été mis sur pied selon les résultats obtenus aux questionnaires afin d'approfondir la réflexion sur les données recueillies.

Six rencontres avec les étudiants ont été tenues entre mars et mai 2020 sur la plateforme numérique Zoom. Les groupes de discussion devaient comprendre en moyenne entre trois et cinq participants à la fois afin de permettre les échanges entre les membres.

Une méthode d'échantillonnage volontaire a été utilisée pour former les groupes de discussion des étudiants. Un courriel de sollicitation a été envoyé aux 791 participants ayant rempli le premier questionnaire afin de les inviter à prendre part aux discussions. De ce nombre, 22 ont répondu positivement à notre invitation (2,8%), mais seulement 8 (uniquement des étudiantes) ont effectivement participé aux groupes de discussion.

Conformément aux recommandations de Yin (2009) et Silverman (2013), deux stratégies ont été utilisées pour éviter des préoccupations telles que les caractéristiques de la demande et d'autres sources d'erreur systématique qui pourraient avoir biaisé les résultats des groupes de discussion. Premièrement, l'objectif de l'étude a été délibérément maintenu vague. Deuxièmement, des questions de réflexion ont été élaborées et testées *a priori* dans les groupes de discussion.

Les rencontres des groupes de discussion ont sensiblement toutes eu la même structure. L'objectif général et les consignes ont été présentés aux participantes. S'en sont suivi des questions sur les caractéristiques d'apprentissage, l'organisation pédagogique, l'accompagnement, la persévérance et l'abandon, le soutien à la persévérance pendant les études, ainsi que l'organisation personnelle et le soutien social. Initialement, les participantes devaient répondre à des questions structurées. Elles ont été invitées à donner des réponses de base, puis à développer leurs propos en fonction des commentaires et réactions des autres participantes. Au cours d'une session, les participantes qui ne répondaient pas directement aux questions ont été invitées à compléter, à modifier ou à clarifier les points soulevés par les principaux intervenants. Le dialogue tenu par les participantes lors des groupes de discussion a été enregistré sur bande audio, transcrit et ensuite présenté à deux assistants de recherche pour interprétation. Les assistants ont reçu des instructions générales. Ils ont été invités à créer des catégories (thèmes) et – si on leur demandait – à expliquer le processus qu'ils ont utilisé pour créer ces thèmes. Le chercheur a ensuite examiné les thèmes créés par les assistants et les a formellement rédigés.

4 LES ANALYSES STATISTIQUES

Les analyses statistiques choisies dans le cadre de ce projet sont de trois ordres. Tout d'abord, des analyses descriptives ont été réalisées sur l'ensemble des résultats obtenus. Les tris à plat donnent une vision globale des résultats et sont en mesure d'indiquer des tendances qu'il convient de chercher à comprendre. Le deuxième type d'analyse concerne les statistiques croisées. En particulier, il inclut des méthodes telles que les tests t de Student, ANOVA et de corrélation de Spearman (ou Kendall, le cas échéant). La plupart des variables de cette étude ont été croisées avec l'abandon des cours par les étudiants, la note finale obtenue dans le cours

ainsi que la persévérance des étudiants dans leurs études (encore inscrits ou diplômés dans l'institution après deux sessions). Enfin, la troisième catégorie d'analyse tente de créer un portrait global d'un ensemble de facteurs ayant un effet sur l'abandon et la persévérance. Il s'agit d'une analyse de *cluster* en deux étapes (*two-step*).

4.1 Les tris à plat et la description sociodémographique de l'échantillon

Les tris à plat permettent d'obtenir une vision d'ensemble des résultats collectés au cours de cette recherche. À cet égard, ils ont permis d'identifier notre échantillon comme étant composé majoritairement de citoyens canadiens (95 %) de sexe féminin (78 %) et ayant le français comme langue maternelle (91,5 %). Plus de la moitié de l'échantillon (67,8 %) vit avec un partenaire (marié ou conjoint de fait). Près d'un étudiant sur deux (51,5 %) vit avec un ou des enfants à sa charge. En ce qui a trait aux études, 56,3 % de l'échantillon indique préparer un certificat et 32 % un baccalauréat. De plus, la majorité des étudiants travaille plus de 30 heures par semaine (62 %) et considère sa situation financière comme acceptable ou bonne (92,4 %).

4.2 Les analyses croisées et la représentativité de l'échantillon

Les analyses de cette deuxième section servent à vérifier le degré d'association entre deux ou plusieurs variables. Les modèles statistiques varient selon les variables qui font l'objet de l'analyse allant d'un test de khi-deux à des tests t de Student et ANOVA, en passant par une analyse de corrélation. Les analyses croisées peuvent se dépeindre en deux catégories. Tout d'abord, elles incluent les croisements entre les données sociodémographiques des participants avec les réponses obtenues dans les différents questionnaires autoadministrés. Elles peuvent inclure, par exemple, un croisement entre le nombre d'heures travaillé de façon hebdomadaire par les participants mis en lien avec la satisfaction déclarée concernant leurs interactions avec leurs proches. Dans un deuxième temps, des mesures composites sont créées selon certains regroupements d'énoncés et croisées avec d'autres variables. C'est le cas, par exemple, des énoncés de la section des stratégies d'apprentissage qui sont regroupés selon la typologie proposée par Zimmerman (2002) avant d'être croisés avec les variables d'abandon et de persévérance.

Les analyses croisées permettent aussi de comparer les données sociodémographiques de notre échantillon avec les données institutionnelles. Ainsi, une analyse de la représentativité de l'échantillon démontrée sur 7 variables sociodémographiques et scolaires des participants à l'enquête inscrits dans 19 cours retenus a confirmé que notre échantillon est représentatif de la population, sauf en ce qui concerne 2 variables (le genre pour 2 cours et le régime d'études pour 4 cours). Des analyses subséquentes ont démontré que ces différences s'avèrent mineures pour expliquer la propension d'abandon du cours.

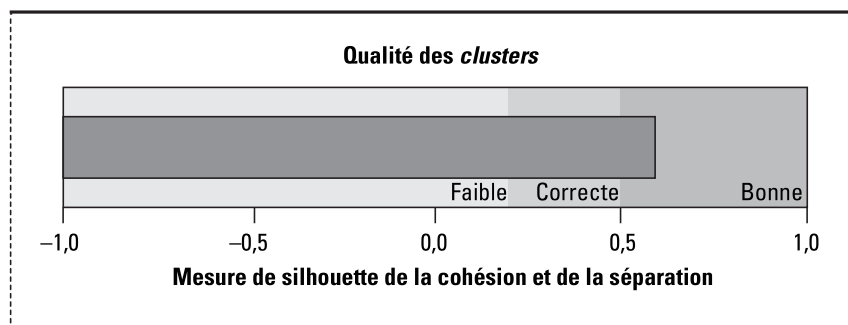
4.3 L'analyse de *cluster*

L'analyse de *cluster* est une analyse exploratoire qui tente d'identifier des structures dans les données (Driver et Kroeber, 1932; Zubin, 1938). Plus précisément, elle cherche à déterminer des groupes de cas homogènes lorsque le regroupement n'est pas connu au préalable. Parce qu'elle est exploratoire, elle ne fait pas de distinction entre les variables dépendantes et indépendantes (Tryon, 1939). Les différentes méthodes d'analyse de regroupements peuvent traiter des données binaires, nominales, ordinales et d'échelle (intervalle ou rapport). L'analyse a été initiée en anthropologie par Driver et Kroeber (1932) et introduite en psychologie par Joseph Zubin (1938) et Robert Tryon (1939). L'analyse de *cluster* est très peu utilisée en éducation (Battaglia *et al.*, 2016). À notre connaissance, aucune étude n'a utilisé cette dernière afin de regrouper les caractéristiques des cours ou de leur encadrement en ligne, et encore moins de les croiser avec l'abandon et la persévérance des étudiants au niveau universitaire.

L'analyse statistique des *clusters* n'est pas un algorithme spécifique, mais plutôt une tâche à résoudre. En ce sens, elle se rapproche plus de la résolution de problèmes. La mise en grappes (*clusters* ou groupes) peut donc être formulée comme un problème d'optimisation multiobjectifs (Driver et Kroeber, 1932). Elle peut être réalisée par divers algorithmes qui diffèrent sensiblement dans leur compréhension de ce qui constitue un *cluster* et de la manière de les trouver efficacement (Driver et Kroeber, 1932; Zubin, 1938). L'algorithme de mise en grappes et les paramètres appropriés (y compris des paramètres tels que la fonction de distance à utiliser, un seuil de densité ou le nombre de grappes attendues) dépendent de l'ensemble de données individuelles

et de l'utilisation prévue des résultats. L'analyse des grappes en tant que telle n'est pas une tâche automatique, mais un processus itératif de découverte de connaissances ou d'optimisation multiobjective interactive qui implique des essais et des échecs. Ainsi, il est souvent nécessaire de modifier le prétraitement des données et les paramètres du modèle jusqu'à ce que le résultat atteigne les propriétés souhaitées.

FIGURE 3.1
Mesure de silhouette de la cohésion et de la séparation



Concrètement, l'analyse de *cluster* consiste à regrouper un ensemble de variables de manière à ce que les objets d'une même grappe soient plus semblables les uns aux autres qu'à ceux d'autres groupes. Pour aider le chercheur, les logiciels statistiques offrant l'option d'analyse des *clusters* indiquent la cohésion et le degré de séparation des grappes par une mesure de silhouette. La figure 3.1 est un exemple de mesure de silhouette fourni par une analyse de *cluster* faite par le logiciel SPSS 21. Cette mesure permet au chercheur d'estimer à quel point les regroupements sont cohérents entre eux et à quel point les grappes sont différentes les unes des autres. Une mesure de silhouette supérieure à 0.5 est considérée comme acceptable (Lee *et al.*, 2012).

4.4 L'analyse de *cluster* en deux étapes (*two-step*)

La procédure d'analyse de *cluster* en deux étapes (*two-step*) est un outil exploratoire conçu pour révéler des regroupements possibles au sein d'un ensemble de données *a priori* disparates (Rundle-Thiele *et al.*, 2015). L'algorithme utilisé par cette procédure présente plusieurs caractéristiques souhaitables qui le différencient des techniques traditionnelles de regroupement,

notamment la possibilité de créer des *clusters* basés sur des variables catégorielles et continues et la possibilité d'analyser efficacement des fichiers de données volumineux (Kent *et al.*, 2014; Rundle-Thiele *et al.*, 2015).

Afin de traiter les variables catégorielles et continues, la procédure d'analyse de grappes en deux étapes utilise une mesure de distance de vraisemblance qui suppose que les variables du modèle de grappes sont indépendantes. En outre, chaque variable continue est censée avoir une distribution normale (Gaussienne) (Rundle-Thiele *et al.*, 2015) et chaque variable catégorielle, une distribution multinomiale (Li et Sun, 2018). Bien que la littérature scientifique indique que la procédure est assez robuste aux violations des hypothèses d'indépendance et de distribution, le chercheur utilisant cette analyse statistique doit néanmoins tenter de savoir dans quelle mesure ces hypothèses sont respectées (Driver et Kroeber, 1932).

Comme son nom l'indique, l'analyse de *cluster two-step* contient deux étapes. Tout d'abord, la procédure commence par la construction d'un arbre des caractéristiques des grappes (CG). Ensuite, l'analyse amène à placer le premier cas à la base de l'arbre. Chaque cas successif est ensuite ajouté à un nœud existant en fonction de sa similarité avec les nœuds déjà présents dans l'algorithme en utilisant la mesure de distance comme critère de similarité. Si le degré de séparation est trop grand, un nouveau nœud sera créé. Un nœud qui rassemble plusieurs cas contient un résumé des informations des variables concernant ces cas. Ainsi, à la fin de la première étape, l'arbre CG fournit un résumé des cas regroupés sous un même nœud. La seconde étape consiste ensuite à regrouper les nœuds de l'arbre CG à l'aide d'un algorithme de regroupement. Le regroupement aggloméré peut être utilisé pour produire une série de solutions, c'est-à-dire les regroupements finaux. Pour déterminer quel est le « meilleur » nombre de grappes, chacune de ces solutions est comparée en utilisant le critère bayésien de Schwarz (Schwarz, 1978) ou le critère d'information d'Akaike (de Leeuw, 1992) comme critère de regroupement. La solution ayant le critère le plus élevé devient alors la solution finale qui sera présentée au chercheur par le logiciel d'analyse statistique.

CONCLUSION

Ce chapitre visait avant tout à présenter la méthodologie appliquée dans le cadre de l'étude *Regard des apprenant.es universitaires sur les modes d'organisation et d'encadrement pédagogique en formation à distance et en ligne*, financée par le Fonds de recherche du Québec – Société et culture (FRQSC). Ainsi, les chapitres subséquents de cet ouvrage y feront référence, sans en répéter les éléments afin de présenter plus finement les principaux résultats de la recherche. Ensuite, il avait pour objectif de servir de référence au lecteur qui désire approfondir ses connaissances statistiques, notamment en ce qui a trait à l'utilisation de l'analyse de *cluster two-step*, méthode statistique encore peu utilisée dans le domaine de l'éducation et offrant des pistes intéressantes pour d'autres études sur la persévérance ou l'abandon notamment.

Alors que l'une des conséquences de la volonté de saisir les problèmes et obstacles rencontrés par les apprenants de manière holistique, telle que proposée par Aydin et ses collaborateurs (2019), est le nombre important de variables à analyser, ce chapitre a mis en lumière la nécessité d'adopter une méthodologie exhaustive et une certaine flexibilité dans les types d'analyses statistiques employées. Cette flexibilité peut cependant s'avérer une limite dans certains cas. De fait, la collection d'un large éventail de données autant qualitatives que quantitatives peut rendre quelques résultats difficilement interprétables. Le problème inverse est également vrai ; le manque de données – que nous avons par exemple rencontré lorsque nous avons soumis des questionnaires aux intervenants ou que nous les avons sollicités pour participer à des entretiens – limite la possibilité d'analyse et de généralisation des résultats. Enfin, bien qu'il soit intéressant de mesurer la persévérance sur le long terme, les subventions de recherche sont généralement limitées dans le temps.

La recherche sur la persévérance et l'abandon dans les études supérieures implique ainsi, comme bon nombre de recherches en sciences humaines, de relever différents défis. Néanmoins, les résultats qui seront présentés plus en détail lors des prochains chapitres de cet ouvrage permettent de faire ressortir la pertinence de la méthodologie employée pour aller plus loin dans la compréhension des phénomènes de persévérance et d'abandon. Nous allons ainsi présenter les principaux résultats obtenus concernant

l'influence des stratégies d'apprentissage (chapitre 4), des modes d'organisation pédagogique (chapitre 5) et de l'accompagnement (chapitre 6) sur la persévérance des étudiants.

RÉFÉRENCES

- Andrà, C., G. Magnano et F. Morselli (2011). «Dropout undergraduate students in mathematics: An exploratory study», *Current State of Research on Mathematical Beliefs XVII*, p. 13-22.
- Arifin, M.H. (2018). «The role of student support services in enhancing student persistence in the Open University context: Lesson from Indonesia Open University», *Turkish Online Journal of Distance Education*, vol. 19, n° 3, p. 156-168.
- Aumond, M. et F. Beaulieu (1994). «En Ontario, on raccroche», *Éducation et francophonie*, vol. 22, n° 2, p. 32-39.
- Aydin, S., A. Öztürk, G.T. Büyükköse, F. Er et H. Sönmez (2019). «An investigation of drop out in open and distance education», *Educational, Sciences: Theory & Practice*, vol. 19, n° 2, p. 40-57.
- Bakker, E.J., K.J. Verhaegh, J.H. Kox, A.J. van der Beek, C.R. Boot, P.D. Roelofs et A.L. Francke (2019). «Late dropout from nursing education: An interview study of nursing students' experiences and reasons», *Nurse Education in practice*, vol. 39, p. 17-25, <<https://doi.org/10.1016/j.nepr.2019.07.005>>, consulté le 3 mars 2021.
- Barr-Telford, L., F. Cartwright, S. Prasil et K. Shimmons (2003). «Access, persistence and financing: First results from the Postsecondary Education Participation Survey (PEPS)», *Culture, Tourism and the Centre for Education Statistics Research Papers*, vol. 595, n° 81, p. 1-33.
- Battaglia, O.R., B. Paola et C. Fazio (2016). «A new approach to investigate students' behavior by using cluster analysis as an unsupervised methodology in the field of education», *Applied Mathematics*, vol. 7, n° 15, p. 1649-1673, <<http://dx.doi.org/10.4236/am.2016.715142>>, consulté le 3 mars 2021.
- Bernard, P.-Y. (2011). *Le décrochage scolaire*, Paris, Presses universitaires de France.
- Blaya, C. (2010). «Décrochage scolaire: parents coupables, parents décrocheurs?», *Informations sociales*, vol. 5, n° 161, p. 46-54.
- Bowers, A.J. (2010). «Analyzing the longitudinal K-12 grading histories of entire cohorts of students: Grades, data driven decision making, dropping out and hierarchical cluster analysis», *Practical Assessment, Research, and Evaluation*, vol. 15, n° 1, p. 1-7, <<https://doi.org/10.7275/r4zq-9c31>>, consulté le 3 mars 2021.
- Burchell, B. et C. Marsh (1992). «The effect of questionnaire length on survey response», *Quality and Quantity*, vol. 26, n° 3, p. 233-244, <<https://doi.org/10.1007/BF00172427>>, consulté le 3 mars 2021.
- Chung, J.Y. et S. Lee (2019). «Dropout early warning systems for high school students using machine learning», *Children and Youth Services Review*, vol. 96, p. 346-353, <<https://doi.org/10.1016/j.chilyouth.2018.11.030>>, consulté le 3 mars 2021.

- Conseil régional de prévention de l'abandon scolaire (CRÉPAS) (2001). « Définition de l'abandon scolaire », <<https://crepas.qc.ca/perseverance-scolaire/portrait-regional/#:~:text=Voici%20quelques%20d%C3%A9finitions%20de%20concepts,%27une%20institution%20d%27enseignement>>, consulté le 3 mars 2021.
- Conseil supérieur de l'éducation (CSE) (2000). *La reconnaissance des acquis, une responsabilité politique et sociale*, Québec, CSE.
- Cooper, D. et P. Schindler (2011). *Business Research Methods*, (11^e éd.), New York, McGraw-Hill International.
- Crain-Dorough, M.L. (2003). « A study of dropout characteristics and school-level effects on dropout prevention », thèse de doctorat, Baton Rouge, Louisiana State University.
- Creswell, J.W. et V.L. Plano Clark (2017). *Designing and Conducting Mixed Methods Research*, Thousand Oaks, Sage.
- Dalipi, F., A.S. Imran et Z. Kastrati (2018). « MOOC dropout prediction using machine learning techniques: Review and research challenges », *IEEE Global Engineering Education Conference (EDUCON)*, p. 1007-1014, <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8363340&tag=1>>, consulté le 3 mars 2021.
- Del Bonifro, F., M. Gabbrielli, G. Lisanti et S.P. Zingaro (2020). « Student dropout prediction », *International Conference on Artificial Intelligence in Education*, p. 129-140, <https://doi.org/10.1007/978-3-030-52237-7_11>, consulté le 3 mars 2021.
- de Leeuw, J. (1992). « Introduction to Akaike (1973): Information theory and an extension of the maximum likelihood principle », dans S. Kotz et N.L. Johnson (dir.), *Breakthroughs in Statistics: Foundations and Basic Theory*, Londres, Springer, p. 599-609.
- Doherty, W. (2006). « An analysis of multiple factors affecting retention in Web-based community college courses », *Internet and Higher Education*, vol. 9, n° 4, p. 245-255, <<https://doi.org/10.1016/j.iheduc.2006.08.004>>, consulté le 3 mars 2021.
- Driver, H.E. et A.L. Kroeber (1932). « Quantitative expression of cultural relationships », *University of California Publication in American Archeology and Ethnology*, vol. 31, n° 4, p. 211-256.
- Figueroa-Cañas, J. et T. Sancho-Vinuesa (2019). « Predicting early dropout student is a matter of checking completed quizzes: The case of an online statistics module », *LASI-SPAIN Learning Analytics in Higher Education*, p. 100-111.
- Fortin, L., D. Marcotte, P. Potvin, É. Royer et J. Joly (2006). « Typology of students at risk of dropping out of school: Description by personal, family and school factors », *European Journal of Psychology of Education*, vol. 21, n° 4, p. 363-383, <<https://doi.org/10.1007/BF03173508>>, consulté le 3 mars 2021.
- Franklin, S., C. Walker et Statistique Canada (2010). « Survey methods and practices », Ottawa, Ministère de l'Industrie, <<https://www150.statcan.gc.ca/n1/pub/12-587-x/12-587-x2003001-eng.pdf>>, consulté le 3 mars 2021.
- Frey, J.H. et A. Fontana (1991). « The group interview in social research », *Social Science Journal*, vol. 28, n° 2, p. 175-187, <[https://doi.org/10.1016/0362-3319\(91\)90003-M](https://doi.org/10.1016/0362-3319(91)90003-M)>, consulté le 3 mars 2021.
- Garrison, D.R. (1987). « Researching dropout in distance education », *Distance Education*, vol. 8, n° 1, p. 95-101, <<https://doi.org/10.1080/0158791870080107>>, consulté le 3 mars 2021.

- Gilar-Corbi, R., T. Pozo-Rico, J.L. Castejón, T. Sánchez, I. Sandoval et J. Vidal (2020). «Academic achievement and failure in university studies: Motivational and emotional factors», *Sustainability*, vol. 12, n° 23, p. 97-98, <<https://doi.org/10.3390/su12239798>>, consulté le 3 mars 2021.
- Grayson, J.P. et K. Grayson (2003). *Les recherches sur le maintien et la diminution des effectifs étudiants*, Québec, Fondation canadienne des bourses d'études du millénaire.
- Gupta, S.K., J. Antony, F. Lacher et J. Douglas (2020). «Lean Six Sigma for reducing student dropouts in higher education – An exploratory study», *Total Quality Management & Business Excellence*, vol. 31, n° 1-2, p. 178-193.
- Holloway, I. (1997). *Basic Concepts for Qualitative Research*, Hoboken, Wiley-Blackwell.
- Honorez, M., J. Therer, R. Cahay, B. Monfort et F. Remy (2000). «Les styles d'apprentissage: mode d'emploi», présenté au 1^{er} Congresso dos pesquisadores em educação, Bruxelles.
- Hsu, C.W. et C.C. Yeh (2020). «Mining the student dropout in higher education», *Journal of Testing and Evaluation*, vol. 48, n° 6, p. 4563-4575, <<https://doi.org/10.1520/JTE20180021>>, consulté le 3 mars 2021.
- Iam-On, N. et T. Boongoen (2017). «Generating descriptive model for student dropout: A review of clustering approach», *Human-centric Computing and Information Sciences*, vol. 7, n° 1, p. 1-24, <<https://doi.org/10.1186/s13673-016-0083-0>>, consulté le 3 mars 2021.
- Kemper, L., G. Vorhoff et B.U. Wigger (2020). «Predicting student dropout: A machine learning approach», *European Journal of Higher Education*, vol. 10, n° 1, p. 28-47, <<https://doi.org/10.1080/21568235.2020.1718520>>, consulté le 3 mars 2021.
- Kent, P., R.K. Jensen et A. Kongsted (2014). «A comparison of three clustering methods for finding subgroups in MRI, SMS or clinical data: SPSS TwoStep Cluster analysis, Latent Gold and SNOB», *BMC Medical Research Methodology*, vol. 14, n° 1, p. 1-14, <<https://doi.org/10.1186/1471-2288-14-113>>, consulté le 3 mars 2021.
- King, C. (2005). *Factors Related to the Persistence of First-Year College Students at Four-year Colleges and Universities: A Paradigm Shift*, Wheeling, Wheeling Jesuit University.
- Kolb, D.A. (1984). *Experiential Learning: Experience as the Source of Learning and Development* (vol. 1), Englewood Cliffs, Prentice-Hall.
- Kotsiantis, S.B., C.J. Pierrakeas et P.E. Pintelas (2003). «Preventing student dropout in distance learning using machine learning techniques», *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, p. 267-274, <https://doi.org/10.1007/978-3-540-45226-3_37>, consulté le 3 mars 2021.
- Krosnick, J.A. (2018). «Questionnaire design», dans D.L. Vannette et J.A. Krosnick (dir.), *The Palgrave Handbook of Survey Research*, Cham, Palgrave Macmillan, p. 43-50, <<https://doi.org/10.1007/978-3-319-54395-6>>, consulté le 3 mars 2021.
- Krstić, K., I.S. Ilć, M. Videnović et O. Tošković (2019). «Should they stay or should they go? Dropout in higher education», *Empirical Studies in Psychology*, vol. 25, p. 68-81.
- Lee, K.M., K.M. Lee et C.H. Lee (2012). «Statistical cluster validity indexes to consider cohesion and separation», *International Conference on Fuzzy Theory and Its Applications (IFUZZY2012)*, Taichung, p. 228-232, <<https://doi.org/10.1109/IFUZZY.2012.6409706>>, consulté le 3 mars 2021.

- Leiva, F.M., F.J.M. Ríos et T.L. Martínez (2006). « Assessment of interjudge reliability in the open-ended questions coding process », *Quality and Quantity*, vol. 40, n° 4, p. 519-537, <<https://doi.org/10.1007/s11135-005-1093-6>>, consulté le 3 mars 2021.
- Li, G. et L. Sun (2018). « Characterizing heterogeneity in drivers' merging maneuvers using two-step cluster analysis », *Journal of Advanced Transportation*, vol. 2018, p. 1-16, <<https://doi.org/10.1155/2018/5604375>>, consulté le 3 mars 2021.
- Li, Y., H. Chen et Y. Han (2010). « Why do they drop out of school? Interview survey of learners in open and distance education », *Open Education Research*, vol. 2, p. 1-11.
- Li, Y., J. Niu et X. Ding (2012). « A follow-up study of the dropouts from the english program of open and distance learning (part 2) », *Open Education Research*, vol. 18, n° 6, p. 80-86.
- López-Goñi, J.J., J. Fernández-Montalvo et A. Arteaga (2012). « Addiction treatment dropout: Exploring patients' characteristics », *American Journal on Addictions*, vol. 21, n° 1, p. 78-85, <<https://doi.org/10.1111/j.1521-0391.2011.00188.x>>, consulté le 3 mars 2021.
- Lottering, R., R. Hans et M. Lall (2020). « A model for the identification of students at risk of dropout at a university of technology », International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD), KwaZulu Natal, p. 1-8, <<https://doi.org/10.1109/icABCD49160.2020.9183874>>, consulté le 3 mars 2021.
- McDermott, E.R., S. Anderson et J.F. Zaff (2018). « Dropout typologies: Relating profiles of risk and support to later educational re-engagement », *Applied Developmental Science*, vol. 22, n° 3, p. 217-232, <<https://doi.org/10.1080/10888691.2016.1270764>>, consulté le 3 mars 2021.
- Millet, M. (2012). « L'échec des étudiants de premiers cycles dans l'enseignement supérieur en France. Retours sur une notion ambiguë et descriptions empiriques », dans M. Romainville et C. Michaut (dir.), *Réussite, échec et abandon dans l'enseignement supérieur*, Bruxelles, De Boeck, p. 69-88.
- Muthukumar, V. et N. Bhalaji (2020). « MOOCVERSITY-deep learning based dropout prediction in MOOCs over weeks », *Journal of Soft Computing Paradigm*, vol. 2, n° 3, p. 140-152, <<https://doi.org/10.36548/jscp.2020.3.001>>, consulté le 3 mars 2021.
- Nichols, A.J. et Y. Levy (2009). « Empirical assessment of college student-athletes' persistence in e-learning courses: A case study of a US National Association of Intercollegiate Athletics (NAIA) institution », *Internet and Higher Education*, vol. 12, n° 1, p. 14-25, <<https://doi.org/10.1016/j.iheduc.2008.10.003>>, consulté le 3 mars 2021.
- No, F. et Y. Hirakawa (2012). « Identifying causes of dropout through longitudinal quantitative analysis in rural Cambodian basic schools », *Journal of International Development and Cooperation*, vol. 19, n° 1, p. 25-39, <<http://doi.org/10.15027/33828>>, consulté le 3 mars 2021.
- Phillion, R., M. Bourassa, R. Leblanc, D. Plouffe et I. Arcand (2010). « Persistance et réussite aux études postsecondaires: étude d'un accompagnement personnalisé pour étudiants en situation d'échec », *Revue de recherche appliquée sur l'apprentissage*, vol. 3, n° 6, p. 1-27.
- Rahmat, A. (2017). « Clustering in education », *European Research Studies Journal*, vol. 20, n° 3A, p. 311-324.

- Ramsdal, G.H., S. Bergvik et R. Wynn (2018). «Long-term dropout from school and work and mental health in young adults in Norway: A qualitative interview-based study», *Cogent Psychology*, vol. 5, n° 1, p. 1-16, <<https://doi.org/10.1080/23311908.2018.1455365>>, consulté le 3 mars 2021.
- Rekkedal, T. (1985). *Introducing the Personal Tutor/Counsellor in the System of Distance Education*, Stabekk, Bærum.
- Roblyer, M.D., L. Davis, S.C. Mills, J. Marshall et L. Pape (2008). «Toward practical procedures for predicting and promoting success in virtual school students», *American Journal of Distance Education*, vol. 22, n° 2, p. 90-109, <<https://doi.org/10.1080/08923640802039040>>, consulté le 3 mars 2021.
- Romainville, M. et C. Michaut (2012). «Conclusion», dans M. Romainville et C. Michaut (dir.), *Réussite, échec et abandon dans l'enseignement supérieur*, Bruxelles, De Boeck, p. 251-262.
- Rubin, H.J. et I.S. Rubin (2011). *Qualitative Interviewing: The Art of Hearing Data*, Thousand Oaks, Sage.
- Rundle-Thiele, S., K. Kubacki, A. Tkaczynski et J. Parkinson (2015). «Using two-step cluster analysis to identify homogeneous physical activity groups», *Marketing Intelligence & Planning*, vol. 33, n° 4, p. 522-537, <<https://doi.org/10.1108/MIP-03-2014-0050>>, consulté le 3 mars 2021.
- Sauvé, L., C. Papi, S. Gérin-Lajoie et G. Desjardins (2020). *Regard des apprenant.es universitaires sur les modes d'organisation et d'encadrement pédagogique en formation à distance et en ligne*, rapport de recherche scientifique, Québec, Université TÉLUQ et Fonds de recherche du Québec – Société et culture.
- Sauvé, L. et R. Viau (2003). *L'abandon et la persévérance à l'université: l'importance de la relation enseignement-apprentissage*, Québec, Université TÉLUQ.
- Schertzer, C.B. et S.M. Schertzer (2004). «Student satisfaction and retention: A conceptual model», *Journal of Marketing for Higher Education*, vol. 14, n° 1, p. 79-91, <https://doi.org/10.1300/J050v14n01_05>, consulté le 3 mars 2021.
- Schwarz, G. (1978). «Estimating the dimension of a model», *Annals of Statistics*, vol. 6, n° 2, p. 461-464.
- Silverman, D. (2013). *A Very Short, Fairly Interesting and Reasonably Cheap Book about Qualitative Research*, (2^e éd.), Thousand Oaks, Sage.
- Simpson, O. (2006). «Predicting student success in open and distance learning», *Open Learning*, vol. 21, n° 2, p. 125-138, <<https://doi.org/10.1080/02680510600713110>>, consulté le 3 mars 2021.
- Spady, W. (1970). «Dropouts from higher education: An interdisciplinary review and synthesis», *Interchange*, vol. 1, p. 64-85, <<https://doi.org/10.1007/BF02214313>>, consulté le 3 mars 2021.
- Stiller, K.D. et R. Bachmaier (2017). «Dropout in an online training for trainee teachers», *European Journal of Open, Distance and E-learning*, vol. 20, n° 1, p. 80-95, <<https://files.eric.ed.gov/fulltext/EJ1187842.pdf>>, consulté le 3 mars 2021.
- Tan, M. et P. Shao (2015). «Prediction of student dropout in e-learning program through the use of machine learning method», *ijET*, vol. 10, n° 1, p. 11-17, <<https://doi.org/10.3991/ijet.v10i1.4189>>, consulté le 3 mars 2021.
- Teijlingen, E.R.V. et V. Hundley (2001). «The importance of pilot studies», *Social Research Update*, vol. 35, p. 1-4, <<https://doi.org/10.7748/ns2002.06.16.40.33.c3214>>, consulté le 3 mars 2021.

- Thanassoulis, E., P.K. Dey, K. Petridis, I. Goniadis et A.C. Georgiou (2017). «Evaluating higher education teaching performance using combined analytic hierarchy process and data envelopment analysis», *Journal of the Operational Research Society*, vol. 68, n° 4, p. 431-445, <<https://doi.org/10.1057/s41274-016-0165-4>>, consulté le 3 mars 2021.
- Tinto, V. (1987). *Leaving College: Rethinking the Causes and Cures of Student Attrition*, Chicago, University of Chicago Press.
- Tinto, V. (2005). «Epilogue: Moving from theory to action», dans A. Seidman (dir.), *College Student Retention: Formula for Student Success*, Westport, ACE/Praeger, p. 1-57.
- Tryon, R.C. (1939). *Cluster Analysis: Correlation Profile and Orthometric (Factor) Analysis for the Isolation of Unities in Mind and Personality*, Ann Arbor, Edwards Brothers.
- Xuan, W., C. Mellis, K. Williams et J.K. Peat (2020). *Health Science Research: A Handbook of Quantitative Methods*, Milton Park, Routledge.
- Yasmin, D. (2013). «Application of the classification tree model in predicting learner dropout behaviour in open and distance learning», *Distance Education*, vol. 34, n° 2, p. 218-231, <<https://doi.org/10.1080/01587919.2013.793642>>, consulté le 3 mars 2021.
- Yin, R.K. (2009). *Case Study Research: Design and Methods*, Thousand Oaks, Sage.
- Yorke, M. et L. Thomas (2003). «Improving the retention of students from lower socio-economic groups», *Journal of Higher Education Policy and Management*, vol. 25, n° 1, p. 63-74, <<https://doi.org/10.1080/13600800305737>>, consulté le 3 mars 2021.
- Zhang, G., T.J. Anderson, M.W. Ohland et B.R. Thorndyke (2004). «Identifying factors influencing engineering student graduation: A longitudinal and cross-institutional study», *Journal of Engineering Education*, vol. 93, n° 4, p. 313-320, <<https://doi.org/10.1002/j.2168-9830.2004.tb00820.x>>, consulté le 3 mars 2021.
- Zimmerman, B.J. (2002). «Becoming a self-regulated learner: An overview», *Theory into Practice*, vol. 41, n° 2, p. 64-70, <https://doi.org/10.1207/s15430421tip4102_2>, consulté le 3 mars 2021.
- Zubin, J. (1938). «A technique for measuring like-mindedness», *Journal of Abnormal and Social Psychology*, vol. 33, n° 4, p. 508-516, <<https://doi.org/10.1037/h0055441>>, consulté le 3 mars 2021.

ANNEXE 3.1
Les outils technologiques répertoriés par la grille des cours

Outils technologiques	Offert par			Modalité			Commentaire
	Cours	Faculté/ Département/ Université	Autres	Emplacement (hors cours)	En ligne	À distance	
1.1. Adaptation à des handicaps physiques (auditif, visuel, etc.)							
1.2. Bureautique (traitement de texte, tableur, etc.)							
1.3. Prise de notes (bloc-notes, cahier de notes, etc.)							
1.4. Gestion des études (calendrier, liste de tâches, feuille de route)							
1.5. Gestion bibliographique (EndNote, Zotero, etc.)							
1.6. Captation de l'information textuelle, audio ou vidéo (capture d'écran en image ou en vidéo, etc.)							
1.7. Espace de dépôt des travaux des étudiant(e)s							
1.8. Portfolio (dossier des productions des étudiants ou outil de réflexion) (Mahara, autres)							
1.9. Réseautage social (Facebook, Twitter, Instagram, etc.)							

ANNEXE 3.1
Les outils technologiques répertoriés par la grille des cours (suite)

Outils technologiques	Offert par				Modalité			Commentaire
	Cours	Faculté/ Département/ Université	Autres	Emplacement (hors cours)	En ligne	À distance	En présence physique	
1.10. Partage de documents en ligne par le principe de l'infonuagique (cloud) (p. ex. Dropbox, iCloud, Google Drive, WeTransfer)								
1.11. Écriture collaborative (wiki, blogue, etc.)								
1.12. Analyse de contenus et planification d'activités (carte conceptuelle, schéma intégrateur, etc.)								
1.13. Gestion de l'apprentissage (métacognition : évaluation, analyse et planification ; p. ex. un journal de bord)								
1.14. Aide à la réalisation des travaux collaboratifs (outils pour choisir ses équipiers, former des équipes de travail ou coordonner le travail collaboratif)								