



## Une analyse de l'utilisation des vidéos pédagogiques des MOOC par les non-certifiés

► **Matthieu CISEL** (STEF, ENS Paris-Saclay)

---

---

■ **RÉSUMÉ** • Nous montrons, sur la base des traces de six MOOC, que les non-certifiés effectuent l'essentiel des actions observables réalisées sur les vidéos du cours. Nous cherchons à quantifier et interpréter divers comportements mis au jour sur la base de ces traces. Certains non-certifiés visionnent l'ensemble des vidéos du cours sur la plateforme ; ils sont minoritaires par rapport aux certifiés. Plus fréquemment, ils en téléchargent la totalité, sans les visionner depuis la plateforme. Selon une approche de méthode mixte, nous cherchons à mettre en lumière les déterminants de tels comportements. Le fait de se concentrer sur les seules vidéos pédagogiques correspond parfois à une décision contrainte, et non nécessairement à une stratégie d'apprentissage. Le téléchargement des ressources relève parfois de logiques d'archivage. Il n'est pas nécessairement suivi d'utilisations effectives des ressources téléchargées.

■ **MOTS-CLÉS** • MOOC, traces d'activité, comportement, méthodes mixtes

■ **ABSTRACT** • *Based on the learning analytics of six MOOCs, we show that non-completers are responsible for most of the actions realized on course videos. We aim at quantifying and interpreting some of the behaviors revealed by learning analytics. Some non-completers view all of the course videos, but these users are usually less common than completers. More frequently, they download all of them without viewing them on the platform. Following a mixed method approach, we carried out semi-structured interviews in order to interpret such behaviors. Focusing on the course videos is sometimes a forced decision, rather than a deliberate learning strategy. Non-completers who download all videos may archive them without further uses.*

■ **KEYWORDS** • *MOOC, learning analytics, behavior, mixed methods*

Matthieu CISEL

Une analyse de l'utilisation des vidéos pédagogiques des MOOC par les non-certifiés

*Sticef*, vol. 24, numéro 2, 2017, DOI: 10.23709/sticef.24.2.7

## **1. Introduction : les faibles taux de certification en question**

De nombreux travaux publiés dans la littérature grise comme dans la littérature scientifique montrent que le pourcentage des inscrits qui obtiennent le certificat ou l'attestation d'un MOOC (Daniel, 2012), autrement appelé taux de certification, est généralement inférieur à 10% (Ho *et al.*, 2014), (Ho *et al.*, 2015). Ce constat a engendré un vif débat au sein de la communauté scientifique quant à la signification de cette forte attrition (Cisel, 2014), l'attrition étant ici définie comme le pourcentage des inscrits qui n'obtiennent pas de certificat, c'est-à-dire, le pourcentage des non-certifiés. Un certain nombre de travaux ont visé à mettre au jour les caractéristiques des dispositifs (Jordan, 2014), (Jordan, 2015) ou des participants - variables sociodémographiques, motivations pour s'inscrire, etc. - dont le lien avec l'obtention du certificat était statistiquement significatif (Champaign *et al.*, 2014), (Colvin *et al.*, 2014), (Kizilcec et Schneider, 2015), (Barak *et al.*, 2016). Ces approches se fondent implicitement sur une vision dichotomique des utilisations qui sont faites des MOOC, entre des certifiés qui réaliseraient l'ensemble des activités prescrites, et des non-certifiés dont l'engagement dans la formation serait minimal.

Un certain nombre de travaux ont nuancé cette vision de l'engagement dans les MOOC en mobilisant diverses analyses de traces d'activité issues des plateformes hébergeant les cours. Par exemple, (Kizilcec *et al.*, 2013) mettent au jour une classe de participants n'obtenant pas le certificat mais visionnant la plus grande partie des vidéos pédagogiques : ils les appellent les auditeurs libres. Ces analyses visent à dépasser une vision dichotomique des utilisations qui sont faites des MOOC, entre certifiés et non-certifiés ; à cet égard, nous nous inscrivons dans la continuité de tels travaux. Notre objectif est ici de mieux comprendre, via des méthodes quantitatives, quelles sont les utilisations développées par les non-certifiés en termes de vidéos pédagogiques, et d'en proposer des interprétations via des méthodes qualitatives.

La question des utilisations des dispositifs faites par les non-certifiés représente un enjeu de taille dans les débats entourant les MOOC. Certains acteurs ont minimisé la portée des faibles taux de certification en arguant du fait que les non-certifiés développeraient de nombreux usages, importants sur le plan quantitatif, que l'on négligerait à tort en utilisant un indicateur de performance comme le taux de certification. L'une des

personnes les plus représentatives de cette tendance est la fondatrice de la plateforme américaine *Coursera* (Koller *et al.*, 2013). Celle-ci avançait en plein cœur des débats que nombre de non-certifiés, loin de se désengager de la formation, se concentraient sur les vidéos pédagogiques du cours. Pour reprendre les termes utilisés par les auteurs, ces « *passive participants* » éprouveraient une certaine « indifférence » pour les certificats. Pour appuyer leurs propos, ils affirment qu'au sein de la plateforme, les non-certifiés qui visionnent la quasi-totalité des vidéos du dispositif sont deux fois plus nombreux que les certifiés et doivent dès lors être pris en compte dans toute réflexion sur les formes d'attrition.

Dans le cadre de cet article, nous contesterons au moins en partie les interprétations faites par ces auteurs. En premier lieu, nous suggérerons, avec l'analyse de traces de six MOOC de la plateforme Coursera, que leur position repose vraisemblablement sur l'absence de distinction entre visionnage d'une vidéo depuis la plateforme et téléchargement de ladite vidéo. Or rien ne permet d'affirmer qu'une vidéo téléchargée est visionnée par la suite. Nous utiliserons le terme consommer une vidéo, pour désigner soit un visionnage depuis la plate-forme, soit un téléchargement. En second lieu, nous critiquerons, via l'analyse d'entretiens, l'interprétation en termes « d'indifférence vis-à-vis du certificat ». En effet, le choix de se cantonner aux vidéos d'un MOOC ne découle pas nécessairement d'une décision positive, il peut être subi.

## **2. Des concepts issus des travaux sur l'attrition en formation**

Nous avons mené ce travail sur la base de notions et concepts issus des travaux sur l'attrition en formation, comprise comme la proportion des inscrits à une formation qui n'obtiennent pas le certificat ou le diplôme associé. La question de l'attrition constitue une thématique de recherche vieille de plusieurs décennies tant au sein de la formation des adultes (Boshier, 1973) - pour mémoire, le public des MOOC est constitué essentiellement de personnes insérées dans le monde professionnel (Cisel, 2016) - qu'au sein de la formation à distance (Zawacki-Richter, 2014). Nous en retraçons ici les grandes lignes, en nous focalisant sur des recherches de nature descriptive. Celles-ci consistent en un développement de typologies de formes d'attrition, puis de motifs avancés pour expliquer un désengagement de la part du participant.

### **2.1.1. Construction de typologies de formes d'attrition**

Les recherches visant à construire des typologies de formes d'attrition remontent, aux États-Unis, au lendemain de la seconde guerre mondiale. Elles sont motivées par les coûts que l'abandon des études fait peser sur les institutions d'enseignement supérieur en particulier, et sur la société en général. On voit d'abord apparaître dans la littérature scientifique la distinction entre le retrait volontaire (Vaughan, 1968) - décision positive d'interrompre sa participation à une formation - et l'échec académique - décision subie, notamment du fait de notes insuffisantes.

Ces typologies évolueront peu par la suite, et ce n'est qu'à la marge des principaux travaux que des notions susceptibles d'être mobilisées ici apparaissent. En particulier, (Tinto, 2006) relève la notion d'*atteindre* proposée par (Terenzini, 1987). Celle-ci n'a jusqu'à présent rencontré que peu d'écho dans les travaux consacrés à l'attrition, et l'étayage empirique correspondant est particulièrement limité. Un *atteindre*, que nous avons traduit par *cibleur*, est un participant qui suit la formation avec un objectif relativement défini - réalisation d'un cours en particulier au sein d'un diplôme ou maîtrise d'une compétence particulière - et qui se désengage de la formation une fois cet objectif atteint. Dans le cadre de cet article, un objectif pourra être défini en termes d'activité à réaliser - le visionnage de l'essentiel des vidéos, par exemple - ou en termes de connaissances ou de compétences, comme la maîtrise de tel ou tel concept d'algorithmique. Cet objectif peut avoir été défini en amont du lancement du cours ou se préciser au fil de la formation.

Cependant, si le choix de se concentrer sur telle ou telle activité est dicté par une contrainte temporelle - manque de temps, etc. - ou de manière générale par ce qui relève d'une *barrière situationnelle* (notion définie par la suite), nous ne pouvons parler de cibleurs selon notre définition, car il ne s'agit pas d'une décision positive, mais d'un choix contraint. À bien des égards, Koller *et al.* (2013) identifient l'ensemble des non-certifiés qui visionnent la totalité des vidéos du cours à ce que nous définissons ici comme étant des *cibleurs*, une interprétation qui nous semble contestable.

Enfin, les recherches sur l'attrition en formation à distance conduisent à l'apparition d'une autre notion, restée jusque-là relativement marginale dans les premiers travaux, la non-participation, que l'on retrouve parfois dans la littérature anglo-saxonne sous le terme *non-start*. Le terme apparaît dès les années 1970 (Orton, 1977) pour désigner la classe des

inscrits qui ne commencent pas les formations, la définition de « commencer » changeant selon les types de formation. Plusieurs études prennent comme critère le non-rendu du premier devoir, pour définir la non-participation (Kempfer, 1996). Dans le cadre d'un MOOC, on dira ici qu'il y a non-participation si l'utilisateur ne visionne ni ne télécharge aucun contenu (vidéo, etc.) et ne réalise aucune activité évaluée. Cet élément est à prendre en compte pour éliminer des biais comme les différences de non-participation d'un cours à l'autre. Cette première typologie nous permet d'appréhender différents types d'attrition. Elle ne nous permet en revanche pas de qualifier les mécanismes qui ont conduit à cette attrition, quelle que soit sa nature. C'est à cette lacune que se sont adressés les travaux sur les *barrières* ou *obstacles* (Cross, 1981), (Garland, 1992), (Garland, 1993).

### **2.1.2. Les barrières : motifs avancés par les participants pour expliquer leur désengagement**

L'objectif des recherches qualitatives sur les « obstacles », ou « barrières », est de qualifier les éléments qui poussent les participants à se désengager au sens large, quelle que soit la forme d'attrition correspondante. Au cours de ses travaux doctoraux, Cross identifie trois barrières potentielles chez les apprenants adultes (Cross, 1981). Les barrières d'ordre *situationnel* sont liées aux circonstances particulières que rencontre l'apprenant : responsabilités familiales, manque de temps, etc. Les barrières d'ordre *dispositionnel* sont liées aux croyances, attitudes, ou valeurs de l'étudiant. C'est par exemple le cas lorsqu'un enquêté déclare ne pas être suffisamment autonome, de manière générale, pour terminer avec succès un MOOC. Les barrières d'ordre *institutionnel* dépendent de l'institution et des caractéristiques de la formation ; c'est le cas, par exemple, si un étudiant pointe une défaillance de l'équipe pédagogique dans la gestion du cours.

(Garland, 1992) développe le travail de Cross en suivant une approche ethnographique, dans le contexte de recherches sur l'enseignement à distance. Elle propose une catégorie supplémentaire, la barrière épistémique, liée notamment aux attentes des étudiants, à la pertinence du contenu de la formation au regard de ces attentes, à la maîtrise des prérequis, etc. La barrière épistémique correspond au contenu pédagogique de la formation plus qu'à la forme qu'elle prend. On s'intéresse ici au contenu de la formation défini en termes de thématiques abordées, d'objectifs pédagogiques, de pédagogies suivies (pédagogie par

projet, etc.), et non aux caractéristiques de la formation qui en sont indépendantes, comme la réactivité d'une équipe pédagogique. La distinction avec la barrière institutionnelle peut être tenue, ce qui explique sans doute l'apparition plus tardive de ce concept. Ces barrières peuvent être mises en jeu dans les différentes formes d'attrition caractéristiques d'un MOOC ou d'un dispositif de formation quel qu'il soit. Cette typologie de barrières nous a permis de structurer les entretiens menés au cours de cette recherche, afin d'éclairer les analyses de traces d'activités que nous avons menées par ailleurs. Ces traces sont issues d'un corpus de six cours sur lesquels nous allons maintenant revenir.

### **3. Méthodes de collecte des données**

#### **3.1. Présentation des cours étudiés**

Notre étude est basée sur six MOOC distincts de l'École Polytechnique et de Centrale-Supelec disponibles sur la plate-forme Coursera, dont deux correspondent à une itération de la formation. Les noms de code utilisés pour ces MOOC, le titre correspondant, les statistiques descriptives et les dates de lancement sont fournis dans le Tableau 1, où figurent également les informations suivantes : #Vid., nombre de vidéos présentes dans le dispositif ; #V-Q, nombre de quiz contenus dans les vidéos (format dit du In-vidéo quiz) ; #Quiz, nombre de quiz présents à l'extérieur des vidéos ; #Dev. Aut., nombre de devoirs évalués automatiquement ; #Dev. E.P., nombre de devoirs évalués par les pairs.

À l'exception du MOOC *AVI*, tous les MOOC étudiés incluaient des devoirs à rendre, qu'ils soient corrigés automatiquement, comme dans le cas du MOOC *ALG*, ou qu'ils soient corrigés exclusivement par les pairs, comme dans les cas du *BP* et *FA*. Nous parlerons d'*activité prescrite*, pour désigner toute tâche prescrite explicitement ou implicitement par les concepteurs du MOOC, et d'*activité évaluée* pour désigner toute activité dont la réalisation permet d'obtenir des points utilisés pour l'obtention d'un certificat. On notera que tous les certificats présentés ici sont strictement gratuits, aucun service payant n'est proposé aux participants. Le MOOC *AVI* ne propose que des vidéos et des quiz, tandis que les certificats des MOOC *BP* et *FA* se basent sur des devoirs évalués par les pairs. Seul le MOOC *ALG* se base sur des devoirs corrigés automatiquement.

La plupart des exercices automatisés proposés par ces cours, autrement appelés quiz, sont contenus dans les vidéos pédagogiques, c'est-à-dire qu'il

n'est possible d'y répondre qu'en visionnant au moins une partie de la vidéo pédagogique qui le contient. Les MOOC sont organisés en différents modules, un nouveau module étant diffusé chaque semaine.

**Tableau 1 • Structure des six MOOC de Coursera étudiés**

Code Cours	Titre du cours	#Vid.	#V-Q	#Quiz	#Dev. Aut.	Dev. E.P.	Date de début	Date de fin
ALG1	Conception et mise en œuvre d'algorithmes	45	29	3	8	0	28/10/2013	01/01/2014
ALG2	Conception et mise en œuvre d'algorithmes	47	31	4	9	0	27/10/2014	27/12/2014
FA1	An Introduction to Functional Analysis	54	32	8	0	9	02/02/2014	07/04/2014
FA2	An Introduction to Functional Analysis	56	38	8	0	7	12/09/2014	06/11/2014
BP1	Business and Philosophy	38	20	4	0	7	14/04/2014	25/05/2014
AV1	Discrete inference And Learning in Artificial Vision	66	37	8	0	0	09/01/2014	06/03/2014

Le terme *certificat* désigne toute forme de validation gratuite permettant d'attester du suivi avec succès d'un MOOC, ce qui correspond sur Coursera au *Statement of Accomplishment*. Le terme *taux d'attrition*, ou *attrition* employée comme quantité, désigne, pour un cours donné, la proportion des inscrits qui n'obtiennent pas la note nécessaire à l'obtention du certificat au moment de la clôture de la formation, sans pour autant s'être désinscrits à ce moment-là, ce qui équivaut à  $(1 - \text{taux de certification})$ . Les participants qui se sont inscrits puis désinscrits avant la

fin de la formation ne sont pas pris en compte dans le calcul de cet indicateur.

Les statistiques précises pour chaque cours sont fournies en Annexe. Le nombre d'inscrits peut dépasser 70.000, et le nombre de certifiés peut atteindre 2700. Les taux de certification ne dépassent pas 5% dans le cas des MOOC étudiés, ils varient entre 1,5% des inscrits pour le MOOC *FA1* et 4,6% pour *AVI*. L'analyse des traces d'activité des MOOC de Coursera montre que le taux de connexion, c'est-à-dire le pourcentage des inscrits qui se connectent au moins une fois au cours, varie selon les six MOOC entre 50,9 et 69,5%. La proportion des participants qui consomment au moins une vidéo pédagogique est comprise entre 48 et 56% des inscrits selon les cours, elle est jusqu'à 20% inférieure à ceux qui se connectent au moins une fois à la plateforme. La proportion des inscrits qui rendent au moins un devoir varie selon les cours entre 5,8 et 11,4% des inscrits. Nous avons cherché dans cette recherche à aller au-delà d'une approche purement descriptive et avons pour cela mobilisé des méthodes mixtes.

### **3.2. Méthodes mixtes et stratégie séquentielle explicative**

Nous avons collecté d'une part des données quantitatives, les traces d'activité, et d'autre part des données qualitatives, au travers des entretiens, que nous avons articulées lors de la présentation des résultats. Les méthodes mixtes connaissent une popularité croissante dans le champ des sciences sociales en général (Mertens, 2003), (Creswell, 2009, p. 203) et dans celui de l'éducation en particulier (Creswell et Clark, 2006). Creswell propose quatre axes pour discriminer les différentes approches, dont l'ordre et l'importance relative de la collecte et de l'analyse des données qualitatives et quantitatives (Creswell, 2009).

Sur la base de ces quatre axes de description, Creswell propose de ranger les approches mixtes en six grandes catégories, dont la stratégie *séquentielle explicative* (Ivankova *et al.*, 2006), que nous avons suivie ici. Dans cette stratégie, les données quantitatives sont collectées et analysées en amont des données qualitatives, celles-ci contribuant à l'interprétation des données quantitatives. Nous décrirons d'abord la manière dont ont été menées les analyses sur les traces d'activité, qui correspondent aux données quantitatives, pour nous pencher ensuite sur les données qualitatives, correspondant aux entretiens semi-directifs.



### **3.3. Traces d'activité**

#### **3.3.1. Tables fournies par Coursera**

Les traces d'activité des MOOC de Coursera étudiés ici ont été fournies par les deux établissements partenaires de la plateforme avec lesquels nous avons collaboré. Sur demande des établissements partenaires, la plateforme *Coursera* envoie, en sus des logs bruts, un fichier SQL comportant trente-trois tables, dont nous exploitons deux. L'une de ces tables comprend les informations sur l'utilisation de vidéos, indiquant notamment si une vidéo donnée a été visionnée depuis la plateforme ou si elle a été téléchargée. La plateforme *Coursera* se distingue d'autres plateformes comme *edX*, qui ne fournissent pas d'informations sur les vidéos téléchargées. La seconde table utilisée comporte les informations relatives aux notes obtenues par le participant, ainsi que l'obtention éventuelle du certificat.

#### **3.3.2. Définitions opératoires**

Sur la table relative à la consommation de vidéos, nous nous sommes intéressés notamment à deux types d'actions liées aux vidéos, avant tout chez les non-certifiés : le téléchargement d'une part et le visionnage d'autre part. Par la suite, nous utiliserons le terme *visionnage* pour désigner le fait de cliquer sur le bouton *Lecture* de la vidéo depuis le lecteur de la plateforme. On dit en revanche que la vidéo a été *lancée* si l'on a appuyé sur ce bouton, que cela soit depuis la plateforme ou depuis l'appareil sur lequel la vidéo a été téléchargée - dans ce dernier cas, l'action n'est pas visible dans les traces. Notons que même lorsqu'il y a *visionnage* - l'action étant relevée dans les traces - l'on ne peut pas savoir si la vidéo est jouée dans son intégralité, ni si la vidéo est affichée à l'écran tandis qu'elle est jouée (l'utilisateur peut se contenter d'écouter la bande audio) ou, en d'autres termes, si elle est *regardée* dans son intégralité, c'est-à-dire avec l'utilisateur physiquement présent devant la vidéo durant toute sa durée. De la même manière, dans le cas du téléchargement, il est impossible de déterminer si l'utilisateur lance ou non la vidéo suite au téléchargement. Par conséquent, la probabilité que la vidéo ait été regardée est a priori plus élevée si la vidéo est lancée depuis la plateforme que si elle a été simplement téléchargée.

Nous parlerons par la suite de *consommation de vidéo* lorsque nous désignerons indistinctement le visionnage ou le téléchargement de vidéos, car c'est une pratique fréquente de Coursera de ne pas faire cette

distinction. En d'autres termes, si le participant a visionné ou téléchargé une vidéo, il l'a *consommée*. L'*assiduité* concerne, dans le cadre de cet article, la proportion de vidéos du dispositif consommées.

Nous parlerons de non-certifié visionneur assidu (NCVA) pour désigner tout non-certifié qui a visionné au moins 90% des vidéos du dispositif, de non-certifié téléchargeur assidu (NCTA) pour désigner toute personne qui a téléchargé au moins 90% des vidéos du dispositif. Ces deux catégories ne sont pas mutuellement exclusives, car il est possible d'être simultanément un NCVA et un NCTA. Nous qualifierons de non-certifié consommateur assidu (NCCA) toute personne qui a visionné et/ou téléchargé au moins 90% des vidéos du dispositif. Enfin, un utilisateur qui a consommé au moins une vidéo est un non-certifié consommateur (NCC). Le Tableau 2 présente la définition des sigles utilisés.

**Tableau 2 • Définition des sigles utilisés pour qualifier les non-certifiés**

Sigle	Décomposition du sigle	Définition opératoire
NCVA	Non-certifié visionneur assidu	Non-certifié qui visionne depuis la plateforme au moins 90% des vidéos présentes dans le dispositif
NCTA	Non-certifié téléchargeur assidu	Non-certifié qui télécharge depuis la plateforme au moins 90% des vidéos présentes dans le dispositif
NCCA	Non-certifié consommateur assidu	Non-certifié qui visionne OU télécharge depuis la plateforme au moins 90% des vidéos présentes dans le dispositif
NCC	Non-certifié consommateur	Non-certifié qui visionne OU télécharge depuis la plateforme au moins une vidéo du dispositif
NCVA-D	Non-certifié visionneur assidu - devoir rendu	Non-certifié visionneur assidu rendant au moins un devoir prescrit (que celui-ci soit évalué par les pairs ou par un programme)
NCVA-Q	Non-certifié visionneur assidu - quiz soumis	Non-certifié visionneur assidu soumettant au moins trois quiz

Nous chercherons à déterminer si ces différentes catégories de participants réalisent des quiz ou des devoirs. Dans cette perspective, nous ajoutons les sigles -Q et -D. Ils fonctionnent par apposition et permettent respectivement de désigner, parmi les catégories de participants préalablement définies, ceux qui ont réalisé au moins trois quiz ou (respectivement) soumis au moins un devoir, indépendamment de la manière dont celui-ci est corrigé (nous n'illustrons dans le Tableau 2 que les cas avec les NCVA). Nous choisissons trois quiz, et non un seul, car pour certains de ces cours, il existe un ou deux quiz, difficiles à identifier dans les traces, dont la fonction est de s'assurer que le participant a bien compris le fonctionnement de la plateforme ; y répondre ne signifie pas que l'on a commencé à suivre le cours, alors que c'est probablement davantage le cas lorsque l'on a réalisé au moins trois quiz.

### **3.3.3. Logiciels et tests statistiques utilisés pour l'analyse**

Les statistiques rapportées dans les tableaux présentés dans le corps de l'article ont été réalisées avec le logiciel R, version 3.2 (<http://www.R-project.org>). Nous avons utilisé plusieurs types de tests statistiques, dont une analyse de survie, et plus précisément le modèle de Cox, méthode statistique adaptée pour analyser les variables qui correspondent à des données de survie. Les données de survie correspondent à des variables correspondant au temps restant avant un événement donné, comme le départ d'une formation. Les données comme le nombre de semaines durant lequel un participant est actif, ou comme la proportion des vidéos d'un dispositif visionnées, ont le format requis pour être considérées comme des données de survie (Reich, 2014). La statistique associée à une analyse de survie est nommée le Hazard Ratio (HR) ; elle est généralement fournie avec l'intervalle de confiance à 95% associée à ce HR, le  $R^2$  du modèle (l'équivalent de la part de variance expliquée pour le modèle linéaire), et la valeur du test z. Cette dernière est plus intéressante que la p-value lorsque celle-ci est très basse. Dans les cas que nous allons présenter, la p-value est systématiquement inférieure à  $10^{-15}$  et présente donc peu d'intérêt. Enfin, rappelons que le nombre précédé du signe  $\pm$  correspond à l'écart-type ; il est calculé sur la base des six cours analysés.

### **3.4. Entretiens semi-directifs**

En sus de l'analyse des traces d'activité, nous avons fait des entretiens semi-directifs. Nous ne présentons ici que succinctement la méthodologie suivie - voir (Cisel, 2016) pour de plus amples développements.

### **3.4.1. Sélection des interviewés**

Plusieurs stratégies ont été employées pour atteindre d'éventuels enquêtés : diffusion d'une annonce via la lettre d'information du MOOC *Monter un MOOC de A à Z*, et diffusion via les réseaux sociaux. Les personnes répondant à nos sollicitations étaient catégorisées selon leurs caractéristiques socioprofessionnelles (PCS) dans un tableur, puis recontactées par la suite, en cherchant à maximiser la diversité de PCS. On notera ici que les enquêtés ont suivi d'autres MOOC que ceux étudiés dans le volet quantitatif de l'article. Ceci tient à plusieurs raisons ; tout d'abord, la réglementation de *Coursera* interdisait de contacter les inscrits d'un cours à des fins de recherche aussi longtemps après son déroulement. Ensuite, l'objectif n'était pas, à travers ces entretiens, d'interpréter les comportements observés dans le cas particulier de notre corpus, mais de proposer des contre-exemples à l'interprétation faite par (Koller *et al.*, 2013).

La population d'enquête, à majorité féminine (24 femmes pour 17 hommes) était en moyenne âgée de 41 ( $\pm 11$ ) ans. Le nombre de MOOC suivis par ces apprenants était plus élevé que la moyenne, qui est de 2,3 selon les données d'inscription de FUN (Cisel, 2016) : 5,8 ( $\pm 7,5$ ) MOOC suivis en moyenne, l'écart-type élevé résultant de l'existence d'utilisateurs inscrits à plusieurs dizaines de cours. Il n'est pas surprenant qu'un utilisateur qui réponde à une sollicitation d'entretien ne soit pas tout à fait représentatif des utilisateurs de MOOC ; le biais est trop important pour que l'on puisse inférer à partir de la fréquence des comportements rapportés au cours des entretiens. S'ils ne visent pas la représentativité, les extraits d'entretien indiquent en revanche des possibles et constituent de ce fait une aide précieuse pour l'interprétation.

### **3.4.2. Condition de réalisation et d'analyse des entretiens**

Nous avons réalisé des entretiens semi-directifs durant en moyenne une quarantaine de minutes avec quarante-et-un enquêtés. Le canevas d'entretien était composé de plusieurs parties, répondant chacune à une fonction particulière. Outre le talon sociologique (âge, catégorie socioprofessionnelle, etc.), un certain nombre de questions étaient posées, relatives notamment aux motivations pour s'inscrire et aux intentions sous-jacentes aux actions réalisées au sein du MOOC.

Les trois questions suivantes sont à la base d'un certain nombre des extraits qui suivent :

- « Quand vous vous inscrivez, est-ce que vous savez avec certitude si vous allez terminer le cours ? » ;
- « Est-ce qu'il vous arrive de butiner dans un MOOC, c'est-à-dire ne regarder que les contenus qui vous intéressent, sans vous contraindre à suivre la séquence prescrite, et si oui pourquoi ? » ;
- « Vous est-il arrivé de suivre un MOOC en ne faisant que consulter les ressources pédagogiques, sans participer aux activités évaluées ? Si oui, pourquoi ? ».

La structure de l'entretien n'étant pas figée, il arrivait fréquemment que les éléments d'intérêt émergent au détour d'une autre question. La majorité des entretiens ont été réalisés via l'outil de visioconférence *Skype* ou par téléphone. Les entretiens réalisés étaient enregistrés par *Camtasia*, puis retranscrits dans leur intégralité. Une fois retranscrits, les entretiens ont été analysés selon le principe de l'analyse thématique (Blanchet et Gotman, 2010), (Beaud et Weber, 2010). L'objectif est avant tout d'assurer une correspondance élevée entre les explications fournies par un enquêté et les comportements observables qu'un extrait d'entretien est supposé éclairer. Nous nous sommes en particulier intéressés aux actions des non-certifiés qui se sont investis dans la formation, notamment pour ne regarder que les vidéos, ou qu'une partie des vidéos du dispositif. C'est sur cette thématique que l'essentiel des extraits d'entretiens présentés ici sont issus.

#### **4. Résultats**

Nous présenterons d'abord une analyse des traces d'activité et poursuivrons par une présentation d'extraits d'entretiens, suivant la logique de la stratégie séquentielle explicative appliquée dans cette étude.

##### **4.1. Analyse des traces d'activité**

L'analyse des traces est décomposée en trois temps. Dans un premier temps, nous comparons les parts respectives de l'activité observable du cours réalisées par les certifiés et les non-certifiés, et chercherons, pour les seconds, à déterminer dans quelle mesure il existe un recouvrement entre visionnage et téléchargement des vidéos. Nous nous intéressons ensuite à l'importance quantitative des NCVA et des NCTA par rapport aux certifiés, afin de réaliser une comparaison avec les résultats de (Koller *et al.*, 2013). Enfin, nous nous penchons sur la participation des NCVA et des NCTA aux activités évaluées, afin, notamment, de trouver des indices relatifs au devenir des vidéos téléchargées.

**4.1.1. La part des non-certifiés dans l'activité observable du cours**

Selon une logique similaire à celle d'une étude menée sur des MOOC de Unow (Cisel et al., 2015), nous avons réalisé une synthèse des parts respectives des certifiés et des non-certifiés dans l'activité observable du cours, mesurée en termes de consommation de vidéos. Cette synthèse est présentée dans le Tableau 3, où la dernière ligne (*Moy.*) indique la moyenne sur les six cours et les écarts-types correspondants. La distinction est faite pour la consommation de vidéos entre le visionnage de vidéos sur la plate-forme et le téléchargement de vidéos. Les deux types d'actions sont confondus pour le calcul des sommes. Les certifiés sont responsables en moyenne de 17% des actions réalisées sur les vidéos, l'action de visionnage dominant. Les non-certifiés sont responsables en moyenne de 83% des actions sur les vidéos, l'action de téléchargement dominant sur l'action de visionnage. Au vu de la taille de notre corpus de six cours, le seul test statistique valide est le test du chi-deux, réalisé sur les deux variables que sont la nature des actions réalisées et l'obtention du certificat. Le test du chi-deux montre qu'à l'échelle d'un MOOC, les non-certifiés sont responsables de davantage de téléchargements que de visionnages ( $\chi^2=49850$ ,  $ddl=1$ ,  $p\text{-value}<10^{-16}$ ), tandis que nous observons l'inverse pour les certifiés. Ceci souligne l'importance de la prise en compte du devenir des vidéos téléchargées.

**Tableau 3 • Part des certifiés et des non-certifiés dans l'activité observable des six cours analysés, en termes de consommation de vidéos**

	Certifiés		Non-certifiés		Total actions
	Visionnage (%)	Téléchargement (%)	Visionnage (%)	Téléchargement (%)	
<i>ALG1</i>	11	9	24	56	116439
<i>ALG2</i>	11	6	31	51	74565
<i>FA1</i>	5	3	65	26	384984
<i>FA2</i>	10	4	36	49	251296
<i>BPI</i>	13	6	43	37	781363
<i>AVI</i>	14	10	17	58	439452
<i>Moy.</i>	11 ( $\pm 3$ )	6 ( $\pm 3$ )	36 ( $\pm 17$ )	47 ( $\pm 12$ )	

L'activité observable des non-certifiés est bien supérieure à celle des certifiés en termes de consommation de vidéos. Les non-certifiés pris dans leur ensemble sont systématiquement responsables de l'essentiel de l'activité observable du cours en termes de consommation de vidéos, que celle-ci soit mesurée en termes de visionnage, de téléchargement, ou indistinctement en termes de consommation. Néanmoins, un tel résultat ne corrobore pas à ce stade l'idée avancée par (Koller *et al.*, 2013) selon laquelle nombreux sont les non-certifiés qui visionnent la plupart des vidéos du dispositif. Tout d'abord, rappelons que les non-certifiés sont beaucoup plus nombreux que les certifiés. Ensuite, leur activité observable peut être concentrée sur les premières vidéos des cours. Enfin, le téléchargement n'implique pas le visionnage.

Nous ne dirons que quelques mots sur la participation aux activités évaluées (Tableau 4), qui ne constitue pas la focale de cet article. On constate que pour deux des trois MOOC qui mobilisent des quiz, les non-certifiés sont responsables de la plus grande partie de l'activité observable en termes de quiz effectués, ce qui est cohérent avec les résultats concernant le visionnage des vidéos, puisque les quiz sont le plus souvent intégrés dans les vidéos. En revanche, les certifiés sont systématiquement responsables de l'essentiel de l'activité mesurée en termes de devoirs rendus, que ces devoirs soient corrigés automatiquement ou par les pairs.

**Tableau 4 • Part des certifiés et des non-certifiés dans l'activité observable des six cours analysés en termes de participation à des activités évaluées**

	Soumission de quiz			Rendu de devoirs		
	Certi- fiés (%)	Non- certifiés (%)	Tot. actions	Certifiés (%)	Non- certifiés (%)	Tot. actions
<i>ALG1</i>	28	71	14364	73	27	30018
<i>ALG2</i>	59	40	18973	60	39	12618
<i>FA1</i>				56	43	3925
<i>FA2</i>				59	40	3163
<i>BPI</i>				84	15	17527
<i>AVI</i>	36	63	51183			

#### **4.1.2. Visionner, télécharger, ou les deux à la fois ?**

Le deuxième élément que nous avons étudié est le degré de recouvrement entre visionnages depuis la plateforme et téléchargements, et ce pour les seuls non-certifiés. On constate sur la base de deux indicateurs que ce recouvrement est au final relativement faible, indicateurs que nous présentons dans les premières colonnes du Tableau 5. Le premier indicateur (*% Vid. Vis. Tél.*) est le pourcentage de vidéos visionnées également téléchargées (indépendamment de l'ordre dans lequel ces deux actions sont réalisées), tandis que  $\rho$  est coefficient de corrélation de Spearman, pour la corrélation entre le nombre de vidéos visionnées et le nombre de vidéos téléchargées.

Le recouvrement est toujours inférieur à 26%, et peut diminuer jusqu'à 14%. Il est en moyenne de 19% ( $\pm 4,5$ ). En somme, soit on télécharge une vidéo, soit on la visionne ; seule une minorité des vidéos, quoique non négligeable, sont à la fois visionnées et téléchargées par les non-certifiés. Pour étayer cette affirmation avec un test statistique, nous avons calculé le coefficient de corrélation de Spearman entre le nombre de vidéos téléchargées et le nombre de vidéos visionnées, pour un NCC, selon une logique de test apparié : la corrélation est presque toujours non statistiquement significative. Dans le seul cours où elle l'est, FA1, le coefficient est négatif, et relativement faible, ce qui suggère que le comportement qui consiste à visionner une vidéo puis à la télécharger, pour une raison ou pour une autre, est relativement peu représenté.



**Tableau 5 • Indicateurs de consommation de vidéos pour les Non-Certifiés  
Consommateurs**

	% Vid. Vis. Tél.	$\rho$ (p-value)	M vid. vis. (décile)	M vid. tél. (décile)	HR	HR Inf	HR Sup	Z
<i>BP1</i>	14	$6.10^{-4}$ (0.96)	1.1	4.7	0.466	0.455	0.478	-60.2 (0.11)
<i>ALG1</i>	26	0.06 (0.03)	1.0	4.4	0.298	0.279	0,219	-36.1 (0.24)
<i>ALG2</i>	19	-0.020 (0.58)	1.0	4.4	0.293	0.271	0.318	-29.8 (0.23)
<i>FA1</i>	18	-0.060 (<0.00 1)	2.6	5.7	0.294	0.283	0.304	-64.1 (0.23)
<i>FA2</i>	14	0.020 (0.34)	1.4	5.1	0.305	0.290	0.321	-46.1 (0.21)
<i>AV1</i>	19	0.015 (0.46)	1.1	5.2	0.272	0.256	0.285	-55.7 (0.27)

#### **4.1.3. Visionnage des vidéos et désengagement rapide**

Nous avons vu que les non-certifiés étaient peu investis dans les activités évaluées, mais responsables d'une part importante de l'activité observable d'un MOOC en termes de consommation de vidéos. Cela ne signifie pas pour autant qu'ils soient nombreux à faire preuve d'assiduité vis-à-vis des vidéos. Nous cherchons ici à contraster les comportements de visionnage et de téléchargement, avant de comparer nos résultats à ceux de (Koller *et al.*, 2013).

Pour chaque cours, nous nous sommes intéressés aux indicateurs que sont les nombres moyens de vidéos qu'un individu visionne ou télécharge, en raisonnant en déciles, pour comparer des cours contenant des nombres différents de vidéos. Si un participant a visionné plus de 10% des vidéos du cours, on considère qu'il a visionné le premier décile. Il est préférable de raisonner en déciles plutôt qu'en pourcentages au vu du

faible nombre de vidéos par dispositif, quelques dizaines au plus. Cette approche est par ailleurs recommandée pour le test statistique mobilisé, à savoir l'analyse de survie, qui ne nécessite pas obligatoirement un modèle dépendant du temps. Nous présentons ces analyses dans la suite du Tableau 5. « *M vid. vis.* » y représente, pour ceux qui visionnent au moins une vidéo, le nombre moyen de déciles de vidéos visionnées. « *M vid. tél.* » correspond, pour ceux qui téléchargent au moins une vidéo, au nombre moyen de déciles de vidéos téléchargées. *HR* est le Hazard Ratio (statistique spécifique des analyses de survie), tandis que *HR Inf* et *HR Sup* sont les valeurs supérieures et inférieures du HR avec un intervalle de confiance à 95%. *Z* est la valeur du test réalisé (modèle de Cox).

On constate que les non-certifiés qui visionnent au moins une vidéo visionnent en moyenne à peine au-delà du premier décile, à l'exception d'un cours, où les non-certifiés visionnent 2,6 déciles en moyenne. La situation est relativement différente pour le téléchargement, où ceux qui téléchargent au moins une vidéo vont en moyenne au-delà du quatrième, voire, pour certains cours, au-delà du cinquième décile.

Pour chacun des cours, nous utilisons une analyse de survie pour vérifier que ces différences sont statistiquement significatives. L'analyse de survie permet d'estimer le risque d'arrêt d'une activité à un instant donné, ou plutôt la différence qui peut exister entre deux activités données vis-à-vis du risque d'interruption. On considère le visionnage et le téléchargement comme deux modalités de la variable « consommation de vidéos ». Le Hazard Ratio calculé est issu de la comparaison de l'activité de téléchargement et de l'activité de visionnage, cette dernière servant de référence : s'il est inférieur à 1, cela signifie que la chute de l'activité de téléchargement (toujours mesurée en proportion de vidéos du dispositif) est plus faible que celle de l'activité de visionnage. On observe (Tableau 5) que le HR est toujours systématiquement inférieur à 0,5, voire à 0,3 et que les valeurs du test *z* sont très négatives, ce qui suggère que les différences sont statistiquement très significatives entre activité de téléchargement et activité de visionnage pour tous les cours que nous avons analysés.

En résumé, un NCC visionne en moyenne peu de vidéos, mais quand il en télécharge, il télécharge en moyenne une partie significative de ces vidéos, voire la majorité. L'activité observable du cours en termes de visionnage de vidéos s'explique largement par l'existence de non-certifiés qui visionnent une faible proportion des vidéos pédagogiques disponibles dans le dispositif. L'activité de consommation de vidéos, constatée dans le

Tableau 3, est due essentiellement au grand nombre de non-certifiés qui visionnent les premières vidéos de la formation et au plus petit nombre de non-certifiés qui téléchargent une partie significative des vidéos téléchargées (sans les avoir visionnées depuis la plateforme, comme nous l'avons vu plus tôt). Le grand nombre de non-certifiés compense leur faible investissement individuel et explique qu'ils génèrent souvent davantage d'activité observable que les certifiés lorsque l'on raisonne à l'échelle du cours.

#### **4.1.4. Calcul de différents ratios et comparaison avec Koller et al. (2013)**

Complétons ces analyses en comparant nos résultats avec les conclusions de l'article de Koller *et al.* (2013). Nous présentons en premier lieu les ratios entre les participants de type NCVA, NCTA ou NCCA et le nombre de certifiés Cf, indicateurs utilisés par ces auteurs. Ils sont calculés sur la base de la moyenne des ratios des six cours étudiés (Figure 1, Gauche). Le ratio entre le nombre de NCVA et le nombre de certifiés est en moyenne égal à 0,33 ( $\pm 0,26$ ) - il varie entre 0,13 et 0,79, respectivement pour le MOOC *ALG1* et pour le MOOC *BPI*. Le ratio entre le nombre de NCTA et le nombre de certifiés vaut en moyenne 2,34 ( $\pm 1,35$ ) - il varie entre 1,38 et 5,06. Enfin, le ratio entre le nombre de NCCA et le nombre de certifiés vaut en moyenne 2,72 ( $\pm 1,44$ ) - il varie entre 1,71 et 5,63.

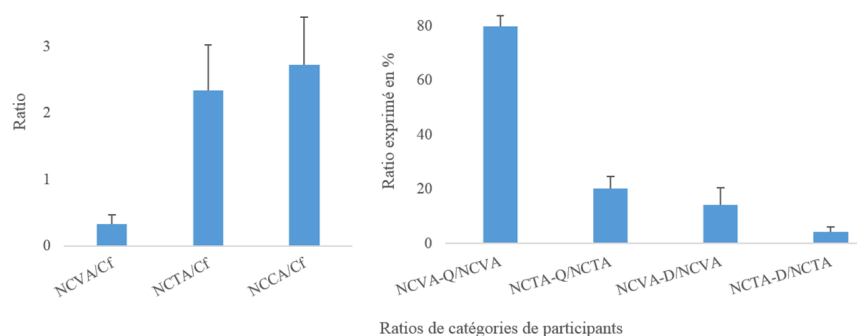
Ces comparaisons n'appellent pas de tests statistiques, car même lorsque la nature de la comparaison pourrait légitimer un tel test, le nombre de cours étudiés est trop faible pour que les conditions de validité soient respectées. On retrouve des statistiques relativement similaires à celles avancées par (Koller *et al.*, 2013) - des non-certifiés assidus plus nombreux que les certifiés - si l'on prend en considération uniquement le ratio ayant pour numérateur les NCCA ou les NCTA. L'emploi de cet indicateur peut conduire à surestimer le nombre de personnes ayant lancé une vidéo, pour les raisons citées en introduction : on ne peut savoir quelle est la proportion des vidéos téléchargées qui ont été visionnées par la suite. Si l'on se base sur l'indicateur qu'est le nombre de NCVA, les seuls inscrits pour lesquels on puisse dire avec certitude qu'ils ont lancé la vidéo, le ratio est beaucoup plus bas que suggéré par ces auteurs.

Les cours mobilisés dans notre article, produits par l'École Polytechnique et l'École Centrale, sont certes postérieurs à l'analyse proposée par (Koller *et al.*, 2013) ; par ailleurs, nous ne pouvons pas écarter le fait que les thématiques abordées - algorithmique, entrepreneuriat,

ingénierie - comme la structure des dispositifs de notre corpus puissent introduire des biais. Mais il est peu probable que les cours de notre corpus constituent autant d'exceptions à la règle énoncée par les auteurs, et que de tels biais suffisent à expliquer les différences observées dans la comparaison des deux corpus.

Nous avons effectué des calculs supplémentaires en prenant comme dénominateur tous les non-certifiés (Cisel, 2016). Les calculs montrent que les NCVA représentent systématiquement moins de 1% des non-certifiés, à l'exception du MOOC *BPI*, où ce taux est 2,9%. Les NCTA constituent une forme d'attrition non négligeable, puisqu'ils peuvent représenter jusqu'à 7,6% de l'attrition. De manière générale, les NCCA ne représentent jamais plus de 10% des non-certifiés dans les six cours étudiés. Ce sont des formes d'attrition marginales.

Si l'on ne peut estimer à partir des traces d'activité dans quelle mesure les NCTA visionnent les vidéos qu'ils téléchargent, on peut en revanche analyser leur participation en termes d'activités évaluées, soumissions de quiz ou de devoirs. Une participation aux activités évaluées suggère que les vidéos sont vraisemblablement visionnées. Pour ce faire, nous calculons le pourcentage, au sein des NCVA ou des NCTA, de ceux qui réalisent au moins trois quiz et de ceux qui rendent au moins un devoir. On constate que, sur les six MOOC étudiés, 79,8%(±7,8) des NCVA réalisent trois quiz, tandis que cet indicateur chute à 20,2%(±8,9) pour les NCTA (Figure 1, Droite). Le différentiel est également frappant en ce qui concerne les devoirs : 14,3%(±12,1) des NCVA ont rendu un devoir, contre 4,3%(±3,8) des NCTA. Pour les six cours étudiés, environ 80% des NCTA ne réalisent donc aucune activité évaluée, quelle qu'elle soit. Cela ne signifie pas qu'ils ne visionnent pas les vidéos, mais les traces liées aux activités évaluées ne suggèrent pas que ces vidéos soient regardées.



**Figure 1 • Importance relative des différentes catégories de participants**  
**Gauche : Ratios NCVA, NCTA et NCCA / nombre de certifiés Cf**  
**Droite : Engagement dans les activités évaluées des NCVA et NCTA**

Notons que nombre de quiz sont inclus dans les vidéos, et qu'il faut visionner la vidéo pour pouvoir y répondre, ce qui peut contribuer à expliquer la forte représentation des NCVA-Q par rapport aux NCVA. Le ratio NCTA-Q/NCTA peut s'expliquer en grande partie du fait du recouvrement, c'est-à-dire par le fait que ce sont des NCTA qui ont également visionné les vidéos depuis la plateforme. Le temps nécessaire à la réalisation des devoirs est bien supérieur à celui nécessaire pour la passation des quiz, ce qui explique sans aucun doute la faible représentation des NCVA-D et des NCTA-D. Ce n'est pas tant les différences entre ces catégories de participants qui méritent un éclaircissement, que les mécanismes qui conduisent à adopter ces comportements. À cet égard, nous atteignons la limite de ce que peut nous apporter l'analyse des traces d'activité, d'où un recours à des entretiens semi-directifs qui nous offriront des pistes d'interprétation.

#### **4.2. Analyse des entretiens**

Nous avons établi dans le Tableau 2 une distinction entre NCVA et NCTA, et, au sein de ces catégories de participants, entre ceux qui réalisent des activités évaluées, devoirs (NCVA-D par exemple), ou quiz notés (NCVA-Q), et ceux qui ne le font pas. Nous allons chercher dans les entretiens des éléments susceptibles d'interpréter ces divers comportements sur la base de la notion de « barrière ». Si ces extraits nous permettront de mieux comprendre la pertinence de la distinction entre NCVA et NCTA, ils n'ont pas vocation à expliquer ce qui mène à l'apparition des divers types de NCVA ou de NCTA. Nous mettrons en évidence une certaine diversité de mécanismes pouvant conduire à des

résultats relativement proches en termes de comportements observables, afin de nuancer l'affirmation selon laquelle ces comportements correspondent essentiellement à des décisions positives. Dans une première partie de l'analyse des entretiens, nous illustrerons les barrières qui peuvent conduire à devenir un NCVA. Dans une seconde partie, nous nous intéresserons au devenir des vidéos téléchargées par les NCTA.

Sur les quarante-et-un enquêtés, quatre ont déclaré s'être contents de visionner la quasi-totalité des vidéos pédagogiques du cours, sans participer aux activités évaluées, et correspondent ainsi à des NCVA, ou à des NCTA ayant visionné toutes les vidéos téléchargées. Huit enquêtés supplémentaires ont dit avoir visionné la majorité des vidéos de cours dont ils n'ont pas obtenu le certificat, sans pour autant en avoir visionné la quasi-totalité. Nous interprétons les justifications apportées par ces enquêtés pour expliquer leur comportement selon les quatre catégories de barrières proposées par (Garland, 1992). Notons que ces extraits illustrent différents types de barrières, mais que ces barrières peuvent correspondre à des cas bien plus divers que ne le laissent supposer les extraits mobilisés.

#### **4.2.1. Une diversité de barrières pour les NCVA**

Nous commencerons par illustrer le cas de la barrière dispositionnelle, qui correspond dans une large mesure à l'hypothèse d'un désintérêt pour le certificat. Dans ce cas de figure, et bien que l'on parle de « barrière », il s'agit d'une décision positive de ne pas chercher à obtenir le certificat, tout en visionnant un certain nombre de vidéos ou en participant à des activités évaluées par intérêt pour les ressources. Cette situation correspond à cet égard au cas des cibleurs. Nous allons illustrer ce cas de figure par deux extraits d'entretiens. Dans les deux cas qui suivent, les interviewés ne correspondent pas à proprement parler à des NCVA, car ils n'ont pas visionné la quasi-totalité des vidéos, mais ils consomment la majeure partie des vidéos de la formation et réalisent éventuellement quelques activités évaluées. Commençons par les propos de cet ingénieur web, qui considère ne pas avoir à suivre exactement ce que propose l'équipe pédagogique, et pour qui le MOOC représente une sorte de « buffet » dans lequel il se sert.

*Je ne vais pas courir après les attestations ou autres, ce n'est pas le but ; ce qui m'intéresse c'est d'apprendre. [...] Si c'est pour me former dans le domaine technique sur un point bien particulier, je vais m'accrocher et je vais essayer de bien creuser le truc. Si c'est juste pour apprendre de nouvelles choses, piocher*

*des idées ou des connaissances, à droite, à gauche, c'est un peu le buffet où je me sers. Je ne suis pas obligé de suivre ce qui est proposé exactement. (E24)*

Nous considérons que cet extrait illustre parfaitement le cas du cibleur, qui sait très bien ce qu'il veut ou ce qu'il ne veut pas, dans la mesure où c'est une attitude par rapport à la formation qui le conduit à « piocher » parmi les éléments disponibles. La logique hédonique, prendre du plaisir à apprendre, prévaut dans certains MOOC sur une logique utilitariste, qui le conduirait à être plus assidu dans la réalisation des activités évaluées. C'est donc une décision positive, fondée sur la motivation qui l'a poussé à s'inscrire au MOOC, qui le conduit à devenir un NCVA.

Néanmoins, le fait de devenir un NCVA ne peut s'interpréter uniquement en termes de motivation, d'intérêt, d'objectif d'apprentissage (Tough, 1971). Certaines caractéristiques de la formation comme le contexte qui entoure l'inscription peuvent conduire à faire le choix d'adopter ce type de comportement, ce qui amène à discuter des barrières institutionnelles et situationnelles. Le premier argument consiste à avancer que les activités évaluées sont trop exigeantes en termes d'investissement temporel. Le participant préfère alors se contenter de visionner les vidéos pédagogiques. Ce cas de figure est illustré par les propos de ce lycéen de dix-huit ans, qui ne fait que visionner les vidéos d'un MOOC de géopolitique du fait de la charge de travail trop importante associée aux devoirs évalués par les pairs :

*Il y a certains MOOC, Géopolitique, de Grenoble Ecole de Management, la charge de travail était très très élevée, parce qu'il fallait faire des dissertations toutes les deux semaines. Je pouvais pas, mais j'ai vraiment regardé les vidéos. (E41)*

Cet argument correspond à une barrière institutionnelle (Garland, 1992, p.126), dans la mesure où ce sont des caractéristiques de la formation qui sont utilisées pour légitimer le fait de ne pas réaliser les activités prescrites, mais aussi une barrière situationnelle, car l'enquête fait référence à un manque de temps. Dans le second exemple, une arrivée trop tardive dans la formation fait obstacle à l'obtention du certificat, les dates limites de réalisation des activités évaluées ne permettant pas de rattraper le retard. Ce cas de figure est illustré par ce formateur à la retraite qui, malgré son incapacité à obtenir le certificat, visionne l'ensemble des vidéos d'un MOOC de télécommunication :

*J'avais essayé aussi celui sur les télécommunications, mais je suis arrivé en cours de route et il ne restait qu'une ou deux semaines donc en fait j'ai suivi les cours, j'ai regardé les vidéos, mais je n'ai pas passé les examens car les deadlines étaient passées. (E7)*

Ce cas de figure correspond davantage à une barrière situationnelle, dans la mesure où la non-obtention du certificat est liée aux circonstances particulières que rencontre l'apprenant, responsabilités familiales (Garland, 1992, p.124), ou, dans le cas présent, la date d'arrivée dans le cours. Les deux cas que nous venons de citer suggèrent que l'on ne peut interpréter uniquement en termes d'objectif d'apprentissage le fait de devenir un NCVA. Dans les paragraphes qui suivent, nous allons chercher à étendre la réflexion au cas des NCTA, pour mettre en lumière les spécificités de ce type de comportement.

#### **4.2.2. Interpréter le téléchargement**

Dans certains cas, la différence entre les deux comportements NCVA et NCTA se réduit au fait que l'un visionne les vidéos depuis la plateforme, et l'autre depuis un autre support, ordinateur personnel ou support mobile. Dans ce cas, les comportements sont relativement analogues, y compris si l'on se place du point de vue des barrières qui les déterminent. La question qui nous intéresse ici est le devenir des vidéos téléchargées, en particulier chez ceux qui ne les visionnent pas dans leur intégralité, bien qu'ils aient téléchargé la quasi-totalité des vidéos du dispositif.

Les entretiens suggèrent au moins deux types d'usages : d'une part un téléchargement suivi d'un visionnage qui intervient peu après, voire immédiatement après, notamment dans les transports en commun, et d'autre part un téléchargement pour un visionnage postérieur éventuel, selon une logique d'archivage. Le premier cas est illustré par cet entrepreneur, qui considère le visionnage de MOOC comme une activité qu'il réalise dans les transports en commun, en substitut d'activités comme la lecture :

*Parce que dans FUN par exemple pour télécharger les vidéos, c'est tout un truc quoi, il faut se préparer, il faut ouvrir le site. Sur Coursera avec l'appli, là j'ai deux cours que je regarde comme ça du coin de l'œil, toc je mets comme ça les vidéos, moi je prends la 13, la 14, et la 13 donc c'est toujours un peu long, bah du coup moi je perds pas de temps c'est l'alternative à la lecture. (E8)*



On peut supposer que ce cas de figure est notamment représenté parmi les participants qui téléchargent les vidéos depuis un appareil mobile. L'autre cas est le stockage dans une logique d'archivage, qui n'implique pas nécessairement ni un visionnage préalable, ni un visionnage postérieur. Ainsi, cette chargée de projet dans une association explique qu'elle ne consulte que rapidement les ressources d'un MOOC, mais qu'elle en archive généralement l'ensemble des contenus téléchargeables pour un éventuel usage ultérieur :

*Non, mais ce que j'ai trouvé génial, c'est le fait de pouvoir télécharger. Enfin voilà quoi. Sur chaque module, j'en ai pris ce que je pouvais dans une première lecture, mais je les ai, j'ai le truc, le jour où je m'y colle, j'ai ça sous le coude. (E12)*

Un participant de ce type pourrait être un NCTA qui a priori visionne peu, voire aucune des vidéos qu'il télécharge, en tout cas pour la durée du cours. À la lueur de cet entretien, nous avançons qu'il est possible qu'une partie des NCTA téléchargent des vidéos à des fins d'archivage, sans réaliser le moindre visionnage, avec comme seul objectif la possibilité de pouvoir accéder aux contenus si le cours venait à ne plus être accessible.

## **5. Discussion**

### **5.1. Le devenir des vidéos téléchargées, une inconnue majeure**

Nos résultats nous amènent à mettre en doute la position selon laquelle les usages des non-certifiés, notamment des non-certifiés visionneurs assidus, seraient importants sur le plan quantitatif, ce qui est la position défendue par (Koller *et al.*, 2013). Ces auteurs basent probablement leur affirmation sur la consommation de vidéos, qui prend en compte tant le visionnage que le téléchargement des vidéos. Nous ne disposons pas des données permettant de confirmer ou d'infirmier cette hypothèse, mais nous avons pu constater, dans les six cours analysés, que la participation des NCTA aux activités évaluées était minime. Certains entretiens suggèrent qu'il est possible que pour certains d'entre eux, il y ait un usage minimal ou inexistant des ressources pédagogiques une fois téléchargées. Des analyses supplémentaires seraient nécessaires pour trancher à ce stade, car l'analyse de traces d'activité montre rapidement ses limites à cet égard.

Dans l'impossibilité de détecter toute utilisation de la ressource postérieure au téléchargement, une approche par enquêtes s'avère selon

nous souhaitable pour appréhender le devenir des ressources. On pourrait en effet demander à toute personne téléchargeant les ressources pédagogiques du cours de répondre à des questions relatives à l'utilisation qu'elle envisage de faire des ressources téléchargées. Si une telle pratique peut être légitimement considérée comme antinomique avec l'esprit des MOOC, les questionnaires obligatoires pour accéder au cours ont représenté une pratique courante pour toute inscription sur Coursera. Un certain nombre de travaux ancrés dans le champ de l'Open Education ont adopté cette démarche. Des enquêtes ont ainsi été diffusées auprès d'individus téléchargeant des ressources de l'*Open University* sur *ItunesU* (Rosell-Aguilar, 2013), mais ces travaux ont jusqu'à présent été centrés sur la question des profils et des motivations des téléchargeurs plus que sur celle du devenir des ressources téléchargées. L'approfondissement de cette question nous semble incontournable pour une meilleure compréhension des usages qui sont faits des ressources du cours.

D'autres pistes peuvent être explorées pour identifier des barrières situationnelles susceptibles de conduire au fait de se cantonner au visionnage ou au téléchargement de vidéos. On peut supposer, sur la base d'un des entretiens, que le temps disponible pour les MOOC se situe pour certains participants sur le temps de trajet pour aller au travail : ils visionneraient alors les vidéos juste après les avoir téléchargées. Le fait de travailler dans les transports constitue une barrière situationnelle pour la réalisation des activités évaluées, dont on pourrait apprécier l'importance par des enquêtes. On pourra s'intéresser par ailleurs aux barrières situationnelles d'ordre technologique (Roberts, 2004). Il n'est pas impossible que certains participants n'aient la possibilité d'accéder au cours que depuis un téléphone mobile, faute d'un accès suffisant à un ordinateur, en particulier dans les pays du Sud. Le visionnage de vidéos depuis le téléphone, s'il ne suffit pas pour obtenir le certificat, permet à défaut de s'appropriier le cours. On pourrait ainsi multiplier les travaux sur l'impact des barrières situationnelles sur les usages qui sont fait du cours, pour nuancer l'interprétation selon laquelle les NCCA sont avant tout des cibles.

## **5.2. Importance du certificat dans la persistance des apprenants**

Le dernier point de notre discussion porte sur le désengagement rapide des participants qui se cantonnent au visionnage des vidéos pédagogiques. On peut imaginer qu'en ne participant pas aux activités

évaluées, ou de manière insuffisante pour obtenir le certificat, les apprenants se privent d'un des principaux mécanismes qui influent le plus significativement sur leur persistance. Il est possible qu'existe alors une forme d'effet circulaire, où le faible engagement initial et le désintérêt pour les activités évaluées conduisent à un affaiblissement de l'engagement. Ce résultat a déjà été observé dans les travaux sur les Open Coursewares, ces cours universitaires filmés diffusés gratuitement en ligne. Dans un travail d'enquête mené auprès de plus de sept cents individus, (Arendt et Shelton, 2009) constatent que le principal obstacle identifié par les utilisateurs d'Open Courseware réside dans l'impossibilité d'obtenir un certificat à l'issue de leur apprentissage.

Le relatif désintérêt pour les certificats a été démontré dans la plupart des enquêtes diffusées dans les MOOC (Cisel, 2016) et, plus généralement, corroboré par nombre de recherches sur l'éducation non formelle des adultes, des enquêtes séminales comme celle de (Johnstone et Rivera, 1965) aux travaux des sociologues des loisirs (Dumazedier, 1962), ainsi que par les travaux qui s'en inspirèrent. Néanmoins, le fait que le certificat ne constitue que rarement une motivation première pour s'inscrire n'exclut pas le fait qu'il puisse jouer un rôle majeur dans la persistance au sein du dispositif.

## **6. Conclusion**

Nous synthétisons les résultats de cet article avant d'évoquer les évolutions des plateformes de MOOC et leurs conséquences.

### **6.1. Synthèse des résultats**

Nous nous sommes intéressés au cours de cet article à la question de l'activité observable des non-certifiés, et plus particulièrement aux questions de visionnage et de téléchargement de vidéos. Nous avons montré sur la base des données de six cours que les non-certifiés étaient responsables d'une partie importante de cette activité, mais que le nombre d'actions diminuait au fil des modules. Nous avons vu en particulier que les NCVA étaient généralement largement minoritaires par rapport aux certifiés, contrairement aux NCTA. La part que les non-certifiés jouent dans l'activité observable est due à leur grand nombre, relativement au nombre de certifiés, qui compense le faible nombre moyen d'actions qu'ils réalisent individuellement au sein de la plateforme.

Lorsque l'on utilise des indicateurs de performance du MOOC tels que le nombre de vidéos vues, il faut prendre en considération le fait qu'une partie significative de cette statistique correspond à l'activité observable liée aux premières vidéos du dispositif. Nous avons dans la suite de ce travail cherché à interpréter le comportement des NCVA et NCTA au prisme des barrières de (Garland, 1992). Nous avons vu au travers d'entretiens que le choix de se cantonner aux vidéos pédagogiques n'était pas nécessairement dû à un désintérêt pour le certificat, ce qui nous amène à contester l'interprétation des NCVA comme relevant principalement de cibleurs. Des barrières situationnelles, comme une inscription tardive, peuvent peser dans la balance. Nous avons suggéré qu'il était possible qu'une partie importante des vidéos téléchargées ne soient pas visionnées.

## **6.2. Conséquences des mutations des plateformes de MOOC**

La question du comportement des non-certifiés est d'autant plus importante que des plateformes aussi centrales que *Coursera* sont en pleine mutation. La plateforme a subi un profond changement en 2016 (Koller, 2015). Les contenus des MOOC sont désormais accessibles en permanence, ce qui constitue un passage du format temporalisé au format atemporel ; en revanche, les activités notées et les certificats ne sont généralement plus disponibles qu'aux utilisateurs ayant payé pour le cours. Seules les vidéos pédagogiques sont pour le moment encore accessibles à tous les inscrits. Cette mutation constitue à certains égards un retour à un modèle traditionnel de formation à distance dont les ressources pédagogiques seraient librement accessibles, et dont l'Open University a constitué l'un des pionniers (Rosell-Aguilar, 2013). Le premier paramètre dont il faudrait étudier l'effet est celui de l'accessibilité permanente des vidéos, variable étudiée sur quelques cours par (Mullaney et Reich, 2014), dont on peut supposer qu'elle est susceptible d'affecter, à terme, les stratégies des participants.

Ensuite, le fait de devoir payer pour participer aux activités évaluées peut conduire un certain nombre d'utilisateurs, qui, dans d'autres conditions, auraient suivi le parcours menant au certificat gratuit, à faire d'autres choix. On peut supposer qu'une partie cessera sa participation aux MOOC, qu'une autre se « convertira » en NCVA. Si le taux de conversion est faible, ou en d'autres termes si les utilisateurs traditionnellement les plus investis mais ne souhaitant pas payer ne

deviennent pas des NCVA, alors il est vraisemblable que l'on assiste à une chute globale du nombre d'utilisateurs de MOOC à l'échelle planétaire, ce qui pose la question suivante: en diminuant le nombre de services gratuits, limiterait-on les usages faits des MOOC, ou contribuerait-on à développer de nouveaux usages? La disparition des certificats gratuits, qui semble concerner un nombre croissant de plateformes, doit nous inciter à étudier la traduction des changements de modèles économiques dans les usages des apprenants. En effet, il est probable que cette mutation sonne le glas de ce qui faisait l'une des spécificités des dispositifs MOOC, à savoir leur caractère « massif » et « ouvert ».

## **REFERENCES**

- ARENDRT, A. M. et SHELTON, B. E. (2009). Incentives and Disincentives for the Use of OpenCourseWare. *The International Review of Research in Open and Distributed Learning*, 10(5).
- BARAK, M., WATTED, A. et HAICK, H. (2016). Motivation to learn in massive open online courses: Examining aspects of language and social engagement. *Computers and Education*, 94, 49-60.
- BEAUD, S. et WEBER, F. (2010). *Guide de l'enquête de terrain*. Paris, France: Armand Colin.
- BLANCHET, A. et GOTMAN, A. (2010). *L'entretien: L'enquête et ses méthodes* (2e éd.). Paris, France: Armand Colin.
- CHAMPAIGN, J., COLVIN, K. F., LIU, A., FREDERICKS, C., SEATON, D. et PRITCHARD, D. E. (2014). Correlating Skill and Improvement in 2 MOOCs with a Student's Time on Tasks. Dans *Proceedings of the First ACM Conference on Learning @ Scale* (p.11-20). New York.
- CISEL, M. (2014). Analyzing Completion Rates in the First French xMOOC. *eLearning papers*, 37, 77-85.
- CISEL, M., MANO, M., BACHELET, R. et SILBERZAHN, P. (2015). A Tale of Two MOOCs: Analyzing Long-Term Course Dynamics. Dans *Proceedings of EMOOCs Conference 2015* (p. 191-199).
- CISEL, M. Utilisations des MOOC: éléments de typologie (Manuscrit de thèse non publié). Université Paris-Saclay, France. Disponible à l'adresse: <http://www.theses.fr/2016SACLN024> (consulté le 9 juillet 2016)
- COLVIN, K. F., CHAMPAIGN, J., LIU, A., ZHOU, Q., FREDERICKS, C. et PRITCHARD, D. E. (2014). Learning in an introductory physics MOOC: All cohorts learn equally, including an on-campus class. *The International Review of Research in Open and Distributed Learning*, 15(4).
- CRESWELL, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (3e éd.). SAGE Publications.
- CRESWELL, J. W. et CLARK, V. L. P. (2006). *Designing and Conducting Mixed Methods Research*. SAGE Publications.
- CROSS, K. P. (1981). *Adults as learners. Increasing participation and facilitating learning*. San Francisco, CA: Jossey-Bass.

DANIEL, J. (2012). Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility. *Journal of Interactive Media in Education*, 2012(3)

DUMAZEDIER, J. (1962). *Vers une civilisation du loisir ?* Paris, France : Seuil.

GARLAND, M. (1992). *Variables affecting persistence in distance education in the natural resource sciences*. Thèse de doctorat non publiée, University of British Columbia, Canada.

GARLAND, M. R. (1993). Student perceptions of the situational, institutional, dispositional and epistemological barriers to persistence. *Distance Education*, 14(2), 181-198.

HO, A. D., REICH, J., NESTERKO, S. O., SEATON, D. T., MULLANEY, T., WALDO, J. et CHUANG, I. (2014). *HarvardX and MITx: The First Year of Open Online Courses, Fall 2012-Summer 2013*. Rochester, NY : Social Science Research Network.

HO, A. D., CHUANG, I., REICH, J., COLEMAN, C. A., WHITEHILL, J., NORTHCUTT, C. G. et PETERSEN, R. (2015). *HarvardX and MITx: Two Years of Open Online Courses Fall 2012-Summer 2014*. Rochester, NY : Social Science Research Network.

IVANKOVA, N. V., CRESWELL, J. W. et STICK, S. L. (2006). Using Mixed-Methods Sequential Explanatory Design: From Theory to Practice. *Field Methods*, 18(1), 3-20.

JORDAN, K. (2014). Initial trends in enrollment and completion of massive open online courses. *The International Review of Research in Open and Distributed Learning*, 15(1).

JORDAN, K. (2015). Massive open online course completion rates revisited: Assessment, length and attrition. *The International Review of Research in Open and Distributed Learning*, 16(3).

KEMPFER, H. (1996). How to Reduce Dropouts in Distance Education. EUNED.

KIZILCEC, R. F., PIECH, C. et SCHNEIDER, E. (2013). Deconstructing Disengagement: Analyzing Learner Subpopulations in Massive Open Online Courses. Dans *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, LAK 2013. Leuven, Belgium.

KIZILCEC, R. F. et SCHNEIDER, E. (2015). Motivation as a Lens to Understand Online Learners. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 22(2).

KOLLER, D., NG, A., DO, C. et CHEN, Z. (2013). Retention and intention in massive open online courses. *EDUCAUSE Review*. Disponible sur Internet : <https://er.educause.edu/articles/2013/6/retention-and-intention-in-massive-open-online-courses> (consulté le 9 juillet 2016)

KOLLER, D. (2015) An update on assessments, grades, and certification. Disponible sur Internet : <http://Coursera.tumblr.com/post/131520811622/an-update-on-assessments-grades-and> (consulté le 9 juillet 2016)

MERTENS, D. M. (2003). Mixed methods and the politics of human research: The transformative-emancipatory perspective. Dans A. Tashakkori et C. Teddlie (dir.), *Handbook of mixed methods in the social & behavioral research* (p. 135-164). Thousand Oaks, CA: Sage.

MULLANEY, T. et REICH, J. (2014). Staggered versus All-at-Once Content Release in Massive Open Online Courses: Evaluating a Natural Experiment (SSRN Scholarly

Paper No. ID 2499729). Rochester, NY: Social Science Research Network. Récupéré à <http://papers.ssrn.com/abstract=2499729>

ONWUEGBUZIE, A. J. et JOHNSON, R. B. (2006). The validity issue in mixed research. *Research in the Schools*, 13(1), 48-63.

ORTON, L. J. (1977). Completion and non-start ratios in correspondence courses. *Canadian Journal of University Continuing Education*, 4, 21-26

REICH, J. (2014). MOOC Completion and Retention in the Context of Student Intent. *EDUCAUSE Review*. Disponible sur Internet : <http://er.educause.edu/articles/2014/12/mooc-completion-and-retention-in-the-context-of-student-intent> (consulté le 9 juillet 2016)

ROBERTS, L. E. (2004). Not now, maybe later, and often not at all: situational, institutional, dispositional, epistemological, and technological barriers to business-based online training courses. Manuscrit de thèse non publié, North Carolina State University.

ROSELL-AGUILAR, F. (2013). Delivering unprecedented access to learning through podcasting as OER, but who's listening? A profile of the external iTunes U user. *Computers & Education*, 67, 121-129.

TERENZINI, P. (1987). Studying student attrition and retention. Dans G. W. McLaughlin and J. A. Muffo (dir.), *A primer of institutional research* (p. 20-35). Tallahassee, FL : Association of Institutional Research.

TINTO, V. (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research*, 45(1), 89-125.

TINTO, V. (2006). Research and Practice of Student Retention: What Next? *Journal of College Student Retention: Research, Theory and Practice*, 8(1), 1-19.

TOUGH, A. M. (1971). *The adult's learning projects: A fresh approach to theory and practice in adult learning*. Ontario Institute for Studies in education.

VAUGHAN, R. P. (1968). College Dropouts: Dismissed vs. Withdrew. *The Personnel and Guidance Journal*, 46(7), 685-689.

ZAWACKI-RICHTER, O. et ANDERSON, T. (2014). *Online Distance Education: Towards a Research Agenda*. Athabasca University Press.

**ANNEXE : Statistiques des six MOOC de Coursera utilisés dans cette étude**

# inscr. : Nombre d'inscrits. # cert. : Nombre de certifiés.

% : pourcentage des inscrits. cert. : inscrits ayant obtenu un certificat, avec ou sans distinction. conn. : inscrits s'étant connectés au moins une fois au cours. vid. : inscrits ayant visionné ou téléchargé au moins une vidéo. dev. : inscrits ayant rendu au moins un devoir.

Code Cours	Titre du cours	# inscr.	#cert.	% cert.	% conn.	% vid.	% dev.
ALG1	Conception et mise en œuvre d'algorithmes	8950	246	2,7	62,3	49,7	6,9
ALG2	Conception et mise en œuvre d'algorithmes	6341	133	2,1	58,0	48,7	11,4
FA1	An Introduction to Functionnal Analysis	25354	382	1,5	57,5	55,8	5,8
FA2	An Introduction to Functionnal Analysis	16691	405	2,4	67,5	55,1	7,3
BP1	Business and Philosophy : what managers	72633	2709	3,7	50,9	43,5	7,0
AV1	Discrete inference And Learning in Artificial Vision	19461	889	4,6	69,5	49,4	NA