

Le choix d'un MOOC, un processus influencé par l'organisation des cours en catalogues ?

Matthieu Cisel

▶ **To cite this version:**

Matthieu Cisel. Le choix d'un MOOC, un processus influencé par l'organisation des cours en catalogues ?. STICEF (Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation), 2019, 10.23709/sticef.25.1.8 . hal-02523650

HAL Id: hal-02523650
https://hal.science/hal-02523650

Submitted on 31 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Le choix d'un MOOC, un processus influencé par l'organisation des cours en catalogues ?

▶ **Matthieu CISEL** (Laboratoire EDA, Université Paris-Descartes)

■ **RÉSUMÉ** • Dans cet article, nous analysons les données d'inscription issues de la plateforme FUN afin d'interpréter les comportements observés au sein des cours qu'elle héberge. Pour étayer nos propos, nous mobilisons, d'une part, une enquête au prisme de la notion de clé d'entrée issue de la littérature sur le e-commerce et, d'autre part, un million d'inscriptions réalisées sur FUN. Selon notre interprétation des données, il existe une forte représentation d'inscriptions issues de visites expérientielles, qui correspondent à un mode de navigation où l'utilisateur parcourt un catalogue sans nécessairement avoir une idée précise du type de MOOC qu'il souhaite suivre. Ces analyses montrent que les comportements observés au sein d'une formation en ligne incluse dans un catalogue peuvent être influencés par les caractéristiques de celui-ci.

■ **MOTS-CLÉS** • CLOM, usages, analyse de traces

■ ***ABSTRACT*** • *In this paper, we analyze registration data from the French MOOC platform France Université Numérique to interpret behaviors observed in individual courses, based on notions drawn from the webmarketing literature, such as the type of visit. Our interpretations are based on both surveys sent in a dozen of courses, and on the quantitative study of more than a million registrations from FUN. We explore various dimensions of users' behavior. We show that learners discover the courses they register to mostly by navigating on the platform catalog, a particular case called the experiential visit. Our results show that behaviors observed in an online course included in a catalogue can be influenced by the characteristics of the said catalogue.*

■ ***KEYWORDS*** • *MOOC, uses, learning analytics*

1. Introduction

Entre 2012 et 2016, des centaines d'établissements d'enseignement supérieur investissent à travers le monde dans la conception de MOOC ou Massive Open Online Courses (Cisel et Bruillard, 2013 ; Daniel, 2012). Parmi les nombreux débats qui virent le jour durant cette période, l'un d'entre eux prit une place considérable, celui des faibles taux de certification des MOOC (Cisel, 2015). Les taux de certification des formations à distance traditionnelles[1] sont en effet plus bas que ceux de leurs pendants en présentiel (Zawacki-Richter et Anderson, 2014). Cette tendance est plus nette encore pour les MOOC. Sur les célèbres plateformes américaines *Coursera* et *edX*, seuls 5 à 10 % des inscrits obtiendraient le certificat (Chuang et Ho, 2016). La propension des utilisateurs à s'inscrire à plus de cours qu'ils ne peuvent en suivre, conséquence de la gratuité selon Koller (2013), constitue l'une des explications avancées de manière récurrente.

Cette considération nous a amené à nous poser une succession de questions auxquelles nous avons cherché à répondre sur la base d'enquêtes et des données d'inscription de la plateforme française *France Université Numérique (FUN).* Nous qualifions de « données d'inscription » l'ensemble des données permettant de qualifier les inscriptions réalisées au sein d'une même plateforme : date de l'inscription, obtention de l'éventuel certificat suite à cette inscription, etc. La première série de questions porte sur les modalités de découverte du cours et leur influence sur le comportement d'inscription. Comment les utilisateurs de MOOC prennent-ils connaissance des cours auxquels ils s'inscrivent ? Quelle est leur propension à s'inscrire à de multiples MOOC, et à s'investir de manière simultanée au sein de plusieurs de ces cours ? Comment la modalité de découverte d'un cours influe-t-elle sur le nombre de cours au sein desquels un utilisateur s'investit ?

La deuxième série de questions porte davantage sur le comportement d'inscription au sein de *FUN* et fait écho aux propos de Koller (2013). Quel est le lien entre nombre d'inscriptions réalisées dans une plateforme donnée et investissement dans les MOOC correspondant, l'investissement étant mesuré sur la base de l'obtention des certificats ? Dans quelle mesure les comportements d'inscription peuvent-ils nous donner des indices

1 Par « traditionnelles », nous entendons des formations à distance payantes et intégrées dans un diplôme ou une certification reconnue sur le marché du travail.

quant à la manière dont l'apprenant a pris connaissance du MOOC ? Enfin, comment les fluctuations des caractéristiques du catalogue d'une plateforme peuvent-elles influer sur la manière dont les utilisateurs s'investissent dans chaque cours pris de manière individuelle ?

Ce travail est réalisé sous l'angle du « comportement d'inscription », terme qui désigne tout comportement observable dont la description se base sur des données d'inscription : nombre d'inscriptions réalisées sur la plateforme, rythme des inscriptions, etc. Sur le plan notionnel, nous avons cherché à qualifier et à quantifier les trajectoires qui conduisent à l'inscription au prisme des notions de clé d'entrée et de visite, toutes deux issues de la littérature sur les processus de sélection de produits en ligne (Isaac et Volle, 2008 ; Stenger et Bourliataux-Lajoinie, 2014).

Nous reprendrons la typologie de visites qu'ils proposent, en distinguant notamment visites exploratoires, expérientielles et expéditives ; nous reviendrons sur leur définition. Les données d'inscription vont nous amener à proposer l'hypothèse selon laquelle il existe une forte représentation d'inscriptions issues de visites expérientielles, qui correspondent à un mode de navigation où l'utilisateur parcourt un catalogue sans nécessairement avoir, en amont, une idée précise du type de MOOC qu'il souhaite suivre. Cette hypothèse découle notamment d'une forte représentation dans le jeu de données de ce que nous appellerons les inscriptions groupées, qui correspondent au fait que les participants réalisent des séries d'inscription dans un intervalle de temps réduit, généralement inférieur à une heure.

Si l'on s'attache à une définition stricte de la notion de visite expérientielle, il faudrait des données supplémentaires pour pouvoir identifier à quel type de visite correspond une inscription en particulier : données de navigation précédant l'inscription, données déclaratives sur les intentions des participants au cours de ce processus de choix du cours. Il est en pratique presque impossible d'obtenir toutes ces données, et à plus forte raison de les croiser par le biais d'identifiants uniques. Nous ne proposons pas ici une équivalence entre visite expérientielle et inscription groupée, d'autant que des stratégies mixtes peuvent exister : la première inscription d'une série d'inscriptions groupées peut être issue d'une visite exploratoire tandis que les suivantes peuvent être issues d'une visite expérientielle. Notre hypothèse postule simplement l'existence d'une corrélation entre un type d'inscription et un type de visite.

En conclusion de cet article, nous proposerons une nouvelle hypothèse à la suite des différentes analyses que nous aurons présentées tout au long de l'article : les caractéristiques des catalogues de cours en ligne (comme le nombre et la nature des cours proposés), payants ou non, influent de manière importante sur les comportements observables au sein des cours qu'ils hébergent. Il peut dès lors être intéressant de se pencher sur ces catalogues si l'on veut prendre pleinement la mesure des dynamiques observées au sein de dispositifs individuels, mais aussi sur les évolutions successives d'un dispositif au fil du temps. Dans cette contribution, nous écartons la question des apprentissages réalisés par les utilisateurs, pour nous focaliser sur celle de leur engagement au sein des dispositifs.

2. Des travaux sur les MOOC aux recherches sur l'attrition en formation

2.1. Des comportements d'inscription peu étudiés

La rareté des études portant sur le comportement d'inscription s'explique sans doute par le caractère stratégique des données d'inscription, de sorte que les premiers travaux sur la question ne sont pas l'œuvre de chercheurs, mais des plateformes elles-mêmes. Ainsi la fondatrice de *Coursera* mobilise ces données avant tout pour minimiser la portée des faibles taux de certification des MOOC (Koller, 2013), c'est-à-dire pour répondre à la critique récurrente selon laquelle la faible rétention des apprenants signifie que les MOOC « ne marchent pas ». Bien que les quelques chiffres avancés par Koller suggèrent que le comportement d'inscription constitue l'une des clés de la compréhension des taux de certification, les statistiques fournies sont maigres. Elles portent sur des millions d'individus et des centaines de cours, mais elles manquent de précision : tout au plus sait-on que les utilisateurs sont inscrits à environ quatre cours en moyenne.

À notre connaissance, le lien entre comportement d'inscription et taux de certification n'a été étudié que dans un nombre limité de publications (Albo, Hernández-Leo et Oliver, 2016 ; Anderson, Huttenlocher, Kleinberg et Leskovec, 2014 ; Banerjee et Duflo, 2014 ; Perna *et al.,* 2016), et le plus souvent sous l'angle du délai entre la date d'inscription et la date de lancement du cours. Plusieurs de ces auteurs cherchent à identifier quelles sont les inscriptions qui sont les plus susceptibles de déboucher sur un investissement des participants. Ils montrent que celles qui sont réalisées plusieurs mois en avance ont nettement moins de chances

d'aboutir à l'obtention du certificat que celles qui sont réalisées dans les jours qui précèdent ou suivent le lancement du cours.

Quant au problème de l'inscription à de multiples cours – par là nous entendons le fait d'avoir réalisé plus d'une inscription sur une plateforme donnée – elle reste largement sous-investie. Si certains rapports (Chuang et Ho, 2016 ; Université d'Edimbourg, 2013) évoquent la question, ils se cantonnent aux MOOC d'un nombre réduit d'établissements, sans que l'analyse ne prenne en compte l'ensemble des cours hébergés par la plateforme attenante. On ne peut ainsi voir qu'une partie réduite du comportement d'inscription. Au-delà de ce problème, les analyses ne se réclament pas d'une quelconque problématique. Il s'agit tantôt de fournir des statistiques descriptives – nombre de participants inscrits à un, deux, trois cours de l'établissement, tantôt de décrire les trajectoires des participants d'une discipline à l'autre, des cours d'informatique aux cours de littérature par exemple.

À notre connaissance, l'analyse des données de *MiriadaX* (Albo *et al.*, 2016) constitue l'une des seules recherches publiées dans la littérature scientifique portant sur les comportements d'inscription à l'échelle de l'ensemble du catalogue de cours d'une plateforme. L'étude porte sur près de 200.000 participants répartis dans 144 cours. Néanmoins, les auteurs se cantonnent à établir des corrélations entre le nombre d'inscriptions par cours et des variables sociodémographiques comme le niveau d'études ou le pays de résidence. Ils cherchent à montrer qu'il n'y a pas de différence majeure dans le comportement d'inscription entre différents pays hispanophones. Nous apprenons par exemple que les argentins s'inscrivent approximativement à autant de cours que les espagnols. S'ils se basent sur les données d'inscription, les auteurs ne cherchent pas à établir de lien entre comportement d'inscription et taux de certification.

2.2. Les travaux sur l'attrition en formation

Avant de présenter les notions que nous avons mobilisées dans cet article, nous proposons une brève revue des travaux sur l'attrition en formation, qui ont largement inspiré les analyses présentées ici, et sur les taux de certification dans les MOOC. Les MOOC constituant un cas particulier de formation à distance, l'attrition que l'on y observe est parfois analysée au prisme des travaux sur l'attrition en formation à distance (Cisel, 2016). Ces travaux prennent leur origine dans les recherches sur l'attrition, ou son contrepoint, la rétention, observée dans les universités nord-américaines (Tinto, 2006). L'attrition est alors considérée au sein

d'une institution donnée et correspond à la part des étudiants inscrits qui n'obtiennent pas le diplôme correspondant (Tinto, 1975). Une vaste littérature s'est développée sur le sujet à partir des années 1970. Deux décennies de recherches sur l'attrition avaient permis d'en distinguer plusieurs formes, c'est-à-dire plusieurs causes immédiates d'une interruption des études : retrait volontaire lorsque l'interruption de la formation correspond à une décision positive et échec académique lorsque cette décision est subie, notamment du fait de mauvaises notes (Vaughan, 1968). Cette première typologie est complétée par des éléments comme le *retour*, ou le *transfert* (Hackman et Dysinger, 1970), pour prendre en compte le fait que les étudiants peuvent se réinscrire à l'établissement initial ou s'inscrire dans un autre établissement.

Si ces quelques notions ont une valeur descriptive certaine, elles montrent leurs limites dès lors que l'on s'intéresse aux mécanismes qui ont conduit au retrait volontaire ou à l'échec académique. C'est dans ce contexte qu'est apparu le modèle de rétention de Tinto (1975), qui constitue le point de départ d'une série de travaux dans ce domaine (Munro, 1981 ; Tinto, 1982, 1987). Il prend notamment en compte l'intégration sociale des étudiants, c'est-à-dire les relations qu'ils établissent avec leurs pairs, mais aussi l'intégration académique, qui se focalise, elle, sur les relations avec le corps enseignant et le rapport aux disciplines étudiées. Ce type de modèle connaît de multiples transformations afin de l'adapter aux caractéristiques de la formation à distance (Bean et Metzner, 1985 ; Boshier, 1973 ; Garrison, 1985 : Kember, 1989 ; Kemp, 2002 ; Sweet, 1986). Ces modèles visent à prendre en compte le fait que l'on étudie bien souvent des cours isolés dont les caractéristiques particulières pèsent de manière considérable dans la décision d'interrompre la formation. Par ailleurs, le public est souvent constitué d'apprenants adultes engagés dans la vie professionnelle.

Une grande partie des travaux réalisés sur l'attrition dans les MOOC se base, explicitement ou implicitement, sur des méthodes ou des concepts qui ont été développés dans cette dernière littérature (Barak, Watted et Haick, 2016 ; Breslow, Pritchard, DeBoer, Stump, Ho et Seaton, 2013 ; Gillani et Eynon, 2014 ; Greene, Oswald et Pomerantz, 2015 ; Kizilcec et Schneider, 2015 ; Rosé *et al.,* 2014). Les auteurs reprennent ainsi souvent des données issues de questionnaires (motivations sous-jacentes à l'inscription, intentions, études passées) qu'ils croisent avec les données relatives à l'attrition. Qu'ils soient focalisés sur les universités, la formation à distance ou les MOOC, ces travaux ont en commun le fait de se

concentrer sur ce qui se passe une fois l'inscription réalisée. Or c'est à ce qui se passe en amont de l'inscription que nous nous intéressons dans le cadre de cet article.

Une remarque faite par Tinto (1975) s'est révélée inspirante à cet égard. L'auteur souligne l'importance de raisonner à l'échelle d'un système éducatif dans son ensemble et non à l'échelle d'un cours ou d'une université. C'est à la suite de cette injonction que nous avons cherché à aborder la question des utilisations des MOOC en raisonnant à l'échelle d'une plateforme. L'analogie a ses limites (après tout, un cours en ligne n'est pas une université, pas plus qu'une plateforme de cours ne représente un système éducatif), mais nous empruntons volontiers à Tinto l'idée selon laquelle on ne saurait appréhender pleinement les dynamiques d'un dispositif qui s'insère dans un réseau d'autres dispositifs, sans se pencher sur les dynamiques qui animent ce réseau.

Dans le cas d'un MOOC, cela signifie chercher à appréhender comment les participants parcourent les catalogues, passent éventuellement d'un cours à l'autre (comme on peut se désinscrire d'une unité d'enseignement pour s'inscrire à une autre, dans une université) et quelles sont les trajectoires qui ont conduit à l'inscription au cours. En d'autres termes, nous nous penchons également sur le parcours de navigation entre les pages Web, qui a précédé l'inscription au MOOC. Ce changement de perspective impose un renouvellement des cadres conceptuels guidant les analyses, ce qui nous a amené à nous pencher sur des notions issues de la littérature consacrée au processus de sélection de produits en ligne.

3. Des notions pour appréhender les modalités de sélection d'un produit sur Internet

3.1. La notion de clé d'entrée

Les notions que nous mobilisons dans cette contribution ont été développées afin de mieux comprendre les comportements d'achat sur Internet, et plus particulièrement les modalités de sélection de toutes sortes de produits, payants ou non, au sein de catalogues. À cet égard, nous nous situons, avec les MOOC organisés en catalogues, dans le domaine de validité de ces notions. Le fait de considérer un cours en ligne comme un produit comme un autre peut questionner le chercheur ou le praticien. Certes, nous avons réalisé ce travail sur des cours gratuits, aux certificats gratuits — nous avons analysé des processus de sélection de produits et non des processus d'achat à proprement parler. Mais cette logique de la

gratuité des MOOC a régressé partout dans le monde au cours de ces dernières années (Cisel, 2016), et la question de l'achat n'a plus grand-chose de métaphorique, au moins outre-Atlantique.

Plusieurs notions ont été mobilisées au cours de cette recherche, notamment « la clé d'entrée ». Elle désigne le processus qui a conduit à la sélection d'un produit, ce qui correspond dans notre cas à l'inscription au cours. Deux modalités y sont associées, la « clé d'entrée plateforme » d'une part, et la « clé d'entrée produit » d'autre part : « Soit l'internaute se rend sur un site avec une vague idée des produits qu'il souhaite acheter, soit il souhaite acheter un ou plusieurs produits sans avoir nécessairement une idée très précise des sites marchands où il peut les acheter. Dans le premier cas, la clé d'entrée est le site marchand à visiter, alors que dans le deuxième cas, la clé d'entrée est le produit à acheter » (Isaac et Volle, 2008).

Les auteurs font l'analogie entre la « clef d'entrée plateforme » et le « lèche-écran ». Un exemple de « clef d'entrée plateforme » est représenté par le système de recommandation d'*Amazon*, qui, sur la base des données récoltées sur des comportements antérieurs, pousse les utilisateurs connectés à sélectionner des produits dont ils ignoraient l'existence. Les auteurs croisent ces deux notions avec le degré de connaissance des platesformes et des produits. Soit l'internaute connaît bien l'offre (les sites, les catégories de produits), soit il la connaît peu. Le croisement de ces deux niveaux avec la clé d'entrée génère quatre types de visites que nous allons définir par la suite : expérientielle et expéditive pour la clef d'entrée plateforme, exploratoire et comparative pour la clef d'entrée produit.

3.2. Quatre types de visites

Dans le cas de la « visite expérientielle », « les motivations de visite sont plutôt hédonistes (découvrir une offre originale, se faire surprendre, se promener dans un vaste assortiment, comme une caverne d'Ali Baba, etc.). Les visites régulières sur PriceMinister ou sur eBay rentrent dans cette catégorie d'achat expérientiel » (Isaac et Volle, 2008, p. 213). Ce cas correspond notamment aux participants qui, après avoir découvert une plateforme y reviennent régulièrement pour vérifier si de nouveaux cours sont susceptibles de les intéresser, sans avoir une idée précise de la thématique sur laquelle ils désirent se former.

Dans le cas de la « visite expéditive », « l'internaute se rend sur un site pour réaliser un achat assez précis. Soit il connaît l'offre de ce site pour ce produit (il a vu une offre sur une publicité ou un ami lui en a parlé, par

exemple) ; soit il sait [...] que ce site est le meilleur pour telle catégorie de produit. [...] À la différence du précédent, ce processus de décision est guidé par un objectif précis : les motivations de visite sont plutôt utilitaires » (Isaac et Volle, 2008). Cela correspond dans notre cas aux personnes qui vont s'inscrire à un MOOC après en avoir entendu parler sur Internet, mais hors de la plateforme, ou après y avoir été incitées par un tiers. Ils ne se connectent à la plateforme que pour s'inscrire à un cours précis identifié auparavant.

Dans le cas de la « visite exploratoire », « l'internaute souhaite réaliser un achat dans une catégorie précise, mais ne connaît pas les sites à visiter. Il choisira probablement de surfer sur des sites connus dans le commerce physique, ou bien de saisir le nom de la catégorie de produits sur un moteur de recherche ou sur un comparateur » (Isaac et Volle, 2008). Dans ce cas, le participant sait de manière relativement précise dans quel domaine il souhaite se former. En revanche, il ne sait pas nécessairement quel site utiliser, ni même quel format d'apprentissage adopter (MOOC, tutoriels, etc.). Une schématisation de deux types de visite (exploratoire et expérientielle) conduisant à une inscription est présentée en figure 1.

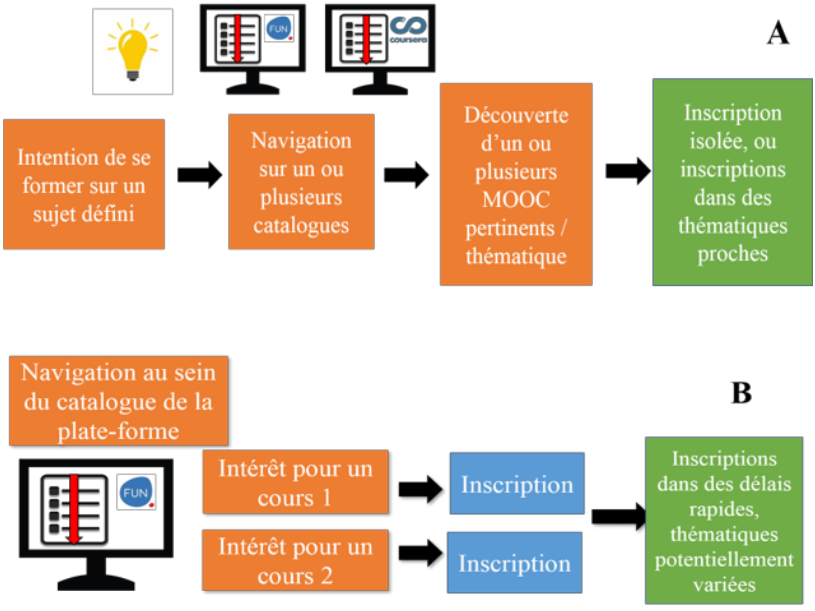


Figure 1 • Schématisation de 2 types de visite menant à une inscription : exploratoire (A), expérientielle (B)

Enfin, dans le cas de la « visite comparative ou évaluative », « l'internaute souhaite réaliser un achat précis et connaît les sites qui commercialisent la catégorie de produits recherchée. [...] L'internaute va donc comparer les différents sites qui font partie de son ensemble de considération. Ce processus de décision est guidé par un objectif précis » (Isaac et Volle, 2008). Pour satisfaire à la définition de la visite comparative, il faudrait que l'utilisateur cherche à se former sur un sujet précis via un MOOC et compare les pages de présentation de ces différents cours portant sur une thématique similaire. Par ailleurs, il faudrait que la comparaison entre ces différents cours puisse déboucher sur la possibilité de s'inscrire à au moins une des formations. Cela impliquerait qu'elles soient organisées sur des périodes similaires.

Pour que l'on puisse parler de visite comparative, nous avons considéré que la similitude entre les cours devait être très élevée. Les cours doivent être dans la même langue, gratuits, avec une date de début et une date de fin séparées de quelques semaines. L'analyse de la structure de l'écosystème MOOC au moment de la réalisation de cet article (Cisel, 2016) montre clairement que FUN constitue l'acteur majoritaire des MOOC francophones et que, dès lors, c'est au sein même de son catalogue que des visites comparatives sont le plus susceptibles d'avoir lieu. Or sur FUN, le nombre de cours ouverts aux inscriptions sur une période donnée est généralement inférieur à une vingtaine (Cisel et Leh, 2017), et les thématiques de ces formations divergent suffisamment pour que l'on puisse exclure la visite comparative. Le caractère délimité dans le temps des MOOC fait qu'il y a en pratique, à un moment donné, peu de compétition entre cours portant sur les mêmes thématiques (Cisel, 2016). Dès lors, nous avons fait le choix d'écarter les visites comparatives de cette étude, pour nous concentrer sur les visites expérientielles, exploratoires et expéditives.

4. Méthodologie

4.1. Une enquête diffusée dans une douzaine de cours

L'enquête a été réalisée à l'automne 2015 auprès de douze MOOC organisés sur FUN et d'un MOOC organisé sur Canvas.net. Elle se base sur la collecte de données recueillies via des questionnaires diffusés par nos soins, ou par des partenaires au sein de leur formation. Elle a permis de

collecter 8545 réponses[2]. Les questionnaires ont été élaborés sur la base des notions que nous venons de présenter et de données obtenues par le biais de quarante entretiens semi-directifs réalisés dans une précédente étude. L'analyse de ces entretiens nous a permis d'orienter la formulation des questions en légitimant la pertinence des catégories de visites issues de la littérature - voir (Cisel, 2016) pour une présentation de ces entretiens. Nous qualifierons de « répondant » toute personne ayant répondu à un questionnaire.

Une partie des questions de l'enquête, notamment celles relatives aux catégories socioprofessionnelles des participants, est issue des nomenclatures de l'INSEE. L'inscription à de multiples cours et le suivi simultané de différents cours est également traitée dans ce questionnaire. Ces questions s'inspirent partiellement d'une traduction du questionnaire diffusé par la plateforme américaine *Coursera*. La mise au point des questions relatives aux comportements d'inscription a fait suite à l'analyse des entretiens (Cisel, 2016) et à celle des notions de clé d'entrée (cf. section 2.1). La formulation exacte des questions sera présentée dans la partie consacrée aux résultats.

En ce qui concerne le type de visite, nous avons cherché à identifier deux paramètres : le média à travers lequel le participant a entendu parler du cours pour la première fois, que ce média soit une plateforme d'hébergement, un mail, un article de blog ou tout simplement une autre personne, et le processus qui a conduit à l'inscription. Nous avons notamment cherché à déterminer dans quelle mesure l'intention de se former à la thématique du cours précédait ou non la découverte du MOOC sur la plateforme. Dans le reste de l'enquête, nous posons des questions relatives à l'inscription à d'autres plates-formes de MOOC. En effet, un certain nombre de participants sont inscrits en plus de FUN à d'autres plates-formes de cours comme *Coursera* ou *edX,* et en nous cantonnant à la seule plateforme française, accessibilité des données oblige, nous ne voyons qu'une partie du comportement des utilisateurs vis-à-vis des MOOC et d'autres types de formations auxquelles ils peuvent être inscrits. L'analyse de l'enquête quantitative permet d'apprécier partiellement l'importance de ce biais.

2 Le détail de ces cours, ainsi que le nombre de réponses obtenues par cours est disponible dans (Cisel, 2016)

Nous avons utilisé *Google Form* au cours de cette enquête, cet outil permettant de réaliser des questions « à tiroir », c'est-à-dire que certaines réponses appellent des questions spécifiques. Il présente l'avantage de pouvoir être intégré directement au sein du cours par la procédure dite de l'*embed*, cette intégration augmentant considérablement les taux de réponse, par rapport à une stratégie qui consisterait à diffuser un lien vers un outil extérieur à la plateforme, comme *Limesurvey*. Tous les questionnaires étaient diffusés au cours de la première semaine de la formation, sans que nous ne puissions maîtriser le jour de la semaine où la diffusion aurait lieu, ce choix étant à la discrétion des équipes pédagogiques partenaires. Dans la mesure où il y a souvent une chute de l'engagement des participants au cours de la première semaine du MOOC (Cisel, 2016), le choix du jour peut affecter légèrement le nombre de réponses obtenues. Néanmoins, rien ne laisse à penser que la répartition des différentes réponses puisse être affectée par la décision des équipes.

L'approche par enquête présente l'avantage d'autoriser l'accès à des éléments du comportement des participants qui ne laissent pas de traces d'interaction, et d'articuler aisément notions mobilisées et collecte de données. Elle souffre néanmoins du biais d'auto-sélection inhérent à toute enquête en ligne ; aussi avons-nous cherché à compléter cette approche par l'analyse d'un jeu de données volumineux dont l'accès nous a été permis grâce à une collaboration étroite avec FUN.

4.2. Données d'inscription de la plateforme FUN

Les données d'inscription de FUN ont été extraites par l'équipe technique de la plateforme au 20/09/2015, et portent dès lors sur deux années d'inscriptions ; elles concernent 168 cours. Ces données comportent pour chaque inscription une date, l'identifiant du cours, l'identifiant anonymisé de l'utilisateur et l'obtention éventuelle du certificat associé au cours. Rappelons que l'inscription aux cours est gratuite et que pendant la période que couvre l'analyse, l'obtention des certificats sur FUN, qui prennent généralement le nom d'attestations de réussite, est elle-même entièrement gratuite.

Le corpus comprend au total 1.047.445 inscriptions. On constate que 308.396 d'entre elles correspondent à des cours qui n'ont pas commencé ou qui ne sont pas terminés au moment de l'extraction des données. Elles ont donc été exclues de toute réflexion sur les taux de certification, le corpus étant réduit à 119 cours.

À partir des données d'inscription brutes, nous avons bâti un certain nombre d'indicateurs visant à décrire chaque inscription. Nous qualifierons une inscription de « certifiante » dès lors qu'elle débouche sur l'obtention d'un certificat, et de « non certifiante » dans le cas contraire. Par ailleurs, nous parlerons de « certifiés » pour désigner les utilisateurs obtenant au moins un certificat sur la plateforme *FUN* et de « non-certifiés » pour désigner les autres utilisateurs.

– Nous parlerons « d'inscriptions multiples » pour désigner les inscriptions de participants qui se sont inscrits à plusieurs MOOC sur la plateforme hébergeant le cours, et « d'inscription unique » si le participant ne s'est inscrit qu'à un seul cours sur la plateforme, sur la période que couvre l'analyse.

– Une inscription sera dite « groupée » dès lors qu'une autre inscription aura été réalisée par le même utilisateur dans la même journée, et « isolée » dans le cas contraire. On dira d'une inscription qu'elle correspond à un « retour » si elle a lieu dans un cours où l'utilisateur s'était déjà inscrit.

La logique qui sous-tend le traitement de ces données diffère sensiblement de celle de l'enquête présentée auparavant, dans la mesure où les analyses menées, si elles sont inspirées de notions comme la clé d'entrée, relèvent d'une approche plus inductive. Nous sommes conscients du fait que le comportement d'inscription n'est que partiellement caractérisé par les données de FUN. Dès lors, dans la discussion (section 6), nous présenterons un certain nombre d'arguments pour justifier en quoi les indicateurs que nous avons mobilisés peuvent être révélateurs des modalités de découverte des cours.

5. Résultats

Nous présenterons en premier lieu les résultats de l'enquête, en croisant différentes variables comme la modalité de découverte du cours et le nombre de MOOC suivis, puis nous reviendrons sur l'analyse des données d'inscription, en fournissant un certain nombre de statistiques descriptives sur des indicateurs d'intérêt.

5.1. L'analyse des types de visite basée sur les enquêtes

Par les enquêtes diffusées principalement sur FUN, nous avons analysé la manière dont les participants avaient entendu parler du cours pour la première fois et, d'autre part, si la volonté de se former à la thématique

correspondant au cours préexistait à sa découverte. La combinaison de ces deux éléments suffit à qualifier de manière relativement précise la nature d'une visite ou, en d'autres termes, la modalité de découverte du cours. Dans un second temps, nous nous sommes penché sur le nombre d'inscriptions réalisées sur la plateforme et au nombre de cours suivis en parallèle. Par la suite, nous avons cherché à établir le lien entre modalité de découverte du MOOC et nombre de MOOC suivis en parallèle. Ces cours sont généralement suivis sur la même plateforme – 80 % du temps, les inscriptions multiples sont réalisées sur *FUN* (Cisel, 2016), mais peuvent être aussi suivis sur d'autres plateformes, comme *Coursera* ou *edX*. Néanmoins, cette considération n'affecte pas outre mesure la nature de notre propos.

Concernant la question portant sur la nature de la visite, formulée en ces termes : « *Parmi les affirmations suivantes, quelle est celle qui correspond le plus à votre situation ?* », nous proposons quatre modalités de réponse :

1. C'est en naviguant sur la plateforme que j'ai découvert ce MOOC et que j'ai décidé de m'inscrire, mais je n'avais pas l'intention de me former à cette thématique avant de découvrir cette formation ;
2. J'ai entendu parler de ce MOOC hors de la plateforme (bouche-à-oreille, article, etc.) et j'ai décidé de m'inscrire de ma propre initiative, mais je n'avais pas l'intention de me former à cette thématique avant d'entendre parler de cette formation ;
3. On m'a contraint à m'inscrire, ou fortement incité ;
4. J'avais l'intention de me former à la thématique de ce MOOC avant de le découvrir, et c'est en faisant des recherches que je l'ai découvert et que je m'y suis inscrit.

La première réponse correspond à une visite expérientielle, dans la mesure où le participant découvre le cours pour la première fois via la plateforme et n'a pas en amont de la visite une idée précise du produit qui l'intéresse. Les réponses 2 et 3 correspondent à une visite expéditive. La seule différence entre ces deux cas de figure réside dans le fait que dans un cas, le cours n'est pas prescrit, tandis qu'il l'est dans l'autre cas. En termes de visite, le résultat est le même, puisqu'il se passe *a priori* peu de temps entre la connexion à la plateforme et l'inscription au cours. La dernière réponse correspond à une visite exploratoire.

Nous labélisons respectivement ces différentes réponses par le type de visite correspondant : expérientiel, expéditif non prescrit, expéditif prescrit, et exploratoire. Les pourcentages suivants portent sur l'ensemble

des cours de l'étude[3] : 54 %(±15) (N = 8545) des répondants considérés n'avaient pas l'intention de se former à la thématique du cours avant de le découvrir. Parmi les réponses, 38 % (±13) correspondent à une visite expérientielle, 16 % (±6) à une visite expéditive non prescrite, 7 % (±11) à une visite expéditive prescrite et 33 % (±10) à une visite exploratoire. Enfin, 6 % (±2) estiment être dans une autre situation.

Au sein des enquêtes, la visite expérientielle est la plus représentée des visites et la clé d'entrée plateforme en général – qui combine visites expérientielle et expéditive – semble l'emporter sur la clé d'entrée produit. Les biais d'auto-sélection limitent néanmoins la représentativité de ces réponses, d'autant que les enquêtes ont été diffusées principalement sur la plateforme FUN, sans qu'il soit aisé de déterminer de quelle manière.

Etudions maintenant le lien entre modalité de découverte du cours et nombre d'inscriptions réalisées dans d'autres MOOC. Pour ce faire, nous avons utilisé les deux questions suivantes :

– « *Combien de MOOC avez-vous commencé jusqu'à présent en dehors de ce MOOC-ci ? Par « commencer », nous entendons avoir consacré au moins une heure au MOOC* » ;

– « *Combien de MOOC êtes-vous en train de suivre en ce moment en plus de celui-ci ? Par « suivre », nous entendons vous investir au moins une heure par semaine sur le MOOC* ».

La plupart des répondants s'étaient déjà inscrits à un autre MOOC, et la majorité – 63 % – suivait au moins un autre cours en parallèle[4] (tableau 1).

Nous avons ensuite croisé la variable que représente la modalité de découverte du cours avec le nombre de MOOC suivis en parallèle. Après avoir regroupé les différents types de visites expéditives dans une même catégorie, on constate que les participants dont la visite est expéditive sont ceux qui suivent le moins de cours en simultané (figure 2), la différence avec les autres formes de visites étant statistiquement significative (chi = 535, ddl = 10, p-value < 10^{-4}). Seuls 27 % des répondants issus

3 Les chiffres introduits par le symbole ± correspondent à des écarts-types, représentant la variabilité entre les différents cours

4 Les apprenants qui répondent aux questionnaires font certes partie des plus investis et ne sont pas représentatifs de l'ensemble des utilisateurs de MOOC. Néanmoins, les données d'inscription fournies par la plateforme FUN vont montrer que l'inscription à de multiples cours constitue le cas majoritaire.

d'une visite expérientielle à ne suivre aucun autre cours en simultané. Cette proportion monte à 68 % pour les participants issus d'une visite expéditive. Le cas de la visite exploratoire est intermédiaire, avec 41 % des participants qui en sont issus qui ne suivent aucun autre MOOC. Il existe une corrélation entre le nombre de MOOC suivis en simultané, ainsi que le nombre de MOOC auxquels le participant s'est inscrit, et le type de visite associé à l'inscription.

Tableau 1 • Nombre d'inscriptions réalisées à d'autres MOOC et nombre de MOOC concomitants avec le cours où est diffusée l'enquête. N = 8545

	Commencé		En cours	
	Moy.	Ecart-type	Moy.	Ecart-type
Aucun	27 %	18 %	37 %	16 %
Un	20 %	5 %	26 %	5 %
Deux	14 %	4 %	17 %	8 %
Trois	11 %	4 %	9 %	3 %
Quatre	8 %	3 %	5 %	2 %
> Quatre	20 %	5 %	5 %	3 %

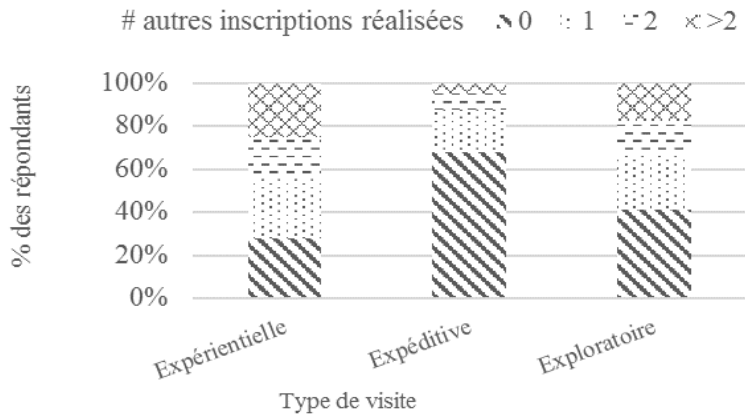


Figure 2 • Croisement entre type de visites ayant conduit à l'inscription et nombre de MOOC suivis en parallèle par le participant, sur la base de l'enquête. N = 7850

Cet élément constitue à nos yeux un indice concordant en faveur de l'hypothèse selon laquelle l'inscription à de multiples cours est plus représentée dans le cas des visites expérientielles. Les données d'inscription de la plateforme FUN vont nous permettre de préciser notre propos en nous permettant de quantifier l'importance des inscriptions groupées.

5.2. Une analyse des données d'inscription de FUN

L'analyse des données de FUN peut être décomposée en trois temps. Dans un premier temps, nous analysons la propension des utilisateurs à s'inscrire à de multiples cours au sein de la plateforme, ainsi qu'au rythme auquel sont réalisées ces inscriptions. Notre interprétation de ces données est que les inscriptions sont, dans une large mesure, issues de visites expérientielles. Dans une deuxième partie, nous nous fondons sur cette interprétation pour montrer que les comportements observés au sein de cours individuels sont influencés par le contexte dans lequel s'inscrit un cours, et en particulier des caractéristiques du catalogue auquel il appartient. Nous verrons que des caractéristiques des inscriptions, déterminantes pour l'engagement dans le cours, peuvent fluctuer au fil du temps, du fait de dynamiques propres aux plates-formes d'hébergement et à leur catalogue. Dans une troisième partie, nous nous pencherons sur la question du retour, c'est-à-dire sur le fait qu'un utilisateur se réinscrit parfois à la deuxième ou à la troisième itération d'un cours après s'être inscrit à une version antérieure de la formation, ce qui peut modifier sensiblement les utilisations qui sont faites du cours.

5.2.1. Inscriptions groupées et taux de certification

L'analyse des données d'inscriptions montre que 51 % des utilisateurs de FUN ne s'inscrivent qu'à un seul cours à l'échelle de la plateforme, ils représentent 20,2 % des inscriptions ; 79,8 % des inscriptions sont donc le fait d'utilisateurs s'inscrivant à plusieurs cours et correspondent dès lors à ce que nous avons défini comme des inscriptions multiples. À l'échelle de FUN, les utilisateurs s'inscrivent à 2,7 cours en moyenne. L'inscription à de nombreux cours constitue un phénomène central au sein de la plateforme : 29,3 % des inscriptions sont réalisées par des participants ayant réalisé dix inscriptions ou plus et, parmi celles-ci, 7,3 % sont le fait de participants qui ont réalisé trente inscriptions ou davantage. Les utilisateurs qui ont obtenu au moins un certificat sur la plateforme, *i.e.* les certifiés, représentent 11,3 % des utilisateurs. Ils s'inscrivent généralement à plus de cours que ne le font les non-certifiés. En effet, ces derniers sont

inscrits à 2,3 cours en moyenne, tandis que les certifiés sont inscrits à 4,5 cours en moyenne.

Sur les 669.423 inscriptions multiples étudiées, 46 % sont groupées, c'est-à-dire qu'une autre inscription a été réalisée dans une même journée. Le reste des inscriptions est isolé. Pour 18 % des inscriptions multiples étudiées, entre un jour et un mois sépare deux inscriptions consécutives et dans 33 % des cas, cet écart est compris entre un mois et un an. Ces résultats suggèrent que près de la moitié des inscriptions de FUN sont le fait d'utilisateurs qui s'inscrivent la même journée à plusieurs cours d'affilée. De plus, 73 % de ces inscriptions groupées sont réalisées le premier jour de l'inscription sur la plateforme.

Enfin, les inscriptions isolées sont celles dont les taux de certification sont les plus élevés – 12,5 % en moyenne, 19 % pour les inscriptions isolées réalisées lors de la première inscription à un cours de la plateforme, tandis que les taux de certification des inscriptions groupées avoisinent 5 %. Une régression logistique publiée précédemment (Cisel, 2017) confirme le lien statistique entre non-obtention du certificat et caractère groupé de l'inscription. Nous interprétons ce résultat comme découlant du fait que les visites expérientielles, où l'inscription est moins synonyme d'investissement, sont surreprésentées parmi les inscriptions groupées.

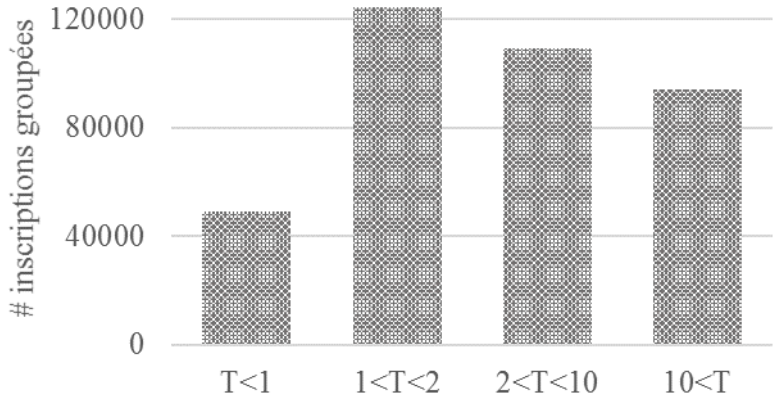


Figure 3 • Pour les inscriptions groupées, répartition des écarts temporels (T) en minutes, séparant l'inscription réalisée de l'inscription précédente. N = 376541

Le deuxième élément qui, selon nous, vient étayer le fait que les inscriptions groupées découlent de visites expérientielles réside dans le faible écart temporel qui les sépare. Pour la plupart des inscriptions groupées réalisées dans la même journée, celui-ci se compte en minutes. Nous avons analysé pour 376.541 inscriptions groupées ne correspondant pas à une première inscription l'écart temporel qui les séparait de l'inscription précédente (figure 3).

Cet écart est inférieur à une minute dans 13 % des cas ; il est compris entre une et deux minutes dans 33 % des cas, entre deux et dix minutes dans 29 % des cas, et supérieur à dix minutes dans seulement 25 % des cas. On conçoit difficilement que des visites exploratoires, qui nécessitent par définition de réaliser un certain nombre de recherches sur Internet, ou même des visites expéditives, puissent déboucher, en un temps aussi court, sur diverses inscriptions au sein de la plateforme. Pour aller plus loin, il serait nécessaire de disposer des traces de la navigation éventuelle sur la plateforme qui a pu précéder les visites ayant conduit à l'inscription (Clow, 2013) ; elles nous permettraient de préciser la manière dont l'offre a été explorée et dont les décisions ont été prises.

L'analyse des données montre que les inscriptions groupées, et de manière plus générale les inscriptions multiples, correspondent à des cours qui sont organisés sur des périodes similaires, ce qui peut faire décroître encore la probabilité qu'ils débouchent sur l'obtention du certificat. Nous constatons que, dans près de 12 % des cas, les cours consécutifs, pour un individu donné, commencent le même jour ; dans 40 % des cas, le délai entre deux cours consécutifs d'un utilisateur est compris entre une journée et deux semaines, et ce délai n'est supérieur à deux mois que dans 21 % des cas. Ceci est à mettre en relation avec le fait que les cours durent en moyenne six semaines sur FUN (Cisel, 2016) ; lorsque le délai entre deux cours est inférieur à six semaines, il y a de fortes chances que ceux-ci se superposent.

Ainsi, pour les inscriptions multiples, le taux de certification est de 7,4 % lorsque deux cours sont lancés le même jour, pour les deux cours considérés ; le taux monte à 14,1 % pour le premier cours lorsque le délai le séparant du suivant est supérieur à deux mois. En d'autres termes, plus les dates de lancement sont rapprochées les unes des autres, plus faible est la probabilité de terminer chacun de ces cours pris individuellement. La régression logistique (Cisel, 2017) confirme que ces différences sont statistiquement très significatives. Il est fort probable que les participants

s'inscrivent sans avoir les moyens temporels de suivre simultanément autant de cours qui se superposent. Ce choix est-il délibéré ou non ? Cette question sera traitée dans la discussion.

Il existe donc un lien fort entre le fait de faire des inscriptions groupées et le fait de ne pas obtenir le certificat. Nous allons voir que la proportion des inscriptions groupées au sein des inscriptions d'un MOOC varie considérablement d'un cours à l'autre, et au fil des mois, pour des raisons qui sont indépendantes des caractéristiques du cours.

5.2.2. Une fluctuation temporelle du nombre d'inscriptions groupées

Pour 25 des 119 cours analysés, la proportion d'inscriptions groupées est inférieure à 20 % ; pour 11 autres d'entre eux, elle est supérieure à 40 %. Compte tenu des faibles taux de certification associés à ces inscriptions, cette proportion d'inscriptions groupées affecte le taux de certification moyen d'un cours, comme nous l'avons vu dans les régressions logistiques. Or la proportion des inscriptions groupées réalisées à un instant donné fluctue au fil du temps (figure 4). Ainsi, au moment du lancement de la plateforme fin 2013, plus de 60 % des inscriptions multiples étaient groupées. Cette proportion a chuté à 30 % juste avant l'été 2014, pour remonter par la suite.

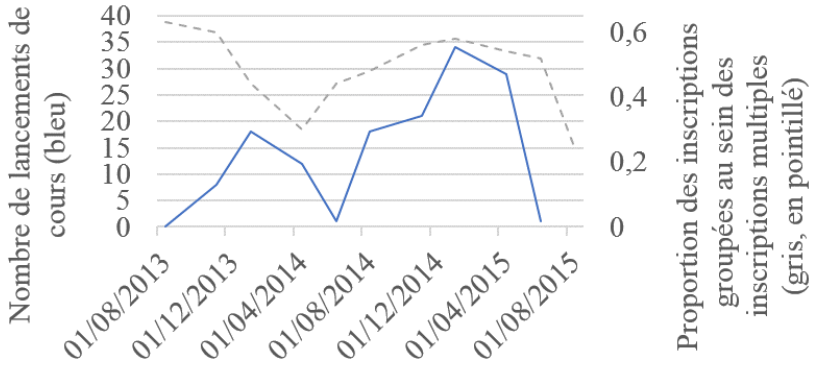


Figure 4 • Évolution de la proportion d'inscriptions groupées parmi les inscriptions de la plateforme en fonction de la date d'inscription

Il faudrait approfondir l'analyse pour mieux appréhender les déterminants de ces évolutions, mais il est vraisemblable que cette proportion soit liée à un certain nombre de caractéristiques de la plateforme, comme le nombre et la nature des cours ouverts aux inscriptions sur une période donnée. Cette hypothèse nous a amené à représenter sur la même figure (figure 4), en bleu, l'évolution du nombre de cours lancés au fil du temps. Mis à part au cours des premiers mois de la plateforme, période durant laquelle aucun cours n'est lancé mais où des inscriptions sont malgré tout engrangées, on observe un certain parallélisme entre le nombre de MOOC lancés et la proportion d'inscriptions groupées. Néanmoins, une analyse statistique plus poussée incluant sans doute davantage de plateformes, et fondée sur des analyses de séries temporelles, serait nécessaire pour établir rigoureusement le lien entre ces deux paramètres.

5.2.3. Le cas du retour des participants d'une itération sur l'autre

À la question de l'environnement dans lequel s'inscrit le MOOC s'ajoute celle de la trajectoire du dispositif, car un MOOC peut connaître de multiples itérations susceptibles d'être analysées globalement (Anderson *et al.,* 2014). C'est là une particularité des MOOC qui ne sont pas ouverts en permanence, il peut y avoir retour d'inscrits d'une itération sur l'autre (Bouchet et Bachelet, 2015). Nous allons quantifier son importance avant d'en tirer les conséquences en termes d'interprétation des analyses de traces.

Pour appréhender l'importance quantitative de la réinscription, nous avons analysé les données d'inscription de quarante-six MOOC de FUN ayant connu au moins une deuxième itération au cours de la période sur laquelle porte l'étude, dont la moitié exactement est terminée au moment de l'extraction des données. Ils correspondent à 101 itérations au total et à 603476 inscriptions, et représentent donc l'essentiel des inscriptions de la plateforme. Comme (Chuang et Ho, 2016), nous constatons une chute quantitative de l'audience au fil des différentes itérations d'un même MOOC. En moyenne, le nombre d'inscrits au cours d'une deuxième itération d'un cours correspond à 74 % (±33 %) du nombre d'inscrits de la première. Par ailleurs, on observe un phénomène de réinscription d'une édition sur l'autre.

Sur ces quarante-six cours, si l'on se cantonne uniquement aux deuxièmes itérations, on constate que 12,8 % des inscriptions correspon-

dent à des réinscriptions de participants à la première itération. La proportion de ces réinscriptions dans l'audience d'un cours ne passe jamais sous la barre des 5 %, et peut atteindre 37 % des inscrits dans certains cas. Les individus qui se réinscrivent représentent en moyenne 7,8 % des participants de la première itération du cours ; la probabilité d'obtenir le certificat pour les participants qui se réinscrivent est sensiblement inférieure à celle des autres utilisateurs du cours : 9 % pour un réinscrit contre 10,5 % pour les autres participants.

Sans être négligeable, le phénomène du retour ne joue que marginalement sur la question des taux de certification. Il est plausible que les comportements de ces réinscrits diffèrent sensiblement de ceux des autres participants, en particulier s'ils ont déjà visualisé au cours de la première inscription une partie des ressources du cours[5]. Pour cette raison, ce n'est pas tant l'existence de nouvelles itérations d'un cours qui nous semble intéressante à prendre en compte pour mieux interpréter l'engagement des participants dans le cours, mais plutôt celle du nombre de cours ouverts à un moment donné. Sans pouvoir le démontrer formellement à partir des données que nous avons présentées, celles-ci nous amènent à émettre l'hypothèse selon laquelle plus le nombre de cours ouverts à un instant donné est élevé, plus un participant a tendance à réaliser des inscriptions groupées, et plus l'engagement moyen dans chaque cours décroît, qu'on le mesure en heures ou en nombre d'actions réalisées.

6. Discussion

Nous introduirons la discussion par une réflexion sur le lien entre les indicateurs issus de l'analyse des données de FUN et les notions que nous avons présentées en introduction. Sur la base de l'analyse de questionnaires et des données d'inscription, nous avons étayé l'hypothèse selon laquelle les MOOC de FUN étaient caractérisés par une prédominance de la clé d'entrée plateforme sur la clé d'entrée produit, via notamment les visites expérientielles. Cette interprétation nécessite néanmoins une démonstration plus poussée que celle que nous avons fournie jusqu'à présent.

5 Néanmoins, de telles différences ne sont pas visibles avec les seules données d'inscription ; il serait nécessaire de mobiliser les traces d'interaction de chaque cours pris individuellement.

6.1. L'inscription groupée, un indicateur de visite expérientielle ?

Nous expliquons l'importance des inscriptions groupées par l'hypothèse selon laquelle la visite expérientielle correspond à la modalité de découverte du cours la plus fréquente ; en d'autres termes, les utilisateurs n'ont généralement pas une idée précise du sujet sur lequel ils veulent se former lorsqu'ils naviguent dans un catalogue. D'autres explications sont néanmoins possibles ; les paragraphes qui suivent visent à démontrer la plausibilité moindre de ces interprétations.

Première explication alternative : l'inscription groupée est associée le plus souvent à des visites exploratoires, où le participant a déjà en tête une idée de ce qu'il souhaite apprendre et il explore Internet à la recherche d'une formation appropriée. Dès lors, une forte représentation des visites exploratoires parmi les inscriptions groupées signifierait qu'au cours d'une même journée, un participant donné ait voulu, sans connaître l'offre de la plateforme, se former sur plusieurs sujets distincts, et qu'incidemment l'offre de FUN ait correspondu à ces différentes attentes, déclenchant une succession d'inscriptions. Le nombre de cours de FUN ouverts aux inscriptions à un moment donné ne dépasse qu'exceptionnellement la trentaine (figure 4). C'est assez pour autoriser des inscriptions groupées, mais c'est vraisemblablement trop peu pour coïncider avec des besoins préexistants d'autant de participants. Il est en effet peu plausible qu'un nombre si élevé de participants, d'une part, ait eu en tête une liste de thématiques sur lesquelles se former et, d'autre part, ait trouvé dans l'offre de la plateforme, relativement réduite à un instant donné, de nombreux cours correspondant à cette liste.

Deuxième cas, celui de la visite expéditive ; rappelons qu'elle implique que l'utilisateur ait identifié en amont, parce qu'un tiers lui aura suggéré un MOOC en particulier, le cours auquel il veut s'inscrire dans le catalogue. Il ne navigue sur la plateforme que pour s'y inscrire. La prépondérance des visites expéditives parmi les inscriptions groupées impliquerait que nombre de participants ne se connectent que pour s'inscrire successivement à plusieurs cours dont ils avaient entendu parler en amont. Là encore, ce scénario paraît peu vraisemblable, si tant est que l'on considère que les courriels envoyés par la plateforme aux utilisateurs pour les informer des nouveaux cours ne sont qu'un prolongement du catalogue. Il faudrait pour cela qu'on leur ait recommandé deux, trois, quatre cours parfois – ou qu'ils les aient découverts dans un article de journal par

exemple – cours dont les inscriptions, incidemment, seraient ouvertes au moment où l'utilisateur se connecte. Le scénario n'est pas impossible, mais on peut douter de sa plausibilité.

Par élimination, la prévalence de visite expérientielle constitue la meilleure explication pour interpréter la forte représentation des inscriptions groupées. Cette hypothèse ne signifie pas qu'il existe une équivalence entre inscriptions groupées et visites expérientielles, mais simplement qu'il existe une forte corrélation entre les deux. En effet, les entretiens que nous avons menés par ailleurs (Cisel, 2016) suggèrent que la première inscription d'une série d'*inscriptions groupées* peut être issue d'une visite exploratoire ou d'une visite expéditive, et que seules la ou les suivantes découlent à proprement parler de visites *expérientielles*. Le participant découvre la plateforme car il a entendu parler d'un cours en particulier, ou parce qu'il cherche à se former sur un sujet défini ; après avoir effectué cette première inscription, il parcourt pour le plaisir l'offre de la plateforme et s'inscrit à de nombreux autres cours. Il est néanmoins impossible de détecter l'existence de telles stratégies mixtes sur la base des seules données d'inscription, sans donnée auto-déclarée. A défaut de pouvoir collecter pour chaque inscription les modalités de découverte du cours, la dimension qualitative (Cisel, 2016) de notre enquête a permis d'éclairer certains des mécanismes sous-jacents à ces inscriptions groupées, et de manière générale, pourquoi les utilisateurs s'inscrivent à plus de cours qu'ils ne peuvent en suivre *a priori*.

6.2. De l'intention de l'utilisateur au moment de l'inscription

Les entretiens nous ont permis en premier lieu de constater que les participants peuvent sous-estimer le temps que prendra chaque MOOC individuellement, réaliser une fois la formation commencée le problème de la contrainte temporelle et se désengager d'au moins un d'entre eux. Ils peuvent vouloir les comparer, pour choisir un cours en particulier après quelques heures d'investissement. Ils peuvent n'avoir aucune intention de s'engager dans la formation, notamment lorsqu'ils cherchent avant tout, par l'inscription à de multiples cours, à bénéficier de la possibilité de retrouver le lien vers la formation plus facilement s'ils veulent en consulter plus tard les ressources (on ne peut accéder aux ressources d'un cours terminé qu'en s'étant inscrit en amont). À cet égard, nos analyses vont dans le sens des propos de Koller (2013), lorsqu'elle affirme que nombre d'utilisateurs s'inscrivent sans véritable intention de suivre l'ensemble des

cours choisis, d'une part, et/ou qu'ils sont dans l'impossibilité de suivre autant de MOOC de manière simultanée, d'autre part.

Il s'agit d'un panel non exhaustif d'hypothèses dont il est bien malaisé de mesurer l'importance respective du fait des limites de l'approche par enquête. Certains auteurs ont montré (Reich, 2014), sur la base d'analyses croisant traces d'interaction et données de questionnaire, que les biais d'auto-sélection limitaient considérablement la possibilité d'appréhender les intentions des répondants au moment de l'inscription.

6.3. Organisation en catalogues, trajectoires de dispositifs et analyses de traces

Rappelons enfin que nous avons saisi l'opportunité offerte par ces analyses pour montrer que les caractéristiques des inscriptions, et donc l'engagement au sein des cours, pouvaient dépendre de dynamiques propres à la plateforme. Nous faisons en particulier l'hypothèse que la visite expérientielle est d'autant plus prépondérante au sein d'une plateforme que l'offre est riche sur le plan quantitatif. Toutes choses restant égales par ailleurs, la multiplication des inscriptions est selon cette hypothèse positivement corrélée au nombre de cours dont les inscriptions sont ouvertes. Cette hypothèse pourrait être testée sur la base d'études corrélationnelles menées à l'échelle de différentes platesformes, visant à établir un lien entre le nombre de cours ouverts à un instant donné et l'inscription multiple.

Une telle considération est sans doute pertinente dès lors que l'on s'intéresse aux traces des cours en ligne organisés en catalogues, qu'ils soient ou non payants et ouverts à tout moment. Il en va de même pour la question de la trajectoire d'un dispositif (que nous avons illustrée avec la question du retour d'utilisateurs revenant de manière régulière).

Le problème du retour et de son impact sur les comportements des utilisateurs n'a ici été qu'évoqué ; il est spécifique des cours aux itérations multiples, mais on peut arguer que la trajectoire du dispositif mérite également d'être explorée pour les cours en ligne ouverts en permanence. Sans parler de retour à proprement parler, puisqu'il n'y a pas nécessairement désinscription puis réinscription, il est vraisemblable que certains apprenants ne s'engagent dans le cours que par à-coups. Ils suivent une partie du cours sur un laps de temps réduit, quelques jours par exemple, pour y retourner plusieurs semaines voire plusieurs mois plus tard afin de bénéficier du reste de la formation, ou afin de réviser des segments de la

formation. Si la question a été dans une certaine mesure étudiée sur les MOOC dits « archivés » (Campbell, Gibbs, Najafi, et Severinski, 2014), c'est-à-dire inactifs mais accessibles aux personnes s'étant déjà inscrites, elle n'a, à notre connaissance, pas été étudiée dans le cas des cours ouverts en permanence.

7. Conclusion

Nous conclurons cet article en discutant de l'une des principales limitations de ce travail, l'impossibilité d'obtenir certaines données cruciales pour étayer plus solidement nos hypothèses, et en soulignant les questions éthiques que soulève ce type d'analyse.

L'une des principales limites de notre démonstration est que nous ne disposons pas des données sur la navigation précédant la session menant à l'inscription au cours. À notre connaissance, seul un article (Clow, 2013) a abordé la question des données de navigation précédant l'inscription, sur la base des données de trois MOOC connectivistes. Ces résultats restent embryonnaires et mériteraient d'être étendus à un travail sur une plate-forme prise dans sa totalité.

Ces données pourraient nous aider à identifier un type de visite donnée. On peut par exemple supposer qu'un utilisateur naviguant sur les pages de présentation de nombreux cours correspondant à des disciplines distinctes, avant de s'inscrire à un cours donné, est probablement en train de réaliser une visite expérientielle. À l'inverse, une visite rapide, pour laquelle le délai entre la connexion à la plateforme et l'inscription est minimal, correspond plus vraisemblablement à une visite expéditive. Enfin, elles permettraient de déterminer si un apprenant a visité en amont la page de présentation du cours, peut-être plusieurs heures ou plusieurs jours auparavant, avant de prendre la décision de s'inscrire. Le principal problème associé à ce type de données réside dans le fait que très souvent, les participants naviguent sur le catalogue sans s'être identifiés – ils ne s'identifient qu'au moment de l'inscription. Alternativement, ils prennent souvent connaissance des cours qui les intéressent par les courriels envoyés par la plateforme et « naviguent » sur une version tronquée du catalogue au sein de leur messagerie. L'un comme l'autre de ces deux phénomènes complexifieraient l'interprétation de ces données de navigation précédant l'inscription.

Par ailleurs, l'utilisation de telles données par les plateformes soulève des questions éthiques, analogues à celles que présente l'existence de

publicités ciblées fondées sur les cookies, lors d'une navigation sur Internet. Le respect de la vie privée n'est pas ici la seule dimension à prendre en compte ; après tout, si les données sont anonymisées rigoureusement, le problème n'est que secondaire. L'utilisation des données de navigation et d'inscription soulève en revanche d'autant plus de questions qu'elles sont utilisées au quotidien par les plates-formes américaines comme *edX* ou *Coursera* à des fins de recommandation. Ces acteurs envoient régulièrement des courriels à leurs utilisateurs proposant des listes de cours susceptibles de les intéresser, listes dont la personnalisation est fondée sur les données d'inscription.

Une bonne compréhension du processus de sélection et de suivi des cours peut certes être utilisée pour mieux répondre aux besoins réels ou supposés des utilisateurs, mais elle peut également être mobilisée pour permettre aux plateformes payantes d'augmenter leurs ventes, sans que la question de l'intérêt de l'apprenant ne soit centrale. Et si l'on ne saurait reprocher aux acteurs de l'écosystème de vouloir assurer leur viabilité économique, il est à craindre que les intérêts économiques de ces derniers ne soient pas systématiquement alignés avec les intérêts des utilisateurs des plates-formes en termes d'objectifs d'apprentissage. Nous voilà à nouveau face aux contradictions des acteurs des MOOC, pris entre la difficulté à établir des modèles économiques et la velléité, présente au moins dans les discours (Koller, 2013), d'accroître toujours plus l'accessibilité de cours en ligne de qualité.

La compréhension des processus de choix au sein d'une offre toujours croissante présente sans doute un intérêt pour la communauté scientifique et pour celle des praticiens. Espérons néanmoins qu'elle ne contribue pas à soutenir une logique essentiellement mercantile de l'apprentissage en ligne, incompatible avec les idéaux qui animèrent sans doute au moins une partie des acteurs impliqués initialement dans le phénomène MOOC.

REMERCIEMENTS

L'auteur tient à adresser ses remerciements à l'équipe de France Université Numérique pour l'accès à des données que peu de plateformes ont accepté de partager. Il remercie également Monique Baron pour sa

relecture des dernières versions de l'article et pour ses retours, toujours précis et pertinents.

RÉFÉRENCES

Albo, L., Hernández-Leo, D. et Oliver, M. (2016). Are higher education students registering and participating in MOOCs? The case of MiríadaX. Dans *Proceedings of the European MOOCs Stakeholders Summit (EMOOCS 2016)* (p. 197-211). Norderstedt, Allemagne : BoD.

Anderson, A., Huttenlocher, D., Kleinberg, J. et Leskovec, J. (2014). Engaging with massive online courses. Dans *Proceedings of the 23rd International Conference on World Wide Web* (p. 687–698). New York, NY : ACM.

Banerjee, A. V. et Duflo, E. (2014). (Dis) organization and success in an economics MOOC. *American Economic Review, 104* (5), 514–18.

Barak, M., Watted, A. et Haick, H. (2016). Motivation to learn in massive open online courses: Examining aspects of language and social engagement. *Computers & Education, 94*, 49–60.

Bean, J. P. et Metzner, B. S. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research, 55* (4), 485–540.

Boshier, R. (1973). Educational participation and dropout: A theoretical model. *Adult Education Quarterly, 23* (4), 255–282.

Bouchet, F. et Bachelet, R. (2015) Do MOOC students come back for more? Recurring Students in the GdP MOOC. Dans *Proceedings of the European MOOCs Stakeholders Summit* (p 174-182).

Breslow, L., Pritchard, D., DeBoer, J., Stump, G., Ho, A. et Seaton, D. (2013). Studying learning in the worldwide classroom: Research into edX's first MOOC. *Journal of Research & Practice in Assessment, 8*, 13–25.

Campbell, J., Gibbs, A. L., Najafi, H. et Severinski, C. (2014). A comparison of learner intent and behaviour in live and archived MOOCs. *The International Review of Research in Open and Distance Learning, 15*(5). Récupéré de : https://goo.gl/hgs83u

Chuang, I. et Ho, A. (2016). *HarvardX and MITx: Four years of open online courses--fall 2012-summer 2016.* Récupéré de : https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2889436

Cisel, M. (2015). Réponses aux détracteurs de l'enseignement en ligne. *Cités, 63*(3), 99–108. https://doi.org/10.3917/cite.063.0099

Cisel, M. (2016). *Utilisations des MOOC, éléments de typologie.* (Thèse de doctorat, ENS Paris-Saclay, France). Récupéré de : http://www.theses.fr/2016SACLN024

Cisel, M. (2017). Analyser les taux de certification des MOOC au prisme du comportement d'inscription. Dans *Actes du 8ème colloque EIAH (EIAH 2017)* (p. 167-177).

Cisel, M. et Bruillard, E. (2013). Chronique des MOOC. *STICEF, 19*. Récupéré de : http://sticef.univ-lemans.fr/num/vol2012/13r-cisel/sticef_2012_cisel_13r.htm

Cisel, M. et Leh, A. (2017). Les MOOC, entre transposition de cours universitaires et adaptation à la massification des audiences. *Revue Internationale des Technologies en Pédagogie Universitaire, 14*(3), 35-46.

Clow, D. (2013). MOOCs and the funnel of participation. Dans *Proceedings of the Third Conference on Learning Analytics and Knowledge* (p. 185–189).

Daniel, J. (2012). Making sense of MOOCs: Musings in a maze of myth, paradox and possibility. *Journal of Interactive Media in Education, 3*(0). Récupéré de : https://jime.open.ac.uk/articles/10.5334/2012-18/

Garrison, D. R. (1985). Predicting dropout in adult basic education using interaction effects among school and nonschool variables. *Adult Education Quarterly, 36*(1), 25–38.

Gillani, N. et Eynon, R. (2014). Communication patterns in massively open online courses. *The Internet and Higher Education, 23*, 18–26.

Greene, J. A., Oswald, C. A. et Pomerantz, J. (2015). Predictors of retention and achievement in a Massive Open Online Course. *American Educational Research Journal, 52*(5), 925–955.

Hackman, J. R. et Dysinger, W. S. (1970). Commitment to college as a factor in student attrition. *Sociology of Education, 43*(3), 311–324.

Isaac, H. et Volle, P. (2008). *E-commerce : de la stratégie à la mise en oeuvre opérationnelle*. Paris : Pearson Education France.

Kember, D. (1989). A longitudinal-process model of drop-out from distance education. *The Journal of Higher Education, 60*(3), 278-301.

Kemp, W. C. (2002). Persistance of adult learners in distance education. *American Journal of Distance Education, 16*(2), 65–81.

Kizilcec, R. F. et Schneider, E. (2015). Motivation as a lens to understand online learners. *ACM Transactions on Computer-Human Interaction (TOCHI), 22*(2), 6:1-6:24.

Koller, D. (2013). Retention and intention in Massive Open Online Courses: In depth. *Educause*. Récupéré de : https://goo.gl/aqcGMu

Munro, B. H. (1981). Dropouts from higher education: Path analysis of a national sample. *American Educational Research Journal, 18*(2), 133–141.

Perna, L. W., Ruby, A., Boruch, R. F., Wang, N., Scull, J., Ahmad, S. et Evans, C. (2014). Moving Through MOOCs Understanding the Progression of Users in Massive Open Online Courses. *Educational Researcher, 43*(9), 421–432.

Reich, J. (2014). MOOC completion and retention in the context of student intent. *EDUCAUSE Review*. Récupéré de : https://goo.gl/Q2cmhu

Rosé, C. P., Carlson, R., Yang, D., Wen, M., Resnick, L., Goldman, P. et Sherer, J. (2014). Social factors that contribute to attrition in MOOCs. Dans *Proceedings of the First ACM Conference on Learning @ Scale Conference* (p. 197–198). New York, NY : ACM. http://dx.doi.org/10.1145/2556325.2567879

Stenger, T. et Bourliataux-Lajoinie, S. (dir.). (2014). *E-marketing et e-commerce. Concepts, outils, pratiques* (2e éd.). Paris, France : Dunod.

Sweet, R. (1986). Student dropout in distance education: An application of Tinto's model. *Distance Education, 7*(2), 201–213.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89–125.

Tinto, V. (1982). Limits of theory and practice in student attrition. *The Journal of Higher Education*, *53*(6), 687–700.

Tinto, V. (1987). *Leaving college: Rethinking the causes and cures of student attrition.* Chicago, IL : University of Chicago.

Tinto, V. (2006). Research and practice of student retention: What next? *Journal of College Student Retention: Research, Theory et Practice*, *8*(1), 1–19.

Université d'Edimbourg (2013). MOOCs @ Edinburgh 2013 – Report #1, Récupéré de : https://goo.gl/zanmmk

Vaughan, R. P. (1968). College dropouts: Dismissed vs. withdrew. *The Personnel and Guidance Journal*, *46*(7), 685–689.

Zawacki-Richter, O. et Anderson, T. (dir.). (2014). *Online distance education: Towards a research agenda.* Athabasca University Press, Canada.