Analyser les taux de certification des MOOC au prisme du comportement d'inscription

Matthieu Cisel

▶ **To cite this version:**

Matthieu Cisel. Analyser les taux de certification des MOOC au prisme du comportement d'inscription. Environnements Informatiques pour l'Apprentisage Humain (8ème édition), Jun 2017, Strasbourg, France. hal-01670280

HAL Id: hal-01670280
https://hal.science/hal-01670280

Submitted on 21 Dec 2017

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.

Analyser les taux de certification des MOOC au prisme du comportement d'inscription

Matthieu Cisel

ENS Paris-Saclay, STEF, 61 av. du Président Wilson, 94230 Cachan, France
matthieucisel@gmail.com

Résumé. L'analyse des comportements d'inscription au sein des plates-formes de MOOC constitue une piste fructueuse pour mieux comprendre la faiblesse de leurs taux de certification. Dans cet article, nous analysons au prisme des concepts de type de visite et de clé d'entrée les données d'un million d'inscriptions réalisées sur la plate-forme française FUN : rythme et nombre d'inscription, délai entre cours consécutifs, etc. Nous suggérons qu'il existe une forte représentation, au sein de la plate-forme, d'inscriptions issues de visites expérientielles. Les visites expérientielles correspondent à un mode de navigation où l'utilisateur parcoure un catalogue sans nécessairement avoir une idée précise du type de MOOC qu'il souhaite suivre.

Mots-clés. MOOC, comportement d'inscription, méthodes mixtes

Abstract. Registration behaviour within MOOC platforms enables to deepen our understanding of MOOC low completion rates. In this paper, we analyze registration data from the French MOOC platform FUN, based on concepts drawn from the webmarketing literature, such as the *type of visit*. We explore various dimensions of the registration behavior: number and of registrations, delay between consecutive courses, etc. We suggest that learners discover the courses they register to mostly by navigating on the platform catalog, a particular case called the *experiential visit*.

Keywords. MOOC, registration behavior, mixed methods

1 Introduction

1.1 Les faibles taux de certification en question

Alors que le débat sur les faibles taux de certification des MOOC battait son plein, la fondatrice de la plate-forme américaine *Coursera* publia sur la question un réquisitoire qui fit date [9]. Les auteurs présentèrent une longue liste d'éléments susceptibles d'expliquer la faiblesse des taux de certification observés. L'un des arguments avancés par l'auteur, sur la base des données de la plate-forme, est qu'un certain nombre de participants s'inscrivent à plus de cours qu'ils ne peuvent en suivre, et qu'une inscription n'a pas valeur d'engagement à suivre le cours. Bien que les quelques chiffres

avancés par Koller suggèrent que le comportement d'inscription constitue l'une des clés de la compréhension des comportements des utilisateurs, la question a fait l'objet de peu de travaux empiriques, et les tentatives de théorisation sur ce sujet restent à notre connaissance inexistantes. Que peut nous apprendre le comportement d'inscription sur la signification d'une inscription pour un utilisateur ? Dans quelle mesure nous permet-il d'interpréter certains comportements, comme les modalités de choix du cours, et partant de là, à appréhender pourquoi si peu persistent jusqu'au terme de la formation ? Ce travail n'a donc pas vocation à avoir des retombées pédagogiques, mais à apporter un regard différent sur les taux de certification. Il vise en particulier à mieux comprendre pourquoi cet indicateur peut varier significativement d'un cours à l'autre. La gratuité de ces cours et l'absence de cadre contraignant semblent prédominer dans le débat sur les taux de certification, mais ces deux facteurs n'expliquent en rien les variations d'un MOOC à l'autre.

1.2 Une littérature embryonnaire sur les comportements d'inscription

Nous qualifions de *données d'inscription* l'ensemble des données permettant de qualifier les inscriptions réalisées au sein d'une même plate-forme : date de l'inscription, obtention ou non-obtention de l'éventuel certificat suite à cette inscription, etc. Nous utilisons le terme *comportement d'inscription* pour désigner tout comportement observable dont la description se base sur des données d'inscription : nombre d'inscriptions réalisées sur la plate-forme, temporalité des inscriptions, etc.

Le lien entre comportement d'inscription et taux de certification n'a été étudié que dans un nombre limité de publications [1], [2], [3], [7], [11], le plus souvent sous l'angle du délai entre la date d'inscription et la date de lancement du cours. Quant à la question de l'inscription à de multiples cours sur une plate-forme, elle reste largement sous-investie. Si les rapports du MIT [7] évoquent la question, ils se cantonnent aux seules formations de Harvard et du MIT, et n'élargissent pas l'analyse à l'ensemble des cours de la plate-forme *edX*.

À notre connaissance, l'analyse des données de *Miriada X* [1] constitue l'une des seules recherches publiées dans la littérature scientifique portant sur les comportements d'inscription à l'échelle d'une plate-forme. 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. S'ils se basent sur les données d'inscription, ils ne cherchent pas à établir de lien entre comportement d'inscription et taux de certification, et ne tentent pas de théoriser la question.

C'est cette lacune de la littérature, tant sur le plan empirique que théorique, que nous souhaiterions contribuer à combler à travers cet article. Nous suivons ici une méthode mixte mêlant données d'entretiens, et des données relatives à plus d'un million d'inscriptions réalisées sur la plate-forme française France Université Numérique (FUN). Nous nous pencherons en particulier sur une dimension du comportement

d'inscription que constituent le nombre et le rythme d'inscriptions réalisées au sein de la plate-forme qui héberge le cours.

1.3 Les concepts de clé d'entrée et de visite

Sur le plan théorique, nous tenterons de qualifier et de quantifier les trajectoires qui conduisent à l'inscription, au prisme des concepts de *clé d'entrée* et de *visite*, tous deux issus de la littérature du webmarketing, et plus particulièrement, de la littérature sur les processus d'achat en ligne [8], [13]. L'utilisation de ces concepts repose sur l'analogie entre l'achat d'un produit sur Internet et l'inscription à un MOOC. Revenons sur leur définition, avant de présenter la méthodologie suivie.

La clé d'entrée désigne le processus qui a conduit à l'achat d'un produit, ou dans notre cas, à l'inscription au cours ; deux modalités y sont associées, la clé d'entrée plate-forme d'une part, et la clé d'entrée produit d'autre part : « Soit l'internaute se rend sur 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 « le site marchand à visiter », alors que dans le deuxième cas, la clé d'entrée est *le produit à acheter*. » [8]

Les auteurs font l'analogie entre la clef d'entrée plate-forme et le « lèche-écran ». Un exemple de clef d'entrée plate-forme est représenté par le système de recommandation d'*Amazon*, qui, sur la base des données récoltées sur le comportement antérieur sur la plate-forme, pousse les utilisateurs connectés à acheter des produits dont ils ignoraient l'existence. Les auteurs croisent ces deux concepts avec le degré de connaissance des plates-formes 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 maintenant : expérientielle et expéditive pour la clef d'entrée plate-forme, exploratoire et comparative pour la clef d'entrée produit.

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 ». Le participant découvre donc le cours en naviguant sur la plate-forme, sans avoir une idée précise, en amont, de la thématique sur laquelle il souhaite 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 ». 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 plate-forme, ou après

y avoir été incités par un ami, un enseignant ou un supérieur hiérarchique. Ils ne se connectent à la plate-forme 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 ». Dans ce cas de figure, 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.).

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 ». Le caractère délimité dans le temps des MOOC et le nombre réduit de cours sur une thématique donnée fait qu'il y a en pratique, pour les cours qui nous concernent, peu de compétition entre cours portant sur les mêmes thématiques, ce qui limite la possibilité d'une visite comparative. Dans le cadre de notre travail, c'est donc uniquement sur les visites expérientielle, exploratoire et expéditive que nous concentrerons notre attention. Maintenant que nous avons précisé les différents concepts que nous mobiliserons dans le cadre de cet article, présentons la méthodologie qui a présidé à la collecte des données.

2 Méthodologie

Le travail de conceptualisation et d'analyse de données s'inscrit dans une logique de méthode mixte [6]. Il fait suite à la réalisation de quarante-et-un entretiens avec des utilisateurs de MOOC, menés entre septembre 2014 et juin 2015. Les données d'inscription de la plate-forme *FUN* ont été extraites par l'équipe technique de la plate-forme au 20/09/2015 [4]. Ces données comportent pour chaque inscription une date, l'identifiant du cours, l'identifiant de l'utilisateur et l'obtention éventuelle du certificat associé au cours. La base de données comprend au total 1.047.445 inscriptions.

Rappelons que l'inscription aux MOOC est gratuite, et que ceux-ci sont délimités dans le temps. Ils durent en général entre quatre et huit semaines. Ils engrangent l'essentiel des inscriptions, pendant une période d'ouverture de plusieurs semaines à plusieurs mois, qui précède le lancement du cours. En moyenne, 73% des inscriptions sont réalisées au moment où le cours débute. On constate que 308.396 des inscriptions correspondent à des cours qui n'ont pas commencé au moment de l'extraction des données, ou qui ne sont pas terminés, et ont donc été exclues de toute réflexion sur les taux de certification.

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 *productive* dès lors qu'elle débouche sur l'obtention d'un certificat, et d'*improductive*

dans le cas contraire. Nous avons créé pour les besoins de l'analyse un certain nombre de variables construites qu'il nous faut maintenant préciser. Nous parlerons d'*inscriptions multiples* pour désigner les inscriptions de participants s'étant inscrit à plusieurs MOOC sur la plate-forme hébergeant le cours, et *d'inscription unique* si le participant ne s'est inscrit qu'à un seul cours sur la plate-forme sur la période que couvre l'analyse. Nous parlerons ici de *certifiés* pour désigner les utilisateurs obtenant au moins un certificat sur la plate-forme *FUN,* à quelque date que ce soit, et de *non-certifiés* pour désigner les autres utilisateurs. Précisons que dans le contexte de l'analyse du comportement d'inscription, un certifié peut être à l'origine de nombreuses *inscriptions improductives*.

Une inscription est dite *groupée* dès lors qu'une autre inscription a été réalisée dans la même journée, le fuseau horaire choisi pour distinguer les journées étant celui de Paris. Rappelons que l'inscription aux cours est gratuite, et qu'au cours de 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. Tous les cours ont une date de début et une date de fin, leur contenu n'étant disponible après la date de fin que pour les seuls utilisateurs s'étant inscrits avant la clôture des inscriptions. Des inscriptions sont *superposées* si elles correspondent à des cours dont les périodes de déroulement se chevauchent. Enfin, nous nous sommes intéressés à la *temporalité* de l'inscription, qui désigne le délai entre la date d'inscription et la date de lancement du cours. C'est selon une approche quantitative que nous allons maintenant présenter les différentes dimensions qui caractérisent une inscription.

3 Résultats

On constate que 51,0 % des utilisateurs de *FUN* ne s'inscrivent qu'à un seul cours, réalisant par là une *inscription unique*, mais que 79.8 % des inscriptions de la plate-forme sont le fait d'utilisateurs s'inscrivant à plusieurs cours, et peuvent donc être considérées à ce titre comme des *inscriptions multiples*. Les inscrits qui réalisent de multiples inscriptions pèsent lourdement à l'échelle de la plate-forme. On constate que 29,3% des inscriptions sont le fait de participants ayant réalisé dix inscriptions ou plus, et 7,3% sont le fait de participants qui ont réalisé trente inscriptions ou davantage.

Les non-certifiés sont inscrits à 2,3 cours en moyenne, tandis que les certifiés sont inscrits à 4,5 cours en moyenne ; 68% de leurs inscriptions sont improductives. On constate que sur les 669.423 *inscriptions multiples* étudiées, 46% sont *groupées*, c'est-à-dire qu'elles ont été réalisées dans une même journée. Le reste des inscriptions sont *isolées*. Dans 18% des cas, entre un jour et un mois séparent deux inscriptions consécutives, et dans 33% des cas, cet écart est compris entre un mois et un an. Ces résultats suggèrent que plus du tiers des inscriptions de *FUN* sont le fait d'utilisateurs qui s'inscrivent la même journée à plusieurs cours d'affilée. Nous pouvons même aller plus loin en constatant que 72,9% de ces *inscriptions groupées* sont réalisées le premier jour de l'inscription sur la plate-forme.

On constate que les inscriptions isolées du premier jour sont celles dont les taux de certification sont les plus élevés (18 %), tandis que les inscriptions isolées postérieures au premier jour correspondent à des taux de certification intermédiaire (11 %). Un tel écart n'existe pas pour les *inscriptions groupées*, dont les taux de certification sont en moyenne de 6,1%. En termes de taux de certification, il n'existe presque aucune différence entre les inscriptions groupées du premier jour et celles qui lui sont postérieures.

On gardera néanmoins à l'esprit, que qui dit inscription groupée, dit MOOC dont les dates d'ouverture des inscriptions se chevauchent. Or il y a une corrélation forte entre la date d'ouverte des inscriptions et la date de lancement du cours. Les MOOC correspondants à ces inscriptions groupées se chevauchent aussi fréquemment, ce que l'on peut observer dans l'analyse du jeu de données. Il peut y avoir un phénomène de compétition pour le temps d'un utilisateur entre cours par la suite, qui fait chuter la probabilité moyenne d'obtenir le certificat d'un cours donné, et qui contribue aux plus faibles taux de certification des visites expérientielles.

Nous avons cherché par conséquent à appréhender de manière quantitative la superposition, et à établir un lien avec les taux de certification. Nous avons vu que celui-ci est d'autant plus faible que la période de superposition entre deux MOOC consécutifs est longue. Plutôt que de mesurer précisément la période de superposition de deux cours consécutifs, tâche qui nécessite des données dont nous ne disposions pas, nous nous sommes intéressés aux délais séparant le lancement de deux cours consécutifs, pour les utilisateurs inscrits à de multiples cours.

Nous constatons que dans près de 12% des cas, les cours consécutifs commencent le même jour ; dans 40% des cas, le délai entre deux cours consécutifs 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 ;* 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 ; ce taux monte à 14,1% lorsque le délai séparant les deux cours 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.

Nous nous sommes enfin penchés sur la temporalité de l'inscription. Le caractère délimité dans le temps des MOOC suppose que la date de lancement du cours est parfois éloignée de la date d'inscription de plusieurs mois. A l'inverse, les inscriptions tardives, en cours de formation, sont souvent permises, ce qui peut conduire à une impossibilité de respecter les dates-limites. Néanmoins, ce paramètre, comme nous l'avons vu central dans la littérature, ne semble avoir qu'une influence modérée sur les taux de certification.

On constate que 40,3% des inscriptions de *FUN* sont réalisées plus d'un mois avant le lancement du cours. L'analyse des données de *FUN* révèle que 31,5% des inscriptions sont postérieures au lancement du cours. On observe que les taux de certification dépendent dans une large mesure de la date d'inscription. Ils atteignent un pic à 12,1% pour les participants s'inscrivant dans la semaine précédant l'ouverture du

cours, et chutent à 10,3% dès la semaine suivant le lancement du cours. Ils se maintiennent à 8,7% deux mois après le lancement du cours. Les inscriptions les plus précoces comme les plus tardives ont moins de chance d'être productives que les inscriptions précédant de peu l'ouverture du cours, mais les écarts sont nettement moins significatifs que ceux constatés dans la littérature, surtout en ce qui concerne les inscriptions les plus tardives.

Nous synthétisons l'ensemble de ces résultats à travers une régression logistique (Table 1). Celle-ci confirme le lien entre obtention du certificat et caractère groupé de l'inscription. L'Odd-ratio, qui caractérise la probabilité de déboucher sur un certificat pour une inscription groupée par rapport à une inscription isolée, est de 0.42. La différence est statistiquement significative (GLM, p-value $<10^{-15}$). C'est le facteur qui explique l'essentiel de la déviance d'une régression logistique qui prend en compte dans le même modèle trois autres variables en plus du caractère groupé de l'inscription : nombre d'inscriptions réalisées sur la plate-forme, délai entre date d'inscription et lancement du cours, superposition des périodes d'organisation des cours suivis par un participant inscrit à de multiples cours.

L'analyse de la déviance suggère que le groupement des inscriptions est le facteur dont le pouvoir prédictif sur les taux de certification est le plus important. Certes, la superposition des cours et la précocité de l'inscription ont des effets non négligeables, mais ceux-ci sont inférieurs à ceux du groupement. Le pseudo-R^2 correspondant à cette régression logistique est de 0,28, ce qui correspond selon McFadden à un très bon pouvoir prédictif du modèle. Par comparaison, des analyses similaires réalisées avec des données sociodémographiques ou avec des données comme la motivation pour suivre le cours conduisent à des pseudo-R^2 de 0,02 [4].

Les entretiens menés par ailleurs suggèrent que ces inscriptions groupées associées à de faibles taux de certification pourraient dans une large mesure découler de visites expérientielles. Nous y reviendrons dans la discussion. Pour conclure, soulignons que la proportion des inscriptions groupées au sein des inscriptions d'un MOOC varie considérablement d'un cours à l'autre. Pour 25 des 119 cours analysés, elle est inférieure à 20%, pour 11 d'entre eux, elle est supérieure à 40% (Figure 1, Gauche). 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 fluctuent au fil du temps (Figure 1, Droite). Ainsi, au moment du lancement de la plate-forme fin 2013, plus de 60% des inscriptions multiples sont groupées. Cette proportion chute à 30% juste avant l'été 2014, pour remonter par la suite. Des analyses supplémentaires montrent que cette proportion est vraisemblablement liée à un certain nombre de caractéristiques de la plate-forme, comme le nombre de cours ouverts aux inscriptions sur une période donnée.

Table 1. Table des Odd-ratios de la régression logistique liant les caractéristiques d'une inscription et la probabilité qu'elle débouche sur un certificat. N=739.049. Modèle Logit, 1 : Obtention du certificat, 0 : non-obtention. Réf : Référence de l'Odd-ratio ; par exemple, pour le délai entre les dates de lancement des cours, la probabilité d'obtenir le certificat est 1,42 fois plus élevé lorsque l'écart qui sépare les deux cours est compris entre une et deux semaines que lorsqu'ils commencent le même jour (utilisé comme la référence). Dans le modèle, les facteurs ont été entrés dans l'ordre correspondant à la table ci-dessous, c'est-à-dire avec la question du délai entre dates de lancement des cours en premier, et de la temporalité en dernier.

Variable explicative	Odd-Ratio	p-value	significativité
Délai entre dates de lancement des cours			
Le même jour	Réf		
Une semaine	0,95	0,0004	***
Entre une et deux semaines	1,42	0.003	**
Entre deux semaines et un mois	1,60	5.62^{-05}	***
Entre un et deux mois	1,61	5.82^{-05}	***
Plus de deux mois	1,62	$2.\ 10^{-16}$	***
Nombre d'inscriptions réalisées sur la plate-forme			
Entre 1 et 2 cours	Réf		
Entre 2 et 3	1,05	0.017	*
Entre 4 et 5	1,08	5.48^{-05}	***
Entre 6 et 10	1,10	9.92^{-08}	***
Entre 11 et 30	1,21	$<2.\ 10^{-16}$	***
Plus de 30	1,11	9.24^{-06}	***
Caractère groupé de l'inscription			
Inscription isolée	Réf		
Inscription groupée	0,42	$<2.\ 10^{-16}$	***
Délai entre inscription et lancement du cours			
Entre deux et quatre mois après	0,30	$<2.\ 10^{-16}$	***
Entre un et deux mois après	0,63	$<2.\ 10^{-16}$	***
Entre deux semaines et un mois après	0,70	$<2.\ 10^{-16}$	***
Entre une et deux semaines après l'ouverture	0,80	3.62^{-15}	***
Semaine après l'ouverture	1,20	2.93^{-11}	***
Semaine précédant l'ouverture	Réf		
Entre une et deux semaines avant	1,00	0.45	*
Entre deux semaines et un mois avant	0,90	3.35^{-07}	***
Entre un et deux mois avant	0,90	$<2.\ 10^{-16}$	***
Entre deux et quatre mois avant	0,80	$<2.\ 10^{-16}$	***
Plus de quatre mois avant	0,40	$<2.\ 10^{-16}$	***

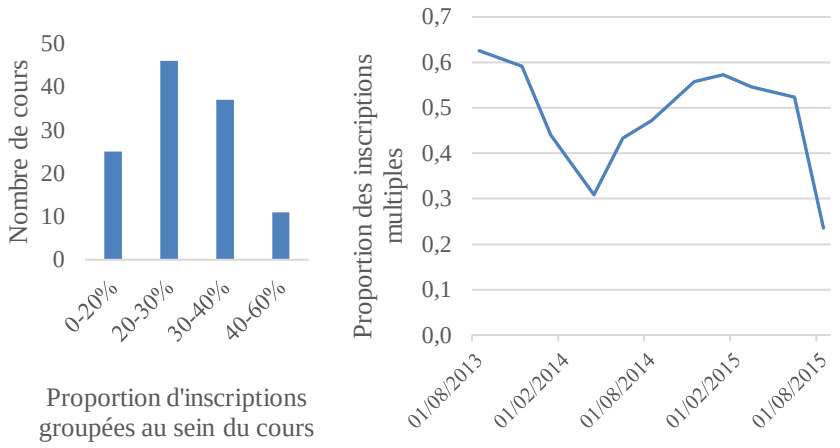

Fig. 1 Gauche : Proportion d'inscriptions groupées au sein du cours, pour 119 MOOC de FUN. Droite : évolution de la proportion d'inscriptions groupées parmi les inscriptions de la plate-forme en fonction de la date d'inscription

4 Discussion

Nos données suggèrent qu'une partie conséquente des inscriptions découlent de visites expérientielles, plus que de visites exploratoires. Cet état de fait expliquerait notamment pourquoi les inscriptions groupées sont si bien représentées dans les données d'inscription. Cette conclusion repose néanmoins sur un certain nombre d'interprétations, sur lesquelles nous nous proposons de revenir maintenant. Nous chercherons en particulier à discuter de la pertinence de la corrélation que nous faisons entre inscription groupée et visite expérientielle. Dans une seconde partie de cette discussion, nous tirerons les conséquences de nos interprétations en termes d'analyse des taux de certification, et de l'utilisation de cet indicateur pour comparer des cours issus de plates-formes distinctes.

Nous avons interprété l'importance des *inscriptions groupées* comme un indicateur de de la visite expérientielle, et donc de la clé d'entrée plate-forme (Isaac & Voole, 2008). Nous avançons que les participants découvrent le plus souvent le MOOC en naviguant sur la plate-forme, et sans avoir eu l'intention de se former sur le sujet avant d'avoir pris connaissance de l'existence de la formation. Appuyons nos propos par une démonstration par l'absurde. Considérons deux cas : soit l'inscription groupée est avant tout un marqueur de visite exploratoire, soit c'est un marqueur de visite expéditive. Comme nous le soulignions plus tôt, une forte représentation des visites exploratoires parmi les *inscriptions groupées* signifierait qu'au cours d'une même journée, un participant donnée ait voulu, sans connaître l'offre de la plate-forme, se former sur plusieurs sujets distincts, et qu'incidemment l'offre de *FUN* ait correspondu à ces différentes attentes. Au vu du faible nombre de cours ouverts aux inscriptions à un

moment donné – tout au plus une vingtaine [4] – une telle situation ne peut qu'être un cas rare.

La prépondérance des visites expéditives parmi les *inscriptions groupées* impliquerait quant à elle que la majorité des participants aient bien connu l'offre de MOOC avant de naviguer sur la plate-forme, et qu'ils ne se connectent que pour s'inscrire successivement à plusieurs cours dont ils avaient entendu parler en amont, par d'autres canaux que la plate-forme d'hébergement. Là encore, ce scénario paraît peu vraisemblable. Nous avons déjà argué que les visites comparatives étaient probablement inexistantes du simple fait que l'offre n'est pas suffisante pour qu'un participant hésite entre plusieurs MOOC pour se former sur une thématique donnée.
D'autres analyses, non présentées ici, nous permettent d'abonder dans ce sens. La première réside dans une enquête ayant récolté plus de 8000 réponses au sein de treize MOOC. Les enquêtés doivent répondre à une question sur les modalités de découverte du cours, qui permet de qualifier la nature de leur visite, et par ailleurs déclarer à combien d'autres cours ils se sont inscrits.

Les participants ayant réalisé une inscription résultant d'une visite expérientielle suivent statistiquement plus de cours en parallèle. Enfin, nous avons analysé les données d'inscription sous l'angle des disciplines académiques associées au cours, celles-ci étant déterminées parmi onze possibles (droit, médecine, etc.). On constate que parmi les inscriptions multiples, les inscriptions groupées sont celles qui sont associées au plus fort *éclectisme disciplinaire*, c'est-à-dire avec un nombre moyen de disciplines différentes plus élevées. Or l'éclectisme disciplinaire est *a priori* davantage associé à la visite expérientielle qu'aux autres types de visites. L'hypothèse inverse signifierait que la recherche concomitante de plusieurs formations dans des disciplines distinctes coïnciderait avec l'offre de FUN à un moment donné.

Nous avons conclu sur le fait que les inscriptions groupées évoluaient au fil du temps, ce qui était susceptible d'affecter les taux de certification d'un cours. Ceci nous amène à affirmer que le taux de certification d'un cours dépend dans une certaine mesure des caractéristiques de la plate-forme qui l'héberge. Cette considération nous pousse à avancer que si l'on souhaite comparer des taux de certification de MOOC de plates-formes différentes, il est préférable de prendre en compte les biais qu'induisent leurs caractéristiques. Ceci passe nécessairement par une réflexion sur les indicateurs susceptibles de rendre compte de la nature des visites effectuées.

En guise de perspective d'analyse, soulignons que l'on pourrait tenter à l'avenir de déterminer l'importance relative des différents types de visites sur la base des données de navigation précédant l'inscription au cours, dans la continuité des travaux de Clow [5] sur le comportement observable précédant l'inscription. 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 davantage qu'une visite comparative, expéditive, ou exploratoire. À l'inverse, une visite pour laquelle le délai entre la connexion à la plate-forme et l'inscription est minimal correspond plus vraisemblablement à une visite expéditive, par définition. Une catégorisation des types de visites sur la base des données de navigation précédant l'inscription pourrait être par

la suite croisée avec le comportement d'inscription et avec les données de certification, afin d'étayer l'hypothèse de la forte corrélation que nous supposons exister entre inscriptions groupées et visite expérientielle. On pourra aussi croiser ces comportements observables avec des intentions déclarées au moment de l'inscription, sur le modèle des travaux de Reich [12].

Concluons sur le fait que la littérature sur les processus d'achat en ligne, et de manière générale, toute littérature consacrée aux comportements des internautes, a sans doute de nombreux autres concepts à offrir à quiconque s'intéresse aux dynamiques de l'apprentissage en ligne. Nous avons tenté d'illustrer cette position avec les concepts de *visite* et de *clé d'entrée*. Peut-être serait-il intéressant de poursuivre cette démarche en allant chercher dans cette littérature d'autres concepts susceptibles d'éclairer d'autres dimensions constitutives de l'apprentissage en ligne. L'offre de cours en ligne gratuits, quelle que soit leur forme, est en évolution rapide. Pour les acteurs toujours plus nombreux qui souhaitent se positionner dans cet écosystème mouvant, il devient de plus en plus fondamental de comprendre comment leurs utilisateurs choisissent ces cours.

Références

1. . Albo, L., Hernández-Leo, D., Oliver, M. Are higher education students registering and participating in MOOCs? (2016) The case of MiríadaX. *Proceedings of EMOOCS 2016,* 197-211.
2. Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2014). Engaging with Massive Online Courses. In *Proceedings of the 23rd International Conference on World Wide Web* (pp. 687–698).
3. Banerjee, A. V., & Duflo, E. (2014). (Dis)organization and Success in an Economics MOOC. *American Economic Review*, *104*(5), 514–18.
4. Cisel, M. (2016). Utilisations des MOOC, éléments de typologie. Manuscrit de thèse non publié.
5. Clow, D. (2013). MOOCs and the funnel of participation (p. 185). ACM Press.
6. Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. *Handbook of mixed methods in social and behavioral research*, 209-240.
7. Ho, A. D., Reich, J., Nesterko, S. O., Seaton, D. T., Mullaney, T., Waldo, J., & Chuang, I. (2014). *HarvardX and MITx: The First Year of Open Online Courses, Fall 2012-Summer 2013*
8. Isaac, H., & Volle, P. (2008). *E-commerce: de la stratégie à la mise en oeuvre opérationnelle*. Pearson Education France.
9. Koller. (2013). Retention and Intention in Massive Open Online Courses: In Depth. Educause.
10. Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers & Education*, *48*(2), 185–204.
11. Perna, L. W., Ruby, A., Boruch, R. F., Wang, N., Scull, J., Ahmad, S., & Evans, C. (2014). Moving Through MOOCs Understanding the Progression of Users in Massive Open Online Courses. *Educational Researcher*, *43*(9), 421–432.

12. Reich, J. (2014) MOOC Completion and Retention in the Context of Student Intent. *EDUCAUSE Review*. Récupéré à l'adresse http://er.educause.edu/articles/2014/12/mooc-completion-and-retention-in-the-context-of-student-intent
13. Stenger, T., & Bourliataux-Lajoinie, S. (2014). *E-marketing & e-commerce - 2e éd.: Concepts, outils, pratiques*. Dunod.