

# The “Idea Advantage”: How Content Sharing Strategies Impact Engagement in Online Learning Platforms

Unnati Narang

Manjit Yadav

Aric Rindfleisch <sup>1</sup>

April 20, 2021

Most updated version available [here](#)

<sup>1</sup>Unnati Narang (unnati@illinois.edu) is an Assistant Professor of Marketing at the Gies College of Business, University of Illinois at Urbana-Champaign. Manjit S. Yadav (myadav@mays.tamu.edu) is JC Penney Chair in Marketing and Retailing Studies, Mays Business School, Texas AM University. Aric Rindfleisch (aric@illinois.edu) is John M. Jones Professor of Marketing at the Gies College of Business, University of Illinois at Urbana-Champaign. We thank Maryalice Wu, Paige Dannielle Cunningham, Sam Cove, Doreen Shen, Dokyun Lee, Emaad Manzoor, Eric Zhou, Jonathan Meer, Adam Bestenbostel, and participants at the 2017 Winter AMA Conference and the 2019 Winter AMA Pre-Conference on Education Marketing for their valuable comments.

## Abstract

In recent years, online learning platforms (e.g., Coursera, edX) have experienced massive growth and have reached over 180 million learners. Although their reach is quite large, the impact of these platforms is constrained by a low level of learner engagement. In traditional face-to-face classrooms, educators seek to engage learners by asking them to participate in class discussions and share information about their identity and ideas. However, the effectiveness of these strategies in online learning platforms is uncertain. The authors examine this issue by assessing the impact of the types of content sharing on learner engagement. The authors conduct a textual analysis of over 12,000 text postings during an 18-month period (Study 1) and a field experiment among over 2,000 learners (Study 2) in a popular Coursera offering by a large U.S. university. The results indicate that asking learners to share ideas related to the course (vs. their identity) has a stronger effect on their video consumption and assessment completion. The authors explain this “idea advantage” by suggesting that learners who share ideas (vs. identity) exhibit a greater degree of elaboration in their postings. This idea advantage is strongest for learners from English-speaking countries and those new to online learning.

**Keywords:** Education, user-generated content, online platforms, engagement, student retention and performance, field experiment, topic modeling, text mining

# 1 Introduction

Higher education is big business. According to recent estimates, the higher education industry earned nearly \$2.5 trillion in global revenue in 2020 and is growing rapidly (HolonIQ 2020). A considerable portion of higher education's future growth is linked to the emergence of new online learning platforms, such as Coursera and edX, which market online courses to over 180 million learners across the globe (Class Central 2020). Over the past decade, the promise of online learning has enticed a growing number of universities to offer not only free courses on these platforms but also paid degrees from prominent institutions such as the University of Colorado, University of Illinois, and University of Michigan (Pickard 2019). These online platforms are poised for further expansion in the wake of the Covid-19 public health crisis (Chronicle 2020).

Despite its increasing prevalence and growing legitimization, online learning is often criticized for a lack of learner engagement (Banerjee and Duflo 2014; Zhenghao et al. 2015). On average, less than 5% of online learners participate in course discussions (Manning and Sanders 2013) and fewer than 10% of those who begin an online course actually complete it (Christensen, Alcorn, and Emanuel 2014). Thus, online learning platforms face a strong need to identify and employ strategies that can enhance learner engagement (Reich and Ruipérez-Valiente 2019). Traditionally, educators have sought to engage learners by asking them to participate in class discussions and share information (i.e., content) about their ideas or their identity (Thompson and Thompson 1996). However, online learning platforms are considerably larger, more diverse, and more impersonal than a face-to-face course (MPI 2018). Therefore, the effectiveness of content sharing in online learning platforms is highly uncertain.

Although online learning platforms are expanding rapidly, little is known about the impact of content sharing on engagement in these platforms in either the marketing or education literature. Research in marketing has focused primarily on understanding why consumers share content online (e.g., Berger and Milkman 2012; Melumad, Inman, and Pham 2019; Toubia and Stephen 2013) However, comparatively less work has examined how sharing

affects the sharers' behavior, such as their level of engagement in an online platform. Moreover, despite substantial diversity among the millions enrolled in online learning platforms, little is known about how the impact of content sharing differs among various types of learners.

Our research seeks to address these knowledge gaps by examining the impact of content sharing on learner engagement in online learning platforms. Specifically, we seek to answer three research questions: How do different types of content sharing (i.e., related to course ideas or learner identity) impact learner engagement? What mechanisms underlie these effects? How do these effects vary across different types of learners? To answer these questions, we conduct a multi-method investigation of learner engagement in a popular Coursera course offered by a large U.S. university. Our first study examines our first research question using data across 18 months of this course and shows that learners who share course-related ideas consume more course videos, complete more assessments, and perform better than learners who share their identity. To obtain measures of the two types of content (i.e., ideas and identity), this study applies textual analyses (i.e., Latent Dirichlet Allocation [LDA] topic modeling and Linguistic Inquiry and Word Count [LIWC]) to over 12,000 learner postings in this course's discussion forum. The study provides a descriptive analysis of the relationship between the learners' textual content and their engagement. However, the causal path could go either way due to a potential latent cause, such as learners' interest in the course. We address causality in our second study.

Our second study is a large-scale field experiment that examines the impact of idea- and identity-related content sharing strategies on a variety of engagement metrics across more than 2,000 learners enrolled in the same course in 2019. This experiment randomly assigned learners to one of three versions of the focal course: the first version invited learners to post ideas related to the course (i.e., idea sharing), the second version invited learners to post their introductions (i.e., identity sharing), and the third (i.e., control) contained no invitation to post.

The results of the two studies show that idea sharing has a stronger relationship with learner

engagement than identity sharing. Specifically, learners asked to share course-related ideas displayed a 31% increase in video consumption and a 30% increase in assessment completion relative to the control group. The results of this study also show that content sharing related to course ideas has a positive relationship with learner performance.<sup>1</sup>

We also investigate the potential mechanism underlying this “idea advantage” through a fine-grained analysis of the learners’ textual responses. This analysis reveals that idea sharing entails a higher degree of elaboration, relative to identity sharing. Hence, learner elaboration appears to underlie the idea advantage. The results of this study also uncover considerable heterogeneity in the idea advantage across learners and indicate that this advantage is strongest for learners from English-speaking countries and those new to online learning.

Collectively, our research offers four contributions to the marketing and education literature on content sharing. First, our research provides the first textual investigation of the types of content (i.e., ideas and identity) shared in online learning platforms. Second, our research offers an investigation of the effect of different types of content sharing strategies on learner engagement metrics, including video consumption, assessment completion, and course performance. Third, our research uncovers an idea advantage and suggests that this advantage may stem from enhanced elaboration when sharing ideas (vs. identity) in an online setting. We also uncover heterogeneity across learners by showing that idea sharing has a stronger impact on engagement among learners from English-speaking countries and those new to online learning. These findings can potentially help institutions and instructors design more inclusive, engaging, and personalized learning approaches in an online setting. Finally, our research demonstrates that inviting learners to share course-related ideas is a low-cost and effective means of enhancing learner engagement. Thus, our research identifies a potentially impactful driver of online learning engagement (i.e., idea sharing), while also uncovering an easily implementable technique (i.e., an invitation to share ideas) for enhancing learner engagement.

---

<sup>1</sup>In both Study 1 and Study 2, learners who post course-related ideas (vs. their identity) show a higher performance. In Study 2, there is no intent-to-treat (ITT) effect of asking learners to post ideas on their course performance. However, learners who engage in posting exhibit results that are quite similar to what was found in Study 1.

## **2 Related Literature and Theory Development**

Our theoretical foundation is grounded in three important building blocks. We start by examining prior research on content sharing in learning environments (both face-to-face and online) and suggest that learners are likely to post two types of content (i.e., ideas related to the course and information about their identity). We then explore insights from (and limitations of) prior research in education and marketing that focuses on content sharing. Finally, we examine how different types of content may influence learner engagement in online learning platforms.

### **2.1 Content Sharing in Learning Environments**

Content sharing is ubiquitous in learning environments. In a traditional face-to-face (F2F) learning environment, content sharing is frequently employed as a strategy to engage learners and advance their learning outcomes. Prior research suggests that sharing various types of content in F2F classrooms strengthens a variety of learning outcomes. For example, Steinberg, Empson and Carpenter (2004) show that asking students to explain how they solved math problems improves their understanding of course content. Likewise, Thompson and Thompson (1996) suggest that informal ice-breaking interactions, such as personal introductions, improve student participation and social integration. In sum, content sharing in F2F learning environments is often highly interactive and positively related to learner engagement. However, the degree to which these findings translate to an online learning context is uncertain.

In contrast to the intimacy, synchronicity and physicality of an F2F learning context, online learning environments are typically more distant, asynchronous and digital. For example, many online courses enroll thousands of learners who view pre-recorded videos in isolation from one another and have little, if any, direct access to the instructor. In this type of setting, content sharing is largely restricted to textual interactions in the form of posts in a discussion forum (Chaturvedi, Goldwasser and Daumé 2014). Moreover, most of this content sharing is one-way and elicits little or no response from either peers or instructors (Manning and Sanders 2013).

Despite these notable differences, an online classroom has many of the same accoutrements (e.g., instructors, students, course materials) of a traditional F2F classroom. Both environments are, in essence, contexts (i.e. classrooms) in which the delivery of education occurs. Hence, students in both types of classrooms are likely to exhibit a common set of expectations and behaviors (Schank and Abelson 1977). Consequently, online learners are likely to form expectations and enact behaviors normally associated with traditional F2F classrooms. Despite the differences between these two learning modalities, content sharing practices in online learning platforms should display some degree of similarity with content sharing in F2F classrooms. In particular, much like F2F learners, online learners are also likely to share information about either themselves (i.e., identity sharing) or about the course (i.e., idea sharing). In our research context, identity sharing refers to the act of disclosing information about oneself such as one's name, age, birthplace, or occupation. In contrast, idea sharing refers to the act of offering an opinion, observation, or insight about course topics.

## **2.2 Prior Literature: Key Insights and Limitations**

*Education Literature on Online Learning.* Research on the relationship between content sharing and engagement in online learning platforms is quickly emerging but is still in an early stage of development. As a result, the literature on online learning is largely descriptive and/or correlational (see Web Appendix Table A1 for details). For example, Reich and Ruipérez-Valiente (2019) document completion rates in online courses across various countries. While a handful of field experiments have attempted to enhance online course engagement using a variety of interventions (e.g., Kizilcec et al. 2014; Lamb et al. 2015; Patterson 2018; Zhang, Allon, and Mieghem 2017), most of these interventions are unrelated to the type of content shared. Some of these studies examine content sharing but are largely agnostic about the specific type of content shared. For example, Kizilcec et al. (2014) study the effect of sending randomized email invitations, which highlight different benefits of content sharing, on online forum participation but do not distinguish between the different

types of content.

The emerging education literature on e-journaling (see Dwyer and Davidson 2020 for review) provides added insights on how content generation impacts those who generate the content. Journaling typically involves a writing activity in which students reflect on a specific set of topics (O'Connell and Dymont 2006). Prior research suggests that these types of reflective writing activities may enhance the self-efficacy of students engaged in journaling their thoughts and ideas (Fritson 2008). The evidence from this literature suggests a connection between content generation (which must occur before any sharing) and subsequent engagement. However, little is known in the e-journaling literature about the impact of generating and sharing such reflections in online learning platforms.

*Marketing Literature on Online Platforms and User-Generated Content (UGC).* Prior research in marketing focuses primarily on understanding the drivers of content sharing and the type of content that is shared (e.g., Berger and Milkman 2012; Melumad, Inman, and Pham 2019; Toubia and Stephen 2013). In contrast, our research examines the outcomes of content sharing, such as video consumption, assessment completion, and course performance, with a focus on the type of content that is shared. Although a few UGC studies assess the outcomes of content sharing, these studies typically do not examine the specific topics being shared (e.g., Babić Rosario et al. 2016; Schweidel and Moe 2014; Tirunillai and Tellis 2012). In addition, our research also assesses how different types of content sharing interventions affect the sharer rather than the reader. Prior research in this domain typically focuses on the impact of content sharing on readers rather than sharers (e.g., Liu, Lee, and Srinivasan 2019; Zhang, Allon, and Mieghem 2017). One notable exception is a recent study by Berman et al. (2019), which examines how content sharing impacts both sharers and receivers, albeit in a much different setting. This research offers a descriptive examination of the content shared on Twitter during and after a live election debate and drivers of the tweets' popularity. This research suggests that sharing content on Twitter during an election debate is positively associated with engagement with the live event among sharers (Berman et al. 2019; see also Houston et al.



2013). Thus, the mere act of online posting about an event appears to enhance sharer engagement. Building on this underlying rationale, we propose that a similar engagement boost may occur among content sharers in an online learning context and test this thesis using a field experiment (i.e., Study 2).

In sum, prior research in education and marketing is useful for framing our investigation but provides few directly applicable insights for our inquiry. In general, prior research views sharing as an outcome with little focus on its impact on other outcomes, such as engagement. Moreover, the extant literature in this domain largely ignores both the type of content shared as well as its impact on the sharer. Our research extends and enriches this literature in marketing and education by providing the first investigation of the effect of types of content sharing on important engagement metrics in online learning platforms.

### **2.3 The Effect of Types of Content Sharing in Online Learning Platforms**

Thus far, our conceptualization has suggested that learners in online learning platforms largely post content focused on either identity or ideas. We have also documented that the literature in education and marketing provides limited insights about the impact of different types of content sharing in online platforms. Next, we focus on these two types of content learners share (i.e., ideas and identity) and how they may relate to engagement in an online setting.

We propose that idea sharing will be more strongly related to learner engagement than identity sharing because the expression of ideas entails a greater degree of elaboration (e.g., Homburg, Ehm, and Arts 2015). We define elaboration as the amount of information contained in a learner's posting (e.g., Melumad, Inman, and Pham 2019; Moon 2000). Prior research suggests that in online settings, individuals typically restrict the disclosure of identity-related content to statements about one's name, age, or birthplace (Berman and Bruckman 2001). These types of statements can be expressed with little need for elaboration. In contrast, the disclosure of idea-related content, such as ideas about a new concept is typically broader in scope and more complex and analytical in nature, which should require greater elaboration

(Smith and DeCoster 2000). Thus, we expect the sharing of idea-related (vs. identity-related) content to spur higher levels of learner engagement and course performance by requiring greater elaboration and forging a deeper connection to course-related topics.<sup>2</sup>

If our theoretical reasoning has merit, idea-related postings should display a higher degree of elaboration compared to identity-related postings and potentially spur more engagement. Therefore, our empirical investigation examines not only the type of content but also the linguistic features of this content (i.e., elaboration). Prior research suggests that linguistic features can reveal important behavioral insights (e.g., Hartmann et al. 2019; Netzer, Lemaire, and Herzenstein 2019). Specifically, we focus on linguistic features related to an individual's degree of elaboration, such as the number of words, words per sentence, and complexity (e.g., Melumad, Inman, and Pham 2019; Moon 2000).

Finally, we expect the effects of idea and identity sharing to vary across different types of learners. As noted earlier, sharing ideas (relative to identity) should require greater elaboration, and hence, a higher level of familiarity and comfort with the language in which a course is offered. Therefore, we expect learners who are more familiar and comfortable with the course language (i.e., English) to engage in greater elaboration and exhibit a higher degree of engagement. We also expect the effects of these two types of content sharing to vary based on a learner's prior experience in online learning. In particular, we consider two competing possibilities regarding the role of prior online learning experience. On the one hand, learners with prior online learning experience should be more adept at navigating an online platform. Thus, these learners should be able to engage in idea sharing more easily, which should enhance their degree of engagement. According to this perspective, prior online learning experience should strengthen the effect of idea sharing on engagement. On the other hand, learners lacking prior online learning experience may display a greater degree of enthusiasm due to the novelty of the experience (Kim, Costello, and Lee 2020). As a result, these learners

---

<sup>2</sup>It is also possible that greater engagement among those immersed in the course content drives greater elaboration in idea sharing.

may be more willing to engage in greater elaboration, which should enhance their engagement with the course. According to this perspective, prior online learning experience should weaken the effect of idea sharing on engagement. Given these competing possibilities, the effect of prior online learning experience on engagement remains an empirical question.

We next present two studies aimed at examining the effects of idea- and identity-related content sharing in online learning platforms. Collectively, these studies assess the relative impact of idea and identity sharing, explore the underlying mechanisms, and examine the underlying heterogeneity of the effects across different types of learners.

### **3 Research Context and Setting**

Given our research goal of understanding the impact of types of content sharing in online learning platforms, the context of our study is Coursera, the world's largest massive open online course (MOOC) platform. Coursera was established in 2012 with the mission of democratizing higher education by lowering cost and increasing access. In order to accomplish this goal, Coursera has established strategic partnerships with over 200 educational partners, including large prestigious universities, such as Stanford University, Yale University, and the University of Pennsylvania. In conjunction with these partners, Coursera currently offers nearly 6,000 courses across a wide range of disciplines (e.g., business, engineering, humanities, law, medicine) to over 75 million learners across the world. Since April 2020, the platform has added 21 million learners and grown rapidly in the wake of the Covid-19 pandemic (Conversation 2020).

We enlisted the help of Coursera to conduct two studies. Both studies focus on a popular Coursera course on digital marketing offered by a large U.S. university. The content of the course covers four broad themes: product, promotion, placement, and pricing. The course is four weeks (modules) long and contains 41 instructional videos and 12 assessments (i.e., cases, exercises and quizzes) that learners need to successfully complete in order to pass the

course. In accord with Coursera policy, in order to successfully complete this course, learners must also pay a small monthly subscription fee. Thus, our analysis focuses on paid learners, as these are the only type of learners eligible to complete the course.<sup>3</sup> Our focus on a single course enhances the internal validity of our research by avoiding confounds due to differences in instructors and course material (Zhang, Allon, and Mieghem 2017). In our first study, we conducted a textual analysis of learner postings in order to quantify idea and identity sharing and assess their relationship with learner outcomes. In our second study, we conducted a field experiment that investigates these relationships by assigning learners to one of three conditions: (1) a prompt to share ideas, (2) a prompt to share identity or (3) no prompt. We next describe the two studies and discuss their findings and implications.

## **4 Study 1: Archival Data**

Our first study provides preliminary (i.e., descriptive) insights into our first research question about the impact of different types of content sharing on learner engagement. To address this research question, we first develop proxy measures of both idea and identity sharing via textual analyses. We then estimate a set of regression models that assess the relationship of the types of content sharing with a set of key learner engagement metrics (i.e., video consumption, assessment completion, and course performance) to understand the association between idea (vs. identity) sharing and learner engagement.

The data for this study come from our focal course across an 18-month period (i.e., March 2016 to August 2017). During this period, this course enrolled 18,695 paid learners. As shown in Table 1, on average, these learners viewed 14.75 videos, completed 4.45 assessments, and earned an average grade of 84%. Demographically, 21% of the learners were full-time students, 39% were females, and 36% had at least a Masters' degree. These demographics for this course closely match the demographics of Coursera's broader learner population across thousands of

---

<sup>3</sup>Paid users spend more time in forums and post more than unpaid users (Goli, Chintagunta, and Sriram 2019).

courses. The textual data for this study come from the course discussion forum. This forum is prominently displayed in the course menu and provides an opportunity for learners to post (and read) a variety of comments. Of the total learners, 4,676 participated in the forum.

## **4.1 Text Mining: Methodology and Results**

We analyze the textual content derived from discussion forum postings via a multimethod approach that employed a combination of probabilistic topic models (i.e., Latent Dirichlet Allocation or LDA) (Blei, Ng, and Jordan 2003) and dictionary-based methods (i.e., Linguistic Inquiry and Word Count or LIWC) (Pennebaker et al. 2014). The goal of the LDA model is to recover the types of content contained in the textual data. This allows us to quantify idea- and identity-related content. We then apply LIWC in order to obtain an indicant of the degree of elaboration using relevant linguistic features (e.g., word count, words per sentence).

We first discuss the LDA model. Following Blei, Ng, and Jordan (2003), we employ an LDA model to uncover and extract key types of content from learner postings. Before applying this model, we removed rarely occurring words (e.g., chess, merrier) from our data (e.g., Huang et al. 2018; Netzer, Lemaire, and Herzenstein 2019). After this removal, we ended up with a final list of 1,771 words that appeared most frequently. As noted by Berger et al. (2020), an LDA model requires the number of topics to be selected by the researcher and this determination should be based on the degree of fit (i.e., perplexity score). The lower the perplexity score, the higher the fit of the model. Thus, the optimal model can be identified by the number of topics that produce the lowest perplexity score. We varied the number of topics from one to five and found that a two-topic model produced the best fit and lowest perplexity score (see Web Appendix Table B1). These two topics reveal two distinct sets of words. The most frequent words for Topic 1 include “product, idea, customer, https, company, video, review” while the most frequent words for Topic 2 include “look, year, busy, hello, current, hope, forward.”

The distinction between these two topics is illustrated by the following two postings drawn from the discussion forums:

Topic 1: “In the telecom sector, some firms have developed a web based after-sales service platforms. A customer can post his problem or request on the platform and other customers who know the solution can respond. Customers can vote on the best solutions published and customers that solve the most problems are rewarded. This type of co-creation, or co-support, allows the firm to reduce support charges: call center, technical support...”

Topic 2: “Hello dudes! Greetings from Mauritius! My reason behind going for Digital Marketing Specialization is to be able to take a new challenge from my current job. I am in a Destination Management Company which is now developing in the digital era. We have a website and currently an application but it is kind of average even though the traditional marketing still working. Therefore, I want to upgrade my skills, look forward to making new friends here.”

Based on our earlier conceptual discussion of online learning platforms, these two topics appear to reflect the two types of content we expected learners to share. Specifically, Topic 1 appears to suggest the sharing of ideas, while Topic 2 seems centered on the sharing of identity. These two types of content and their most frequently occurring keywords are reported in Table 2 and depicted in Figure 1.<sup>4</sup> In addition to words from the LDA model, Table 2 also reports the linguistic features of these two types of content using LIWC. As shown in this table, idea-related content displays a higher number of words per sentence, dictionary words, cognitive words, and lower levels of authenticity and pronoun usage relative to identity-related content. In contrast, identity-related content focuses more on the present and on personal experiences (Pennebaker, Booth, and Francis 2007). In sum, Topic 1 appears to indicate a focus on course ideas, while Topic 2 appears to indicate a focus on learner identity and personal introductions. Based on these linguistic differences, idea-related content appears to entail a greater degree of elaboration (e.g., word count, words per sentence) than identity-related content, consistent with our thesis.

To validate our interpretation of these two types of content, we conducted a coding analysis of a randomly selected subset of 120 forum postings from our focal course and compared the outcome of this analysis with our LDA model results (Huang et al. 2018). Specifically, we asked two graduate students who were unfamiliar with our research objectives to independently

---

<sup>4</sup>The frequencies of the top words for each topic appear in Web Appendix Figures B1 (Study 1) and B2 (Study 2).

read this set of postings and label each posting as an example of a learner sharing either an idea or their identity. The level of agreement among these coders was high (93%) and this coding analysis displayed a 96% level of concurrence with the LDA assessment.

To verify the topics recovered by the LDA model, we also conducted an online validation study that invited 298 participants (via Prolific) to share words or phrases associated with either idea or identity sharing (see Web Appendix E). Demographically, 29% of the participants were full-time students, 44% were females, and 19% had a Masters' degree. A descriptive analysis of these data reveals that 74% of the words shared in the idea condition appear in posts classified by the LDA as ideas (Topic 1), while 85% of the words shared in the identity condition appear in the posts classified by the LDA as identity (Topic 2) (see Web Appendix Figure B3). These robustness checks lend added confidence to our interpretation of the LDA results.

## 4.2 Impact of Type of Content on Learner Engagement

In order to obtain an initial assessment of the degree to which the two types of content (i.e., idea vs. identity) relate to learner engagement, we conduct a set of ordinary least square (OLS) regressions.<sup>5</sup> Specifically, our regressions employ the following form:

$$Y_i = \beta_0 + \beta_1 IdeaSharing_i + \beta_2 X_i + Cohort_t_i + \epsilon_i \quad (1)$$

where  $Y$  is the dependent variable denoting engagement (operationalized as video consumption, assessment completion, and course performance) by individual learner  $i$  during the duration of the course.  $IdeaSharing$  is the proportion of content related to ideas (relative to identity) across a learners' postings based on the LDA model.  $X$  represent a vector of covariates, including the learner's status as a full-time student or not, gender (1 = female, 0 otherwise), and highest level of education (1 = Masters' level, 0 otherwise). We control for the total number of posts made by a learner. We include cohort-level fixed effects to control for

---

<sup>5</sup>For this analysis, we use the aggregate number of postings and aggregate engagement for each individual who posted in the course discussion forum over the entire duration of the course.

unobservables specific to each cohort. We also include country and language fixed effects. Since we are interested in the relative effect of idea and identity sharing for those who post, this regression is limited to learners who posted one or more comments in the course discussion forum.

The results of these regressions are reported in Table 3. As shown in this table, learners who post ideas (relative to those who post identity) display significantly higher levels of video consumption ( $\beta_1 = 2.27$ ,  $t = 3.35$ ,  $p < .001$ ), assessment completion ( $\beta_1 = 1.07$ ,  $t = 3.67$ ,  $p < .001$ ) and course performance ( $\beta_1 = .09$ ,  $t = 4.24$ ,  $p < .001$ ). These results translate to a 15% increase in video consumption, 24% increase in assessments completed, and an 11% increase in course performance under idea sharing relative to identity sharing. These results suggest that the sharing of ideas is associated with greater levels of engagement than the sharing of identity.<sup>6</sup>

Although this analysis provides some support for our thesis regarding the superior effect of idea sharing, the lack of random assignment limits our ability to derive causal inference. Since learners who are highly engaged in a course may be more likely to post ideas about the course, the direction of causality is uncertain. Thus, we conducted a field experiment (Study 2) that randomly assigned learners to an idea- vs. identity- sharing (vs. control) condition in order to investigate the effect of the types of content on engagement and its underlying mechanisms.

## 5 Study 2: Field Experiment

Our second study investigates the effect of different types of content sharing on learner engagement (i.e., video consumption, assessment completion, and course performance) via a randomized field experiment. This study also explores the underlying mechanism behind these effects as well as the degree to which our findings vary across different types of learners. This experiment was conducted during September and October 2019 in the same focal course that was employed for Study 1. During this period, this course was offered twice and enrolled a

---

<sup>6</sup>The percentages are computed as the coefficient of idea sharing in Table 3 divided by the average value of the outcome in Table 1. Thus, the effect for video consumption is  $2.27/14.75 = 15\%$  for idea (vs. identity) sharing.



total of 2,122 paid learners. Summary statistics for this study are reported in Table 4. On average, learners consumed 6.70 videos, completed 1.92 assessments, and earned a grade of 82%. Demographically, 51% of the learners in this course were full-time students, 36% were females, and 24% earned at least a Masters' degree.

## **5.1 Experimental Design**

This field study employed a between-subjects design in which learners were randomly assigned to one of three conditions: (1) an idea-sharing condition that invited learners to post ideas related to the course, (2) an identity-sharing condition that invited learners to introduce themselves, or (3) a control condition without any invitation to post. In essence, we created three separate versions of the course that were identical, except for their treatment condition (i.e., type of sharing). In the idea sharing condition, learners were presented with a 30-second video prompt near the start of the course in which the instructor encouraged them to share ideas related to the course. In the identity-sharing condition, learners were presented with a 30-second video prompt near the start of the course in which the instructor encouraged them to introduce themselves. In both conditions, the prompt was only shown once, and learners were provided the opportunity to share content with other learners. The wording of the video prompts in the two conditions is provided in Figure 2. In the control condition, no prompt was shown.

This type of experimental approach (i.e., randomized encouragement design) is commonly used in the economics, education and medical literature (Duflo and Saez 2003; Hirano et al. 2000; Rubin 1974). In this type of design, participants are randomly invited to engage in different types of behaviors. While assignment to the different conditions is random, the choice to comply with the requested behavior (i.e., compliance) is determined by the participants (Rubin 1974). In our experiment, 29% of learners in the identity sharing condition and 24% of learners in the idea sharing condition complied with our invitation by posting a response to the prompt. This degree of compliance is consistent with other studies that employ

this approach (e.g., Duflo and Saez 2003).

Since only a subset of our participants complied with our requested behavior, we present two sets of analyses for our experiment. First, we compare the outcomes of learners assigned to each condition. These estimates represent the intent-to-treat (ITT) effect of the prompt, rather than the treatment effect of the actual act of posting. Second, we compare only those learners who complied with each prompt (i.e., posted a response), which allows us to assess the results for those who complied and the possible underlying mechanism by analyzing the textual content of their posted response.<sup>7</sup>

## 5.2 Validation Tests

In order to ensure that our experimental manipulation worked as intended, we conducted a series of validation tests. First, to verify that our conditions were randomized, we compared learner demographics (i.e., student status, gender, education level) and past behaviors (i.e., prior video consumption, assessment completion, course performance) across the three conditions. As shown in Figure 3, these indicants are similar across all conditions, which suggests that randomization was successfully achieved.

Second, to ensure that the content posted within each condition conformed to the type of prompt learners were assigned, we analyzed the content of learners' postings in the idea vs. identity conditions using both LDA and LIWC. Similar to Study 1, the LDA model identified two as the optimal number of topics, which offers additional validation for our initial results. The topics and keywords for the LDA are reported in Table 5 and depicted in Figure 4. As shown in this table and figure, the most frequent words for Topic 1 included "consume, product, world, customer, online, firm, company," while the most frequent words for Topic 2 included "course, work, year, current, learn, manage, India, image, hello." In essence, Topic 1's keywords appear to focus on the course, while Topic 2's keywords appear to focus on the learner. Table 5 also reports the differences in the linguistic features of these two topics using LIWC. As shown

---

<sup>7</sup>Since learners self-select into compliance, we caution against a causal interpretation of these latter estimates.

in this table, idea-related content displays higher word count, words per sentence, dictionary words, analytical thinking, cognitive words, and lower authenticity and use of pronouns relative to identity-related content (consistent with Study 1). These linguistic features associated with Topic 1 are suggestive of a greater level of elaboration compared to those associated with Topic 2. Overall, the LDA model successfully classified 95% of the posts in the idea-sharing condition as ideas (Topic 1) and 96% of the posts in the identity-sharing condition as identity (Topic 2). This pattern of results is congruent with our conceptualization and indicates that learner postings are closely aligned with what they were prompted to post.

### 5.3 Empirical Analysis of the Impact of Idea vs. Identity Sharing

We begin our assessment of the impact of idea- vs. identity-sharing prompts on learner engagement by conducting a model-free analysis across the three different conditions. As shown in Figure 5, learners in the idea-sharing condition consumed 21% more video content (Mean<sub>(Idea)</sub>: 7.33, Mean<sub>(Identity)</sub>: 6.07,  $p = .04$ ), and completed 25% more assessments (Mean<sub>(Idea)</sub>: 2.13, Mean<sub>(Identity)</sub>: 1.71,  $p = .04$ ) compared to learners in the identity-sharing condition. These learners also consumed 36% more video content (Mean<sub>(Idea)</sub>: 7.33, Mean<sub>(Control)</sub>: 5.38,  $p < .001$ ) and completed 32% more assessments (Mean<sub>(Idea)</sub>: 2.13, Mean<sub>(Control)</sub>: 1.61,  $p = .01$ ) compared to learners in the control condition. In contrast, there were no significant differences between the identity vs. control conditions ( $p > .1$ ) for any of the three outcomes (i.e., video consumption, assessment completion, course performance).

Next, we formally model the impact of idea vs identity sharing on learner engagement by employing an OLS regression of the following form:

$$Y_i = \gamma_0 + \gamma_1 \text{IdeaPrompt}_i + \gamma_2 \text{IdentityPrompt}_i + \gamma_3 X_i + \text{Cohort}_i + \epsilon_i \quad (2)$$

where Y is the dependent variable denoting engagement (operationalized as video consumption, assessment completion, and course performance) by individual learner i. IdeaPrompt refers to

learners assigned to the idea condition, while IdentityPrompt refers to learners assigned to the identity condition. The baseline (or control) condition denotes learners who were not prompted to share. We also include country, language and cohort fixed effects in this estimation. The results of this regression represent the intent-to-treat (ITT) effects. This approach is consistent with Chetty, Hendren, and Katz (2016), who examine ITT effects by comparing outcomes for families assigned to vouchers with families not assigned to vouchers irrespective of their actual usage. Similarly, Jacob and Ludwig (2012) evaluate the ITT effect of assigning specific housing assistance vouchers on employment rather than the actual use of vouchers.

## 5.4 Results and Discussion

The results of our regression model are provided in Table 6. As shown in this table, the idea-sharing prompt has a positive effect on video consumption ( $\gamma_1 = 2.08$ ,  $t = 3.38$ ,  $p < .001$ ) and assessment completion ( $\gamma_1 = .57$ ,  $t = 2.80$ ,  $p = .005$ ) relative to the control condition. These results translate to a 31% increase in video consumption and 30% increase in assessments completed under idea sharing relative to the control condition. In contrast, there is no significant effect of the identity-sharing prompt for any of the three indicants of engagement. Furthermore, neither type of content sharing condition appears to impact course performance. The results of this analysis align with our model-free evidence and suggest that a brief prompt asking learners to share ideas has a stronger impact on engagement than not having such a prompt. In addition, a prompt asking learners to share ideas seems superior to a prompt to share their identity.<sup>8</sup>

In addition to examining the effects for those assigned to idea vs. identity sharing, we also examine the effects for the subset of learners who posted in response to our prompt (i.e., compliers). In general, a comparison of means indicates that compliers (vs. non-compliers) display higher engagement in terms of video consumption ( $Mean_{Complier}$ : 16.01,

---

<sup>8</sup>Only 24% of learners responded to the idea prompt, while our intent-to-treat effects are computed for all learners. Since the effects are localized to a small proportion of learners, the complier average causal effect (CACE) for those who post is likely to be higher (Angrist, Imbens, and Rubin 1996; Stuart et al. 2008). The CACE for idea sharing can be computed by scaling the ITT by the percentage of compliers (i.e., 2.08 divided by 24% = 8.67).

$Mean_{(Non-complier)}$ : 3.34,  $p < .001$ ), assessment completion ( $Mean_{(Complier)}$ : 4.65,  $Mean_{(Non-complier)}$ : .93,  $p < .001$ ), and course performance ( $Mean_{(Complier)}$ : .84,  $Mean_{(Non-complier)}$ : .83,  $p > .10$ ).

In order to provide a finer-grained investigation of the relative effect of idea vs. identity sharing among compliers, we re-estimated the regression model from Equation 2 among this set of learners. The results of this analysis are reported in Table 7. As shown in this table, learners who posted in response to the idea-sharing prompt (relative to an identity-sharing prompt) displayed significant positive effects across all three metrics of learner engagement: video consumption ( $\gamma_1 = 3.89$ ,  $t = 2.32$ ,  $p = .02$ ), assessment completion ( $\gamma_1 = 1.57$ ,  $t = 2.66$ ,  $p = .01$ ), and course performance ( $\gamma_1 = .10$ ,  $t = 2.18$ ,  $p = .03$ ). Compared with the effects across all the learners (reported earlier in Table 6), the results among compliers reveal that idea sharing has a positive relationship with not only video consumption and assessment completion, but also course performance. Thus, the actual posting of ideas (vs. identity), rather than simply being assigned to an invitation to post, is significantly associated with performance. This result is consistent with the descriptive evidence from Study 1.

So far, we have examined our first research question about the effects of different types of content sharing and uncovered a distinct advantage for idea sharing relative to identity sharing. Now that we have identified this idea advantage, we seek to answer our second and third research questions by examining the potential mechanisms underlying this advantage as well as the degree to which the effect varies across different types of learners.

## 5.5 Underlying Mechanisms

According to our conceptualization, sharing ideas should elicit greater elaboration compared with sharing identity. This enhanced elaboration, in turn, is likely to have a beneficial effect in terms of a higher degree of learner engagement (e.g., Valsesia, Proserpio, and Nunes 2020). The descriptive results from the textual analyses conducted in Study 1 provide preliminary evidence for this thesis. In this study, we provide further assessment of this proposed mechanism by

conducting a series of additional analyses.

To directly investigate whether learners who share ideas (relative to identity) engage in greater elaboration, we compare the length of their postings in response to our prompt across the two conditions (i.e., ideas and identity). This analysis reveals that learners who shared ideas display significantly greater elaboration compared to those who shared identity ( $Mean_{(length)} = 66$  words,  $Mean_{(length)} = 37$  words,  $p < .001$ ). Thus, learners prompted to share ideas appear to post more elaborate (i.e., longer) responses. As noted in our conceptualization, this greater degree of elaboration could be a potential explanation for their higher degree of engagement.

In addition to examining the length of idea- and identity-related posts, we also examine elaboration using the relevant linguistic features of learners' textual postings based on a LIWC analysis (Pennebaker, Booth, and Francis 2007). As shown in Table 8, idea-related postings tend to be more elaborate (e.g., higher dictionary words, higher words per sentence) and more analytical; 84% of the content in idea-related postings reflect analytical thinking (i.e., logical, structured thoughts), while only 75% of the content in identity-sharing postings reflects this type of thinking. In contrast, identity-related postings are more authentic and use more pronouns in their linguistic style. Finally, the cognitive processes associated with idea- and identity-related postings also vary greatly. Specifically, the words employed in idea postings appear to be more cognitive (idea = 9.5 vs. identity = 4.2) in nature. This pattern of results is congruent with our conceptualization that idea-related postings are more complex and analytical in nature, and thus, entail more elaboration.

So far, our analysis has focused on learners who post in response to an idea or identity prompt (i.e., compliers) and the level of elaboration in their responses. In addition to analyzing how and why idea sharing impacts participants who complied with our prompts, we also examine the outcomes of learners who did not comply (i.e., those who did not engage in any type of content sharing activity). The goal of this analysis is to examine if the act of posting and a higher degree of elaboration is necessary to benefit from a prompt, or if there is an effect of simply being exposed to a content-sharing prompt. To conduct this analysis, we

re-estimated our OLS regression among the subsample of learners assigned to either the idea or the identity prompt who did not post a response (i.e., non-compliers). The results of this analysis appear in Table 9. As shown in this table, while the idea prompt exerts a positive effect on video consumption ( $p = .01$ ), it has no impact upon either assessment completion or course performance among this set of learners. This pattern of results suggests that there is some value (at least for video consumption) of simply exposing learners to a prompt even though the effects appear to be mainly driven by those who comply (i.e., actually post).

## **5.6 Effect of Learner Heterogeneity**

Next, we explore the degree to which the idea advantage varies across different types of learners. Specifically, we test our proposition that idea sharing should have differential effects based on learners' language (i.e., English speaking or not) and prior online learning experience.

**Language.** Since sharing ideas requires a greater ability to communicate and articulate complex thoughts compared to sharing information about one's identity, language fluency should play a greater role for idea sharing than for identity sharing (e.g., Lee and Kronrod 2020; Packard and Berger 2017). Hence, learners from countries where English is the primary language (e.g., US, Canada, England) should exhibit a greater idea advantage (for a U.S.-based course) compared with learners from countries where English is not the primary language (e.g., Brazil, China, Germany). In essence, learners who are native English speakers should be more willing and able to express course-related ideas (Conrad 2002; Thompson and Thompson 1996). To assess this proposition, we examined the interaction of the sharing condition (idea vs. identity) with a dummy variable that indicates whether a learner is from an English-speaking country or not (1 or 0); the outcome variables are video consumption, assessment completion, and course performance. As shown in Table 10, learners from English-speaking countries exhibit stronger effects for idea sharing relative to non-English-speaking countries. Thus, the idea advantage appears to be more pronounced among learners fluent in English.

Experience. Learners with prior online learning experience may be more adept at navigating the online learning platform due to their familiarity with the platform from prior experience in other courses. However, the impact of prior experience on the idea advantage could be either positive or negative. If experience helps learners gain greater comfort in terms of sharing ideas, it could have a positive effect. On the other hand, if inexperienced learners are more enthusiastic about idea sharing due to novelty, experience could have a negative effect. To assess these competing possibilities, we examined the interaction of sharing condition (idea vs. identity) with a dummy variable that indicates whether a learner has prior online learning experience or not (1 or 0); the outcome variables are video consumption, assessment completion, and course performance. As shown in Table 10, the interaction between prior online experience and idea sharing is negative for all three indicants of learner engagement (i.e., video consumption, assessment completion, and course performance). Thus, the idea advantage appears to be more pronounced for learners who are new to online learning platforms.

## **5.7 Robustness Checks and Additional Analyses**

In this section, we outline three additional analyses that assess the robustness of our results for both Study 1 and Study 2. First, we used a deep learning-based text mining algorithm, the Focused Concept Miner (FCM) (Lee, Manzoor, and Cheng 2018) to assess the robustness of the textual analysis in both studies. Compared to an LDA model, this emerging method generates concepts that are more interpretable (Chang et al. 2007). In contrast with LDA, which is an unsupervised method that identifies topics probabilistically, FCM offers a supervised learning approach that identifies concepts in the text that are correlated with and predictive of the outcome. For example, when applied to the text associated with product reviews, FCM extracts concepts that are predictive of sales (e.g., service quality, price). In our setting, FCM should be able to identify the type of postings associated with higher levels of learner engagement.

The results of the FCM analysis are shown in Web Appendix Table C1 (for Study 1) and



Table D1 (for Study 2). As shown in these tables, the keywords associated with Concept 1 are “product, company, media, experience, inform, look, people, easier” in Study 1 and “reach, search, need, purchase, just, help, brand, want” in Study 2. The keywords associated with Concept 2 are “help, age, engage, consume, time, year, easier, custom” in Study 1 and “work, business, experience, manage, custom, current, look, year” in Study 2. When compared with the results of the LDA model, several words for Concept 1 overlap with words from posts classified as ideas by the LDA (e.g., product, company, brand). Likewise, several words for Concept 2 overlap with words from posts classified as identity by the LDA (e.g., year, look, current). This additional analysis provides enhanced confidence in our results.

Second, we assess the robustness of our estimates by testing alternative model specifications. Because two of our outcomes (i.e., video consumption and assessment completion) are count variables, we re-estimate our regression models using Poisson models that are better suited for count data. The results of these analyses (which are reported in Web Appendix Tables C2 and D2) are similar to the results of our main models.

Finally, we examine the robustness of our results over time. The majority of the postings in both Study 1 (86%) and Study 2 (73%) occur in the first week of the course. This pattern helps mitigate confounds about the timing of the posts and the diminishes the possibility that learners who post may already be more immersed in the course. However, this finding also raises questions about the degree to which the idea advantage persists beyond the first week. Thus, we examine degree of persistence by reporting the results for the coefficients of idea sharing across all four weeks of the course (see Web Appendix Tables C3 and D3). As shown in these tables, the effects of idea sharing on video consumption and assessment completion are also significant in the second, third, and fourth weeks. Interestingly, the effects for video consumption in Study 1 start appearing only in the second week. Finally, the weekly results from Study 2 suggest that although an idea-sharing prompt enhances engagement throughout the duration of the course, its influence is strongest at the start of the course.

## 6 General Discussion

Our research focuses on the important issue of bolstering engagement in online learning platforms. Specifically, we answer three key questions regarding the effect of content sharing on learner engagement: How do different types of content sharing (i.e., related to course ideas or learner identity) impact learner engagement? What mechanisms underlie these effects? How do the effects vary across different types of learners?

We address these questions via both a textual analysis across 18 months (Study 1) and a field experiment (Study 2) among over 20,000 combined learners in a popular Coursera course offered by a large public U.S. university. The results from Study 1 suggest that learners generally share content about either the course (i.e., ideas) or themselves (i.e., identity), and that idea sharing is more positively related to learner engagement than identity sharing. The results from Study 2 show that an idea-sharing prompt leads to a 31% increase in video consumption and a 30% increase in assessment completion relative to a control condition with no prompt. Across both studies, idea sharing has a stronger effect on engagement than identity sharing.

We investigate the potential mechanism underlying this “idea advantage” through a fine-grained analysis of the textual content of learners’ response to the idea and identity prompts. The results of this analysis suggest that enhanced elaboration in textual postings may explain this idea advantage. Moreover, there is significant heterogeneity in the idea advantage across learners; the effect is strongest for learners from English-speaking countries and those new to online learning. We believe that these findings offer important insights for theory and practice and lay a foundation for future research in this emerging domain.

### 6.1 Implications for Theory

The results of our two studies provide a number of insights for enriching theory regarding engagement in online learning environments. First, we offer and empirically examine a novel

conceptualization of the effect of content sharing on engagement in online learning platforms. This conceptualization is supported by a large-scale textual analysis which reveals that learners generally share content about either the course (i.e., ideas) or themselves (i.e., identity). Second, we conduct a randomized field experiment to assess the impact of asking learners to share these two types of content upon a variety of metrics of engagement, including video consumption, assessment completion, and course performance. The results of this investigation suggest that a 30-second idea-sharing prompt is a low-cost and effective way to enhance learner engagement. The effectiveness of this minimal content intervention lends added support to the surprisingly strong power of simple interventions to shape individuals' thoughts and actions (e.g., Thaler and Sunstein 2008) and extends this tactic to a new domain (i.e., online learning platforms).

From the perspective of theory development, our findings suggest that some practices for encouraging engagement in face-to-face learning environments may not readily translate to online contexts. Specifically, we find that identity-sharing prompts are less effective than idea-sharing prompts in online contexts. While this type of personal engagement strategy may work well in a face-to-face classroom, our research suggests that its efficacy in online environments appears to be weaker relative to an engagement strategy related to course ideas.

More generally, while online learning platforms and face-to-face classrooms appear to share a set of commonalities, some practices that are common in face-to-face classrooms (e.g., the role of learner introductions) may need to be reimagined for digital environments. To develop a richer understanding of these differences, we suggest that script theory may provide a useful theoretical lens (Schank and Abelson 1977). An important tenet of script theory is that individuals enact behavioral patterns from learnings based in similar contexts (e.g., Bitner, Booms, and Mohr 1994; Leigh and Rethans 1984). Although our approach did not allow for a direct test of this theory, our results cohere with its key tenets. In particular, we find learners in online learning platforms appear to follow the script of face-to-face classrooms by primarily sharing content that either relates to themselves (i.e., identity) or to the course (i.e., ideas).

Our research also provides fresh insights for enriching the understanding of engagement as

a construct. In the new digital age, engagement has become an increasingly important marketing objective and is carefully tracked using a variety of metrics (Rajaram and Manchanda 2020). Likewise, our study employs a number of key metrics of learner engagement such as number of videos consumed, number of assessments completed, and course performance. These metrics can help scholars assess granular behavioral patterns in large-scale online settings and are more actionable than traditional measures of educational engagement such as degree of student attendance and participation in school activities (e.g., Finn 1989; Goodenow 1993). Overall, our research suggests that scholars interested in examining learner engagement in online settings have an opportunity to significantly expand and enrich the metrics they employ.

Finally, our research opens the door for marketing scholars to apply existing marketing knowledge and thought to important issues facing the education industry, particularly the role of technology and the design and delivery of online education. Our research leverages the rich body of marketing scholarship on user-generated content via online platforms to study the important challenge of enhancing engagement in online learning platforms. As recently illustrated by Sen and Tucker (2020), established marketing concepts such as product quality can help shed light in important issues in the education domain. Likewise, our research offers new insights regarding the interplay of content sharing and engagement in online learning platforms.

## **6.2 Implications for Practice**

Our research also has implications for practice. First, our uncovering of the idea advantage indicates that the sharing of ideas may be an effective way to spur engagement in online learning environments. Specifically, our results suggest that idea sharing enhances video consumption by 31% and assessment completion by 30%. In contrast, identity sharing has substantially weaker effects. Hence, while encouraging learners to disclose information about their identity may be effective in a traditional face-to-face classroom setting, the wisdom of this strategy is

limited relative to idea sharing in large online classroom contexts where anonymity may be valued. In online learning platforms, a prompt to share one's ideas seems to instill a stronger degree of engagement than a prompt to share one's identity. Thus, our results suggest that educators interested in enhancing online learning engagement should place more emphasis on encouraging idea sharing. Online course designers can easily put this new finding into practice by simply adding a short video-based prompt like the one we employed in Study 2.

Second, our finding that elaboration appears to explain increased engagement among learners who share course-related ideas (vs. identity) suggests that platform designers and course providers may be able to use simple design interventions to test the effects of varying levels of elaboration. For example, educators could seek to control the degree of elaboration by varying the allowable length of learning postings (e.g., by altering the word limit or altering the size of a text box). Our research suggests that this type of design intervention may be testable for effectively managing learner engagement in an online learning context.

Third, from an engagement perspective, our results suggest that a simple course intervention that invites learners to share ideas can result in an average of 10.38 minutes of additional video consumption per learner across a four-week course. For our focal course, this amounted to 367 additional hours of video engagement across our set of 2,122 learners.<sup>9</sup> In essence, the insertion of a simple 30-second idea-sharing prompt at the start of a course appears to enhance the stickiness of its content. The number of minutes viewed (i.e., watch time) is an important engagement metric for video-based platforms such as YouTube, but has received scant attention from marketing scholars. In many marketing contexts, such as retail stores and video games, time spent engaging in an activity is positively associated with increased revenue (Huang, Jasin, and Manchanda 2019; Hui et al. 2013). Thus, our finding that a 30-second prompt to share ideas can result in over 10.38 minutes of engagement per course could have important implications for the marketing and financial outcomes of online

---

<sup>9</sup>This estimate is based on 31% more videos in the idea prompt. An average learner watches 6.7 videos in 4 weeks, of about 5 minutes each, so additional video watching = 31% \* 6.7 \* 5 minutes, or 10.38 minutes in the course.

educational offerings. Although our findings regarding the impact of idea-sharing on watch time need to be confirmed by collecting additional data from other online courses, our focal course shares many similarities with online courses in general (e.g., composition of international students). Therefore, we expect this type of increased engagement to apply across a wide set of online courses.

### **6.3 Limitations and Future Research Directions**

Our research also has some limitations that present several promising opportunities for future research. First, our data are drawn from a single course on the Coursera platform. Thus, we are unable to examine the degree to which the idea advantage persists across courses in other disciplines, offered on alternative platforms, or taught by other instructors from other universities. We are also unable to track the effects for the readers of discussion postings. Thus, replications and extensions that involve a broader range of online courses and on other learning platforms (e.g., EdX, Udemy) and learners would help assess the generalizability of our findings.

Second, while our field experiment (Study 2) is able to achieve randomization, our estimates are best interpreted as the intent-to-treat (ITT) effects. Moreover, recall that only 24% of learners actually shared ideas when prompted in our experiment. Future research could employ prompts that encourage a higher percentage of learners to engage in idea sharing. For example, learners could be offered an incentive, such as course credit for posting content in a discussion forum. Likewise, future research could also examine the effects of the quality of ideas shared in order to ascertain the specific nature of ideas that is most effective in encouraging learner engagement. More generally, future research efforts could seek to expand our line of inquiry by employing a broader range of interventions, such as planning and commitment prompts in conjunction with content sharing (Patterson 2018).

Third, our findings regarding learner heterogeneity present additional opportunities for future research. For instance, our research suggests that the positive effects of idea sharing are

stronger for learners from English-speaking vs. non-English speaking countries. Thus, there may be an opportunity to develop interventions capable of leveraging this heterogeneity. One possibility would be for scholars to collaborate with online learning platforms to add new features and technologies that can implement “scalable customization” by automating which learners receive (or do not receive) certain types of engagement-enhancing prompts. Future research that explores outcomes associated with these types of customized prompts could help glean valuable insights about some of the weak or mixed findings reported in this initial investigation (e.g., limited effects of identity sharing and mixed effects on course performance).

## **7 Conclusion**

The digital revolution has transformed nearly every industry and is currently reshaping higher education. Face-to-face classes that usually enroll a few dozen largely homogenous students are now being replaced by online courses that enroll thousands of diverse learners from across our planet. As a result, educators may need to reassess many longstanding pedagogical assumptions as they transform their course design and delivery for online learning platforms. The research challenge is significant: Approximately 90% of learners who begin a course on an online learning platform drop out within the first week. Our research suggests that platforms, institutions and instructors can help address this engagement problem through an easily implementable and low-cost intervention rooted in content sharing strategies. Engagement can be enhanced by nearly a third by simply prompting learners to share their ideas. Unfortunately, online education remains under-researched by marketing scholars despite its close connection to the marketing literature on digital platforms, customer engagement, and user-generated content. We hope that our inquiry provides marketing educators and scholars with a set of concepts and ideas for understanding and leveraging this new digital learning environment.

## References

- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin (1996), "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91 (434), 444-455.
- Babić Rosario, Ana, Francesca Sotgiu, Kristine De Valck, and Tammo HA Bijmolt (2016), "The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors," *Journal of Marketing Research*, 53 (3), 297-318.
- Banerjee, Abhijit V., and Esther Duflo (2014), "(Dis)organization and Success in an Economics MOOC," *American Economic Review*, 104 (5), 514-518.
- Berger, Jonah, Ashlee Humphreys, Stephan Ludwig, Wendy W. Moe, Oded Netzer, and David A. Schweidel (2020), "Uniting the Tribes: Using Text for Marketing Insight," *Journal of Marketing*, 84 (1), 1-25.
- Berger, Jonah, and Katherine L. Milkman (2012), "What Makes Online Content Viral?" *Journal of Marketing Research*, 49 (2), 192-205. Berman, Joshua, and Amy S. Bruckman (2001), "The Turing Game: Exploring Identity in an Online Environment," *Convergence*, 7 (3), 83-102.
- Berman, Ron, Shiri Melumad, Colman Humphrey, and Robert Meyer (2019), "A Tale of Two Twitterspheres: Political Microblogging During and After the 2016 Primary and Presidential Debates," *Journal of Marketing Research*, 56 (6), 895-917.
- Bitner, Mary Jo, Bernard H. Booms, and Lois A. Mohr (1994), "Critical Service Encounters: The Employee's Viewpoint," *Journal of Marketing*, 58 (4), 95-106.
- Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003), "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3, 993-1022.



- Chang, Jonathan, Sean Gerrish, Chong Wang, Jordan Boyd-Graber, and David Blei (2007), “Reading Tea Leaves: How Humans Interpret Topic Models,” *Advances in Neural Information Processing Systems*, 22, 288-296.
- Chaturvedi, Snigdha, Dan Goldwasser, and Hal Daumé (2014), “Predicting Instructor’s Intervention in MOOC Forums,” *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, 1, 1501-1511.
- Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz (2016), “The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment,” *American Economic Review*, 106 (4), 855-902.
- Christensen, Gayle, Brandon Alcorn, and Ezekiel J. Emanuel (2014), “MOOCs Won’t Replace Business Schools—They’ll Diversify Them,” *Harvard Business Review Blog*, (June 3), <https://tinyurl.com/2ehnezcw>.
- Chronicle (2020), “The Coronavirus is Upending Higher Ed. Here are the Latest Developments,” (March 3), <https://tinyurl.com/vd8wqyf>.
- Class Central (2020), “By the Numbers: MOOCs in 2020,” (November 30), <https://tinyurl.com/54tx83j2>.
- Conrad, Dianne L (2002), “Engagement, Excitement, Anxiety, and Fear: Learners’ Experiences of Starting an Online Course,” *American Journal of Distance Education*, 16 (4), 205-226.
- Conversation (2020), “Massive Online Open Courses See Exponential Growth During COVID-19 Pandemic,” (July 23), <https://tinyurl.com/ycb6o8n4>.
- Duflo, Esther, and Emmanuel Saez (2003), “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment,” *The Quarterly Journal of Economics*, 118 (3), 815-842.

- Dwyer, Karen Kangas and Marlina Davidson (2020), "E-Journaling for All Communication Classes," *Communication Teacher*, 34 (3), 191-197.
- Finn, Jeremy D. (1989), "Withdrawing from School," *Review of Educational Research*, 59 (2), 117-142.
- Fritson, Krista K. (2008), "Impact of Journaling on Students' Self-Efficacy and Locus of Control," *Insight: A Journal of Scholarly Teaching*, 3, 75-83.
- Goli, Ali, Pradeep K. Chintagunta, and S. Sriram (2019), "Effect of Payment on User Engagement in MOOCs," *Journal of Marketing Research*, Forthcoming. Goodenow, Carol (1993), "Classroom Belonging among Early Adolescent Students: Relationships to Motivation and Achievement," *Journal of Early Adolescence*, 13 (1), 21-43.
- Hartmann, Jochen, Juliana Huppertz, Christina Schamp, and Mark Heitmann (2019), "Comparing Automated Text Classification Methods," *International Journal of Research in Marketing*, 36 (1), 20-38.
- Hirano, Keisuke, Guido W. Imbens, Donald B. Rubin, and Xiao-Hua Zhou (2000), "Assessing the Effect of an Influenza Vaccine in an Encouragement Design," *Biostatistics*, 1 (1), 69-88.
- HolonIQ (2020), "Education in 2030", ((accessed July 1, 2020), <https://tinyurl.com/c2bkww3s>.
- Homburg, Christian, Laura Ehm, and Martin Artz (2015), "Measuring and Managing Consumer Sentiment in an Online Community Environment," *Journal of Marketing Research*, 52 (5), 629-641.
- Houston, J. Brian, Joshua Hawthorne, Matthew L. Spialek, Molly Greenwood, and Mitchell S. McKinney (2013), "Tweeting During Presidential Debates: Effect on Candidate Evaluations and Debate Attitudes," *Argumentation and Advocacy*, 49 (4), 301-311.

- Huang, Allen H., Reuven Lehavay, Amy Y. Zang, and Rong Zheng (2018), "Analyst Information Discovery and Interpretation Roles: A Topic Modeling Approach," *Management Science*, 64 (6), 2833-2855.
- Huang, Yan, Stefanus Jasin, and Puneet Manchanda (2019), "'Level Up': Leveraging Skill and Engagement to Maximize Player Game-Play in Online Video Games," *Information Systems Research*, 30 (3), 927-947.
- Hui, Sam K., Yanliu Huang, Jacob Suher, and J. Jeffrey Inman (2013), "Deconstructing the 'First Moment of Truth': Understanding Unplanned Consideration and Purchase Conversion using In-Store Video Tracking," *Journal of Marketing Research*, 50 (4), 445-462.
- Jacob, Brian A., and Jens Ludwig (2012), "The Effects of Housing Assistance on Labor Supply: Evidence from a Voucher Lottery," *American Economic Review*, 102 (1), 272-304.
- Kizilcec, René F., Andrew J. Saltarelli, Justin Reich, and Geoffrey L. Cohen (2017), "Closing Global Achievement Gaps in MOOCs," *Science*, 355 (6322), 251-252.
- Kizilcec, Rene F., Emily Schneider, Geoffrey L. Cohen, and Daniel A. McFarland (2014), "Encouraging Forum Participation in Online Courses with Collectivist, Individualist and Neutral Motivational Framings," *eLearning Papers*, 37, 13-22.
- Lamb, Anne, Jascha Smilack, Andrew Ho, and Justin Reich (2015), "Addressing Common Analytic Challenges to Randomized Experiments in MOOCs: Attrition and Zero-Inflation," In *Proceedings of the Second (2015) ACM Conference on Learning@Scale*, ACM, 21-30.
- Lee, Dokyun, Emaad Manzoor, and Zhaoqi Cheng (2018), "Focused Concept Miner (FCM): Interpretable Deep Learning for Text Exploration," working paper, Tepper School of Business, Carnegie Mellon University.

- Lee, Jeffrey K., and Ann Kronrod (2020), "The Strength of Weak-Tie Consensus Language," *Journal of Marketing Research*, 57 (2), 353-374.
- Leigh, Thomas W. and Arno J. Rethans (1984), "A Script-Theoretic Analysis of Industrial Purchasing Behavior," *Journal of Marketing*, 48 (4), 22-32.
- Liu, Xiao, Dokyun Lee, and Kannan Srinivasan (2019), "Large-Scale Cross-Category Analysis of Consumer Review Content on Sales Conversion Leveraging Deep Learning," *Journal of Marketing Research*, 56 (6), 918-943.
- Manning, Jane, and Marc Sanders (2013), "How Widely used are MOOC Forums? A First Look," (accessed October 20, 2020).
- Melumad, Shiri, J Jeffrey Inman, and Michel Tuan Pham (2019), "Selectively Emotional: How Smartphone use Changes User-generated Content," *Journal of Marketing Research*, 56 (2), 259-275.
- Moon Youngme (2000), "Intimate Exchanges: Using Computers to Elicit Self-Disclosure from Consumers," *Journal of Consumer Research*, 26 (4), 323-339. MPI (2018), "International Students in the United States," (May 9), <https://tinyurl.com/cuxunrjs>.
- Netzer, Oded, Alain Lemaire, and Michal Herzenstein (2019), "When Words Sweat: Identifying Signals for Loan Default in the Text of Loan Applications," *Journal of Marketing Research*, 56 (6), 960-980.
- O'Connell, Timothy, and Janet Dymont (2006), "Reflections on Using Journals in Higher Education: A Focus Group Discussion with Faculty," *Assessment and Evaluation in Higher Education*, 31 (6), 671-691.
- Packard, Grant, and Jonah Berger (2017), "How Language Shapes Word of Mouth's Impact," *Journal of Marketing Research*, 54 (4), 572-588.

- Patterson, Richard W. (2018), "Can Behavioral Tools Improve Online Student Outcomes? Experimental Evidence from a Massive Open Online Course," *Journal of Economic Behavior and Organization*, 153, 293-321.
- Pennebaker, James W., Cindy K. Chung, Joey Frazee, Gary M. Lavergne, and David I. Beaver (2014), "When Small Words Foretell Academic Success: The Case of College Admissions Essays," *PLoS One*, 9 (12), e115844.
- Pennebaker, James W., Roger J. Booth, and Martha E. Francis (2007), "Linguistic Inquiry and Word Count: LIWC [Computer software]," Austin, TX: liwc. Net 135.
- Pickard, Laurie (2019), "35+ Legit Master's Degrees You Can Now Earn Completely Online," *Class Central*, (March 4), <http://tinyurl.com/y6o5x2cw>. Rajaram, Prashant, and Puneet Manchanda (2020), "Video Influencers: Unboxing the Mystique," working paper, Ross School of Business, University of Michigan.
- Reich, Justin and José A. Ruipérez-Valiente (2019), "The MOOC Pivot," *Science*, 363 (6423), 130-131.
- Rubin, Donald B (1974), "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," *Journal of Educational Psychology*, 66 (5), 688-701.
- Schank, Roger C. and Robert P. Abelson (1977), *Scripts, Plans, Goals, and Understanding*, New York, NY: John Wiley Sons.
- Schweidel, David A., and Wendy W. Moe (2014), "Listening in on Social Media: A Joint Model of Sentiment and Venue Format Choice," *Journal of Marketing Research*, 51 (4), 387-402.
- Sen, Ananya, and Catherine Tucker (2020), "Product Quality and Performance in the Internet Age: Evidence from Creationist-Friendly Curriculum," *Journal of Marketing Research*, Forthcoming.

- Smith, Eliot R. and Jamie DeCoster (2000), "Dual-Process Models in Social and Cognitive Psychology: Conceptual Integration and Links to Underlying Memory Systems," *Personality and Social Psychology Review*, 4 (2), 108-131.
- Stark, Robin, Heinz Mandl, Hans Gruber and Alexander Renkl (2002), "Conditions and Effects of Example Elaboration," *Learning and Instruction*, 12 (1), 39-60.
- Steinberg, Ruth M., Susan B. Empson, and Thomas P. Carpenter (2004), "Inquiry into Children's Mathematical Thinking as a Means to Teacher Change," *Journal of Mathematics Teacher Education*, 7 (3), 237-267.
- Stuart, Elizabeth A., Deborah F. Perry, Huynh-Nhu Le, and Nicholas S. Ialongo (2008), "Estimating Intervention Effects of Prevention Programs: Accounting for Noncompliance," *Prevention Science*, 9 (4), 288-298.
- Thaler, Richard, and Cass R. Sunstein (2008), *Nudge: The Gentle Power of Choice Architecture*. New Haven, CT: Yale University Press.
- Thompson, Holly B., and Thompson, George H. (1996), "Confronting Diversity Issues in the Classroom with Strategies to Improve Satisfaction and Retention of International Students," *Journal of Education for Business*, 72 (1), 53-55.
- Tinto, Vincent. (1987). *Leaving College: Rethinking the Causes and Cures of Student Attrition*. Chicago: University of Chicago Press.
- Tirunillai, Seshadri, and Gerard J. Tellis (2012), "Does Chatter Really Matter? Dynamics of User-Generated Content and Stock Performance," *Marketing Science*, 31 (2), 198-215.
- Toubia, Olivier, and Andrew T. Stephen (2013), "Intrinsic vs. Image-Related Utility in Social Media: Why do People Contribute Content to Twitter?," *Marketing Science*, 32 (3), 368-392.

Valsesia, Francesca, Davide Proserpio, and Joseph C. Nunes (2020), “The Positive Effect of Not Following Others on Social Media,” *Journal of Marketing Research*, 57 (6), 1152-1168.

Zhang, Dennis J., Gad Allon, and Jan A. Van Mieghem (2017), “Does Social Interaction Improve Learning Outcomes? Evidence from Field Experiments on Massive Open Online Courses,” *Manufacturing Service Operations Management*, 19 (3), 347-367.

Zhenghao, Chen, Brandon Alcorn, Gayle Christensen, Nicholas Eriksson, Daphne Koller, and Ezekiel Emanuel (2015), “Who’s Benefiting from MOOCs, and Why,” *Harvard Business Review*, 25, 2-8.

Table 1. Variable Descriptions and Summary Statistics (Study 1)

Variable	Operationalization	Mean	SD
Video consumption	Number of videos completed	14.75	16.04
Assessment completion	Number of assessments completed	4.45	5.35
Course performance	Average percentage grade in assessments conditional on submitting an assessment	.84	.12
Idea sharing	Proportion of idea-related content (relative to identity) across a learner's posts	.10	.28
Student	Dummy = 1 if full-time student, else 0	.21	.41
Female	Dummy = 1 if female, else 0	.39	.48
Education	Dummy = 1 if has Masters' degree, else 0	.36	.48

*Notes:* Student refers to those enrolled as full-time students. Education refers to highest attained educational qualification. Video consumption, assessment completion, and course performance are cumulative measures over the entire course duration.

Table 2. Textual Measures for Types of Content (Study 1)

Topic	Top Words from LDA	Top LIWC Differentiators
Topic 1 (Ideas)	Product, idea, customer, https, company video, review, design, local, brand	Words per sentence, dictionary words, cognitive words, past-focused words
Topic 2 (Identity)	Look, year, busy, hello, current, to identity) hope, forward, knowledge, manage, live	Authenticity, pronouns, personal pronouns, present-focused words

*Notes:* For each topic, top LIWC differentiators are the linguistic features that are more prevalent in that topic, e.g., idea-related content has higher words per sentence, etc. relative to identity-related content, which has more authenticity, pronouns, etc. LIWC = Linguistic Inquiry and Word Count; LDA = Latent Dirichlet Allocation.



Table 3. Probability of Idea Content and Learner Engagement (Study 1)

Variable	Video Consumption	Assessment Completion	Course Performance
Idea sharing (vs. identity sharing))	2.27** (.87))	1.07*** (.29)	.09*** (.02)
Student	1.31 (1.33)	.62 (.45)	.05 (.03)
Female	.93* (.47)	.36* (.16)	.02* (.01)
Education	3.21*** (.80)	1.14*** (.26)	.08*** (.02)
No. of posts	2.07*** (.50)	.81*** (.16)	.07*** (.01)
Intercept	21.38** (6.38)	7.02** (2.34)	.53** (.19)
R <sup>2</sup>	.17	.15	.11
Partial $\eta^2$	.001	.003	.004

Notes: \*\*\* p < .001, \*\* p < .01, \* p < .05. Cohort, country and language fixed effects are included; idea sharing is the proportion of idea (vs. identity) content across a learner's posts obtained from the LDA model. Robust standard errors clustered by learner are in parentheses; partial  $\eta^2$  is the effect size for idea sharing. N = 4,676. LDA = Latent Dirichlet Allocation.

Table 4. Variable Descriptions and Summary Statistics (Study 2)

Variable	Operationalization	Mean	SD
Video consumption	Number of videos completed	6.70	11.82
Assessment completion	Number of assessments completed	1.92	3.89
Course performance	Average percentage grade in assessments conditional on submitting an assessment	.82	.13
Idea-sharing prompt	Proportion of idea-related content (relative to identity) across a learner's posts	.33	.49
Identity-sharing prompt	Proportion of idea-related content (relative to identity) across a learner's posts	.33	.49
Student	Dummy = 1 if full-time student, else 0	.51	.50
Female	Dummy = 1 if female, else 0	.36	.48
Education	Dummy = 1 if has Masters' degree, else 0	.24	.42

Notes: Student refers to those enrolled as full-time students. Education refers to highest attained educational qualification. Video consumption, assessment completion, and course performance are cumulative measures over the entire course duration.

Table 5. Textual Measures for Types of Content (Study 2)

Topic	Top Words from LDA	Top LIWC Differentiators
Topic 1 (Ideas)	Consume, product, world, customer, online, firm, company, people, internet, retail	Word count, words per sentence, dictionary words, analytical thinking, cognitive words, past-focused words
Topic 2 (Identity)	Course, work, year, current, learn, manage, India, image, hello, universe	Authenticity, Authenticity, pronouns personal pronouns, present-focused words

*Notes:* For each topic, top LIWC differentiators are the linguistic features that are more prevalent in that topic, e.g., idea-related content has a higher word count, etc. relative to identity-related content, which has more authenticity, pronouns, etc. LIWC = Linguistic Inquiry and Word Count; LDA = Latent Dirichlet Allocation.

Table 6. Idea (vs. Identity) Sharing and Learner Engagement (Study 2)

Variable	Video Consumption	Assessment Completion	Course Performance
Idea sharing (relative to no sharing))	2.08** (.62)	.57** (.20)	.01 (.02)
Identity sharing (relative to no sharing))	.50 (.58)	.03 (.19)	-.02 (.02)
Student	-3.20*** (.84)	-.80** (.31)	-.03 (.03)
Female	-.09 (.74)	-.01 (.25)	.00 (.02)
Education	1.11 (1.09)	.29 (.36)	.06* (.03)
Intercept	4.38 (2.44)	1.29 (.83)	.10 (.06)
R <sup>2</sup>	.11	.09	.09
Partial $\eta^2$	.006	.004	.000

*Notes:* \*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$ . Cohort, country and language fixed effects are included; robust standard errors clustered by learner are in parentheses; idea sharing is 1 for learners assigned to the idea-sharing condition and 0 for learners assigned to the identity-sharing condition. Partial  $\eta^2$  is the effect size for idea sharing.  $N = 2,122$ .

Table 7. Idea (vs. Identity) Sharing and Learner Engagement for Compliers (Study 2)

Variable	Video Consumption	Assessment Completion	Course Performance
Idea sharing (vs. identity sharing))	3.89* (1.68)	1.57* (.59)	.10* (.04)
Student	-9.04* (4.38)	-2.21 (1.54)	-.01 (.12)
Female	-1.32 (2.39)	-.42 (.84)	-.02 (.06)
Education	-3.01 (2.76)	-1.00 (.97)	-.01 (.07)
Intercept	8.00 (14.20)	.00 (4.99)	.00 (.37)
R <sup>2</sup>	.26	.23	.21
Partial $\eta^2$	.018	.023	.016

Notes: \*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$ . Cohort, country and language fixed effects are included; robust standard errors clustered by learner are in parentheses; idea sharing is 1 for learners assigned to the idea-sharing condition and 0 for learners assigned to the identity-sharing condition. Partial  $\eta^2$  is the effect size for idea sharing. N = 2,122.

Table 8. Mechanisms: LIWC Analysis for Textual Response (Study 2)

Top LIWC Differentiators	Example Words	Mean for Idea Sharing	Mean for Identity Sharing	Absolute Difference
<b>Summary variables</b>				
Authentic	-	23.37	37.09	13.72
Analytical thinking	-	84.17	75.54	8.63
Dictionary words	-	67.83	59.74	8.09
Words per sentence	-	23.22	16.80	6.42
<b>Linguistic dimensions</b>				
Pronouns	I, them, itself	6.95	11.97	5.02
Personal pronouns	I, them, her	3.54	9.58	6.04
First person singular	I, me, mine	.98	8.70	7.72
<b>Cognitive processes</b>				
Cognitive	cause, know, ought	9.56	4.22	5.34
Differentiation	hasn't, but, else	1.72	.52	1.20
Tentative	maybe, perhaps	1.50	.64	.86
Insight	think, know	2.07	1.35	.72
Certainty	always, never	1.13	.71	.42
Discrepancy	should, would	.92	.54	.38
<b>Time orientation</b>				
Focus on past	ago, did, talked	1.59	.78	.81
Focus on present	today, is, now	8.55	9.51	.96

*Notes:* LIWC variables listed are those with the most significant difference between idea and identity sharing. LIWC = Linguistic Inquiry and Word Count.

Table 9. Idea (vs. Identity) Sharing and Learner Engagement for Non-compliers (Study 2)

Variable	Video Consumption	Assessment Completion	Course Performance
Idea sharing (vs. identity sharing))	1.33* (.52)	.32 (.18)	.03 (.01)
Student	-.87 (1.27)	-.10 (.43)	-.01 (.03)
Female	.36 (.79)	.21 (.27)	.02 (.02)
Education	1.43 (1.03)	.42 (.35)	.04 (.03)
Intercept	-2.05 (4.05)	-.62 (1.37)	-.05 (.11)
R <sup>2</sup>	.18	.15	.17
Partial $\eta^2$	.007	.004	.004

Notes: \*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$ . Cohort, country and language fixed effects are included; robust standard errors clustered by learner are in parentheses; non-complier refers to learners who do not respond to the sharing prompt; partial  $\eta^2$  is the effect size for idea sharing. N = 1,038.

Table 10. Heterogeneity in the Impact of Idea (vs. Identity) Sharing (Study 2)

Variable	Video Consumption	Assessment Completion	Course Performance
Idea sharing	3.28*	1.37**	.11**
(x English-speaking country)	(1.86)	(.64)	(.05)
Idea sharing	-20.05***	-6.82***	-.45***
(x Prior online experience)	(4.22)	(1.43)	(.11)
Idea sharing	1.32**	.41*	.02
	(.66)	(.22)	(.02)
English-speaking country	-.06	-.05	-.01
	(1.16)	(.40)	(.03)
Prior online experience	12.13***	4.35***	.30***
	(3.75)	(1.30)	(.09)
Student	-2.79**	-.75	-.03
	(1.41)	(.50)	(.04)
Female	.62	.21	.02
	(.92)	(.31)	(.03)
Education	.84	.27	.04
	(1.19)	(.41)	(.03)
Intercept	1.90***	.41*	.05**
	(.71)	(.24)	(.02)
R <sup>2</sup>	.06	.05	.06

Notes: \*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$ . Cohort, country and language fixed effects are included; robust standard errors clustered by learner are in parentheses; idea sharing is 1 for learners assigned to the idea-sharing condition and 0 for learners assigned to the identity-sharing condition;  $N = 1,414$ .

Figure 1. Word Cloud of LDA Topics (Study 1)



Figure 2. Video Script for the Identity- and Idea- Sharing Conditions

Hi. Before we dive into the course, I'd like to invite you to personally introduce yourself to the other members of our learning community. You can easily do this here in the Discussion Prompt by posting your photo or a brief video of yourself. Or you tell us more about you, such as where you are from or what you do by simply typing a couple of sentences. I look forward to learning more about you and I'm sure that your fellow learners are eager to meet you. See you in class!

(a) Identity sharing

Hi. Before we dive into the course, I'd like to invite you to personally reflect on the Digital World and how its impacted the way in which both firms and consumers engage in marketing activity. For example, please give some personal thought about such issues as how the Internet has affected physical retailers or how smartphones have altered how consumers search for information. I look forward to our learning adventure. See you in class!

(b) Idea sharing

Figure 3. Randomization Checks (Study 2)

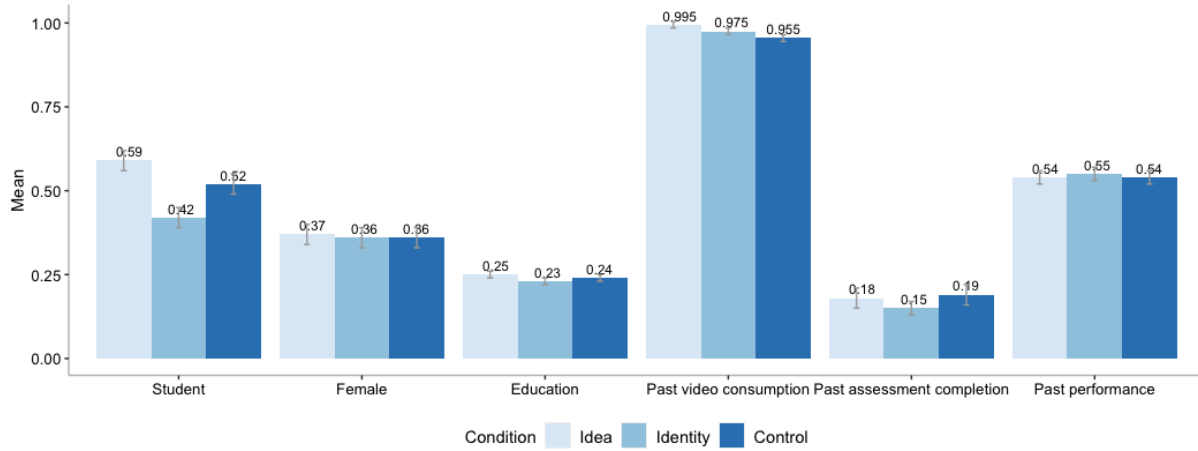
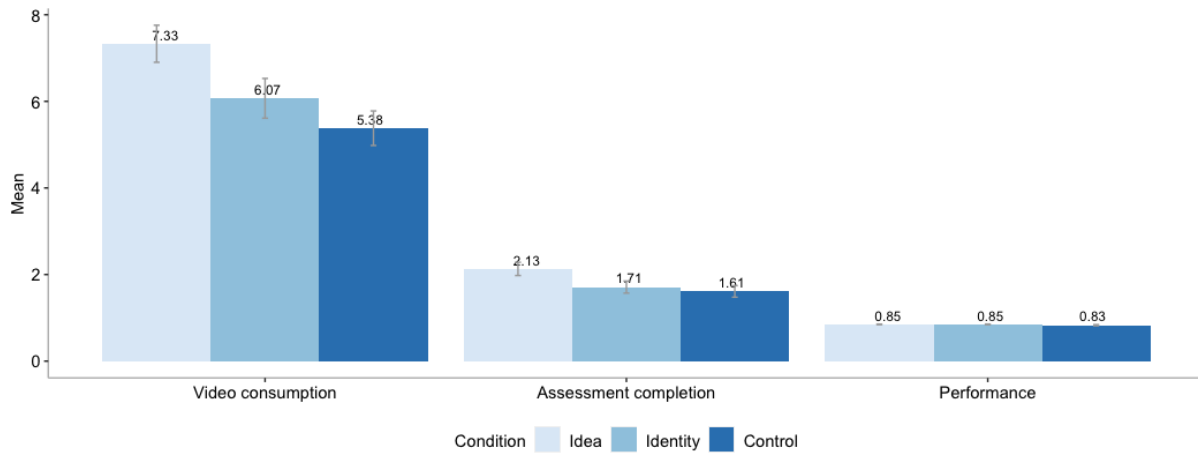


Figure 4. Word Cloud of LDA Topics (Study 2)





Figure 5. Model-Free Evidence for Identity, Idea and Control Conditions (Study 2)



# Web Appendix

Available on this [link](#).