"Why did you enroll in this course?" Developing a Standardized Survey Question for Reasons to Enroll

Emily Schneider
Graduate School of Education
Stanford University
elfs@cs.stanford.edu

René F. Kizilcec
Department of Communication
Stanford University
kizilcec@stanford.edu

ABSTRACT
Understanding motivations for enrolling in MOOCs is key for personalizing and scaling the online learning experience. We develop a standardized survey item for measuring learners' reasons to enroll, based on a corpus of open-ended responses from previous course surveys. Online coders were employed in the iterative development of response options. The item was designed to minimize response biases by adhering to best practices from survey design research.

Author Keywords
Survey research; motivation; MOOC learner goals

ACM Classification Keywords
H.5.2 User interfaces: Evaluation and Methodology

INTRODUCTION
Over the past few years, millions of people across the world have enrolled in hundreds of massive open online courses (MOOCs) across many platforms. With a globally distributed population and a great diversity of learner backgrounds, there are a wide number of reasons that inspire people to enroll in MOOCs. But with no cost to entry or exit, enrollment numbers are merely an indication of interest, consistently overestimating the number of actively participating learners in the course; among active learners, participation rates tend to decline steadily as the course progresses. These variable levels of engagement are likely influenced by learners' reasons for enrolling in the course. Prior work in educational psychology and higher education have shown that learners' goal orientation and attitudes towards the value of achieving their goals are intimately tied to their engagement with learning experiences (see [1] for a review).

Kizilcec, Piech, & Schneider [3] found learners' patterns of course engagement to be associated with self-reported reasons to enroll. Breslow and colleagues [2], however, saw no correlations between motivations for enrollment and certain course success metrics. Even without the issue of different success metrics, it is unclear whether these results are contradictory, because the motivations for enrollment in each of these studies were generated by different survey items.

There is a strong need for a standardized question on reasons for enrollment to create comparability between courses and enable coordinated research efforts across institutions and platforms. Reliably ascertaining learners' reasons to enroll is instrumental for scaling and personalizing the online learning experience; for instance, personalized pathways through available materials, recommended supplementary resources, or providing the option to take the course in different modalities (e.g. a self-paced 'library' of available resources rather than enrollment with a cohort of other learners). Real-time analytics for personalized learning experiences are an important, yet unreached milestone for MOOCs. Moreover, given the relative novelty of MOOCs, there is a general interest in learners' motivations for enrolling and staying engaged in the courses–and understanding how these motivation structures differ from those of learners in other types of online learning environments.

DESIGNING A GOOD QUESTION
The importance of this survey item has not gone unrecognized–most course surveys to date contain one or more items to determine learners' reasons for enrolling in a course. The two major question types to measure this construct have been open response and multiple choice. While open response is a rich source of nuanced information, it is more challenging to adequately analyze textual data in this context. Although simpler to analyze, multiple choice items–with the option to select a single, a certain number, or as many options as apply–can be problematic unless they are designed with careful attention to known survey biases. The development of an item that minimizes induced biases is the goal of this work.

The two core aspects for the optimal design of a multiple choice item for why learners enroll are selection constraints and response options. Selection constraints like 'select one' or 'select three' place an arbitrary limit on the number of reasons respondents can report and coerces them to select a certain number. This tends to induce satisficing behavior, as respondents who intended to select a different number of options become less invested in making an effort to respond accurately [4]. And although 'select all that apply' lets respondents decide how many options to choose, its unguided nature does not require learners to consider each answer option and after selecting a few options they might feel like "it's enough". As a result, leaving an option unselected does not

have a clear and consistent interpretation. An item design that avoids these issues asks respondents to consider each response option in turn and choose whether it applies to them or not. Note that "Applies/Does not apply" scale labels should be used instead of "True/False" or "Yes/No" to avoid inducing acquiescence bias–respondents' tendency to agree with questions independent of their content [4].

The choice of response options is critical for the validity with which a question can measure a certain construct. Response options should be mutually exclusive and collectively exhaustive, which means that options should not overlap and all possible responses should be covered. The latter condition is difficult to satisfy, but can be approached in this context by asking learners to describe in their own words their reasons for enrolling. The textual data from these open response questions can then be systematically analyzed to develop a final list of response options. Another advantage of this approach is that the resulting response options' phrasing will be closer to how learners express their reasons.

ITERATIVE DEVELOPMENT OF RESPONSE OPTIONS

The iterative process of response option development was crowdsourced to 'classification experts' on Amazon Mechanical Turk (MTurk), who manually coded random samples of open response answers from learners in three different MOOCs (on topics in Political Science, Computer Science, and Economics). Before the MTurk coding, a preliminary codebook was developed by two independent volunteer coders using learners' open response texts and a previously developed set of reasons provided in course surveys of a major MOOC platform. For each open response text, MTurk coders were instructed to select all appropriate reason from the codebook. An "other" option was provided and coders were strongly encouraged to choose this option if some aspect of the response was not reflected in the existing response options. There also was a "spam" option.

In the first iteration, 300 randomly chosen responses were *each* coded by four MTurk coders. Each option's frequency and intercoder reliability, and correlations between response options were evaluated; all responses that were coded as "other" by more than one coder were also examined individually. These analyses revealed some holes in the codebook, as well as some categories which did not line up with participants' characterizations of their reasons to enroll. Based on these insights, the codebook was modifed and applied by MTurk coders to a similarly large random sample of responses. A third iteration followed a similar procedure. The final product is the survey item shown in Table 1.

Response Option Ordering

Response options are frequently presented in random order, as presentation order can influence the respondent's choice (order effect). Simple randomization, however, can be problematic when a more general question or response option is preceded by a more specific question. This can bias responses to the general question, a phenomenon known as the "subtraction effect" [5]. In our case, the first response option Table 1, "General interest in topic", is more general than the

Table 1. Final 'Why Enroll' Survey Item

Why did you enroll in this course?	Applies	Does not apply
General interest in topic	○	○
Relevant to job	○	○
Relevant to school or degree program	○	○
Relevant to academic research	○	○
For personal growth and enrichment	○	○
For career change	○	○
For fun and challenge	○	○
Meet new people	○	○
Experience an online course	○	○
Earn a certificate/statement of accomplishment	○	○
Course offered by prestigious university/professor	○	○
Take with colleagues/friends	○	○
To improve my English skills	○	○

other options and should always be placed first; if possible, the remaining items should be presented in random order to address order effects.

FUTURE DIRECTIONS
The proposed survey item has been included in all Stanford MOOC surveys since September 2013 and will remain in the survey template for future courses. We encourage other institutions to adopt this item to facilitate comparative research and unified metrics. Within the courses where the item was deployed, we are currently investigating the associations between engagement patterns and learners' goals, laying the groundwork for developing learner profiles based on self-reported intentions as well as actual behavior in the course.

ACKNOWLEDGEMENTS

REFERENCES

1. Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., and Norman, M. K. *How learning works: Seven research-based principles for smart teaching.* John Wiley & Sons, 2010.

2. Breslow, L. B., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., and Seaton, D. T. Studying learning in the worldwide classroom: Research into edx's first mooc. *Research & Practice in Assessment 8* (2013), 13–25.

3. Kizilcec, R. F., Piech, C., and Schneider, E. Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, ACM (2013), 170–179.

4. Krosnick, J. A. Survey research. *Annual review of psychology 50*, 1 (1999), 537–567.

5. Tourangeau, R., Rasinski, K. A., and Bradburn, N. Measuring happiness in surveys: A test of the subtraction hypothesis. *Public Opinion Quarterly 55*, 2 (1991), 255–266.