Demographic data of MOOC learners: Can alternative survey deliveries improve current understandings?

van de Oudeweetering, K.; Agirdag, O.

DOI
10.1016/j.compedu.2018.03.017

Publication date
2018

Document Version
Final published version

Published in
Computers and Education

License
Article 25fa Dutch Copyright Act

Link to publication

Citation for published version (APA):
van de Oudeweetering, K., & Agirdag, O. (2018). Demographic data of MOOC learners: Can alternative survey deliveries improve current understandings? *Computers and Education*, *122*, 169-178. https://doi.org/10.1016/j.compedu.2018.03.017

Computers & Education 122 (2018) 169–178

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Demographic data of MOOC learners: Can alternative survey deliveries improve current understandings?

Karmijn van de Oudeweetering[*,1], Orhan Agirdag[2]

Department of Educational Sciences, University of Amsterdam, Nieuwe Achtergracht 127, 1018 WS Amsterdam, The Netherlands

ARTICLE INFO

Keywords:
Massive open online courses
MOOC
Demographic survey
Demographics
Response rate
Response errors

ABSTRACT

Although demographic data in Massive Open Online Courses (MOOCs) have regularly been reported, these data are mainly retrieved through email-based surveys with very low response rates. This indicates an increased risk of misrepresentation. This study examined whether a survey embedded in the MOOC environment could yield higher response rates, could affect the representation of demographics and influence the estimated effects of demographics on learning outcomes. In six MOOCs, learners (N = 3834) were randomly assigned to receive a demographic survey only by email or to receive the embedded survey too. Results showed that the inclusion of the embedded survey caused response to increase from 6.9% to 61.5%. Although survey delivery barely affected the representation of demographics, it did influence the estimated effects of parental education and country of residence on learning outcomes. The findings raise awareness about the importance of survey delivery for response rates and data quality in MOOCs.

1. Introduction

The launch of the first globally recognized Massive Open Online Course (MOOC) in 2008 created expectations about new possibilities for the higher education sector. MOOCs are online, mainly university-level courses that can integrate learning materials like video-lectures, assessments and interaction tools and that allow unrestricted and unconditional enrollment against no or small payments. Considering that MOOCs are less expensive, less selective and more time- and place flexible than traditional higher education, it has been expected that they could favor less privileged populations to engage in further learning (e.g., Kay, Reimann, Diebold, & Kummerfeld, 2013). Still, it has also been questioned whether distance education programs, like MOOCs, are capable of attracting and serving underprivileged individuals (Lee, 2017). In order to understand their reach, various studies have surveyed demographic data of learners in MOOCs. These data have generally shown that well-educated, well-paid individuals from an ethnic majority or from developed countries are overrepresented in MOOCs (Christensen et al., 2013; Stich & Reeves, 2017). Moreover, these data suggest that learners who are actually less educated or live in less developed countries are less likely to complete the course or to achieve relevant learning goals (Greene, Oswald, & Pomerantz, 2015; Hood, Littlejohn, & Milligan, 2015; Kizilcec, Saltarelli, Reich, & Cohen, 2017).

Notwithstanding, demographic data are mainly retrieved through email-based surveys with very low response rates (van de Oudeweetering & Agirdag, 2018). For example, one of the most frequently cited studies to present demographics of MOOC learners received response from 4.3 percent of the targeted population of learners (Christensen et al., 2013). The main problem with such low

* Corresponding author.
E-mail address: karmijn.vandeoudeweetering@kuleuven.be (K. van de Oudeweetering).
[1] Methodology of Educational Sciences Research Group, University of Leuven, Tiensestraat 102, 3000 Leuven, Belgium.
[2] Laboratory for Education and Society, University of Leuven, Andreas Vesaliusstraat 2, 3000 Leuven, Belgium.

https://doi.org/10.1016/j.compedu.2018.03.017
Received 26 September 2017; Received in revised form 22 March 2018; Accepted 24 March 2018
Available online 27 March 2018
0360-1315/

response rates is that they decrease the likelihood of representative results (Fan & Yan, 2010). Especially since socially relevant demographics such as income level, educational attainment and ethnic background could influence response to web-surveys (Couper, Kapteyn, Schonlau, & Winter 2007), it is probable that MOOC data resulting from surveys with low response rates misrepresent the actual population. In turn, the misrepresentation could distort the estimated effects of demographics on completion or other learning outcomes. There have been studies that noticed this issue and adopted strategies to enhance response and representativeness in MOOC surveys. However, these strategies have not been ethically desirable, since these either obligated participation (Ho et al., 2015) or favored respondents over non-respondents in course-related activities (Gates, Wilkins, Conlon, Mossing, & Eftink, 2013). This shows that there are several issues that research needs to tackle in order to gain accurate knowledge about the demographic characteristics of the MOOC population and, consequently, the degree of equality in MOOC participation.

First, it is important to raise attention for the issue of response bias. It is an issue that is barely reported in studies using surveys in MOOCs, while it is a necessary aspect to acknowledge when interpreting the validity of research findings. Consequently, research is needed to understand how response rates might influence knowledge on the demographic background of MOOC learners and completers. It is further relevant to explore strategies that help to increase the response rates and enhance representativeness in MOOCs surveys, without violating ethics of voluntary research participation.

2. Background

2.1. Increasing response rates

The low response to MOOC surveys may be typical for their online delivery, as it is seen that response rates are generally lower for surveys that are web-based than for surveys using other delivery modes (Fan & Yan, 2010). One significant theory that may help to understand these low response rates, and how they may be raised, is the social exchange theory. This theory assumes that survey response is a rational decision that is based on cost-benefit considerations (Porter, 2004). Studies based on this theory have mainly focused on raising benefits towards survey participation, especially financial in nature (e.g., Galesic & Bosnjak, 2009; Laguilles, Williams, & Saunders, 2011). However, applying monetary rewards can be too costly in large samples and might even induce unethical stresses for those in underprivileged circumstances. Hence, focusing on the minimization of costs for respondents might be a more appropriate approach in MOOC research.

Considering that costs of survey response are mainly related to time and effort, reducing inconvenience in survey participation seems to be a good measure to increase response rates (Han, Albaum, Wiley, & Thirkell, 2009). Some of the most evident inconveniences of default email-based MOOC surveys are the necessary access through hyperlinks with extended loading times, the overburdening of email inboxes or simple disappearance of the emails in spam folders (Fan & Yan, 2010; Reips, 2000, pp. 95–104; Sax, Gilmartin, & Bryant, 2003). These hindrances can easily be overcome my changing the delivery mode. That is, the MOOC learning environment provides great opportunities to deliver surveys in a user-friendly way. For example, it is possible to embed ungraded and optional quizzes in the course environment. In this way, learners are presented with the survey while they are actually active in the course, they do not have to open extra screens and the size and content of the survey is directly visible to them.

2.2. Response errors and represented demographics of MOOC learners

Raising response to demographic surveys in MOOCs is mainly important as it reduces the risk of *response error.* Response error occurs when samples do not reflect the actual population because respondents share specific characteristics that are not generally shared among non-respondents, including demographics (Grandcolas, Rettie, & Marusenko, 2003). In other web-surveys, for example, it is seen that respondents are more likely to be male, relatively young, well-educated and member of an ethnic majority (Bandilla, Couper, & Kaczmirek, 2014). There has been some empirical evidence to expect response errors in MOOC surveys as well. For example, one study found that female and well-educated learners were more likely to be in the survey sample than would be expected based on the platform population (Stich & Reeves, 2017). Other studies, using purposive samples, found that learners who were formal students at a university or followed the course in other languages than English, French or Spanish were more inclined to respond to a survey (Annear et al., 2015; Colas, Sloep, & Garreta-Domingo, 2016). Thus, these studies suggest that socially relevant variables like gender, educational attainment and cultural background could play a role in response to web-surveys and MOOC surveys in specific. This means that current conclusions about the demographic composition of MOOC populations could be inaccurate. However, as these findings are not replicated across different studies and MOOCs, it remains uncertain what groups might be over- or underrepresented in the current demographic surveys. Even though statistical methods are advocated to detect misrepresentations on a larger scale, the sensitivity and accuracy of these methods deteriorates with the magnitude of non-response and the size of the associated response error (Kizilcec, 2014; Schouten, Cobben, & Bethlehem, 2009). For this reason, consistent research on strategies to increase response remains highly important for the validity of research results.

2.3. Estimated association between demographic background and completion

Response errors might not only influence knowledge on the demographic background of MOOC learners. In addition, the measured effects of demographic background on MOOC learning outcomes are possibly confounded by response errors in demographic surveys (van de Oudeweetering & Agirdag, 2018). There are even grounded reasons to expect distortions in these findings. For example, several studies indicated that survey response rates were higher among those who complete a course than among the noncompleting population (e.g., Cisel, Bachelet, & Bruillard, 2014, pp. 403–404; Rizzardini, Gütl, Chang, & Morales, 2014). In turn, response is lower for those who are less successful in assignments and quizzes (Gates, Wilkins, Conlon, Mossing, & Eftink, 2014). Response rates also appeared to be higher in undergraduate and graduate level courses in comparison to high school level courses (Kizilcec, Piech, & Schneider, 2013). As non-completers and less educated learners are less likely to be part of the monitored sample, the omission of their data might skew the estimated association between demographics and learning outcomes. To distinguish the effect of demographics more accurately, it is further important to consider that some learners never attempt assignments or exams, do not have the intention to complete the MOOC or might have different levels of background knowledge (Belanger & Thornton, 2013; Greene et al., 2015; Phan, McNeil, & Robin, 2016). Accounting for respondents' learning behavior, motivations and background knowledge is therefore also relevant when investigating the association between demographic backgrounds and learning outcomes.

2.4. The present study

This study aims to examine whether delivering a survey in the MOOC environment in addition to the regular email-based survey will yield higher response rates and alternative outcomes than solely distributing email-based surveys. The assumption is that the additional survey will make survey response more convenient, specifically because it is presented in the MOOC environment, which could cause a decrease in non-response and in the likelihood of response error. Based on this research purpose, the questions that guide this study are:

1. To what extent does a demographic MOOC survey yield a higher response rate when a twofold survey delivery is used, which combines a survey embedded in the MOOC environment with an email-based survey, in comparison to using solely an email-based survey delivery?
2. To what extent does the twofold survey delivery influence the represented demographics in the survey sample, including age, gender, education, highest parental education, country of residence and ethnicity?
3. To what extent does the twofold delivery influence the estimated effect of education, parental education and country of residence on (a) completion and (b) grade averages of MOOC learners?

This study is of specific relevance to researchers in the context of MOOCS, as it can raise awareness about the issue of misrepresentation in surveys. Moreover, the research results could advance knowledge on superior strategies to deliver MOOC surveys and therewith raise the quality of MOOC research. Especially, improved research tools could benefit knowledge on the participation and completion of disadvantaged versus advantaged learners and, consequently, on the status of social equality in MOOCs.

3. Methods

3.1. Sample

The study focused on six MOOCs of the University of Amsterdam. In each MOOC, samples were formed within a specific session, which is a cohort of four or eight weeks (see Table 1). Learners in these sessions were included in the sample following four basic criteria. These criteria were (a) they watched the first video lecture of the course, as an indication they attempted to participate in the course, (b) they did not actively disagree with study participation, (c) they received the survey as planned in the experimental procedure and (d) learners were only selected once, which implied that responders were included only for the first response they had given and non-responders were included for the course that was given first in time.

For the second and third research question, additional selection criteria for the samples were applied. As both questions focused on the representation of demographics in survey responses, only responders were included. Furthermore, as the third research question focused on learning outcomes, only learners who had attempted at least one graded activity (e.g., quiz, peer-graded assessment) were included in the sample. For the second sub-question, which focused on grade average, only learners who completed the course were included in the sample. This was done to preclude overlap with the analysis for the first sub-question and to avoid a skewed distribution for values on course grade average, which could impede the analyses.

Table 1
Session dates and sample sizes for the selected MOOCs.

Course	Session	Sample (N = 3834)
Quantitative Methods	Oct 24, 2016–Dec 19, 2016	n = 665
Qualitative Research Methods	Oct 24, 2016–Dec 19, 2016	n = 327
Classical Sociological Theory	Oct 31. 2016–Dec 26, 2016	n = 377
Intr. To Communication Science	Oct 31, 2016–Nov 28, 2016	n = 499
Basic Statistics	Dec 5, 2016–Jan 30, 2017	n = 1795
Inferential Statistics	Dec 5, 2016–Jan 30, 2017	n = 171

3.2. Experiment

A randomized experiment was implemented in the six selected MOOCs. Learners who were enrolled in one of the courses during the time of the experiment were randomly assigned to one of two *branches*. These branches were different versions of the course that were distributed simultaneously. Thus, each leaner could only see one version of the course. The two branches were labeled as the control condition and the experimental condition and the only difference between them was the survey delivery. Sample sizes for the control condition ($n = 2004$) and experimental condition ($n = 1830$) appeared due to chance and do not represent potential biases.

Control condition – email-based survey delivery In the control condition, learners received an email during the second week of the session with a short invitation text to participate in the survey. Through a hyperlink in the email, learners could access the survey website in which the survey was presented. The first page of the survey included an information note on the goal of the study, the voluntary nature of participation, confidentiality of learners' data and contact information of the research team. Learners could indicate whether they agreed or disagreed with study participation. On the next page, there were 17 questions about learners' demographic background, prior knowledge, learning objectives and reason for participating in the MOOC. To enhance participation, learners received a reminder email with a similar short text and the same link to the survey, 14 days after the first email was sent.

Experimental condition – twofold survey delivery In the experimental condition, the same two emails were sent out as in the control condition. In addition, the demographic survey was delivered as an ungraded quiz embedded in the course environment. This survey is labeled as the *embedded survey.* To maximize its visibility to the learners, this additional survey was placed just before the first video lecture. The embedded survey was introduced with the same information note, consent question and included the same 17 questions on learners' background as the email-based survey. This meant that the experimental condition enabled learners to respond to the same survey through two different channels. Therefore, the survey delivery in the experimental condition is labeled as the *twofold survey delivery*. Responses to the embedded as well as the email-based survey were recorded. In case a learner would use both survey deliveries, their response to the embedded survey would be included in the analyses, while the response to the email-based survey would be eliminated.

3.3. Measures

3.3.1. Response rate

The response rate was estimated as the proportion of learners in the sample that responded to the survey, separate for both experimental conditions. Learners who answered at least 50 percent of the questions in the survey were identified as responders. There are two reasons for this threshold. First, it was assumed that those with largely incomplete responses share characteristics with non-responders and could therefore best be considered as non-responders. Second, techniques to account for missing responses are less likely to yield valid results with increasing amounts of missing data. Power to detect valid effects specifically declines when the amount of missing information exceeds 50 percent (Graham, Olchowski, & Gilreath, 2007). The 50 percent threshold is therefore intended to protect the relevance and accuracy of the results.

3.3.2. Represented demographics

A specific set of demographic indicators was selected to assess the representation of demographics in the learning population.

Gender was assessed through three categories, including *male, female* and *other*, which were the response options in the survey. The group that indicated to neither belong to the male nor to the female category was very small ($n = 6$). As this raised statistical issues, these responses were eliminated in the analyses that focused on gender.

Age was based on respondents' self-reported year of birth and was estimated as the difference between year of birth and the year the analyses were conducted (2017). *Educational attainment* was measured on a continuous scale with eleven response options (0 = *no schooling completed* and 10 = *doctorate degree*).

Highest parental education was measured using the same eleven response options as for the indicator 'educational attainment', yet distinguished into two questions targeting both mother and father and with an additional response option for those who did not know the level of schooling of the parent. In case the respondent did not know the education level of one parent, the education level of the other parent was adopted to represent this indicator. If the respondent did not know the education level of both parents, this was recorded as a missing response.

Country of residence was estimated based on respondents' self-reported country of residence. The responses were recoded to represent whether the learner lived in a developed country or developing country, based on the indicators of the UN Statistics Division (United Nations, 1999).

Ethnic minority status was measured through a question about the respondents' subjective minority status. Respondents could indicate whether they considered themselves to be a member of an ethnic minority, or not, or whether they did not know. Responses for those who did not know the answer were recorded as missing responses.

3.3.3. Learning outcomes

Learning outcomes were assessed with two indicators. Data on both indicators were retrieved through data exports.

Course completion reflected whether learners achieved a passing course grade (0 = *no completion* and 1 = *completion*).

Grade average represented learners' average grade based on all assignments, quizzes and exams in the course, measured on a discrete continuous scale from 0 to 10.

3.3.4. Control variables

Several factors were anticipated to confound either all examined associations or only the estimated association between demographics and learning outcomes. To account for these factors, they were adopted as control variables.

Course Learners within the same course could share specific characteristics that are not shared among other learners, since they were attracted to the same topic and interacted with the same course content. To control for these potential differences between courses, dummy variables for the six courses were included in the analyses with the course Quantitative Methods as the reference category.

Learning objective was measured as a dummy variable that represented the objective to earn a certificate, while other objectives were integrated to represent the reference category.

Subject matter experience was adopted as a continuous variable that controlled for the positive influence of prior subject knowledge on MOOC learning outcomes. It was measured through learners' self-reports on a Likert-type four-point scale (0 = *no experience* and 3 = *degree or job in the field*).

ICT experience was used as a continuous variable to control for the potential positive impact of higher levels of digital literacies on learning outcomes. It was measured on a Likert-type four-point scale as well (0 = *no prior experience* and 3 = *advanced knowledge* and *degree or job in the field of ICT*).

English proficiency was an index of responses to three separate questions. Two questions targeted learners' English proficiency in reading and writing, measured on a Likert-type five-point scale (0 = *no proficiency* and 4 = *equivalent to native speaker*). In addition, the frequency of speaking English was reported on a five-point scale (0 = *never* and 4 = *every day*). The average of the responses was estimated to compose the index for English proficiency. The Cronbach's alpha = 0.78 of this index indicated that the internal consistency was satisfactory.

3.4. Data analysis

For the first two research questions, descriptive statistics were estimated to gain intuitive insight in the differences in the response rate and the representation of demographics between the two survey deliveries. In addition, for all three research questions, logistic regression or Ordinary Least Squares (OLS) analyses were conducted to estimate the size and significance of the effect of the twofold survey delivery. For each analysis, models with the relevant control variables were assessed. However, the models without the control variables were examined as well, in order to detect potential mediating or suppressing effects of the control variables. To account for responses that were treated as missing, including blank responses or responses indicating 'I don't know', multiple imputation was used. Five ($m = 5$) imputations were conducted for the three separate samples for the second and third research question. An overview of the descriptive statistics for the samples, based on the imputed data, is given in Table 2.

4. Results

4.1. The influence of survey delivery on response

Following the first research question, it was examined whether the twofold survey delivery generated a higher response rate than the email-based survey delivery. Descriptive statistics showed that the email-based survey attained a response rate of 6.9 percent ($n = 138$), while the response rate for the twofold survey delivery was 61.5 percent ($n = 1125$). This means that adding the embedded survey to the email-based survey substantially increased the response rate by approximately a factor of ten. Furthermore, the statistics showed that 59.2 percent of the learners assigned to the twofold survey delivery condition responded to the embedded

Table 2
Relevant descriptive statistics for research question 2, 3a and 3b.

	Range	*M (SD)* Question 2 ($N = 1263$)	*M (SD)* Question 3a ($N = 801$)	*M (SD)* Question 3b ($N = 163$)
Twofold survey delivery	1/0	89.1%	86.4%	82.8%
Female	1/0	48.6%		
Ethnic minority	1/0	17.5%		
Developing country	1/0	41.7%	40.0%	38.7%
Age	16–82	34.45 (11.44)		
Educational attainment	0–10	7.31 (1.83)	7.21 (1.92)	7.12 (1.84)
Highest parental education	0–10	5.95 (2.67)	6.04 (2.66)	5.98 (2.68)
Learning objective	1/0		23.0%	54.6%
Subject experience	0–3		.82 (.75)	.88 (.75)
ICT experience	0–3		1.64 (.85)	1.73 (.82)
English proficiency	0–4		3.40 (.67)	3.42 (.62)
Completion	1/0		20.3%	
Grade average	0–10			9.01(.56)

Note. Descriptive statistics based on imputed data sets. For each variable, < 8% is imputed.

Table 3
Descriptive statistics on demographics per survey delivery.

	Range	*M (SD)*	
		Email-based survey (N = 138)	Twofold survey (N = 1125)
Female	1/0	41.3%	49.5%
Ethnic minority	1/0	13.0%	18.0%
Developing country	1/0	39.1%	42.0%
Age	16–75	38.01 (12.80)	34.01 (11.19)
Educational attainment	0–10	7.43 (1.60)	7.30 (1.86)
Parental education	0–10	6.27 (2.76)	5.91 (2.65)

Note. Statistics are based on imputed data sets. For each variable, < 8% is imputed.

survey and 3.7 percent filled out the survey in both deliveries. This means that the high response to the twofold survey delivery is mainly obtained through the embedded survey.

To single out the effect of the survey delivery and estimate the size of the effect, a logistic regression model that controlled for course-level differences was estimated. This model explained significantly more variance in the outcome than a null model, $\chi^2(6) = 1507.3$, $p < .001$, Nagelkerke $R^2 = 0.45$, and yielded accurate predictions for 71.8 percent of the sample. The outcomes showed that the effect of the twofold survey delivery, $B = 3.18$, $p < .001$, deviated positively and significantly from zero, controlling for course-level differences. More specifically, the results showed that the odds of survey response versus non-response were 23.95 times as large for those who received both the twofold survey in comparison to those who only received the email-based survey. This substantiated that the additional embedded survey increased the survey response considerably.

4.2. The influence of survey delivery on represented demographics

Focusing on the representation of demographics, differences between the estimated descriptive statistics for the two survey deliveries were examined (see Table 3). These showed that the sample based on the twofold survey delivery represented larger proportions of women and ethnic minorities, a slightly larger proportion of learners from developing countries. Furthermore, the twofold survey resulted into lower estimated averages for learners' age, educational attainment and parental education than the results of the email-based survey.

Again, regression models that controlled for course-level differences were used to analyze the effect of the survey delivery (see Table 4). As reflected in the model results, the effect of survey delivery on the represented average age, $B = -3.59$, $p < .001$, deviated significantly from zero. This indicates that learners who responded to the twofold survey delivery were, on average, 3.59 years younger than learners who responded to the email-based survey, taking into account between-course differences. For the representation of other demographics, however, the effect of survey delivery did not appear to be significant. Moreover, the low proportions of explained variance in each model reflects that survey delivery, as well as between-course difference, did not adequately explain the representation of the demographic indicators. The models without the control variables showed that course-level effects barely changed the estimated effect of the survey delivery. First, this means that variables other than survey delivery or course are at play when predicting the representation of demographics. Moreover, it implies that the differences in represented demographics between the two survey deliveries that were found in the descriptive statistics cannot be generalized to a larger MOOC population.

4.3. The influence of survey delivery on the estimated association between demographic background and learning outcomes

In order to understand how survey delivery might influence our understanding of the effect of demographic background on completion and grade average, separate logistic and OLS regression models were assessed (see Table 5). There was specific attention

Table 4
Logistic and OLS regression models for represented demographics (N = 1263).

	Female		Ethnic minority		Developing country		Age		Education		Parental education	
	B	SE	*B*	SE	*B*	SE	*B*	SE	*B*	SE	*B*	SE
Intercept	−0.71**	0.21	−2.07**	0.32	−0.63**	0.20	38.30**	1.15	7.72**	0.18	6.33**	0.26
Twofold survey delivery	0.29	0.19	0.40	0.28	0.12	0.19	−3.59**	1.05	−0.11	0.16	−0.36	0.24
Nagelkerke R^2	0.06		0.01		0.01							
χ^2_{diff}	56.9**		6.8		13.4*							
R^2							0.02		0.06		< 0.01	

Notes. Fixed effects for between course-differences are included in the models, but not presented.
*$p < .05$, **$p < .001$.

Table 5
Logistic and OLS regression models explaining course completion and grade average.

	Course completion (N = 801)						Grade average (N = 163)					
	Interaction Education		Interaction Parent Education		Interaction Country		Interaction Education		Interaction Parent Education		Interaction Country	
	B	SE	*B*	SE	*B*	SE	*B*	SE	*B*	SE	*B*	SE
Intercept	−2.25**	0.67	−2.24**	0.68	−2.55**	0.71	8.83**	0.28	8.85**	0.28	9.07**	0.30
Education	−0.03	0.14	< 0.01	0.05	< 0.01	0.05	0.06	0.07	−0.02	0.02	−0.03	0.02
Parent education	−0.01	0.03	−0.16*	0.08	−0.01	0.03	< 0.01	0.01	0.02	0.03	< 0.01	0.01
Country of residence	−0.20	.20	−0.21	0.21	0.50	0.48	−0.07	0.08	−0.07	0.08	−0.44*	0.19
Twofold Survey delivery	−0.19	0.25	−0.18	0.26	0.15	0.34	−0.21	0.10	−0.25**	0.10	−0.45**	0.14
Interaction terms												
Education x Twofold survey	0.03	0.15					−0.09	0.07				
Parent education x Twofold survey			0.18*	0.09					−0.03	0.04		
Country residence x Twofold survey					−0.83	0.52					0.45*	0.21
Model evaluation												
−2 log likelihood	753.1		749.2		750.5							
χ^2_{diff}	56.2**		60.1**		58.8**							
Nagelkerke R^2	0.10		0.11		0.10							
R^2							0.37		0.37		0.39	
Adjusted R^2							0.32		0.31		0.33	

Notes. Fixed effects for control variables learning objective, subject experience, ICT experience, English proficiency and course differences are included in the models, but are not presented.
$^*p < .05$, $^{**}p < .001$.

for the interaction effects between survey delivery and three demographic indicators. The interaction terms with educational attainment, highest parental education and country of residence were included in three separate models, in order to avoid the confounding impact of other interaction terms in the model. It should be noted that more restricted samples were used for the analyses. First, those who did not attempt graded assignments were not included in the analyses on course completion and second, those who had not completed the courses were eliminated in the analysis focused on the influence on grade average. Although this might reduce the power to detect effects, it was relevant as to control for other factors that might play a role in the estimation of the effects.

For course completion, logistic regression models were assessed. There were no significant issues with respect to multicollinearity (VIF > 2). With regards to model fit, χ^2 tests indicated that the models with control variables explained significantly more variance in the outcome than a null model. Still, Nagelkerke R^2 reflected that the percentage of explained variance remained quite small. After checking these assumptions, the statistics that were relevant for answering the research question were checked. For educational attainment, the independent effect and the interaction effect with the twofold survey delivery were both insignificant and practically negligible. This suggested that educational attainment did not particularly affect course completion, and that the twofold survey delivery would not provide a different conclusion. For parental education, the interaction term with the twofold survey delivery, $B = 0.18$, $p < .05$, was significantly larger than zero. In turn, the independent effect of parental education was significant in a negative direction, $B = -.16$, $p < .05$. This meant that the email-based survey would estimate a negative association between parental education and the likelihood of completion, controlling for other variables in the model. However, this association would be practically negligible if the results were based on the twofold survey delivery. Hence, survey delivery appeared to play a considerable role in the estimation of this association. For the final predictor country, the independent effect and its interaction effect with the twofold survey delivery were practically large, yet insignificant. The main explanation for the insignificance is the relatively large variance for both coefficients. Although the effects can therefore not be generalized, the size and negative direction of the interaction effect suggests that the twofold survey delivery might reflect a smaller likelihood of completion for learners from developing countries than the email-based survey.

For the models predicting grade average, assumptions for linear regression models were checked. Multicollinearity (VIF > 2) did not appear to be an issue and assumptions of linearity and homogeneity of variance were satisfied. However, the assumption of normality could not be satisfied in the models with control variables, due to high levels of kurtosis. This was an indication that the power to detect effects could be attenuated (Stevens, 2009).

Outcomes of the models demonstrated that the independent effects of educational attainment and parental education on grade average, as well as their interactions with survey delivery, were insignificant and negligible (see Table 5). This meant that neither education nor parental education seemed a relevant predictor for average grades, and that survey delivery did not alter the representation of this effect. However, the interaction term between country of residence and survey delivery, $B = 0.48$, $p < .05$, was significant and substantial. It demonstrated that, controlling for other variables in the model, the estimated average grade for learners from developing countries would be 0.48 points higher if only the twofold survey was used. Another important finding was that the independent effect of survey delivery was significantly negative. This indicated that the twofold survey delivery, in comparison to the email-based survey, reached learners who received on average lower grades.

5. Discussion

5.1. Key findings

The purpose of this study was to examine to what extent a particular alternative for the email-based delivery of demographic MOOC surveys yields higher response rates, whether the survey delivery has an impact on the representation of MOOC learners' demographics and on the estimated association between demographics and learning outcomes.

The findings indicate that the twofold survey delivery yields a substantially and significantly higher response rate than the email-based survey. The higher response rate was attributable to the relatively large proportion of response to the survey that was embedded in the MOOC environment. In line with the social exchange theory, the higher response rate could be explained by the fact that the survey in the learning environment holds fewer inconveniences than the email-based survey (Porter, 2004). As previous research has considered several inconveniences related to email surveys, including the fact that they are time-consuming and difficult to process or notice in full email inboxes (Fan & Yan, 2010; Reips, 2000, pp. 95–104; Sax et al., 2003), it remains unknown what specific factors were most important for raising response. Nevertheless, the substantial higher response rate that is detected for using the twofold survey delivery, and specifically to the embedded survey delivery, suggests that an alternative for the email-based delivery is a helpful strategy to raise response rates.

Regression analyses showed that the twofold survey delivery only yields significantly different results with regards to the age of MOOC learners. Still, the results do not suffice to reject the possibility of response errors. When higher response rates do not approximate complete response like in this study, they merely reduce the likelihood of unrepresentative results (Schouten et al., 2009). As it is previously shown that men, ethnic minorities and those without higher education degrees are generally under-represented in different types of web-surveys due to their non-response (Bandilla et al., 2014), this tendency could still cause a misrepresentation in the twofold survey delivery. Nevertheless, the increased response bears relevance for the representativeness of the demographic data, since datasets with lower proportions of missing data constitute a more solid basis for statistical methods that account for non-response to yield reliable results (Schouten et al., 2009).

The results further indicated that the twofold survey delivery yields a different estimated association between parental education and completion and estimates, on average, higher grades for learners from developing countries than the email-based survey. This substantiates that response errors might complicate the validity of evidence on the association between the demographic background of MOOC learners and their learning outcomes (e.g., van de Oudeweetering & Agirdag, 2018). Hence, representativeness seems to be an issue that should also be considered when using the data of demographic MOOC surveys in as predictors in statistical analyses. Furthermore, the twofold survey reached learners with lower average grades than the email-based survey. As previous findings indicated that response was especially low among those with lower grades (Gates et al., 2013), the twofold survey delivery might be a solution to reach these learners.

5.2. Limitations and suggestions for future research

Although the study has provided relevant insights into the importance of survey delivery for the outcomes of MOOC studies, there were some limitations that need to be considered when interpreting the results. Furthermore, these limitations may help to formulate valuable directions for future research.

First, the results are not sufficient to conclude or dismiss a response error in either one of the survey deliveries, due to the considerable degree of non-response. Although some degree of non-response can be expected when surveys remain voluntary, there are still research opportunities to advance knowledge on the representativeness of non-mandatory MOOC surveys. For example, background characteristics, like their profile settings and activity patterns, of respondents and non-respondents could be compared in order to make inferences about representativeness. Statistical methods like response-propensity models could further help to estimate and account for response errors (Kizilcec, 2014). Especially since increasing response can improve the reliability of these techniques (Schouten et al., 2009), the outcomes for different survey deliveries may be useful for assessing the value of survey delivery in advancing these procedures. However, in-depth consideration of the ethical concerns, approval of an Ethics committee on the use of non-respondents' background data, as well as the adoption of a passive informed consent, is needed in order to conduct such analyses. Future research is therefore encouraged to anticipate these ethical concerns.

Furthermore, there were some technical difficulties that had consequences for the results. First, it was impossible to send the email-based survey to only one specific branch, i.e., experimental condition. Therefore, the alternative embedded survey had to be combined with the email-based survey. Although the findings indicate that mainly the embedded survey yielded higher response rates, it could be that its integration with the email-based survey had specific effects. For example, it could be that email invitations reminded the learners in the experimental condition to respond to the embedded survey. Hence, as soon as the technical functionalities in MOOC platforms allow for the possibility of sending emails to distinct branches, it will be relevant to examine the effect of the embedded survey delivery separately. A second technical issue was that a substantial proportion of learners in the selected MOOC sessions received no or only one email. In order to ensure that all learners in the sample experienced the same research conditions, a substantial share of the sample had to be eliminated. In turn, this could have reduced the power of the outcomes. Finally, the study initially intended to include employment status as a learner demographic. However, due to technical difficulties, the data did not reveal which response was selected and employment status as a demographic indicator could not be included. Hence, future research is recommended to anticipate these technical limitations and look for solutions to enhance the precision, power and comprehensiveness of the results.

A final limitation concerns the scope of the study, as only MOOCs facilitated by the University of Amsterdam on the platform Coursera were included. This means that the findings cannot be generalized to other MOOCs and other platforms. As it is still relevant to know how MOOCs and platforms may differ in their ability to reach learners from different backgrounds, the findings appeal to more large-scale and cross-platform research on demographics of MOOC learners and the issue of representativeness.

5.3. Conclusion

This study provides a unique contribution to the literature on demographic data in MOOCs and on web-surveys in general. This study was innovative in nature, as no previous research has examined the impact of alternative survey deliveries on response and response errors in demographic MOOC surveys. A specific quality of this study was its research design. Using a randomized controlled experiment rather than a cross-sectional design helped to improve the validity of the results by ruling out history and maturation effects. Furthermore, the study was able to account for differences between MOOCs in different subject domains.

By focusing on survey deliveries that might make response more convenient, this study provided insight on a new strategy to raise response in MOOC surveys without violating the ethical consideration of voluntary research participation. Hence, the study could inform research in MOOCs and other online contexts to replace email-based surveys with other deliveries that are less time-consuming, more visible and more persuasive. Although the response rate to the examined alternative survey delivery was not sufficiently high to assure representative results, the results did generate knowledge that is valuable for the research context of MOOCs. First, this study breaks new ground by raising awarenessabout possibilities to increase response rates and to examine representativeness in MOOC studies. Moreover, the study presents a data collection strategy that generates higher responses rates than conventional strategies used in current MOOC research, which can benefit statistical methods to account for missing information to generate more precise results. Overall, the study provides well-informed reasons and strategies to scrutinize and promote the validity of research in MOOCs. This may help to gain more accurate knowledge on the potential of MOOCs to reach learners in disadvantaged positions and experience barriers to enter the traditional higher education system. In this way, doubts or expectations about MOOCs as tools to enhance social equality in education and to improve social mobility may receive more valid answers.

References

Annear, M. J., Toye, C. M., Eccleston, C. E., McInerney, F. J., Elliott, K. E. J., Tranter, B. K., et al. (2015). Dementia knowledge assessment scale: Development and preliminary psychometric properties. *Journal of the American Geriatrics Society, 63*(11), 2375–2381.

Bandilla, W., Couper, M. P., & Kaczmirek, L. (2014). The effectiveness of mailed invitations for web surveys and the representativeness of mixed-mode versus internet-only samples. *Survey Practice, 7*(4), 1–9.

Belanger, Y., & Thornton, J. (2013). *Bioelectricity: A quantitative approach Duke University's first MOOCs.* http://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/6216/Duke_Bioelectriciy_MOOC_Fall2012.pdf/ Accessed October 17 2016.

Christensen, G., Steinmetz, A., Alcorn, B., Bennett, A., Woods, D., & Emanuel, E. J. (2013). *The MOOC phenomenon: Who takes massive open online courses and why ?* (SSRN Working paper No. 10) http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2350964/ Accessed October 17 2016 .

Cisel, M., Bachelet, R., & Bruillard, E. (2014). Peer assessment in the first French MOOC. In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.). *Proceedings of the 7th international conference on educational data mining*London, UK: EDM Analyzing assessors' behavior.

Colas, J. F., Sloep, P. B., & Garreta-Domingo, M. (2016). The effect of multilingual facilitation on active participation in MOOCs. *International Review of Research in Open and Distance Learning, 17*(4), 280–313.

Couper, M. P., Kapteyn, A., Schonlau, M., & Winter, J. (2007). Noncoverage and nonresponse in an Internet survey. *Social Science Research, 36*(1), 131–148.

Fan, W., & Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior, 26*(2), 132–139.

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opinion Quarterly, 73*(2), 349–360.

Gates, K., Wilkins, D., Conlon, S., Mossing, S., & Eftink, M. (2014). Maximizing the value of student ratings through data mining. In A. Peña-Ayala (Ed.). *Educational data mining* (pp. 379–410). Cham: Springer International Publishing.

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science, 8*(3), 206–213.

Grandcolas, U., Rettie, R., & Marusenko, K. (2003). Web survey bias: Sample or mode effect? *Journal of marketing management, 19*(5–6), 541–561.

Greene, J. A., Oswald, C. A., & Pomerantz, J. (2015). Predictors of retention and achievement in a massive open online course. *American Educational Research Journal, 52*(5), 925–955.

Han, V., Albaum, G., Wiley, J. B., & Thirkell, P. (2009). Applying theory to structure respondents' stated motivations for participating in web surveys. *Qualitative Market Research: An International Journal, 12*(4), 428–442.

Ho, A. D., Chuang, I., Reich, J., Coleman, C. A., Whitehill, J., Northcutt, C. G., et al. (2015). *Harvardx and MITx: Two years of open online courses fall 2012-summer 2014.* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2586847/ Accessed May 28 2016.

Hood, N., Littlejohn, A., & Milligan, C. (2015). Context counts: How learners' contexts influence learning in a MOOC. *Computers & Education, 91*, 83–91.

Kay, J., Reimann, P., Diebold, E., & Kummerfeld, B. (2013). MOOCs: So many learners, so much potential. *IEEE Intelligent Systems, 28*(3), 70–77.

Kizilcec, R. F. (2014). Reducing nonresponse bias with survey reweighting: Applications for online learning researchers. In M. Sahami, A. Fox, M. A. Hearst, & M. T. H. Chi (Eds.). *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 143–144). Atlanta, GA: ACM.

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In D. Suthers, K. Verbert, E. Duval, & X. Ochoa (Eds.). *Proceedings of the third international conference on learning analytics and knowledge* (pp. 170–179). Leuven, Belgium: ACM.

Kizilcec, R. F., Saltarelli, A. J., Reich, J., & Cohen, G. L. (2017). Closing global achievement gaps in MOOCs. *Science, 355*(6322), 251–252.

Laguilles, J. S., Williams, E. A., & Saunders, D. B. (2011). Can lottery incentives boost web survey response rates? Findings from four experiments. *Research in Higher Education, 52*(5), 537–553.

Lee, K. (2017). Rethinking the accessibility of online higher education: A historical review. *The Internet and Higher Education, 33*, 15–23.

van de Oudeweetering, K., & Agirdag, O. (2018). MOOCs as accelerators of social mobility? A systematic review. *Journal of Educational Technology and Society, 21*(1), 1–11.

Phan, T., McNeil, S. G., & Robin, B. R. (2016). Students' patterns of engagement and course performance in a massive open online course. *Computers & Education, 95*, 36–44.
Porter, S. R. (2004). Raising response rates: What works? *New Directions for Institutional Research, 121*, 5–21.
Reips, U. D. (2000). Context effects in web-surveys. In B. E. Batinic, U. D. E. Reips, & M. E. Bosnjak (Eds.). *Online social sciences*. Göttingen: Hogrefe & Huber Publishers.
Rizzardini, R. H., Gütl, C., Chang, V., & Morales, M. (2014). MOOC in Latin America: Implementation and lessons learned. In L. Uden, Y. H. Tao, H. C. Yang, & I. H. Ting (Eds.). *The 2nd international workshop on learning technology for education in cloud* (pp. 147–158). Amsterdam: Springer Netherlands.
Sax, L. J., Gilmartin, S. K., & Bryant, A. N. (2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education, 44*(4), 409–432.
Schouten, B., Cobben, F., & Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Survey Methodology, 35*(1), 101–113.
Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences.* London: Routledge.
Stich, A. E., & Reeves, T. D. (2017). Massive open online courses and underserved students in the United States. *The Internet and Higher Education, 32*, 58–71.
United Nations (1999). *Standard country or area codes for statistical use (M49)*. https://unstats.un.org/unsd/methodology/m49/ Accessed March 23 2017.