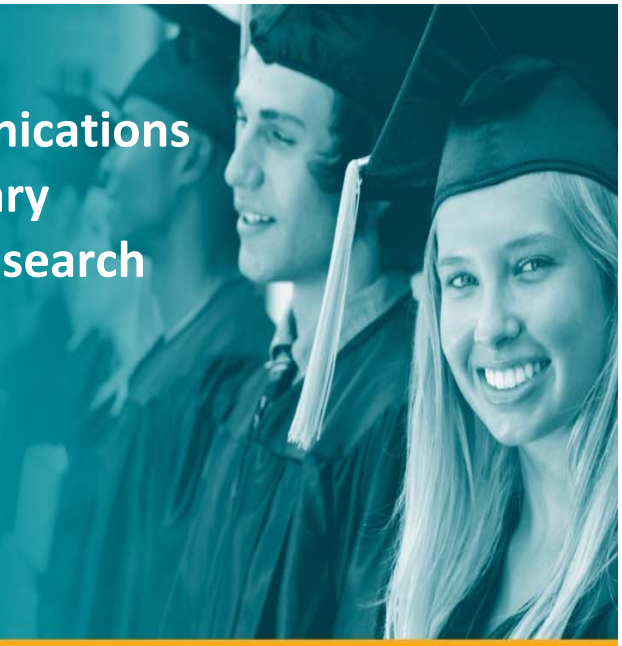
Higher Education Quality Council of Ontario

An agency of the Government of Ontario

The Impact of Incentives, Communications and Task Demand on Postsecondary Student Participation in Online Research



Julie Peters, Chris Hall and Rod Skinkle

Published by

The Higher Education Quality Council of Ontario

1 Yonge Street, Suite 2402
Toronto, ON Canada, M5E 1E5

Phone: (416) 212-3893
Fax: (416) 212-3899
Web: www.heqco.ca
E-mail: info@heqco.ca

Cite this publication in the following format:

Peters, J., Hall, C. and Skinkle, R. (2017) *The Impact of Incentives, Communications and Task Demand on Postsecondary Student Participation in Online Research*

Toronto: Higher Education Quality Council of Ontario.

Table of Contents

List of Tables

List of Figures

Executive Summary

Participation rates and quality of response are two of the biggest factors influencing the results of research involving postsecondary education (PSE) students. Like many other stakeholders in the PSE sector, the Higher Education Quality Council of Ontario (HEQCO) seeks to better understand what best motivates students to participate in voluntary studies. It is for this reason that HEQCO commissioned a study to examine the impact of varying incentives, communications strategies and task demands on postsecondary student participation in online research. HEQCO commissioned this study with the stated aim of informing its ongoing work with the Essential Adult Skills Initiative (EASI), which uses an online test to measure changes in literacy, numeracy and problem-solving skills in college and university students from the time they begin their programs to the time they graduate. Participation rates and quality of response for online research tasks will play a major role in EASI, and having a clear understanding of how task demand, incentives and communication strategies impact these factors can help HEQCO ensure the success of this project.

To examine whether messaging, incentives and task difficulty would have an effect on response rates, email open rates, data quality and sample composition, an experiment was conducted that included two incentive types (monetary and non-monetary), two communication types (regular and enhanced) and two different tasks (high-demand and low-demand). The low-demand task was a short, online survey that took approximately 15 minutes to complete, while the high-demand task was a lengthy assessment that took upwards of 90 minutes to complete. This provided eight conditions (two task demands × two incentive types × two communication types). A sample of 8,000 current postsecondary students in Ontario was selected from two sample sources, and all participants were randomly assigned to one of the eight possible conditions. The monetary incentive was a $10 Amazon gift card. The non-monetary incentive, chosen for its potential salience for students, was premium access to Paddle, an online motivation and career exploration service.

Overall, it was found that students who were sent regular — as opposed to enhanced — email communication opened the email invitation at a higher rate (although the effect size was small and the higher open rate only translated into a small increase in the response rate). Response rates did, however, vary depending on the incentive offered. Those who were offered a monetary incentive were more likely than those who were offered a non-monetary incentive to complete either the high- or low-demand task. Response rates were also significantly higher for the low-demand task compared to the high-demand task. In terms of response quality, low proportions of respondents for both tasks completed the task faster than expected and this did not vary across the eight conditions.

An analysis of demographic and academic characteristics found few significant differences. However, additional analysis examining results separately by incentive type, communication type and task demand found a higher proportion of respondents aged 20 years or older and in their third year of study or higher among those offered the non-monetary incentive compared to the monetary incentive. Given that the non-monetary incentive was premium access to a career exploration site, it is likely that this tool was more

appealing to students who were closer to graduation. Comparing the characteristics of students who completed the high-demand and low-demand tasks, the data showed that a larger proportion of students who spoke Cantonese or Mandarin as their first language completed the high-demand task than the low-demand task, as did students in their third year of study or higher and students in science programs.

These results support the notion that monetary incentives will yield a higher response rate than non-monetary incentives, regardless of the type of messaging or the task demand. They also suggest that email messages with a clear and concise subject line will have a better chance of being opened compared to messages delivered with a more casual subject line, and that communications using a more professional tone may yield slightly higher response rates than those with more casual or conversational messaging. Finally, the very low completion rates on the high-demand task suggest that in addition to traditional incentives, particularly demanding tasks are likely to require more innovative ways of encouraging student participation. We conclude by offering areas for future research.

Introduction

HEQCO regularly sponsors and participates in a variety of studies designed to examine PSE students' experiences and/or outcomes. The success of these studies, which frequently involve multiple institutions and may take place over weeks, months or even years, often depends on high levels of student participation. Across all sectors, encouraging participation in research is an ongoing challenge. Obtaining high participation levels in PSE can be particularly challenging, as students are often asked to participate in multiple studies each year, ranging from course evaluations, to policy studies, to provincially mandated Key Performance Indicator (KPI) assessments on top of their regular course load. These multiple requests to participate in research are frequently cited as causing participation fatigue, which is thought to be a key cause of declining participation rates. Yet research focused on the administration of surveys has found that the relationship between response rates and the number of surveys students are invited to participate in is not linear, suggesting that the number of studies students are invited to is important, but not the only consideration influencing research participation (Porter et al., 2004).

Incentives are frequently used in research studies involving students as a means of increasing participation. While there is an extensive body of literature documenting the impact of using various incentives in research more generally (e.g., Fox, Crask & Kim, 1988; Göritz, 2006; Heberlein & Baumgartner, 1978; O'Neil, Abedi, Miyoshi & Mastergeorge, 2005), there is little research regarding the efficacy of incentives to boost research participation in higher education, particularly among Canadian students. Non-monetary incentives, which are sometimes offered by institutions as a means of providing an incentive that is related to campus or student life, have been the focus of even less research. Given the quantity and variety of research that is ongoing across the PSE sector and its importance to supporting the overall PSE planning and assessment landscape, this study was conducted to examine the influence of incentives on student participation in online research. In addition to the influence of incentives, varying email communications and task demands were also explored.

Literature Review

This literature review focuses on research related to incentives, communications and task demand for surveys and skills assessments, as these were the two types of research that were the focus of our study. It is important to note, however, that there is far more literature on the administration of surveys than the administration of skills assessments. Therefore, much of the literature reviewed and much of the discussion in this section focuses on surveys.

Incentives

Impact on Research Participation

In two systematic reviews of the literature on incentive use in survey research (not specific to the PSE context), Simmons and Wilmot (2004) and Singer and Ye (2013) both conclude that offering monetary

incentives increases response rates across many different survey types. However, most research on the impact of incentives has been conducted for mail, telephone and in-person surveys, in which prepaid incentives (i.e., incentives paid whether or not respondents complete the task) are the most common incentive type used.[1] Studies of prepaid incentives in web surveys have had mixed results, with some finding that they have a positive effect on response rates compared to no incentive (Messer & Dillman, 2011; Millar & Dillman, 2011; Parsons & Manierre, 2014), and others finding that they have no impact on response rates (Bosnjak & Tuten, 2003; Coopersmith et al, 2016).

Studies of other incentive types, such as promised,[2] charity and lottery incentives, have generally been found to be ineffective in increasing response rates, particularly when compared to prepaid incentives (Cantor, O'Hare, & O'Connor, 2008; Church, 1993; Porter & Whitcomb, 2003; Singer & Ye, 2013). Again, these studies have primarily been in relation to mail, telephone or in-person surveys. Research on the effectiveness of incentives in web surveys is much more limited.

Göritz (2006) summarizes research published prior to 2006 in a meta-analysis, finding that across the studies examined, incentives do have a positive impact on web survey completion; yet the effect is quite small. It is important to note that most of the studies in the meta-analyses used lottery incentives. More recent research relating to studies administered online have had mixed results. For example, Hsieh and Kocielnik (2016) investigated the effect of offering fixed, lottery or charity rewards on participation rates in a crowdworking[3] context using Amazon Mechanical Turk, an online marketplace for work. They found that a fixed monetary incentive was most effective in incentivizing participation, followed by a lottery monetary incentive, no incentive and then a charity incentive. In contrast to these findings, Marcus, Bosnjak, Lindner, Pilischenko and Schütz (2007) conducted an online survey of 2,152 owners of personal websites and found that lottery incentives did not have a significant effect on response rates compared to no incentive. Differing populations of interest as well as differing incentive amounts could account for the divergence in findings.

Studies that specifically target postsecondary student populations have also had mixed findings. Porter and Whitcomb (2003) examined the impact of lottery incentives on a survey of prospective college applicants using an experimental design that compared no incentive to four incentive conditions, which were all postpaid lottery incentives ranging from $50 to $200 in value. Response rates across the conditions were similar, with the only significant difference being found between the response rate for the $100 lottery condition compared to the no incentive condition. However, the authors note that the difference was not substantive (2.3 percentage points) and that greater lottery amounts did not result in higher response rates.

1 Drawing on social exchange theory (Emerson, 1976), it is posited that a prepaid incentive, even of a small amount, conveys a sense of trust to the person receiving the incentive and exerts pressure to conform to the norm of reciprocity.
2 A promised incentive refers to an incentive where respondents are notified they will receive it after they have completed the research task.
3 Crowdworking is a type of crowdsourcing in which tasks are outsourced to a large pool of online workers. The tasks are typically small, require little to no specialized skills and payment is generally low.

A more recent study of incentive use in a web survey of U.S. college students examined two conditions, a $10 prepaid incentive only (condition 1) and a $2 prepaid incentive paired with a $10 promised incentive that would be received upon survey completion (condition 2). The researchers found that the combination of the $2 prepaid incentive with the $10 promised incentive resulted in a higher survey response rate and more complete responses than the $10 prepaid incentive. The response rate difference was 47.5% for condition 2 compared to 42.3% for condition 1 (Patrick, Singer, Boyd, Cranford & McCabe, 2013). Finally, van Veen, Göritz and Sattler (2016) conducted an experiment among German university students that involved five conditions:

1) An email invitation to the survey only
2) Mailed pre-notification and an email invitation
3) Mailed pre-notification, an email invitation and a voucher to be received after survey completion
4) A prepaid voucher in the mail, mailed pre-notification and an email invitation
5) Prepaid cash in the mail, mailed pre-notification and an email invitation

Only the prepaid cash incentive (condition 5) was found to increase response rates above the pre-notification only condition (condition 2). It is important to note that these studies all used different incentive types and amounts; therefore, the results are not directly comparable.

Compared to survey-based research, there have been far fewer studies focusing on the use of incentives and participation in research tasks such as skills or learning assessments. Steedle (2010) examined the interrelationships between students' preferred incentives for taking a test, their performance motivation and their observed performance on the Collegiate Learning Assessment.[4] Results indicate that cash was the most desirable of all incentives for Steedle's group, followed by academic incentives such as early course registration. Research by Sims and Hiatt (2003) found that a financial incentive was essential for getting students to participate in test-based assessments, as 50% of respondents said that they would only take the given test for pay. In addition to financial incentives, academic incentives have also proven to have value in improving student response. Chevalier, Dolton and Lührmann (2017) found that students in a PSE setting are not motivated by the prospect of a prize going to the top performer on a test. However, making the assessment count for in-course grades induced a participation rate that was close to the one achieved when the quizzes were academically mandatory. Brase (2009) has observed that performance on moderately difficult tasks (compared to very easy and very hard tasks) is most sensitive to incentives.

4 The Collegiate Learning Assessment is a standardized testing initiative used for higher educational evaluation and assessment in the United States. Rather than only measuring the individual student's performance, it uses a "value-added" outcome model to examine a college or university's contribution to student learning, making the institution, rather than the individual student, the primary unit of analysis.

Impact on Data Quality and Sample Composition

While much of the literature on the use of incentives in research focuses on the effect of incentives in encouraging participation, there has also been some consideration of the impact of incentives on other outcomes, such as data quality and sample composition. Numerous theories are put forward as to why and how incentives may impact the quality of data respondents provide and the type of respondents who participate in incentivized research. For example, some have suggested that an incentive may induce respondents to spend a greater amount of time answering a survey (Singer et al., 2000), while others posit that incentives may lead respondents to fill in inaccurate data in order to quickly complete the survey and obtain the incentive (Göritz, 2006). Similarly, differing perspectives have been put forward regarding the effect of incentives on sample composition. Some suggest that incentives could work to attract groups that would otherwise be less likely to respond (Porter & Whitcomb, 2003), while others caution that incentives may motivate only those who value that reward and could have little impact or a negative impact on those not interested in the reward (Hsieh & Kocielnik, 2016).

Studies on the relationship between incentives and survey data quality have generally examined item nonresponse as the primary data quality measure, finding that incentives are related to a reduction in the amount of missing data (Singer et al., 2000; Simmons & Wilmot, 2004). Item nonresponse is less relevant in web surveys, however, since advancing to the next question in the survey can be made conditional upon responding to the previous item. In web surveys, speeding and straightlining are more commonly used as indicators of data quality (Revilla, 2016). The relationship between these indicators and incentive use, however, is unknown.

For higher demand tasks such as learning assessments, Steedle (2010) found that preferred incentives showed practically no significant relationship with test performance on the Collegiate Learning Assessment. Similarly, Sims and Hiatt (2003) found no correlation between students' self-reported effort levels on test-based assessments and whether they had indicated that they would only do the assessment for pay. The effect of the stakes attached to testing situations has also been examined. Attali (2016) compared the performances of the same group of students in low-stakes and high-stakes testing situations. In this study, students who had just finished taking their GRE were invited to take a low-stakes test, the results of which were compared against the high-stakes GRE. The research found that for more than 80% of the total sample, the low- and high-stakes scores were highly correlated.

With regard to sample composition, Singer and Ye (2013) conclude, based on their meta-analysis, that incentives can affect the composition of a sample. In particular, they cite studies showing that monetary incentives can result in a larger proportion of traditionally harder-to-recruit respondents, including less-educated, low-income and minority respondents as well as individuals who are less interested in the topic of the research. Thus, these studies suggest that offering incentives can reduce nonresponse bias by motivating individuals who are usually underrepresented in research. However, the cited studies were primarily based on mail surveys; there has been far less research related to web surveys or skills assessments. The few available studies on the impact of incentives on sample composition in online research have generally found no difference in the demographic characteristics of respondents across

different incentive conditions (Hsieh & Kocielnik, 2016; Porter & Whitcomb, 2003; Marcus et al, 2007). The one exception is Parsons and Manierre (2014), who found that prepaid monetary incentives were more effective in soliciting responses from women than men, which they argued can exacerbate the overrepresentation of women in research.

An abbreviated version of Singer and Ye's (2013) six key conclusions, which they describe as "The Use and Abuse of Incentives," offers an excellent summary of key findings from this field of research:

1) Incentives increase response rates to surveys in all modes.
2) Monetary incentives increase response rates more than gifts, and some forms of incentives are not suitable for web-based surveys (i.e., providing the incentive prior to beginning the survey).
3) There is no good evidence for how large an incentive should be.
4) Few studies have evaluated the effect of incentives on quality of response, and more research is needed on what effect, if any, they have on reliability and validity.
5) Research on the effect of incentives on sample composition and response distribution is sparse, and those who have studied this have found no significant effects.
6) Incentives have the potential to both increase and reduce response bias. If they can be targeted at those who would otherwise not respond, they may reduce nonresponse bias, but if they attract more participants who would have participated anyway, nonresponse bias may increase.

Communications and Task Demand

In addition to incentives, the impact of research communications and task demand are also of interest to this study. With regard to research communications, Leverage-Saliency Theory (Groves, Singer & Corning, 2000) is often used to explain how participants make decisions about whether or not to participate in research. This theory posits that individuals will assign different importance to different aspects of a research request, and that their decision will also depend on the salience of each aspect. Aspects of a research request could include, for example, the individual or organization inviting potential respondents to participate in the research, details on how the data will be used, the length of time the research will likely take to complete and the incentive offered. The invitation is thus vitally important to participation decisions, as it communicates the key attributes of the research and can make different attributes more or less significant. While there are a number of studies that have examined the impact of survey invitations on response rates, these have primarily focused on issues around personalization and which content to include, as well as the number of invitations and reminders that are sent (Fan & Yan, 2010). Further, research performed by Trouteaud (2004) has found that in email communication, subject lines worded as a "plea" tend to outperform subject lines that pose an "offer" when soliciting participation in web-based surveys. Consistent with the Leverage-Saliency Theory posed above, Porter and Whitcomb (2005) also found that when groups are aware of a study's sponsor, there can be significant differences based on the respondents' levels of attachment to the sponsor. Those with a high level of attachment to the research sponsor did not respond differently to different subject lines.

Considerably more research is available regarding the impact of task demand on participation, though the vast majority has been conducted in the context of mailed surveys. A meta-analysis of 292 randomized controlled trials involving mail surveys conducted by Edwards et al. (2002) found a strong negative linear relationship between survey length and response rates. However, in two meta-analyses of non-experimental studies involving web surveys, no significant relationship between questionnaire length and response rates was found (Cook, Heath & Thompson, 2000; Sheehan, 2001). Galesic and Bosnjak (2009) argue that these may not be accurate representations of the relationship between survey length and response, as neither of the meta-analyses controlled for whether respondents were informed of the survey length in advance. In studies involving web surveys where respondents were informed of the survey length in advance, shorter survey lengths were associated with significantly higher response rates (Crawford, Couper & Lamias, 2001; Marcus et al, 2007; Galesic & Bosnjak, 2009).

Summary

Given the volume of studies regularly undertaken by Canadian postsecondary institutions and stakeholders, paired with the ongoing challenge of declining participation rates, researchers will need to find innovative ways to recruit potential participants. While the existing literature provides a useful starting point for understanding the impact of incentives, there is little known about the impact of incentives among postsecondary students in Canada. In addition, non-monetary incentives that are related to a students' education have not been examined, and there has been very little examination of the influence of different communication types in soliciting student participation. This research project aims to provide insight about incentive types and the tone of research invitation messaging to determine whether these factors could influence the rate of response among Canadian postsecondary students for varying task demands.

Method

Framework

An experimental design was executed that included two incentive types (monetary and non-monetary), two communication types (regular and enhanced), and two different tasks (high-demand and low-demand). The high-demand task was a lengthy skills assessment that took upwards of 90 minutes to complete, while the low-demand task was a short, online survey that took approximately 15 minutes to complete.[5] This provided eight conditions (two task demands × two incentive types × two communication types). A sample of 8,000 current postsecondary students in Ontario was selected, and all participants were randomly assigned to one of the eight possible conditions, as outlined in Table 1.

5 The high-demand task was chosen due to HEQCO's ongoing Essential Adult Skills Initiative work, which involves recruiting college and university students to complete the OECD Education and Skills Online (ESO) assessment. To the best of our knowledge, there are no comparable, validated low-demand skills assessments. Therefore, a survey was chosen for the low-demand task because relatively short online surveys are often used by Ontario colleges and universities.

Table 1: Experimental Design

Task Demand	Incentive Type				Total
	Monetary		Non-monetary		
	Communication Type		Communication Type		
	Enhanced	Regular	Enhanced	Regular	
High	1,000	1,000	1,000	1,000	4,000
Low	1,000	1,000	1,000	1,000	4,000
Total	2,000	2,000	2,000	2,000	8,000

Procedure

The sample of 8,000 current Ontario postsecondary students was identified and invited to participate from two separate sample sources. The first source was Academica Group's proprietary Future Research Pool and student panel.[6] The second source was a list of registered student users from EDge Interactive.[7] The sample sources were filtered to include current Ontario students only and then 4,000 students were randomly selected from each source. The purpose of using two sources was to provide a more diverse sample to help control for possible biases present in either of the individual sample sources.[8] It must be noted, however, that the sample is not representative of the entire PSE population in Ontario given that a random selection of all current PSE students was not undertaken.

Survey invitations to all 8,000 invitees were sent from an Academica Group email address (surveys@academicagroup.com) via MailChimp[9] on February 16, 2017. Reminder messages were sent on March 1, March 9 and March 15. Survey campaigns were closed on March 19.

A letter of information and consent appeared on the landing page of the survey. Respondents were informed that the study was being conducted by Academica Group, with funding from HEQCO, but were not informed that the objective of the study was to assess the effectiveness of incentive types and communications styles until after they had completed the task.

6 Academica Group regularly invites research participants to future research opportunities. They are stored either in a future research pool database (n=100,000), or engaged regularly within the StudentVu research panel (n=4,500).
7 EDge Interactive has over 750,000 users who are recruited through their family of education websites: SchoolFinder.com, StudyinCanada.com and ScholarshipsCanada.com.
8 The primary difference between the two sample sources is that the pool from Academica had previously participated in a research study and consented to be invited to future research, whereas the EDge Interactive pool had signed up through their family of education websites and consented to third party communications.
9 MailChimp is an email service and marketing platform.

Tasks

Low-demand

The low-demand task was an instrument designed by Academica Group in consultation with HECQO to represent a relatively brief online survey that postsecondary students are commonly asked to complete. It was described in the email invitation as a 15-minute survey on the relationship between student life goals, well-being and learning. The survey instrument was programmed on Academica's survey software, which is compatible with computer, tablet and mobile devices. The start of the survey had a brief set of screener questions to ensure that respondents qualified to participate.

High-demand

The high-demand task was the OECD Education and Skills Online (ESO) assessment. This was chosen as the high-demand task given HEQCO's ongoing Essential Adult Skills Initiative (EASI), which involves recruiting college and university students to complete the ESO assessment. It was described in the email invitation as a 90-minute assessment that would give the student valuable information about the skills needed to succeed in their studies, at work and in life. The ESO assessment, which tests literacy, numeracy and problem-solving skills, is hosted on the OECD website and can only be accessed via a laptop or desktop through the Mozilla Firefox web browser. Mobile and tablet devices, as well as other browsers, are not supported. To access the test, respondents needed to click on a survey link that took them to a short screener survey hosted by Academica to ensure that they qualified to participate. They were then directed to a study landing page and invited to enter their email address to be sent an authorization code that could be used to complete the ESO assessment. The authorization code was automatically emailed along with details about how to access and complete the ESO assessment. The assessment itself consisted of three components: a background questionnaire, a core assessment, and a problem-solving assessment in technology-rich environments. Respondents needed to complete all three components in order to be classified as having finished the task.

Communication Types

The regular communication emails were representative of the type of invitation that Academica typically sends to potential survey respondents. The enhanced communications, including the subject lines, were modified from the regular communications by a professional communications firm using best-practice criteria, such as reducing the amount of copy, making the copy more conversational and using icons and images to convey key messages. The regular communications were personalized by addressing the invitation to the individual's first name, while the enhanced communications were not. For both the regular and enhanced communication groups, one initial invitation was sent as well as three reminder messages. Examples of all invitation and reminder messaging can be found in Appendix A and Appendix B.

Incentive Types

Respondents were offered one of two different incentive types for task completion: a monetary incentive or a non-monetary incentive. The monetary incentive was a $10 Amazon gift card. The non-monetary incentive, chosen for its potential salience for students, was premium access to Paddle, an online motivation and career exploration service (mypaddle.co). The incentive was provided to the respondent via email after the task was completed.

Outcome Measures

Four outcome measures were analyzed: response rates, email invitation open rates, quality of participation, and sample composition. The response rate[10] was the primary outcome measure of interest, as the key aim of the study was to examine the impact of different incentives, task demand levels and communication types on response rates. Response rates were calculated following the Marketing Research and Intelligence Association standards, which have been recommended by the federal government's Advisory Panel on Online Public Opinion Survey Quality (2008).[11]

As an additional measure of the effectiveness of the email communications, open rates were tracked. The open rate is the percentage of email survey invitations that were opened by recipients. To track open rates, MailChimp loads a small transparent image into each email and then counts each time the image is loaded by a recipient. Given that tracking the open rate relies on the image being loaded, it is not completely accurate, as a recipient's email client could have image loading turned off, in which case the image would not load and the email would not be registered as opened. However, if a recipient clicks on a link within the email, this is factored into the open rate even if the tracking image did not load, since a recipient would have to open the email to click on a link. Given the potential inaccuracies in calculating the open rate, it is more useful to compare open rates as a relative measure across conditions than as an absolute measure of the number of recipients who opened the email invitation.

The study also tested whether the experimental conditions had an impact on the quality of response. Two measures of response quality were used: time to complete task, and straightlining. The low- and high-demand data sets included timestamps indicating the date and time (hours and minutes) a respondent started and finished each task. Respondents to both the low- and high-demand tasks had the option of breaking the task up into various stages of work, as long as they completed the task before the fieldwork period of the study ended. They could reload the survey or assessment at a later time and continue from where they had left off. This meant that the length of time to complete, in some instances, was not simply the amount of time they spent on the task. To determine instances of speeding, we calculated the survey completion time for those who completed the survey within one day. Following Rossmann (2010), speeding was defined as completing the survey in less than 60% of the median completion time. For the low-demand

10 Response rate = (completed task + disqualified) / (invitations sent - bouncebacks - undeliverable)
11 Full response rate calculations are provided in Appendix C.

task, 547 of the 586 respondents (93%) completed the task within one day and the median duration was 10.5 minutes. Instances of speeding were thus categorized as any responses completed in under 6.3 minutes. For the high-demand task, 74 of the 109 respondents (68%) completed the task within one day and their median completion time was 90.5 minutes. Respondents taking less than 54.3 minutes were classified as speeders.

Straightlining refers to a respondent completing an array of questions on a screen by selecting the same response option for all items in order to complete the survey quickly. The low-demand task had seven sets of agreement statements that were used for assessing instances of straightlining. To determine whether straightlining had occurred, the variance for each set of statements was calculated. Since there are seven sets of agreement scale questions, there are seven opportunities in total to straightline. The number of times a respondent's data exhibited zero variance for any one of these sets of seven agreement items were counted in a new variable. This analysis of straightlining could only be conducted on the low-demand data, as the high-demand task data did not include any measurable data points. As a result, no response patterns could be assessed.

Given the previous literature suggesting that incentives may be more or less effective in recruiting certain demographic groups and could result in biased samples, the study also assessed sample composition. The present study is not ideal for investigating this issue, as we do not have demographic information for the entire sampling frame. However, we can compare the demographic characteristics of those who responded to the surveys across the eight survey conditions. Demographic information examined included gender, age, born in Canada, first language, program of study, institution type and year of study.

We examined the data in four different ways in all analyses where it was possible to do so.

1) Across all eight conditions
2) Between those offered a monetary incentive and those offered a non-monetary incentive
3) Between the high-demand task and low-demand task
4) Between those sent the regular communications and those sent the enhanced communications

This allowed for an assessment of differences between the various combinations, as well as the opportunity to separately assess any overall impacts of the incentive, task demand and communication strategy. Chi-square tests and z-tests were conducted. For analyses where multiple statistical comparisons were being made, a Bonferroni correction was applied to reduce the chances of obtaining false positive results (Type 1 error). Effects sizes are reported using Phi and Cramer's V.

Results

Response Rates

Response rates were calculated for all eight survey conditions separately, as well as grouped by task demand, communication type and incentive type, and a series of chi-square tests were conducted. Response rates varied widely, ranging from a low of 1% for the high-demand, enhanced communication, non-monetary incentive group to a high of 28% for the low-demand, regular communication, monetary incentive group. Overall, response rates were found to be significantly related to the survey condition (χ^2=767.381, d.f.=7, p=.000, Cramer's V=.312).

When the three factors were examined individually, task demand and incentive type had the largest impact on response rates. Students invited to the low-demand task were almost six times more likely than those invited to the high-demand task to complete the task (17% compared to 3%, respectively; χ^2=379.234, d.f.=1, p=.000, Φ=.220), and those offered the monetary incentive were about four times more likely to complete the task compared to those offered the non-monetary incentive (16% compared to 4%, respectively; χ^2=276.590, d.f.=1, p=.000, Φ=.188). Communication type was also found to be associated with response rate, but the difference was very small (χ^2=10.939, d.f.=1, p=.001, Φ=.037). The response rate among students sent the regular communications was 11%, compared to 9% for those sent the enhanced communications. Full response rate calculations are provided in Appendix C.

Table 2: Response Rates

		Response Rate
All Eight Conditions (n=1000 per condition)	Low-demand, Regular, Monetary	28%
	Low-demand, Regular, Non-monetary	8%
	Low-demand, Enhanced, Monetary	23%
	Low-demand, Enhanced, Non-monetary	7%
	High-demand, Regular, Monetary	7%
	High-demand, Regular, Non-monetary	2%
	High-demand, Enhanced, Monetary	5%
	High-demand, Enhanced, Non-monetary	1%
Task Demand (n=4000 per demand type)	Low-demand	17%
	High-demand	3%
Communication Type (n=4000 per communication type)	Regular	11%
	Enhanced	9%
Incentive Type (n=4000 per incentive type)	Monetary	16%
	Non-monetary	4%

Email Invitation Open Rates[12]

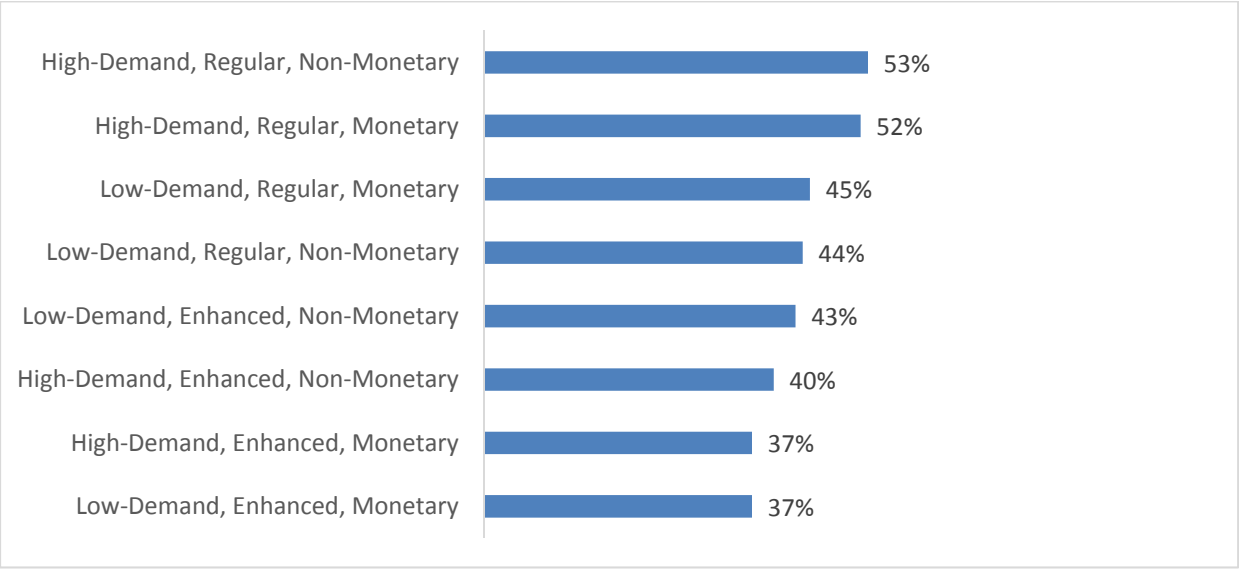
The high-demand regular survey invitation communication, offering either the monetary or non-monetary incentive, achieved the highest open rates of 52% and 53%. A chi-square test of independence found that there was an association between survey condition and whether the invitation email was opened, though the effect size was small (χ^2=112.966, d.f.=7, p=.000, Cramer's V=.120). Of particular interest was whether open rates differed between the regular and enhanced communications. Overall, regular communication emails were opened at a rate of 49% compared to the enhanced communication emails, which were opened at a rate of 39%; this was a statistically significant difference but the effect size was small (χ^2=74.667, d.f.=7, p=.000, Φ=.097). While email open rates do not necessarily translate to more completed surveys for either task type, they do suggest that using simpler, more straightforward subject-line wording increases the likelihood that students will open an email message.[13]

One puzzling finding, however, was that the regular communications for the high-demand task had higher open rates than the regular communications for the low-demand task for both the monetary (52% compared to 45%, respectively) and non-monetary (53% compared to 44%, respectively) conditions. This was puzzling because these email invitations used the same subject lines. It is not clear why this is the case; however, one possible explanation is the different wording used in the first lines of the email. Most email clients display a brief preview of the email text, sometimes known as a pre-header, after the email subject line. In the first line of the high-demand communication, students were invited to participate in a "research study"; in the first line of the low-demand communication they were invited to participate in a "survey." It is possible that the term "research study" was more appealing to students. Nonetheless, it is difficult to know for certain whether this contributed to the difference in open rates, and further research is needed.

12 This section focuses on open rates for the initial email invitation only. This is because the open rates for subsequent emails are influenced by the number of completes that have been obtained to date, since reminder emails are sent to non-completes only. Open rates for all communications sent are provided in Appendix D.

13 The subject lines for the enhanced communications were: "Discover your path" (high-demand) and "You help us. You get stuff." (low-demand). Subject lines for the regular communications were: "Participate in research for premium access to a career exploration service" (non-monetary) and "Participate in research for a $10 Amazon.ca gift card" (monetary).

Figure 1: Initial Email Invitation Open Rates (n=7866)[14]



Quality of Participation

Quality of participation, for the purposes of this study, was measured in two ways: 1) the amount of time to complete the task and 2) incidences of straightlining.

Completion Time

Overall, about one in 10 respondents to the low-demand task were classified as speeders, as they had completed the survey in less than 6.3 minutes (11%) (Table 3**Error! Reference source not found.**); a similar proportion completed the high-demand task in less than 54.3 minutes and were classified as speeders (12%) (Table 4).[15] This generally low incidence of speeding indicates that the vast majority of respondents were taking the time to read the survey questions and provide thoughtful answers. There were no statistically significant differences between the groups, meaning that there was no particular communication or incentive type that led respondents to behave differently from a task completion time standpoint.

14 Full email invitation administration data is provided in Appendix D.
15 As described in the methods section, the completion time analysis uses a subset of respondents who completed the task within one day. Results for the high-demand task should be interpreted with caution given the low n size.

Table 3: Time to Complete Low-demand Task

	Total	Regular Monetary	Regular Non-monetary	Enhanced Monetary	Enhanced Non-monetary
n size	546	224	68	200	54
Less than 6.3 minutes	11%	9%	9%	14%	6%
6.3 minutes or longer	89%	91%	91%	86%	94%

Table 4: Time to Complete High-demand Task

	Total	Regular Monetary	Regular Non-monetary	Enhanced Monetary	Enhanced Non-monetary
n size	74	41	7	24	2
Less than 54.3 minutes	12%	15%	0%	8%	50%
54.3 minutes or longer	88%	85%	100%	92%	50%

Straightlining

As described in the methods section, straightlining was measured in the low-demand task in seven sets of agreement scale questions.[16] Therefore, there were seven opportunities to straightline in the survey. It should be noted that an instance of straightlining on a survey page (selecting the same response option for each question on the page) does not necessarily indicate that the respondent was providing inaccurate answers, as they may have truly had the same level of agreement for each statement on the page. Therefore, one or two instances of straightlining within the survey are not necessarily concerning, while more instances raise concerns regarding data quality. Table 5 shows the proportion of respondents who straightlined, ranging from no instances of straightlining to seven instances of straightlining. Nearly two-thirds of respondents to the low-demand task did not exhibit any straightline behaviour (62%), while 31% straightlined one or two times, and 7% straightlined three or more times. There were no significant differences between the experimental conditions.

16 The straightlining analysis could only be completed for the low-demand task, as the high-demand task data did not contain item responses.

Table 5: Incidence of Straightlining in Low-demand Task

	Total	Regular Monetary	Regular Non-monetary	Enhanced Monetary	Enhanced Non-monetary
n size	586	241	72	212	61
No straightlining	62%	62%	62%	58%	74%
1 instance	22%	21%	28%	24%	11%
2 instances	9%	10%	8%	10%	7%
3 instances	4%	5%	1%	5%	-
4 instances	1%	0%	-	2%	2%
5 instances	1%	0%	-	1%	3%
6 instances	1%	1%	-	1%	2%
7 instances	1%	1%	-	0%	2%

Sample composition

The final outcome measure of interest was sample composition. Since the 8,000 invitees to the study were randomly assigned to one of the eight invitation/incentive/task combinations, it can be reasonably assumed that there should be no significant demographic differences across the conditions unless some combination of incentive type, communication type and/or task demand was more effective in recruiting a particular demographic group of respondents. This possibility was examined by comparing the demographic characteristics across six of the eight conditions,[17] as well as by comparing the monetary to non-monetary incentive, regular to enhanced communication type and high-demand to low-demand task. Chi-square tests were used to identify significant differences, followed by z-tests for larger contingency tables to determine which pairs of cells were significantly different.

Few significant differences were found when the demographic profiles of respondents across the six conditions were compared. A significantly greater proportion of the high-demand, monetary incentive with regular communication spoke Cantonese or Mandarin as their first language (16%) compared to the low-demand, monetary incentive with regular communication (3%). Also, a significantly greater proportion of the low-demand, monetary incentive with regular communication were in first year (58%) compared to the high-demand, monetary incentive with regular communication (28%), while significantly fewer were in third year or higher (19% compared to 39%, respectively).

When examined separately by incentive type, communication type and task demand, the data yielded a few more differences (Table 6). There was a significantly larger proportion of respondents aged 19 or younger in the monetary incentive group compared to the non-monetary incentive group (50% versus 39%, respectively), and significantly fewer who were in their third year of study or higher (23% versus 33%, respectively). No significant demographic differences were found between those who responded to the

17 Due to small n sizes in the two high-demand, non-monetary incentive groups, these could not be included in the analysis.

regular and enhanced communications. Comparing the high- and low-demand task groups, there was a smaller proportion of English first-language speakers in the high-demand group compared to the low-demand group (71% versus 79%, respectively), whereas the high-demand group had a larger proportion of respondents who spoke Cantonese or Mandarin as their first language (11% versus 6%, respectively). The composition of the high- and low-demand respondents also differed in that there was a larger proportion of first-year students in the low-demand group (50% versus 32%, respectively), and a smaller proportion of students in third year or higher (23% versus 36%, respectively). Finally, science students were overrepresented in the high-demand group compared to the low-demand group (38% versus 27%, respectively).

Table 6: Sample Composition by Incentive, Communication and Task

		Incentive		Communication		Task	
		Monetary	Non-monetary	Regular	Enhanced	High	Low
n size		546	149	382	313	109	586
Gender	Man	29%	21%	25%	30%	32%	26%
	Woman	70%	76%	74%	68%	66%	72%
	Other gender identity	1%	3%	2%	1%	2%	1%
Age	19 or younger	50%*	39%*	49%	46%	39%	49%
	20–24	38%	46%	38%	42%	47%	39%
	25 or older	12%	15%	13%	12%	14%	12%
Born in Canada	Yes	83%	80%	83%	81%	82%	82%
	No	17%	20%	17%	19%	18%	18%
First Language	English	79%	74%	79%	77%	71%*	79%*
	French	2%	3%	2%	2%	0%	3%
	Cantonese/Mandarin	6%	9%	6%	7%	11%*	6%*
	Other	14%	13%	13%	14%	18%	12%
First Generation Student	Yes	17%	11%	17%	13%	15%	16%
	No	83%	89%	83%	87%	85%	84%
Sources of Funding**	Government loans/grants	62%	67%	63%	63%	65%	62%
	Scholarships/bursaries	53%	56%	55%	52%	56%	53%
	Other source	83%	89%	84%	85%	89%	84%

		Incentive		Communication		Task	
		Monetary	Non-monetary	Regular	Enhanced	High	Low
Institution Type	College	34%	32%	33%	35%	28%	35%
	University	66%	68%	67%	65%	72%	65%
Year of Study	First year	49%	42%	49%	45%	32%*	50%*
	Second year	28%	25%	26%	30%	32%	27%
	Third year or higher	23%*	33%*	25%	25%	36%*	23%*
Primary Subject Area	Arts & Humanities	12%	11%	12%	12%	14%	12%
	Sciences	29%	27%	30%	27%	38%*	27%*
	Social Sciences & Human Services	38%	36%	38%	38%	31%	39%
	Health Sciences	15%	19%	14%	19%	15%	17%
	Other	5%	6%	6%	4%	3%	6%

**Indicates a statistically significant difference, p<.05; tests are adjusted for all pairwise comparisons using the Bonferroni correction*
*** Multiple response question, therefore numbers will add to greater than 100%*

Conclusion

Research studies are widely used among student populations to collect information related to areas such as students' experiences, satisfaction, learning outcomes, engagement and use of campus resources. Within these research studies, incentives are often used as a means of boosting response rates. However, there is little research available assessing the impact of incentives on research participation for web-based tasks among postsecondary students.

This study set out to determine whether messaging, incentives and task difficulty would have an effect on survey response rates, email open rates, data quality and sample composition. Although it was found that those who were sent regular, as opposed to enhanced, email communication opened the email invitation at a rate of 49% compared to 39%, those open rates only translated into a small increase in the survey response rate (11% compared to 9%). Response rates, however, did vary substantially by the incentive offered. Those who were offered a monetary incentive were more likely than those who were offered a non-monetary incentive to have completed either task. Response rates were also notably higher for the low-demand task than for the high-demand task. In terms of response quality, low proportions of respondents for both survey tasks completed the survey faster than expected (11% in the low-demand task, and 12% in the high-demand task), and there were few respondents who exhibited noteworthy straightlining behaviour. Neither measure of response quality was found to be associated with the survey conditions.

An analysis of demographic and academic characteristics (age, gender, born in Canada, first language, current institution type, current year of PSE and primary subject area) by the survey conditions found few significant differences. However, additional analysis examining results separately by incentive type, communication type and task demand found a higher proportion of respondents aged 20 years or older and in their third year of study or higher among those offered the non-monetary incentive compared to the monetary incentive. Given that the non-monetary incentive was premium access to a career exploration software, it is likely that this tool was preferable for students who were closer to graduation. Comparing the characteristics of students who completed the high-demand and low-demand tasks, the data showed that a larger proportion of students who spoke Cantonese or Mandarin as their first language completed the high-demand task than the low-demand task, as did students in their third year of study or higher and students in science programs. Because demographic data was not available for the entire population, however, it is not known which incentive, task and communications combination produced a sample composition that most closely resembled the population.

These results support the suggestion that monetary incentives will yield a higher survey response rate than non-monetary incentives, regardless of the type of messaging or the task demand. They also suggest that email messages with a clear and concise subject line will have a better chance of being opened compared to messages delivered with more casual subject lines, and that communications that are more professional in tone may yield slightly higher response rates than those with more casual or conversational messaging. Finally, the very low completion rates on the high-demand task suggest that in addition to traditional

incentives, particularly demanding tasks are likely to require more innovative ways of encouraging participation.

When reflecting on the research limitations of this study, as well as the direction for further research on this topic, the following should be considered:

- Starting with a population where the demographic and academic characteristics are known would provide an opportunity to undertake a deeper analysis of the results and provide insight into subgroups that would be more or less likely to open email messages, provide quality responses and complete either task, depending on the type of message and incentive they were offered. Ideally, potential research participants would be randomly selected from the population, whether it be students at a particular institution or a group of institutions.

- Including a group of respondents who were offered no incentive would provide insight into the likelihood of respondents completing the task based on interest in the task itself, and would act as a control group by which to compare other incentive types and amounts.

- Varying the sender of the email invitation could provide additional insight for improving email open rates and response rates. All email invitations for the study were sent from "Academica Group," which would have been a familiar name to some, but not all of those invited. Testing different sender names would help provide insight into the extent to which the sender impacts a students' decision of whether or not to participate in web research.

- Using a larger sample or choosing a high-demand task that is less onerous for respondents to access should also be considered for future research.[18] A key limitation of this study was that two of the eight conditions had a very low number of respondents (the high-demand, enhanced communication, non-monetary incentive group had only four completions and the high-demand, regular communication, non-monetary incentive group had only 12), which limited the analyses that could be conducted.

- More research on response quality is needed. While this study was able to measure the incidence of speeding for both the high- and low-demand task and straightlining for the low-demand task, there was no item-level data available for the high-demand task by which to examine response quality. A deeper understanding of respondent behaviour during the research process is required.

18 As noted in the method section, the high-demand task was the ESO assessment which is available through the OECD website. The assessment can only be accessed via a laptop or desktop through the Mozilla Firefox web browser. Mobile and tablet devices, as well as other browsers, are not supported. Respondents needed to enter their email in a study landing page and were then sent an authorization code that could be used to complete the ESO assessment. The authorization code was automatically emailed along with details about how to access and complete the ESO assessment.

References

Advisory Panel on Online Public Opinion Survey Quality. (2008). *The advisory panel on online public opinion survey quality: Final report*. Ottawa: Public Works and Government Services Canada.

Attali, Y. (2016). Effort in low-stakes assessments: What does it take to perform as well as in a high-stakes setting? *Educational and Psychological Measurement, 76*(6), 1045–1058. https://doi.org/10.1177/0013164416634789

Bosnjak, M. & Tuten, T. L. (2003). Prepaid and promised incentives in web surveys: An experiment. *Social Science Computer Review*, 21(2), 208–217.

Brase, G. L. (2009). How different types of participant payments alter task performance. *Judgment and Decision Making*, *4*(5), 419–428.

Cantor, D., O'Hare, B., & O'Connor, K. (2008). The use of monetary incentives to reduce nonresponse in random digit dial telephone surveys. In J. M. Lepkowski, C. Tucker, J. M. Brick, E de Leeuw, L. Japec, P. J. Lawrakeas, M. W. Link & R. L. Sangster (Eds.), *Advances in Telephone Survey Methodology.* (471–98). New York, NY: Wiley.

Chevalier, A., Dolton, P. & Lührmann, M. (2017) "Making it count": Incentives, student effort and performance. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*.

Church, A. (1993). Estimating the effect of incentives on mail survey response rates: A meta-analysis. *Public Opinion Quarterly. 57(1)*, 62–79.

Cook, C., Heath, F., & Thompson, R. L. (2000). A meta-analysis of response rates in web- or internet-based surveys. *Educational and Psychological Measurement, 60*(6), 821–836.

Coopersmith, J., Vogel, L., Bruursema, T., & Feeney, K. (2016). Effects of incentive amount and type of web survey response rates. *Survey Practice, 9*(1).

Crawford, S. D., Couper, M. P., & Lamias, M. J. (2001). Web surveys: Perceptions of burden. *Social Science Computer Review, 19*(2), 146–62.

Edwards, P., Roberts, I., Clarke, M., DiGuiseppi, C., Pratap, S., Wentz, R., & Kwan, I. (2002). Increasing response rates to postal questionnaires: Systematic review. *Bmj, 324*(7347), 1183.

Emerson, R. M. (1976). Social exchange theory. *Annual Review of Sociology, 2*(1), 335-362.

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

Fox, R. J., Crask, M. R., & Kim, J. (1988). Mail survey response rate: A meta-analysis of selected techniques for inducing response. *Public Opinion Quarterly, 52*(4), 467–491.

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.

Göritz, A. S. (2006). Incentives in web studies: Methodological issues and a review. *International Journal of Internet Science, 1*(1), 58–70.

Groves, R., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation. *Public Opinion Quarterly*, *64*(3), 299–308.

Heberlein, T. A. and Baumgartner, R. (1978). Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. *American Sociological Review, 43*(4), 447–462.

Hsieh, G. & Kocielnik, R. (2016, February). You get who you pay for: The impact of incentives on participation bias. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing.* (pp. 823–835). ACM.

Marcus, B., Bosnjak, M., Lindner, S., Pilischenko, S., & Schütz, A. (2007). Compensating for low topic interest and long surveys: A field experiment on nonresponse in web surveys. *Social Science Computer Review, 25*(3), 372–383.

Messer, B. L., & Dillman, D. A. (2011). Surveying the general public over the Internet using address-based sampling and mail contact procedures. *Public Opinion Quarterly, 75*(3), 429–457.

Millar, M. M., & Dillman, D. A. (2011). Improving response to web and mixed mode surveys. Public *Opinion Quarterly, 75*(2), 249–269.

O'Neil, H. F., Abedi, J., Miyoshi, J., Mastergeorge, A. (2005). Monetary incentives for low-stakes tests. *Educational Assessment, 10*(3), 185–208.

Parsons, N., & Manierre, M. (2014). Investigating the relationship among prepaid token incentives, response rates, and nonresponse bias in a web survey. *Field Methods*, *26*(2), 191–204. http://dx.doi.org/10.1177/1525822x13500120

Patrick, M. E., Singer, E., Boyd, C. J., Cranford, J. A., & McCabe, S. E. (2013). Incentives for college student participation in web-based substance use surveys. *Addictive Behaviours. 38*(3), 1710–1714.

Porter, S. R., & Whitcomb, M. E. (2003). The impact of lottery incentives on student survey response rates. *Research in Higher Education, 44*(4), 389–407.

Porter, S. R., & Whitcomb, M. E. (2005). E-mail subject lines and their effect on web survey viewing and response. *Social Science Computer Review, 23*(3), 380–387.

Porter, S. R., Whitcomb, M. E., & Weitzer, W. H. (2004). Multiple surveys of students and survey fatigue. *New Directions for Institutional Research*, 2004(121): 63–73.

Revilla, M. (2016). Impact of raising awareness of respondents on the measurement quality in a web survey. *Quality & Quantity, 50*(4), 1469–1486.

Rossmann, J. (2010). Data Quality in Web Surveys of the German Longitudinal Election Study 2009. In *3rd ECPR Graduate Conference* Dublin, Ireland.

Sheehan, K. B. (2001). E-Mail survey response rates: A review. *Journal of Computer-Mediated Communication, 6*(2), 0–0.

Sims, W. H. & Hiatt, C. M. (2003). Attitudes, incentives, and test performance (Vol. 7502). Tech. Rep. CAB. Retrieved from https://www.cna.org/CNA_files/PDF/D0007502.A2.pdf

Simmons, E. & Wilmot, A. (2004). Incentive payments on social surveys: A literature review. *Social Survey Methodology Bulletin.* 1–11.

Singer, E., Van Hoewyk, J., & Maher, M. P. (2000). Experiments with incentives in telephone surveys. *Public Opinion Quarterly, 64*(2), 171-188.

Singer, E., & Ye, C. (2013). The use and effects of incentives in surveys. *The ANNALS of the American Academy of Political and Social Science*, *645*(1), 112–141. http://dx.doi.org/10.1177/0002716212458082

Steedle, J. (2010) Incentives, motivation, and performance on a low-stakes test of college learning. Paper presented at the Annual Meeting of the American Education Research Association, Denver, CO. https://www.researchgate.net/publication/265000997_Incentives_Motivation_and_Performance_on_a_Low-Stakes_Test_of_College_Learning

Trouteaud, A. R. (2004). How you ask counts: A test of internet-related components of response rates to a web-based survey. *Social Science Computer Review, 22*(3), 385–392. doi:10.1177/0894439304265650.

van Veen, F., Göritz, A.S., & Sattler, S. (2016). Response effects of prenotification, prepaid cash, prepaid vouchers, and postpaid vouchers: An experimental comparison. *Social Science Computer Review, 34*(3), 333–346.

Higher Education Quality Council of Ontario

An agency of the Government of Ontario