

## **An Alternative Approach: Using Survey Panels to Inform Assessment**

Shimon A. Sarraf  
Sarah Fernandez  
Mark Houlemarde  
Xiaolin Wang

Center for Postsecondary Research  
Indiana University Bloomington

Paper presented at the annual conference of the Association for Institutional Research in Denver, Colorado, May 2015

## Introduction

Survey researchers in higher education are engaged in an uphill battle with declining student response rates. Student cynicism, fatigue, and disinterest combine to pose a significant threat to optimal survey participation (Porter, 2005). Despite significant challenges with the collection of student feedback, demands for evidence-based decisions in higher education have increased (Zhang, 2010). In light of these conflicting trends, identifying potential alternatives to standard survey practices may prove useful for a variety of higher education constituents, including institutional researchers and higher education scholars. Some scholars have already seriously considered this issue. For instance, Stern, Bilgen and Dillman (2014) already proposed survey panels as a potential solution to declining response rates. Though used by some academic disciplines, government agencies, and businesses for decades, higher education researchers and administrators have not employed, generally speaking, survey panels to help with their data collection needs (Zhang, 2010).

As a provider of assessment data to hundreds of colleges and universities across North America, and having witnessed its own response rates decline over the past decade, the National Survey of Student Engagement (NSSE) experimented with survey panels at five diverse institutions in spring 2014. The panel administration included eight surveys of about nine questions each administered over an eight-week time period. Using standard NSSE administration results as a benchmark for each institution, this study documents student interest in joining a NSSE panel, their commitment to completing all eight surveys, and an assessment of panel data quality.

## Literature Review

### Brief History of Survey Panels

For decades, survey panels have been used in market and medical research (Callegaro, Baker, Bathlehem, Goritz, Krosnikc, & Lavrakas, 2014; Callegaro & Diogra, 2008). They have historically been conducted in person, by phone, and by mail, but have more recently transitioned to the internet (Callegaro et al., 2014). It is estimated that the first online survey panel was administered in the mid-1980s in Europe. About a decade later, this method became popular in the United States. Unsurprisingly, annual conferences dedicated to public opinion and survey research professionals, such as the one held for the American Association for Public Opinion Research, are rife with presentations from for-profit, not-for-profit, and government organizations (AAPOR, 2015). According to Callegaro et al. (2014) their benefits are threefold: 1) quick data collection, 2) low administration cost, and 3) sampling efficiency. With a heavy reliance on online student surveys, these benefits may resonate with the needs of institutional researchers and higher education scholars.

National education datasets developed by the federal government have also used survey panels for better understanding the backgrounds and academic experiences of college students and their career trajectories. These include such studies as the *Baccalaureate and Beyond* and *Beginning Postsecondary Students Longitudinal Study* (for more information, see <http://nces.ed.gov/surveys/SurveyGroups.asp?group=2>). These survey panels are distinct from many of the online survey panels previously discussed because they only contact subjects every few years,

not on a weekly or monthly basis. Based on a review of annual conference programs for the Association for Institutional Research over the past four years, members appear to be using these datasets, though in relatively small numbers (AIR, 2015). Even fewer presentations reflected the existence of institution-specific survey panels to inform campus discussions about student issues. A relatively recent presentation at the 2014 AIR annual conference by Sharkness and Miller (2014) is one such example.

## **Panel Definitions & Design Issues**

Various researchers define survey panels differently, so it is important to clarify definitions. The traditional “panel” definition refers to a longitudinal survey panel that involves asking the same individuals the same questions across different points in time (Callegaro et al., 2014; Goritz, Reinhold & Batinic, 2000; Hsiao, 2014) with each point in time referred to as a wave. This approach is inherently suited for studying particular changes among subjects. Many survey panels have since evolved into access panels, which are essentially a “database of potential respondents who declare that they will cooperate for future data collection if selected” (International Organization for Standardization, 2012, p.1). Another form of panel survey utilizes a split questionnaire approach by dividing longer questionnaires into smaller ones (Raghunathan & Grizzle, 1995). Overall, these various panel methods have been used in an effort to minimize missing data (Vriens, Wedel & Sandor, 2001).

*Types.* There are two primary survey panel types based on how members join: probability and nonprobability (Callegaro et al., 2014). Nonprobability panels are open and allow anyone to volunteer to become a panel member. When members self-select into a panel, one does not know the probability of selection, hence the nonprobability label. In contrast, individuals cannot join probability panels unless they have been invited. They require that “all members of the population of interest have a known, non-zero probability of receiving an invitation to join” (Callegaro et al., 2014, p. 7). Prerecruited probability panels refer to panel members being randomly selected to participate while volunteer panel members are not randomly selected (Callegaro & Disogra, 2008).

*Sampling.* Segers and Frases (2014) found that it was difficult to survey every panel member at each point in time because of cost and the potential for non-response. To address these concerns, they utilized a rotating sampling method whereby panel members participated for a fixed amount of time with new panel members integrated at each wave. Other methods include continuous, time, randomized, and matrix sampling. Continuous sampling means that one surveys all individuals at each wave for the extent of the project. Time sampling refers to panel members being surveyed on a regular basis (e.g. biweekly) but not for every wave; they are rotated into waves, helping to ensure data is collected weekly, for instance, or on some other predetermined schedule. Randomized sampling refers to panel members being chosen at random for each particular wave. Lastly, matrix sampling calls for a survey to be split with panel members only getting one part (Segers & Frases, 2014).

*Incentives.* Incentives are frequently used in survey research in an effort to increase response rates, and panel studies are no exception. There are various types of incentives (e.g. pre-paid, post-paid, and lottery). In relation to panel studies, there is some conflicting evidence regarding

different types and their effectiveness. Goritz (2006) found that cash lotteries (both one large prize and multiple smaller prizes) did not reliably increase panel response or retention rates. In a study on the impact of individual payment on a three-wave longitudinal experiment, Goritz, Wolff, and Goldstein (2008) found that guaranteed payments had a negative effect on the first wave and a positive effect on the second wave. Additionally, after a review of the literature, Callegaro et al. (2014) stated contingent incentives are most impactful when the members have prior experience or knowledge of the organization administering the survey. None of these studies used a college student population however so the current study should be helpful for shedding light on this issue.

## **Data Quality Issues**

As with all survey research methods, various issues may affect survey panel data quality. Three issue in particular should be recognized and addressed, if possible, by researchers in their design process: self-selection, non-response, and conditioning.

*Self-selection.* Individuals who commit to a panel may not be fully representative of the original population (Bianchi & Biffignandi, 2013). This stems from selection bias. For example, individuals that are willing to commit to a seemingly longer process like a panel may be substantially different from others in various unknown ways, thus potentially biasing survey population estimates. Somewhat related to this issue is that the perceived commitment of joining a panel makes recruitment often more difficult than for a cross-sectional survey administration (Goritz, Reinhold, Batinic, 2000).

*Non-response.* There are multiple points during a survey panel where non-response emerges among its members, whether it relates to a recruitment email or completing an individual survey item. Once again, these concerns relate to all types of survey research, but especially panels since each individual survey invitation presents panel members with an opportunity to ignore their commitment to participate and/or complete all survey questions. When a participant does not complete one or more individual waves we refer to this as wave non-response while attrition refers to participants dropping out of the panel study entirely (Segers & Franses, 2014). The potential for attrition presents itself during all waves, which can lead to decreased sample sizes and weakened efficiency estimates (Vndecasteel & Debels, 2007). Apodaca, Lea and Edwards (1998) found that participants are reluctant to partake in surveys with multiple components because they were burdensome. They concluded that decreased response rates should be expected with panel studies, and taken into consideration during the study design phase. Individual item non-response is when an individual chooses to not answer one or more individual items (Segers & Franses, 2014). Missing data can impact the quality of panel survey results, therefore making it important to take all forms of non-response and attrition into account during any study's design phase.

*Conditioning.* A third issue for panel studies is conditioning. Conditioning occurs when participants become familiar with an instrument and survey experience, and then this familiarity impacts participant responses (Goritz, Reinhold, & Batinic, 2000; Segers & Franses, 2014). Toepel, Das, & Van Soest (2008) state:

Panel members may learn from taking surveys and become familiar with the question-answering process, learn how to interpret questions, and make fewer errors than new respondents. Or, conversely, experienced respondents may also learn to reduce the burden of their task and complete the survey more rapidly, at the cost of accurately reading and answering each question, thereby making more errors than fresh respondents (p. 987).

It has been found that conditioning is of biggest concern for items related to knowledge, and less of a factor for those that are related to expectations, behaviors and attitudes (Das, Toepoel & Van Soest, 2011). This finding has obvious positive implications for the NSSE instrument and other college surveys that focus on behaviors.

### **Theoretical Framework**

To better understand student survey response behaviors as they relate to panels we briefly review two important theories: Leverage-Saliency Theory (LST) and Social Exchange Theory (SET). Given the lengthier commitment of a panel study compared to a single survey administration, understanding the factors that play into the motivation of an individual to participate should prove helpful for understanding this study's results.

According to LST (Groves, Singer and Corning, 2000) individuals decide to participate in a survey based on the leveraging of a specific attribute. In other words, this theory predicts that if a participant has an interest or connection to the survey topic they will be more likely to respond. Groves and colleagues say the saliency of the particular attribute contributes to the "activation of the potential leverage" (p. 307). According to SET, however, there are three factors that influence the survey response decision (Dillman, 1978). Potential respondents evaluate any survey's available rewards like incentives with their costs, such as how long a survey will take to complete. The third factor is whether or not a participant trusts that the reward will be greater than the cost. If so, an individual is more likely to participate in the survey.

NSSE is often marketed to students by campus administrators as a means for articulating and improving the campus and the quality of the student experience. Through the perspective of LST, students may feel more committed to the survey because of its content and ability to effect change on campus. Students who have a strong affective connection (either positive or negative) to the content of the survey may be more likely to respond as well. Through the perspective of SET, students may weigh the various potential rewards (e.g. guaranteed incentive and providing a service to the campus) with the amount of time and effort they interpret to be required of panel survey participation. This may influence both initial response to the panel invitation as well as non-response and attrition if a student no longer feels the reward outweighs the cost.

### **Study Rationale & Research Questions**

Though response rates for NSSE have slowly declined for years now (the 2014 administration being one recent exception), they are still high enough to reliably estimate institution-level engagement for the vast majority of participating schools. As a recent study using NSSE data has shown, low response rate estimates of student engagement are very similar to estimates based on

high response rates (Fosnacht, Sarraf, Howe, & Peck, 2013). Nevertheless, it is incumbent upon NSSE and other higher education researchers to look closely at alternative data collection methods to see if there are more effective ways to collect student feedback. Whether or not minimal response rates and respondent counts can estimate institution-level engagement sufficiently, low response rates matter to skeptical audiences who believe higher response rates are required for decision making. If response rates drop precipitously for a significant number of institutions in the future, those interested in student opinion may need to rely more heavily on other data collection methods, such as focus groups, non-probability sample surveys and/or survey panels, to supplement their survey administrations.

The following research questions guide this study and help us assess whether survey panels can serve NSSE and the wider institutional research community. Questions are grouped into three general areas—recruitment, survey data quality, and scale results and properties—for assessing the viability of panels. Survey panels and NSSE’s standard cross-sectional administration approach using one long instrument constitute the administration types cited in the research questions. With the exception of the sixth question, we review institution-by-institution results, looking for differences both between and within institutions, rather than aggregating all schools together.

### **Recruitment**

1. What percentage of students try to register for the NSSE survey panel? How do panel registration rates compare to first invitation response for the standard administration?
2. What were claim rates for the panel’s guaranteed incentives?
3. How do 2014 panel members compare to 2014 non-panel members and standard administration respondents in terms of demographics?

### **Survey Data Quality Indicators**

4. What are response rates to individual panel surveys?
5. How many panel surveys and items per survey did students complete?
6. How much time did students spend on completing surveys by administration type?

### **Scale Results and Properties**

7. How do NSSE scale scores compare by administration type?
8. Does the factor structure of NSSE scales vary by administration type?

## **Methods**

*Data Source & Sample.* In order to answer these research questions, we combined NSSE standard administration survey data with spring 2014 panel data for five colleges and universities. These institutions differed by public-private status, college/university designation, and total undergraduate enrollment size. To preserve the anonymity of the five study institutions, we named them using a combination of these characteristics: Small Private College, Small Private University, Medium Public University, Medium Private University and Large Public University. The labels small, medium, and large signify that total undergraduate enrollment is either less than 5,000 (small), between 5,000 and 15,000 (medium), or more than 15,000 (large). Each institution participated in the 2013 standard administration or, in the case of the Large

Public University, an experimental 2014 standard administration that used a smartphone optimized version of NSSE. Given other survey commitments at the Large Public University, we randomly selected 25% of all first-year and senior students to participate in either the panel or smartphone experiments, and then followed up by randomly assigning them to either the panel or standard administrations. The four other institutions provided all first-year student and senior records for sampling. Though unintended, the concurrent panel and standard administrations for the Large Public University allow for stronger claims about panels, whereas we compare results between different times and student populations for the other institutions.

We analyzed a total of 3,331 respondents from both administration types: 2,832 standard administration and 499 panel respondents. 67% were female, 83% underrepresented minority, and 6% part-time students. With the exception of gender, these demographics reflect missing data for some institutions as indicated in forthcoming results tables. We also used 12,273 non-panel members from the 2014 panel administration to assess recruitment success and panel composition.

*Panel Administration Details.* In fall 2013, staff from the National Survey of Student Engagement (NSSE) contacted a set of twelve diverse colleges and universities to express interest in an experimental, no cost survey panel administration during the following spring 2014 semester. NSSE selected five of six institutions that expressed interest.

For each institution, we aimed to recruit 50 first-year students and 50 seniors. We drew repeated random student samples by institution and class level in order to send panel registration invitation emails. Invitations emphasized that each of the eight surveys would take about one minute to complete. Given planning constraints, we only sent students one panel registration email and concluded the registration process over about two days. Those that attempted to register after all one-hundred openings had been filled were told that it would no longer be possible to join the project. As promised in the registration invitation email, all panel members could immediately retrieve a \$10 Amazon.com gift card using an online portal that was created for the study. Additionally, the registration email informed students that their name would be included in a drawing for one \$250 Amazon.com gift card at each school if they completed all eight panel surveys.

Upon registration, students received an email requesting they complete the first survey, followed by a reminder email two or three days later if they did not respond. Approximately every week we delivered another email invitation and follow-up reminder for another panel survey. As the administration for each survey began the survey link was also posted to the online portal. It took approximately nine weeks to administer all eight surveys. During the last week, a final reminder was delivered to students that did not complete all eight surveys, encouraging them to log into the online portal to complete additional surveys.

*Standard Administration Details.* Each year over 500 colleges and universities across the United States and Canada participate in NSSE. As part of a standard administration, typically all first-year students and seniors are asked to complete the survey sometime between early February and June 1. As with all institutions, the five study schools had five email recruitment messages sent on their behalf by NSSE. Institutions decided to use survey incentives to boost response rates,

and, when appropriate, related text was included in all recruitment messages. Each recruitment email included a URL that links to the online survey and informed consent statement. The informed consent statement informed students that the survey takes between 15 to 18 minutes to complete. Though we do not use available data for this study, institutions could append topical module item sets ranging from academic advising to civic engagement, as well as participate in a consortium that administers additional items.

*Survey Content.* The 2013 standard administration survey included approximately 104 survey items (for a survey facsimile, see [nsse.iub.edu/html/survey\\_instruments.cfm](http://nsse.iub.edu/html/survey_instruments.cfm)) whereas the eight experimental panel surveys had about 70 items combined (see table below). The panel administration incorporated arguably the most important survey items used on the standard instrument, including items from ten primary scales used for official NSSE reporting, called Engagement Indicators, items related to high-impact practices for student development, and several important student background and demographic items, like academic major and parental education. Panel survey content order did not follow the standard core instrument’s layout. For a complete list of panel survey items, readers may contact the authors.

**Panel Survey Content (numbers in parantheses indicate how many items)**

Survey #	Topic 1	Topic 2	Topic 3	Student Background/Demographics	# of Items
1	Collaborative Learning (4)	Student-Faculty Interaction (4)			8
2	Reflective & Integrative Learning (7)			Academic Major (2), Class level	10
3	Higher Order Learning (4)	Writing (3)		Age	8
4	Quantitative Reasoning (3)	Effective Teaching Practices (5)		Parental Education	9
5	Discussions with Diverse Others (4)	Learning Strategies (3)	Course Challenge (1)	Living location	9
6	High-Impact Practices (7)	Class Preparation (1)	Reading (1)		9
7	Quality of Interactions (5)			Courses taken, Online Course taken, Race	8
8	Supportive Environment (8)			Transfer Status, Other Schools Attended	10

*Analyses.* The first research questions attempt to assess the attractiveness of panel membership between and within the five study institutions. We present individual institution panel registration results in the aggregate and by class level, using chi-square tests to analyze the data. We also make comparisons to standard administration first recruitment message response rates without any formal statistical testing. To gauge true interest, panel registration rates reflect both those who registered and those who attempted to register (but were unable to because of quota limitations). We calculated initial response to the first standard administration email invitation by using both partial and complete survey submissions five days from delivery.

Given that incentives can affect behavior, and have budgetary implications, we investigate for the second research question the proportion of panel members at each institution retrieving their \$10 Amazon.com gift cards, seeing if these rates vary significantly by institution using a chi-square test.

For the third question, using three key demographic variables for NSSE (gender, full-time enrollment, and underrepresented minority status) we ask whether 2014 panel members are representative of 2014 non-panel members at each institution using chi-square tests. We also ask

whether panel members at each institution are significantly different from corresponding standard administration survey respondents.

The second set of questions focus attention on various data quality issues, namely response rates, the extent of survey item completion, and survey completion times (or duration). Using response rate results for individual panel surveys, we assess the level of participation exhibited by students across study institutions for each panel survey. In addition, we look for evidence of attrition as suggested by the literature. Next, survey item completion behavior is looked at in three different ways. First, we analyze how many of the eight surveys students completed on average at each institution and, using an ANOVA test, determine whether any statistical differences exist between institutions. Second, we run chi-square tests to evaluate completion rates by administration type at each institution, with completion defined as having completed at least 90% of all panel or standard administration survey items. Lastly, we review the average number of panel survey items completed and compare results to the standard administration; this analysis aggregates all institutions together given the homogenous nature of the panel results seen across institutions.

Given the desire to minimize survey burden on respondents, the importance of time spent taking individual panel surveys in relation to the longer survey serves as an important outcome. We evaluate survey duration by calculating the median time in minutes that it took to complete each panel survey and all of them combined, excluding any time spent registering for the panel and retrieving the \$10 Amazon.com gift card. By calculating the median duration, we minimize the impact of outliers resulting from students logging into surveys, getting distracted for a significant amount of time, and then returning to complete the survey later. The duration for the longer standard survey only includes students that completed 90% or more of all core items, but removes time spent on the informed consent and any additional questions appended to the core survey.

For projects like NSSE that rely on scales for various purposes, knowing a panel administration would not have any serious negative consequences is critically important to assessing overall survey panel viability. In order to shed light on this issue, we conducted two sets of analyses using scales, otherwise known as Engagement Indicators (EIs). We first compared scale results by administration type for both first-year students and seniors to assess reliability. For each institution, we calculated twenty effect sizes (ten for each class) to gauge the meaningfulness of the difference. Effect sizes equal the difference between panel and standard administration EI scores divided by the EI standard deviation of the standard administration. We then evaluated effect sizes by using a rule of thumb developed by NSSE staff suggesting that an effect size of .1 is considered small, .3 medium, and .5 large (Rocconi & Gonyea, 2015). *t*-tests determined whether any differences were statistically significant at the .1 alpha level after applying a correction for false discovery rates, a concern when conducting multiple tests concurrently (Benjamini & Hochberg, 1995). Though a year elapsed between data collection for four of the five institutions, one would not expect meaningful differences unless an institution undertook major programmatic changes across their campuses. The ten EIs analyzed included: Higher-Order Learning (HO); Reflective & Integrative Learning (RI); Quantitative Reasoning (QR); Learning Strategies (LS); Collaborative Learning (CL); Discussions with Diverse Others (DD); Student-Faculty Interaction (SF); Effective Teaching Practices (ET); Quality of Interactions (QI);

and Supportive Environment (SE). (For details about EIs, see [nsse.iub.edu/html/engagement\\_indicators.cfm](http://nsse.iub.edu/html/engagement_indicators.cfm).) Differences can originate from the panel administration itself, the fact that relatively few students participated, or unknown changes occurring on campuses.

For the second EI analysis, we conducted multi-group confirmatory factor analysis (MG-CFA) for each EI in order to test measurement invariance across panel and standard administration groups. Confirming measurement invariance ensures that “psychometric test scores can be attributed to differences in the properties that such tests measure” and that a score relates “to the same set of observations in the same way in each group.” (Borsboom, 2006) The MG-CFA for each EI followed several steps. First, a CFA was run separately for each administration type until the same model fit both groups well. If no model fit both groups, measurement invariance was rejected and we pursued no additional testing. Second, assuming a model fit both groups well, we then ran tests for configural, metric, and scalar invariance sequentially. Once a lower level of invariance was tested and rejected, we did not proceed with running tests for higher levels of invariance. Scalar invariance signifies the highest level of invariance, while configural is the lowest. Criteria used for determining acceptable model fit is RMSEA <.06, Chi-square p-value >.05, and CFI/TLI >.90. An even higher level of scalar invariance could be achieved when the chi-square difference test p-values were greater than .05 and  $\Delta$ CFI was less than .01.

## Results

*Recruitment.* Based on one registration email, between 6% and 20% of invited students registered, or tried to register, for the panel administration at the five participating institutions (see Table 1). The proportion that responded varied significantly by institution, whether one reports all first-year students and seniors together ( $p < .001$ ) or separately ( $p < .001$ ). At all but one institution, it was more difficult to get first-year students to register for the panel study. Differences in registration rates favored seniors between 2 and 8 percentage points. Compared to the standard administration response to the first recruitment message, interest in the panel was appreciably higher at all five institutions. The Small Private College showed the greatest difference with 8 percentage points (12% standard administration rate versus 20% panel rate) while the Small Private University and Medium Private University the least with a 2 percentage point difference.

-----

Insert Table 1 about here

-----

Overall, 78% of the 499 panelists claimed their \$10 Amazon.com gift card. With institution rates varying between 75% and 81% we found no statistically significant difference. Claim rates were unrelated to class level as well in the aggregate: 76.2% of first-years and 80.4% of seniors claimed their incentives.

Panel members appeared very similar to non-panel members for two of the three demographic variables we analyzed: full-time enrollment and underrepresented minority status (see Table 2). We found no statistically significant differences for these two demographics across the seven chi-square tests run. However, for four of the five institutions (Large Public University being the exception), females were more represented among panel members: there was a 10 to 20 percentage point greater proportion of females within panel groups compared to non-panel members.

-----

Insert Table 2 about here

-----

In terms of gender, full-time enrollment status and underrepresented minority status, panel members were generally comparable to standard administrations respondents (Table 3). One sees a single statistically significant difference in the proportion of females by administration type across the five schools. Large Public University panel members were 51% female while 62% of its standard administration respondents were female ( $p < .05$ ).

-----

Insert Table 3 about here

-----

*Survey Data Quality Indicators.* Panel survey response rates at the five institutions ranged from 95% for the Small Private College's first survey to 71% for the Medium Private University's seventh survey (see Table 4). All institutions had response rates over 90% for the first survey while rates for the eighth survey ranged from 72% to 86%. Statistical tests suggest that response rates to individual panel surveys do not vary by institution with the second survey being an exception ( $p < .05$ ). As is common with panel administrations, response rates declined across individual surveys. The difference between response rates at the beginning and end of the survey field period for each institution ranged from a low of 11% for the Small Private College to a high of 21% for the Medium Private University. In all cases, response rates for panel surveys far exceeded those of standard administrations for each institution. Standard administration response rates ranged between 34% for the Small Private College and 12% for the Large Public University. Even when compared to the relatively lower rate of the eighth survey, the gap with the standard administration response rates at institutions ranged from 50% (Small Private University) to 63% (Large Public University).

-----

Insert Table 4 about here

-----

Students completed about 6.5 surveys on average across all panel members and institutions varied as well on this outcome (See Table 5). Students at the Large Public University completed 5.6 surveys out of eight on average while Small Private College students completed 7.1 ( $p < .01$ ). By looking at completion rates by administration type we see statistically significant differences in favor of the standard administration for two of four institutions (8 percentage points for the Small Private University and 14 percentage points for Medium Private University). Though statistically insignificant, the 8% difference in favor of the panel administration for the Small Private College contrasts with these two institutions, while the Medium Public University showed no difference. We could not test the Large Public University completion results because the progress indicator for the experimental standard administration did not function properly, which very likely impacted student behaviors.

-----

Insert Table 5 about here

-----

As shown in Table 6, once a panel member completes one survey item they almost always complete the remaining items. Item completion results show that for the 71 panel survey items 71 items on average are completed by survey respondents, compared to 85 of 103 core survey items from the standard administration.

-----

Insert Table 6 about here

-----

As expected, each panel survey took a little more than a minute to complete; all surveys combined took 9.2 minutes (see Table 7). Standard administration survey duration was 11.7 minutes, about two and half more minutes than the time it took to complete all panel surveys.

-----

Insert Table 7 about here

-----

*Scale Results and Properties.* Panel and standard administration EI effect sizes by class level suggest that both survey administration approaches generally yield similar results (see Table 8). Out of the 100 comparisons only nine showed both meaningful and statistically significant differences, ranging from .36 to .81. The Large Public University, which has the advantage of a stronger research design (random assignments into both administration types), showed no differences. For the ten comparisons by class level for each institution, we generally see only one notable difference. For the senior populations at the Small Private University and Medium Public University we see two notable differences, representing the highest level of discrepancy. In terms

of certain scales that may present particular challenges, Quantitative Reasoning appears to have the greatest difference between the two administration approaches.

-----

Insert Table 8 about here

-----

In terms of EI measurement invariance, all first-year EIs met the criteria for scalar invariance with only two out of ten—Higher Order Learning (HO) and Supportive Environment (SE)—not meeting the very highest criteria (see Table 9). Senior year results indicated that eight of the ten EIs were scalar invariant with five meeting the very highest criteria. Only senior HO and SE results indicated variant factor structures between panel and standard administrations.

-----

Insert Table 9 about here

-----

## **Limitations**

There are several limitations associated with this study’s design and analyses. First, despite our findings, results for four of the five study institutions came from comparing potentially different student populations during the spring 2013 and spring 2014 semesters, which could have unknown effects on study outcomes. (For this reason, we place more weight on certain Large Public University findings.) Second, regarding our recruitment questions, we guaranteed that all panel members would receive a substantial incentive along with a chance to win a \$250 Amazon.com gift card, whereas standard administration sample members at three institutions did not get any incentive and the other two conducted lotteries. Making any firm conclusions about the relative attractiveness of panels is therefore confounded; the incentives could have impacted data quality outcomes in unpredictable ways as well, especially as they relate to survey completion. Related to this point is whether any selection bias exists given that panel members chose to join the panel after just one registration email. Although we reviewed several key demographics and determined panel members are generally representative, they may still be different in unknown ways that influence engagement results. Lastly, given the relatively few panel respondents at each institution, we could not explore the impact that panel administrations might have on assessing engagement by academic major fields, only for first-year student and senior populations. For those that typically review survey results by various academic units, or other sub-populations, this study’s results may not offer sufficient information.

## **Discussion**

This study provides various insights about college student survey panels, one possible alternative to the standard survey administration approach used by NSSE, other national survey projects, and many institutional research offices. Foremost among our findings, students from all five study institutions responded to panel surveys at very high rates (70% or more), far above rates for corresponding standard administrations. These findings support the idea of Stern and their colleagues (2014) suggesting survey panels may be one solution to declining response rates.

When considering these results however, keep in mind that we guaranteed participants a \$10 Amazon.com gift card before they completed a single survey as well offered them a chance to win a \$250 gift card if they answered all eight surveys. As social exchange theory might suggest, these high response rates likely stem from the minimal cost of participation (about a minute answering relatively innocuous questions) in relation to several perceived benefits, including the incentives offered and the opportunity for students to provide helpful feedback about their college experience. Another valid explanation is that students may feel a strong ethical obligation to participate after getting their \$10 gift card. A guaranteed incentive of this amount is generally unheard of among NSSE standard administrations and may not be a particularly realistic amount for some institutions to offer students, especially if they want more than 100 panel members. For this reason, we encourage others to experiment lowering the incentive's dollar value to minimize expenditures but without sacrificing too much on response rates. The survey methodology literature on college student panel incentives is thin, if not entirely missing, so further investigation is certainly needed.

With the exception of gender, we also found panel members to be representative of the larger student population and standard administration respondents using enrollment (full time versus part time) and underrepresented minority status. Females traditionally respond to NSSE survey requests at higher rates, so their over-representation is not especially alarming. These results bode well for the use of panel studies as administrators may not have to worry too much about self-selection bias.

Consistent with similar studies completed in market and medical fields, results indicate, unsurprisingly, that panel studies do offer a rapid way to collect information from students. Panel members answered each of the eight- to ten-item surveys in about a minute with nearly zero item skipping/non-response. The short duration and practically nonexistent item nonresponse are other reasons that the panel survey approach may be a viable alternative to standard administration practices. This study contradicts Apodaca et al (1998) conclusion that individuals are reluctant to participate in surveys with multiple components. This is not surprising since length of panel surveys and corresponding completion times likely matter a great deal for minimizing attrition and item nonresponse.

In terms of NSSE scales, the vast majority of results indicated scores originating from panels are comparable to those from standard administrations, even with a limited number of respondents. One EI that appeared unreliable was quantitative reasoning. Given the influence of academic major on this measure (Roconni, Dumford, McCormick, & Sarraf, 2014), one explanation is that panel members' academic major were substantially different from those of our standard administration, weakening our earlier claims of representativeness. Furthermore, with the exception of two EIs for seniors, scale factor structures do not appear to be affected by the data

collection method, something that would be very important for analyzing panel and standard administration results longitudinally. The reasons for higher order learning and supportive environment scales failing to show measurement invariance are unclear at this time. At a minimum, this finding complicates combining results from these two different survey administrations.

In contrast to these favorable panel results, this study also highlights several issues of concern with using survey panels. Based on our recruitment findings, it does not appear that students perceive panel membership to be significantly more attractive than a standard administration approach, even with the \$10 guaranteed Amazon.com gift card and promises of a minute long survey each week. There are several possible explanations such as students weighing the protracted involvement over eight weeks against the incentives provided and determining the costs actually outweigh the benefits. Secondly, we saw that panel member attrition drove overall missing data levels for some institutions to a point that resembles or surpasses standard administration results. When one looks at the proportion of students that complete at least 90% of all survey items by administration type, it is somewhat difficult to generalize results across schools though the data appears to support the point that completion rates are not necessarily improved by splitting up a longer survey. Missing data rates for a long survey like NSSE is a serious issue to be addressed but given these results, survey panels implemented as they were in this experiment do not appear to be a panacea. Variations of this study's panel design, like rotating panel members so they do not get fatigued, may yield more promising results in terms of missing data levels.

Amidst a growing need for survey data to inform decisions in higher education, students are at increased risk for survey fatigue. Survey panels may offer a heuristic solution by limiting the need to administer surveys to entire campus populations. While 100 participants joined the panel and received eight survey invitations, the vast majority of students at each institution did not receive more than a single registration email, thus reducing any potential frustration with multiple requests. High panel response rates may effectively limit the need to send several reminders to sample members. Additionally, we provided students the option to go back and complete previous waves of the survey using an online portal, thereby eliminating the need for several survey reminder messages that could increase perceptions of burden and actual survey fatigue.

## **Conclusion**

Being asked to complete a relatively long 12- to 15-minute survey is often perceived as burdensome for college students, especially considering the many other survey requests they receive. This challenges survey administrators to think about more effective ways to collect student feedback than the customary cross-sectional survey design that NSSE and other survey research projects currently employ. Survey panels represent one data collection alternative that is worthy of further investigation, as our results point to higher response rates, short completion times, and minimal impact on measurement scales. Panel member attrition and associated missing data levels, however, makes this option potentially problematic. The price for encouraging panel membership by using incentives may be cost prohibitive to some as well, especially if one needs significantly more panel members for analyzing campus sub-populations.

Our hope is that others will start experimenting with panels—such as with rotating panel members—to see if there are ways to address these issues. Obviously, we do not know what the future holds for the standard survey administration approach, but investing time, money, and energy in finding some viable alternatives sooner rather than later may provide a good safety net.

## References

- AAPOR—American Association for Public Opinion Research. (2015, May). Conference Program. Retrieved from [http://www.aapor.org/AAPORKentico/AAPOR\\_Main/media/AM15/AAPOR-15-FP\\_FNL.pdf](http://www.aapor.org/AAPORKentico/AAPOR_Main/media/AM15/AAPOR-15-FP_FNL.pdf)
- AIR—Association for Institutional Research. (n.d.). AIR Annual Conference. Retrieved from <https://www.airweb.org/EducationAndEvents/AnnualConference/Pages/default.aspx>
- Bianchi, A. & Biffignandi, S. (2013). Web panel representativeness. In P. Giudici et al. (Eds.). *Statistical Models for Data Analysis, Studies in Classification, Data Analysis, and Knowledge Organization*. Switzerland: Springer.
- Borsboom, D. (2006). When does measurement invariance matter? *Medical Care*, *44*, 176-181.
- Callegaro, M. & Disogra, C. (2008). Computer response metrics for online panes. *Public Opinion Quarterly*, *72*(5), 1008-1032.
- Callegaro, M. & Disogra, C. (2008). Computer response metrics for online panes. *Public Opinion Quarterly*, *72*(5), 1008-1032.
- Callegaro, M., Baker, R., Bethlehem, J., Goritz, A.S., Krosnick, J.A., Lavrakas, P.J. (2014). Online panel research: History, concepts, applications a look at the future. In M. Callegaro, R. Baker, J. Bethlehem, A.S. Goritz, J.A. Krosnick and P.J. Lavrakas (Eds.). *Online Panel Research: A Data Quality Perspective* (1-22). United Kingdom: Wiley & Sons, Ltd.
- Das, M. , Toepoel, V., & van Soest, A. (2011). Non parametric tests of panel conditioning and attrition bias in panel surveys. *Sociological Methods & Research*, *40*(1), 32-56.
- Dillman, D.A. (1978). *Mail and telephone surveys: The total design method*. New York: Wiley.
- Goritz, A.S. (2006). Cash lotteries as incentives in online panels. *Social Science Computer Review*, *24*(4), 445-459.
- Goritz, A.S., Reinhold, N., & Batinic, B. (2000). Online panels. In B. Batinic, U. Reips, M. Bosnjak (Eds.). *Online Social Sciences* (29-51). Gottigen, Ger: Hogrefe & Huber.
- Goritz, A.S., Wolff, H., & Goldstein, D.G. (2008). Individual payments as a longer-term incentive in online panels. *Behavioral Research Methods*, *40*(4), 1144-1149.
- Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-Saliency Theory of survey participation. *Public Opinion Quarterly*, *64*, 299-308.
- Hsiao, C. (2014). *Analysis of panel data* (Vol. 54). Cambridge university press.
- International Organization for Standardization (2012). ISO 20252 Market, opinion and social research: Vocabulary and service requirements (2<sup>nd</sup> ed.). Geneva: ISO.
- Porter, S. R. (2005). Survey research policies: An emerging issue for higher education. *New Directions for Institutional Research*, *127*, 5-15
- Raghunathan, T.E. & Grizzle, J.E. (1995). A split questionnaire survey design. *Journal of the American Statistical Association*, *90*(429), 54-63.

- Rocconi, L. & Gonyea, R. M. (2015). *Contextualizing Student Engagement Effect Sizes: An Empirical Analysis*. Paper presented at the Association for Institutional Research Annual Conference, Denver, CO.
- Segers, R. & Franses, P. H. (2014). Panel design effects on response rates and response quality. *Statistica Neerlandica*, 68(1), 1-24.
- Sharkness, J. & Miller, K. (2014). *Who fills out multiple surveys? Tracking response using online panels*. Paper presented at the Association for Institutional Research Annual Conference, Orlando, FL.
- Stern, M. J., Bilgen, I., & Dillman, D. A. (2014). The State of Survey Methodology Challenges, Dilemmas, and New Frontiers in the Era of the Tailored Design. *Field Methods*, 26(3), 284-301.
- Toepoel, V., Das, M., & Van Soest, A. (2008). Effects of design in web surveys: Comparing trained and fresh respondents. *Public Opinion Quarterly*, 72(5), 985-1007.
- Vndecasteel, L. & Debels, A. (2007). Attrition in panel data: The effectiveness of weighting. *European Sociological Review*, 23 (1), 81-97.
- Vriens, M., M. Wedel, and Z. Sandor (2001), "Split-Questionnaire Designs: A New Tool for Survey Design and Panel Management", *Marketing Research*, Summer, pp. 15-19.
- Zhang, L. (2010). The use of panel data models in higher education policy studies. In J.C. Smart (Ed.). *Higher Education: Handbook of Theory and Research* (307-349). New York: Springer.

**Table 1. Initial response to participation request by administration type**

	Standard Administration <sup>a</sup>	Panel Registration <sup>b</sup>				
		Overall	First-Year Students	Seniors	FY vs. SR ( $X^2$ )	Panel Gift Card Retrieval Rate
Small Private College	12.0%	20.0%	17.4%	23.0%	+	78.0%
Small Private Univ.	10.0%	12.4%	9.4%	17.4%	***	75.0%
Medium Public Univ.	2.0%	6.9%	6.0%	8.0%	+	78.0%
Medium Private Univ.	4% <sup>c</sup>	6.0%	6.9%	5.4%		79.8%
Large Public Univ.	2.0%	9.3%	7.0%	11.4%	**	81.0%
$X^2$	n/a	***	***	***		

+ p< .1; \* p< .05; \*\* p<.01; \*\*\* p<.001

<sup>a</sup> Standard administration results reflect response rate five days after initial invitation to complete survey.

<sup>b</sup> Panel registration results reflect total registered panel members and unsuccessful registration attempts divided by the total number of registration invitations delivered. Approximately 105 students across all five schools tried to register but were not included in the panel because of the 100 student quota per institution.

<sup>c</sup> Institution offered an incentive for survey completion.

**Table 2. Student characteristics by institution and panel membership status (2014 data only)**

		Female	Sig.	Full-time	Sig.	Minority	Sig.
Small Private College	Non-panel member	56.9%		98.7%		10.1%	
	Panel member	<b>74.0%</b>	*	99.0%		11.0%	
Small Private Univ.	Non-panel member	57.3%		97.1%		n/a	
	Panel member	<b>70.0%</b>	*	97.0%		n/a	
Medium Public Univ.	Non-panel member	53.4%		90.4%		18.9%	
	Panel member	<b>63.0%</b>	*	94.0%		18.0%	
Medium Private Univ.	Non-panel member	60.3%		88.6%		27.8%	
	Panel member	<b>74.7%</b>	*	90.9%		25.3%	
Large Public Univ.	Non-panel member	49.2%		n/a		n/a	
	Panel member	51.0%		n/a		n/a	

\*  $p < .05$ ;  $X^2$  test investigates whether student characteristics vary by panel membership status at each institution.

**Table 3. Student characteristics by institution and administration type (2013 and 2014 data)**

	Administration Type	Female	Sig.	Full-Time	Sig.	Minority	Sig.
Small Private College	2013 Standard	78.4%		99.7%		9.3%	
	2014 Panel	74.0%		99.0%		11.0%	
Small Private Univ.	2013 Standard	62.5%		97.4%		5.3%	
	2014 Panel	70.0%		97.0%		n/a	
Medium Public Univ.	2013 Standard	64.7%		90.6%		17.8%	
	2014 Panel	63.0%		94.0%		18.0%	
Medium Private Univ.	2013 Standard	71.3%		92.9%		24.1%	
	2014 Panel	74.7%		90.9%		25.3%	
Large Public Univ.	2014 Standard	62.0%		n/a		n/a	
	2014 Panel	<b>51.0% *</b>		n/a		n/a	

\*  $p < .05$ ;  $X^2$  test investigates whether student characteristics vary by survey administration type at each institution.

**Table 4. Response rates by administration type**

	<b>Panel Surveys</b>									<b>Standard Administration</b>	
	1	2	3	4	5	6	7	8	Panel Members (n)	Response Rate	Respondents (n)
Small Private College	95%	93%	85%	85%	84%	84%	84%	86%	100	34%	333
Small Private Univ.	92%	83%	80%	77%	78%	78%	78%	78%	100	28%	341
Medium Public Univ.	92%	88%	84%	83%	81%	80%	79%	76%	100	15%	753
Medium Private Univ.	93%	78%	79%	74%	72%	73%	71%	72%	99	21% <sup>a</sup>	718
Large Public Univ.	91%	83%	81%	80%	77%	78%	77%	75%	100	12% <sup>b</sup>	308
$\chi^2$		*								n/a	

\* p<.05

<sup>a</sup> Institution offered an incentive for responding to the standard administration survey invitation.

<sup>b</sup> Institution offered incentive after initial invitation; a random sample of students was used for this school's administration.

**Table 5. Average number of panel surveys completed and survey completion rates by administration type**

	Panel Surveys Completed (of 8)	Completion Rate <sup>a</sup>		Sig.
		Panel	Standard Administration	
Small Private College	7.1	81.0%	73.0%	
Small Private Univ.	6.8	73.0%	81.2%	+
Medium Public Univ.	6.7	76.0%	76.6%	
Medium Private Univ.	6.3	65.7%	79.9%	**
Large Public Univ.	5.6	72.0%	n/a	
$X^2$	**	-	-	

+ p< .1; \* p< .05; \*\* p<.01; \*\*\* p<.001

<sup>a</sup> Completion rate defined as the percentage of respondents completing 90% of all survey items; panel survey results incorporate non-response to individual survey requests.

**Table 6. Item completion by administration type**

	Panel Surveys									Standard Administration
	1	2	3	4	5	6	7	8	All	
Average items completed	8 of 8	10 of 10	7.99 of 8	8.98 of 9	9 of 9	9 of 9	8 of 8	10 of 10	71 of 71	85 of 103

**Table 7. Survey duration by administration type**

	Panel Surveys								Standard Administration (core items only)	
	1	2	3	4	5	6	7	8	All	
Median duration (in minutes)	0.70	1.20	1.05	1.03	1.02	1.33	0.87	0.99	9.2 <sup>a</sup>	11.7 <sup>b</sup>

<sup>a</sup> Only includes students that completed all eight mini-surveys; excludes time to register and retrieve gift cards.

<sup>b</sup> Only includes students that completed all survey items; excludes time spent on informed consent page and additional comments.

**Table 8. Panel and standard administration engagement indicator effect size differences by class level and institution**

Class level	School	Reflective and		Quantitative Reasoning	Learning Strategies	Collaborative Learning	Discussions with Diverse Others	Student- Faculty Interaction	Effective Teaching Practices	Quality of Interactions	Supportive Environment
		Higher-Order Learning	Integrative Learning								
First Year	Small Private College	0.05	0.15	<b>0.81 *</b>	0.11	0.10	0.24	0.14	0.20	0.00	0.04
	Small Private Univ.	<b>0.55 *</b>	0.28	0.48	0.29	0.11	0.07	0.07	0.02	0.00	0.01
	Medium Public Univ.	0.01	0.18	0.29	0.11	0.09	0.13	0.34	0.22	0.29	<b>0.45 *</b>
	Medium Private Univ.	0.28	0.16	0.01	0.29	0.13	0.23	0.14	0.22	0.03	0.14
	Large Public Univ.	0.01	0.08	0.31	0.28	0.14	0.22	0.44	0.16	0.14	0.05
Senior	Small Private College	0.01	0.08	0.31	0.28	0.14	0.22	<b>0.44 *</b>	0.16	0.14	0.05
	Small Private Univ.	0.03	0.02	0.33	0.15	<b>0.46 *</b>	0.03	0.10	0.24	<b>0.53 *</b>	0.09
	Medium Public Univ.	0.28	0.19	<b>0.36 *</b>	0.25	0.00	0.08	0.18	0.32	0.09	<b>0.52 *</b>
	Medium Private Univ.	0.24	0.06	<b>0.42 *</b>	0.30	0.14	0.21	0.00	0.30	0.10	0.22
	Large Public Univ.	0.01	0.19	0.02	0.14	0.16	0.01	0.09	0.21	0.09	0.13

\*p<.1; statistical significance test results reflect a corrected level of significance to address false discovery rate concerns when conducting multiple tests concurrently. See Benjamini & Hochberg (1995) for additional details.

**Table 9. Multi-Group Invariance Results by Engagement Indicator**

	First-Year Students	Seniors
Higher-Order Learning	scalar	<i>variant</i>
Reflective and Integrative Learning	scalar +	scalar
Quantitative Reasoning	scalar +	scalar +
Learning Strategies	scalar +	scalar +
Collaborative Learning	scalar +	scalar
Discussions with Diverse Others	scalar +	scalar +
Student-Faculty Interaction	scalar +	scalar +
Effective Teaching Practices	scalar +	scalar +
Quality of Interactions	scalar +	scalar
Supportive Environment	scalar	<i>variant</i>

Note: “scalar” implies the scale achieved strong measurement invariance as indicated by most model fit indices. “scalar +” implies that strong measurement invariance is supported by all model fit indices. ‘*variant*’ implies the scale did not achieve measurement invariance since configural invariance was supported by less than half of model fit indices.