Impacts of Implementing an Automatic Advancement Feature in Mobile and Web Surveys

Stacey Giroux*, Kevin Tharp†, Derek Wietelman‡

Keywords: mobile design, data quality, web surveys, mobile surveys, automatic advancement

DOI: [10.29115/SP-2018-0034](https://doi.org/10.29115/SP-2018-0034)

Survey Practice

Vol. 12, Issue 1, 2019

As more surveys are administered online and taken on mobile devices, researchers need to continue to work toward defining best practices. A relatively new technical feature is an automatic advancement feature, which automatically scrolls the screen to the next item on a page of a survey instrument once a respondent selects an answer to the current question. This and other related forms of automatic advancement methods, such as the horizontal scrolling matrix, have so far produced mixed effects in terms of their impact on data quality, break off rates, and other survey measures. We present findings from an experiment conducted in 2017 using a survey of law school students in which we randomly assigned an automatic advancement feature to half of the respondents in a sample of nearly 40,000 students in the United States and Canada. We do not find significant differences in survey duration time, straight-lining, breakoff rates, or item nonresponse (for mobile users) between the two experimental groups, but desktop users without the automatic advancement feature had higher item nonresponse. More significantly, we find that respondents receiving the automatic advancement treatment on average changed about 50% fewer answers across the survey instrument than those who did not receive the automatic advancement design. We incorporate qualitative data to help us understand why this might have happened. Finally, respondents rated the automatic advancement feature as being easier to use and having a better visual design. This research extends the limited amount of work on automatic advancement features for surveys, helps move the field closer to best practices for survey design for mobile devices, and provides context and understanding about a particular respondent behavior related to data quality.

BACKGROUND

As more people use the Internet and smartphones to take surveys, understanding these users' experiences becomes increasingly important (Couper and Peterson 2016). One recent development in mobile and web survey design includes ways of helping the respondent automatically move through a survey. These types of measures are referred to as automatic advancement, auto-advance, or auto-forwarding features. Automatic advancement techniques vary in their application, with some presenting a

* **Institution:** Indiana University–Bloomington (IN) **Department:** Center for Survey Research

† **Institution:** Indiana University–Bloomington **Department:** Center for Survey Research

‡ **Institution:** Indiana University–Bloomington **Department:** School of Public and Environmental Affairs

single question on a page and automatically pushing users to the *next page* of the survey upon selecting an answer, without having to click a 'Next' button (Hays et al. 2010). Other forms of automatic advancement, such as the one that we apply here, automatically move users through *questions* on a page as each question is answered (Norman and Pleskac 2002). Such designs are intended to help reduce respondent burden and ensure that respondents answer each question presented to them, since the advancing function relies upon, though is not strictly dependent upon, questions being answered. A related feature is the horizontal scrolling matrix (HSM), which combines features of a matrix or grid of questions with a single-question format. It keeps a common set of response options fixed on the page while automatically cycling respondents through a series of questions, one at a time, as they answer each (de Leeuw, E.D. et al. 2012).

Antoun and colleagues' (2017) work toward design heuristics for surveys on smartphones provides a review of efforts to optimize surveys for smartphones, including use of the HSM and automatic advancement features (Antoun et al. 2017). Antoun et al. (2017) did not recommend use of an automatic advancement feature for a one-time survey on a smartphone, because automatic advancement seems to require learning on the part of the respondent to use it properly. Selkälä and Couper's (2017) work on the use of auto-forwarding compared to manual forwarding focused on the cognitive response processes underlying respondents' experiences and how these interact with the design feature. Their literature review covers much of the limited work in the area of automatic advancement designs. In Table 1, we summarize findings for both HSM and automatic advancement features, illustrating not only the dearth of published work on this topic but also the lack of clarity as to effects that such designs have produced.

Table 1. Summary of findings regarding HSM and automatic advancement design features.

	Auto-advance or HSM improves	Auto-advance or HSM harms	Auto-advance or HSM makes no significant difference
Completion time	Rivers (2006) Hays et al. (2010) Klausch et al. (2012) Selkälä & Couper (2017): for PCs only Lugtig et al. 2018	Roberts et al. (2013)	Arn et al. (2015) de Bruijne (2015) Selkälä & Couper (2017): for phones only
Item nonresponse		de Bruijne (2015)	Hammen (2010)
Data quality	Roberts et al. (2013) Klausch et al. (2012)	Hammen (2010)	Hays et al. (2010)
Breakoff rate	Rivers (2006) Roberts et al. (2013) Klausch et al. (2012)		Selkälä & Couper (2017)
User satisfaction	Rivers (2006) Roberts et al. (2013) Klausch (2012)		de Bruijne (2015)

(Rivers 2006; Hays et al. 2010; Klausch et al. 2012; Selkälä and Couper 2017; Lugtig et al. 2018; Roberts et al. 2013; de Bruijne, M.A. 2015; Hammen 2010; Arn, Klug, and Kolodziejcki 2015)

In this paper, we focus on the design aspects of an automatic advancement feature used in a survey of law school students in the United States and Canada. Our research extends the limited work in the area of automatic advancement for surveys. By exploiting a randomized assignment of automatic advancement to half of our respondents and administering traditional manual forwarding to the other respondents, we test for differences in survey duration, straightlining behavior, item nonresponse, breakoff rate, counts of changed answers, and user satisfaction between the two groups.¹

METHODS

We implemented our experiment in the 2017 administration of the annual Law School Survey of Student Engagement (LSSSE). The LSSSE is designed to measure the effects of legal education on law students.² The first six pages of the survey are the same for all students with the exception of a few small variations in questions appropriate to the Canadian students. Each page asked 15 questions on average, with a minimum of nine questions and a maximum of 20 appearing on a single page. We will refer to these first six pages as the core of the survey in the analysis that follows.³

The majority of questions in the LSSSE take the form of a stem, a set of items, and the same set of response options. We broke out the stem and items with what we call a "sticky stem" for each set of questions, so that the respondent always sees the question stem as the items change with the automatic advancement feature (Figure 1). In Figure 1, as the respondent answers the item "Asked questions in class or contributed to class discussions," the item advances upward, out of sight, while the next item, "Prepared two or more drafts of a paper or assignment before turning it in," advances to beneath the question stem, taking the place of the item just answered.⁴

1 T-tests revealed no significant differences between the two groups across a number of demographic characteristics, indicating that the randomization was successful.

2 The survey is part of the Indiana University Center for Postsecondary Research, where its project manager and research analysts also reside. The Indiana University Center for Survey Research is a partner providing programming, testing, recruitment, administration, and data compilation support for the survey.

3 The seventh page is demographic questions. A school may then elect to append up to two additional question sets, which sometimes change year to year, on different topics such as library use or student stress. The survey was custom coded in Adobe ColdFusion 11, SQL Server 2012, and the automatic advancement function was coded in JavaScript using jQuery framework (2.2.4). The survey is administered only in English.

4 For a visualization of our automatic advancement feature, follow this link: (<http://go.iu.edu/21Ar>)

In your experience at your law school during the current school year, about how often have you done each of the following?

Asked questions in class or contributed to class discussions

☐ Very often ☐ Often ☐ Sometimes ☐ Never

Prepared two or more drafts of a paper or assignment before turning it in

☐ Very often ☐ Often ☐ Sometimes ☐ Never

Figure 1. Example of "sticky stem" question set from the LSSSE 2017.

The LSSSE was in the field from February 21, 2017, to April 23, 2017, with a sample size of 39,350 students across 77 law schools in the United States and Canada. All law students at institutions that choose to participate in the study are invited to take the survey. Up to five electronic contacts are made in an attempt to gain cooperation from sampled students. The response rate for 2017 was 52.28%, or 20,574 students (AAPOR RR2). A post-survey evaluation to understand aspects of respondents' satisfaction or dissatisfaction with the experimental design included seven questions and was conducted with students at 44 of the participating schools. The post-survey evaluation was programmed in Qualtrics (Provo, UT) and linked at the end of the LSSSE so that it would not appear to be part of the law school survey itself. When reviewing the characteristics of the schools that were selected for the post-survey evaluation, we ended up with slightly higher percentages of students attending private schools, schools in the United States, schools with smaller enrollments, and religiously affiliated schools than the sample as a whole. The 44 selected schools yielded 11,759 students eligible for this post-survey evaluation, of which 11.06%, or 1,300 students, responded. Of those 1,300 students, 642 received the automatic advancement design.

RESULTS

SURVEY DURATION, STRAIGHTLINING, ITEM NONRESPONSE, AND BREAKOFFS

To test whether there were differences in duration, that is, whether auto-advance encouraged speeding, we could not simply compare the total time to survey completion because, as described previously, schools elect different topical modules to add to the core survey. Therefore, we examined duration

on a per-page basis for respondents completing each of the survey's core pages. We first removed page duration outliers ($> 1.5 \times$ interquartile range). Table 2 provides the upper limit in seconds for each page, along with the number of cases exceeding that limit and the resulting N for analysis for each page. Table 3 presents results of significance tests for differences in survey duration by time by page.⁵

Table 2. Prevalence of cases exceeding duration time cutoffs for first six survey pages

Page	Upper limit, seconds (1.5 x interquartile range)	Number of cases exceeding limit	N for analysis
1	236.9	1,837	17,888
2	148.8	1,736	17,522
3	116.2	1,549	17,436
4	151.8	1,405	17,320
5	127.9	1,594	16,977
6	142.4	1,660	16,633

Table 3. Differences in mean survey duration time between control and experimental groups

Page	Mean difference of manual advance relative to auto advance (seconds)	t statistic	p-value ^a
1	-0.591	-0.994	0.32
2	-1.819	-4.588	.000***
3	0.459	1.686	.092*
4	-1.711	-4.274	.000***
5	-0.292	-0.892	0.372
6	1.702	4.748	.000***

a - Note: All p-values reported are two-tailed p-values.

* Indicates significance at the $\alpha = .10$ level, ** at $\alpha = .05$ level, *** at $\alpha = .01$

Results for the six page were mixed. For pages 1, 3, and 5, there were no significant differences in duration (at $\alpha = .01$, for this and for all tests in this "Results" section) between those with auto-advance and those with manual advance. However, for pages 2 and 4, those with auto-advance spent about 1.7–1.8 seconds longer on average on the page, and for page 6, those with manual advance spent 1.7 seconds longer. Taken together, we judge these results as being evidence of no consistent, practically meaningful difference in duration between those with auto-advance and those with manual advance.

We considered straightlining to mean that respondents chose the same response for all questions in a group that shared the same stem and response

⁵ We urge caution when interpreting the results for differences in duration times. Since our automatic advancement feature scrolled the screen down for the respondent rather than instantly displaying a new question, a given treated respondent's duration time is partly a function of how fast our automatic advancement measure scrolled the screen down for them. Our finding is consistent with the broader literature that automatic advancement measures do not have a consistently clear effect on survey duration times.

set (eight groups total). We included all respondents who completed the core of the survey ($N=18,532$). There was no significant difference in the average number of question sets straightlined between those with automatic advancement (.57 sets) and those with manual advancement (.55 sets).

For those respondents who completed the core of the survey, we tested for differences between the experimental and control groups in the number of items not answered, separating those using a mobile device from those using a desktop or tablet. We measured device use as the last one recorded for taking the survey, with 18.1% of respondents using a mobile device. The difference in the number of questions unanswered for those using a mobile device was not significant (.75 questions on average for those with automatic advancement versus .71 unanswered questions for those with manual advancement), but for those using a desktop, the difference was significant: .83 questions for those with automatic advancement, and 1.06 questions for those without.

In terms of breakoff rates, that is, respondents who consented to the survey and reached at least the first page of questions but did not reach the survey closing page, there were fewer breakoffs for those with automatic advancement than for those with manual advancement, but this difference was not significant ($N=21,144$, 13.99% breakoff rate for automatic advancement versus 14.68% for manual advancement).

USER SATISFACTION

We asked all students who participated in the post-survey evaluation ($N=1,300$) two questions. For the first question, "Considering factors like ease of navigation, ease of reading the screen, and ease of selecting responses (but not question content), please rate how easy it was for you to complete this survey" (7-point scale with 1 = Not at all easy, 7 = Very easy; $N=1,246$), those exposed to the automatic advancement design rated it significantly easier for completing the survey: 6.49 for those with automatic advancement versus 6.36 for those with manual advancement ($p = .035$). For the second question, "How would you rate the visual design of the survey?" (4-point scale with 1 = Excellent, 4 = Poor; $N=1,292$), those exposed to the automatic advancement design rated it higher in terms of visual design: 1.54 versus 1.64, respectively ($p = .008$).

We asked respondents who received the automatic advancement feature ($N=642$) two additional open-ended questions: "What, if anything, did you like about the automatic advance feature?" ($N=197$), and "What, if anything, did you dislike about the automatic advance feature?" ($N=170$). These responses were coded by one of the authors by hand, with 10% of those codes then coded again and checked for agreement by another of the authors. There was agreement for all codes checked. Coding yielded a total of 10 categories for features respondents liked about the design, and 11 categories for what they disliked. The top four facets of the design that respondents liked were that the survey went faster or flowed better; that the survey was easier to answer,

convenient, or efficient; that they did not have to scroll; and that the design helped them focus on and answer questions, ensuring they did not miss any questions. The aspect of the automatic advancement design that respondents stated they liked the least was that scrolling back to previous questions was a problem, or that they could not go back and change their answers. Over one-third of those who responded commented along these lines of having difficulty changing answers (N=61). However, there was nothing in the programming that prevented respondents from going back and changing any answers. If a respondent did want to change an answer on the same page, s/he simply had to scroll up to the question. If they wanted to change an answer on a previous page, s/he had to use the browser's back button and scroll within that page. It appears that the design itself convinced some respondents that changing earlier answers was not a possibility. A few of these respondents did show evidence of figuring out that they could go back and change answers, but even then, some found it cumbersome to do.

CHANGING ANSWERS

As a result of our open-ended evaluation, we decided to examine the effect of the automatic advancement treatment on the number of answers changed by respondents. Our decision to investigate this was further bolstered by the recent findings of Selkälä and Couper (2017), who reported that respondents who received their automatic advancement experimental condition changed answers at a significantly reduced rate compared to those in the control group.

Our paradata provided data on all answer changes made by respondents. As a result, we were able to construct a dummy variable for each question that indicated whether respondents changed their initially selected answers for that particular question. This includes those respondents who changed an answer but ultimately decided to return to their initial choice.

We modeled respondent-level counts of the number of changed answers. Our specifying equation is of the form:

$$Y_i = \alpha + \beta_{AutoAdvance_i} + \gamma_{Response_i} + \delta_{Cognitive_i} + \theta X_i + \epsilon_i$$

where Y_i is a count of the number of changed answers for each respondent i ; *AutoAdvance* is a dummy variable indicating experimental treatment assignment; *Response* is a vector of variables related to the survey response itself that are expected to affect changed answer counts, including the choice to take the survey on a mobile device and the maximum number of pages submitted per respondent; *Cognitive* is a vector of proxy variables for aptitude and familiarity with technology, including LSAT scores and undergraduate grade point average; and X is a collection of fixed effects for ethnicity, gender, year in law school, and the law school attended.

We tested two different response variables: (1) a count of the number of changed answers per respondent for the core and (2) a count of the number

of changed answers per respondent for all pages submitted by the respondent. Since our response variables are count variables, we constructed models using both Poisson regression and the less restrictive negative binomial regression procedure.

Table 4. Results of changed answer count models

Models of respondent-level changed answer counts (n=13,393)				
Variable	Model description			
	1	2	3	4
	Poisson; All Pages	Poisson; Core Pages	Negative Binomial; All Pages	Negative Binomial; Core Pages
Auto-advance	-.72829***	-.77900***	-.72924***	-.77825***
	0.00751	0.0084	0.01218	0.01333
Mobile	.32158***	.34263***	.31704***	.33655***
	0.00832	0.0098	0.01494	0.01627
Pages submitted	.15050***	.12886***	.15941***	.13703***
	0.00232	0.00234	0.00303	0.00314
LSAT Score	-.00408***	-.00358***	-.00507***	-.00466***
	0.00064	0.00071	0.00117	0.00126
Undergraduate GPA	0.00071	0.0012	0.00108	0.00147
	0.00105	0.0011	0.0019	0.00203

Standard errors reported below coefficient estimates in ().

* - significant at 10% level; ** - significant at 5% level; *** - significant at 1% level

All models include fixed effects for gender; ethnicity; year in law school; and school.

As seen in Table 4, coefficient estimates for the automatic advancement treatment remained stable across the four models. The coefficient estimate for the automatic advancement variable represents the causal impact of receiving the feature on the count of changed answers per respondent. The coefficient estimates suggest a large depressing effect on the number of changed answers per respondent, approximately 45%–48% fewer changed answers. Figure 2 demonstrates that the average numbers of changed answers remained consistently lower for those who received the automatic advancement treatment as the survey progressed.

Figure 2: Mean changed answers per page by experimental treatment.

DISCUSSION

Overall, we are encouraged by the results of our automatic advancement experiment. There was little significant impact in terms of the data quality indicators of straightlining, item nonresponse, and breakoff rates, and for the most part, respondents found the design to be more satisfying. We note that although the difference in satisfaction between designs was statistically significant, for practical purposes, it is not particularly great. As users become savvier and technology advances, survey researchers have more tools of increasing sophistication at their disposal. Our finding of depressed changed answer counts, however, cautions us that application of new features must be thoroughly tested and that it is possible to move too quickly in implementing new features in a web survey context. These considerations are in line with Antoun et al.'s (2017) recommendation about automatic advancement designs.

The depressed changed answer count finding is consistent with Selkälä and Couper's work (2017:15), which concluded that "the extra effort of returning to a previous item is higher for respondents in the auto forwarding condition and may dissuade them from going back." However, our open-ended responses offer an alternative explanation. The problem might not be the extra mental or scrolling effort, but rather that the design of the automatic advancement feature led some respondents to believe that they could not go back and change a previously answered question. Future designs in this area may need to include clear instructions or a warning to confirm that respondents know that they can change answers if they wish.

Our data also illustrate the issue raised by Antoun et al. (2017) of the importance of user learning. The American Association for Public Opinion's (AAPOR) report on the use of mobile technologies notes that societal learning will play a key role in how best practices are developed. AAPOR also notes that studies even just 1–2 years apart may find vastly different results as the general populace becomes more comfortable with both mobile devices and taking surveys online (American Association for Public Opinion Research 2014). While our qualitative data indicated that some users did get a better sense of the automatic advancement feature as they progressed, Figure 2 shows a consistent gap in changed answers throughout the core six pages of the survey. Future research should further investigate how respondent experiences with automatic advancement or similar features evolve over the course of a given survey or over repeated administrations of a survey instrument.

Finally, it is important to note that our experiment was conducted exclusively with law students. There could be certain unobserved characteristics within a person (motivation, attention to detail, etc.) that could be correlated with both the decision to attend law school and our measures of interest. We look to future experiments in which automatic advancement features can be tested with a more generalizable population of survey respondents.

REFERENCES

- American Association for Public Opinion Research. 2014. "Mobile Technologies for Conducting, Augmenting, and Potentially Replacing Surveys." Report of the AAPOR Task Force on Emerging Technologies in Public Opinion Research. 2014. <http://www.aapor.org/Education-Resources/Reports/Social-Media-in-Public-Opinion-Research.aspx>.
- Antoun, C., J. Katz, J. Argueta, and L. Wang. 2017. "Design Heuristics for Effective Smartphone Questionnaires." *Social Science Computer Review* 36 (5): 557–74. <https://doi.org/10.1177/0894439317727072>.
- Arn, B., S. Klug, and J. Kolodziejski. 2015. "Evaluation of an Adapted Design in a Multi-Device Online Panel: A DemoSCOPE Case Study." *Methods, Data, Analyses* 9 (2): 185–212.
- Couper, M.P., and G. Peterson. 2016. "Why Do Web Surveys Take Longer on Smartphones?" *Social Science Computer Review* 35 (3): 357–77.
- de Bruijne, M.A. 2015. "Designing Web Surveys for the Multi-Device Internet." PhD thesis., The Netherlands: Tilburg University: Center for Economic Research.
- de Leeuw, E.D., J.J. Hox, T. Klausch, A. Roberts, and A. de Jongh. 2012. "Design of Web Questionnaires: Matrix Questions or Single Question Formats. Paper Presented at the Annual Conference of the American Association for Public Opinion Research." In . Orlando, FL, May 17-20, 2012.
- Hammen, K. 2010. "The Impact of Visual and Functional Design Elements in Online Survey Research." In . Paper presented at the General Online Research Conference (GOR) Mannheim, Germany, May 26-28, 2010.
- Hays, R.D., R. Bode, N. Rothrock, W. Riley, D. Cella, R. Gershon, and et al. 2010. "The Impact of next and Back Buttons on Time to Complete and Measurement Reliability in Computer-Based Surveys." *Quality of Life Research* 19 (8): 1181–84.
- Klausch, T., E.D. de Leeuw, J.J. Hox, A. Roberts, and A. de Jongh. 2012. "Matrix vs. Single Question Formats in Web Surveys: Results from a Large Scale Experiment." In . Paper presented at the General Online Research Conference, Mannheim, Germany, March 5-7, 2012.
- Lutig, P., V. Toepoel, M. Haan, R. Zandvliet, and L.K. Kranenburg. 2018. "Does Smartphone-Friendly Survey Design Help to Attract More and Different Respondents?" In . Paper Presented at the Annual Conference of the American Association for Public Opinion Research. Denver, CO, May 16–19 2018.
- Norman, K., and T. Pleskac. 2002. "Conditional Branching in Computerized Self-Administered Questionnaires: An Empirical Study." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 46 (14): 1241–45.
- Rivers, D. 2006. "Web Surveys for Health Measurement." In . Paper Presented at Building Tomorrow's Patient-Reported Outcome Measures: The Inaugural PROMIS Conference. Gaithersburg, MD.
- Roberts, A., E.D. de Leeuw, J. Hox, T. Klausch, and A. de Jongh. 2013. "Leuker Kunnen We Het Wel Maken. Online Vragenlijst Design: Standaard Matrix of Scrollmatrix?" In *Ontwikkelingen in Het Marktonderzoek: Jaarboek Marktonderzoek Associatie, D1*, edited by A.E. Bronner, P. Dekker, E. de Leeuw, L.J. Paas, K. de Ruyter, A. Smidts, and J.E. Wieringa, 38:133–48. Haarlem: SpaarenHout.
- Selkälä, Arto, and Mick P. Couper. 2017. "Automatic Versus Manual Forwarding in Web Surveys." *Social Science Computer Review* 36 (6): 669–89. <https://doi.org/10.1177/0894439317736831>.