

An Experiment Comparing Concurrent and Sequential Choice+ Mixed-Mode Data Collection Protocols in a Self-Administered Health Survey

Survey Methods: Insights from the Field

Taylor Lewis | Naomi Freedner | Charlotte Looby

How to cite this article : Lewis, T., Freedner, N. & Looby, C. (2022). An Experiment Comparing Concurrent and Sequential Choice+ Mixed-Mode Data Collection Protocols in a Self-Administered Health Survey. *Survey Methods: Insights from the Field*. Retrieved from <https://surveyinsights.org/?p=17134>

DOI : 10.13094/SMIF-2022-00004

Copyright : © the authors 2022. This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0)

Abstract : In self-administered web and paper mode surveys, the Choice+ protocol, which offers both modes concurrently but with a differentially higher incentive for responding by web, has been shown to increase the share of responses obtained by web and to increase response rates overall relative to a concurrent design with equal incentive amounts. This protocol was initially adopted for the 2020 Healthy Chicago Survey. After observing fewer than 20% of the 2020 completes coming via paper, however, an experiment was fielded in the 2021 administration to evaluate a sequential variant of the Choice+ protocol in which the paper mode was only offered to nonrespondents after an initial web-only invitation. This article reports on the key findings from this experiment. Although the sequential Choice+ protocol sacrifices 1.4 percentage points in the yield rate, it comes with a 15.2% lower cost-per-complete and produces a respondent set that looks very similar to the concurrent Choice+ protocol with respect to both demographics variables and key health outcomes measured by the survey. As a result of these findings, the sequential Choice+ variant has supplanted the concurrent Choice+ design as the primary data collection protocol in the Healthy Chicago Survey.

Background

A common concurrent mixed-mode design (de Leeuw, 2018) for mail contact surveys is to offer respondents two methods for completing the survey: (1) return an enclosed paper questionnaire or (2) log into a web-based version of the instrument. When provided the choice, research has shown respondents are more likely to choose paper over web (e.g., Biemer et al., 2016), even though web is the more cost-effective, environmentally friendly method, and produces more timely, higher quality data, since proper branching and data consistency checks can be programmed directly into the electronic instrument. Respondents' preference of paper over web assumes, however, equivalent incentive amounts across the two modes. Biemer et al. (2018) coined the term Choice+ for a concurrent mixed-mode data collection protocol that assigns a higher incentive for the respondent to complete the survey via the web mode relative to the paper mode. Biemer and colleagues evaluated the Choice+ method on the 2015 Residential Energy Consumption Survey (RECS) National Pilot with a \$10 promised incentive for completing by paper but a \$20 promised incentive for completing by web, finding that the majority of respondents chose the web mode over paper. An added benefit was that the Choice+ protocol resulted in

a nearly 3 percentage point response rate increase over an analogous concurrent mixed-mode protocol without the differential incentives. Similar findings have been obtained in other surveys—for an example, see the discussion by Medway and Megra (2022) of mode experiments conducted during the 2019 National Household Education Survey.

In the wake of these findings, the Choice+ protocol has been adopted in numerous mixed-mode mail contact surveys, including the 2018 National Recreational Boating Safety Survey (Duffy et al., 2020), the 2020 New York State Problem Gambling Prevalence Survey (New York State Office of Alcoholism and Substance Abuse Services, 2020), and the 2020 Healthy Chicago Survey (HCS) (RTI International, 2020). As reported in RTI International (2020), the Choice+ protocol implemented in the 2020 HCS resulted in a mix of 81% web respondents and 19% paper respondents. Given the low proportion of paper respondents and the non-trivial printing and postage costs of the 20-page survey booklet, the HCS administration team experimentally evaluated a sequential version of the Choice+ protocol in the 2021 HCS. In the first contact attempt, respondents were offered the web mode exclusively, with the Choice+ protocol offered to nonresponding households three weeks later. The purpose of this article is to report results from this experiment, which we found to produce a nearly equivalent yield rate, but at a significantly lower cost-per-complete and without introducing substantive differences with respect to demographic variables or key health outcome variables.

Data and Methods

Data from this article are drawn from the HCS, an annual health survey of the adult population of the City of Chicago sponsored by the Chicago Department of Public Health. HCS data are used to develop and evaluate numerous public health interventions and policies to promote health equity. First launched in 2014 as a dual-frame, random-digit dial (DFRDD) telephone survey, HCS transitioned to a self-administered format in 2020 using an address-based sampling (ABS) frame (Unangst et al., 2022). The households in the target population are stratified into 77 community areas, and those sampled are asked to have the adult next to celebrate a birthday complete the survey (Olson et al., 2014). The survey sample is fielded in two subsequent releases to target between 4,000 – 4,500 completes overall with a minimum of 35 completes from each of the 77 distinct community areas.

The Choice+ design utilized in the 2020 HCS began by sending households a survey packet containing a \$2 pre-incentive, an invitation letter with log-in credentials to complete the survey via web, and a paper copy of the survey for respondents who wish to complete it using that mode. While the survey only takes about 20 minutes to complete, a 20-page booklet is required to print all of its questions. The invitation letter informed the respondent that a \$10 incentive would be paid for completing via paper but a \$20 incentive would be paid for completing via web. A reminder postcard was sent one week later, and two weeks thereafter another full survey packet was sent to nonresponding households. The final contact attempt was a reminder postcard sent one week after the second full survey packet.

In the 2020 HCS administration, the concurrent Choice+ protocol resulted in a yield rate of 30.5%, but with only 19% of respondents completing the survey by paper. By comparison, Biemer and colleagues (2018) reported that approximately 35% of respondents completed the 2015 RECS National Pilot by paper in the Choice+ condition. The relatively lower share of respondents completing the 2020 HCS via paper led the survey administration team to consider whether the paper booklet could be excluded from the first mailing to save on data collection costs and improve data quality without having a noticeable impact on respondent demographic distributions or distributions of key health outcomes. To test this hypothesis, an experiment was fielded during the first sample release of the 2021 HCS administration whereby a

randomly designated 25% of the sampled addresses ($n = 1,750$) was designated to receive the same Choice+ protocol as was administered in the 2020 HCS, while the complementary 75% ($n = 5,250$) received the sequential variant. Specifically, the sequential data collection protocol involved offering web response exclusively in the first mailing for a \$20 incentive, and then introducing the paper mode as an alternative avenue of response three weeks later for a \$10 incentive. Table 1 summarizes the four contact attempts for this sequential Choice+ design alongside the contact attempts for the concurrent Choice+ design.

Table 1: Timeline of Contact Attempts for the Concurrent Choice+ and Sequential Choice+ Data Collection Protocols Experimentally Compared during the 2021 Healthy Chicago Survey.

	Concurrent Choice+	Sequential Choice+
Day	<u>Mailing Description</u>	<u>Mailing Description</u>
0	Survey Packet	Web Invitation Letter
7	Reminder Postcard	Reminder Postcard
21	Survey Packet	Survey Packet
28	Reminder Postcard	Reminder Postcard

Results

Table 2 presents disposition codes and percentages of the 7,000 sampled addresses allocated to the two Choice+ experimental conditions during the first sample release of the 2021 HCS. The yield rate for the sequential Choice+ protocol, at 24.7%, is only 1.4 percentage points lower than the 26.1% yield rate for the concurrent Choice+ protocol. As would be expected, the sequential Choice+ protocol also yielded fewer paper responses: only 10.1% of completes were obtained by paper, whereas 20.6% of the completes came via paper for the concurrent Choice+ protocol. Although the higher portion of web responses resulted in a higher average incentive payment, the sequential Choice+ variant still resulted in a cost-per-complete that was 15.2% lower. In essence, the printing and postage costs saved by excluding the 20-page survey booklet from the first mailing outweighed the increased average incentive amount.

Table 2: Dispositions of Sample Cases Allocated to the Two Mixed-Mode Data Collection Protocols - Concurrent Choice+ and Sequential Choice+.

Code	Meaning	Description	Concurrent Choice+		Sequential Choice+	
			Count	Percent	Count	Percent
CW	Web Complete	Answered by web with at least 4 weighting variables	363	20.7	1,165	22.2
CP	Paper Complete	Answered by paper with at least 4 weighting variables	94	5.4	131	2.5
PW	Web Partial Complete	Answered by web with at least 1, but fewer than 4, weighting variables	6	0.3	55	1.0
PP	Paper Partial Complete	Answered by paper with at least 1, but fewer than 4, weighting variables	2	0.1	2	0.0
UD	Undeliverable	Mail correspondence returned by USPS (and not a complete or partial complete)	188	10.7	478	9.1
RF	Known Eligibility Nonrespondent	Explicit refusal or blank questionnaire returned	2	0.1	4	0.1
NR	Unknown Eligibility Nonrespondent	All other cases not assigned one of the other codes	1,095	62.6	3,415	65.0
Totals			1,750	100	5,250	100
Yield Rate			$100 * (CW + CP) / (CW + CP + PW + PP + UD + RF + NR)$		26.1	24.7
Paper Completion Rate			$100 * CP / (CW + CP)$		20.6	10.1

Cost savings aside, another key analysis of interest to the HCS administration team was to compare household- and individual-level demographic variable distributions across the two respondent sets as well as distributions for certain key health outcomes measured by the survey. Results of these comparisons are provided in Tables 3 and 4, respectively. The specific variables chosen are many of the same variables investigated as part of the DFRDD-ABS mode transition analyses conducted in Unangst et al. (2022). Statistical significance tests were carried out for the variable as whole using a chi-square test, and for specific categories of a variable using a two-sample *t*-test.

Table 3 shows how the demographic distributions at both the household and individual level are very similar across the two Choice+ protocols. There are only two categories – 25-29-year-olds and Hispanic/Latino individuals – where the percentage of respondents differs significantly. As judged by the chi-square test *p*-values, none of the ten demographic variables, on the whole, has a distribution that differs significantly. We viewed this as a positive finding considering that the sequential Choice+ protocol generated one-half the rate of paper completes than the concurrent Choice+ protocol.

With respect to the key health outcomes, Table 4 conveys a similar story of highly compatible distributions across the two Choice+ approaches. However, we must acknowledge that distributions were significantly different at conventional levels for the following two health outcomes: (1) the indicator for whether the respondent currently has asthma and (2) whether the respondent had experienced psychological distress in the past 30 days. Respondents from the sequential Choice+ protocol were more likely to currently have asthma (40.0% vs. 25.8%) and less likely to have experienced mild or moderate psychological distress in the past 30 days (77.3% vs. 71.4%). It is certainly plausible these differences occurred by chance, as there were no statistically significant differences observed across any of the other eleven health outcomes, and neither protocol appears to result in a respondent set that looks

systematically different.

Table 3: Comparison of Household- and Individual-Level Demographic Distributions of Respondents from the Two Mixed-Mode Data Collection Protocols - Concurrent Choice+ and Sequential Choice+.

	Concurrent Choice+		Sequential Choice+		Statistical Tests	
	Count	Percent	Count	Percent	Percent Difference (SE)	Chi-Sq Statistic (p-value)
Age category						4.5 (0.3485)
18-24	27	5.9	82	6.3	-0.4 (1.30)	
25-29	67	14.7	141	10.9	3.8 (1.87)	*
30-44	132	28.9	378	29.2	-0.3 (2.47)	
45-64	134	29.4	414	32.0	-2.6 (2.50)	
65+	96	21.1	279	21.6	-0.5 (2.23)	
Sex						2.3 (0.1259)
Male	159	35.3	502	39.4	-4.0 (2.64)	
Female	291	64.7	773	60.6	4.0 (2.64)	
Race-Ethnicity						8.6 (0.0721)
Non-Hispanic White	236	52.2	631	48.9	3.3 (2.73)	
Non-Hispanic Black	106	23.5	268	20.8	2.7 (2.29)	
Non-Hispanic Asian or Pacific Islander	36	8.0	103	8.0	-0.0 (1.48)	
Non-Hispanic Other or Multiple Races	17	3.8	63	4.9	-1.1 (1.08)	
Hispanic/Latino	57	12.6	226	17.5	-4.9 (1.89)	**
Educational attainment						1.1 (0.5713)
High school or less	74	16.3	202	15.6	0.7 (2.01)	
Some college or technical school	111	24.5	290	22.4	2.1 (2.33)	
College graduate	268	59.2	800	61.9	-2.8 (2.68)	
Marital status						1.1 (0.7779)
Married	159	34.8	466	36.0	-1.2 (2.60)	
Divorced or separated	69	15.1	170	13.1	2.0 (1.92)	
Widowed	30	6.6	89	6.9	-0.3 (1.36)	

Never married	199	43.5	569	44.0	-0.4 (2.70)	
Household tenure						0.3 (0.6112)
Owns home	222	48.6	645	50.0	-1.4 (2.72)	
Rents or other arrangement	235	51.4	646	50.0	1.4 (2.72)	
Household has kids						0.7 (0.3879)
Yes	106	26.6	295	24.4	2.2 (2.54)	
No	292	73.4	912	75.6	-2.2 (2.54)	
Employment status						3.6 (0.6076)
Employed for wages	253	56.2	725	56.2	-0.0 (2.72)	
Self-employed	34	7.6	85	6.6	1.0 (1.43)	
Out of work or unable to work	48	10.7	154	11.9	-1.3 (1.71)	
Homemaker	14	3.1	34	2.6	0.5 (0.93)	
Student	21	4.7	40	3.1	1.6 (1.11)	
Retired	80	17.8	251	19.5	-1.7 (2.11)	
Sexual identity						0.4 (0.5252)
Heterosexual or straight	394	87.4	1105	86.2	1.2 (1.84)	
Gay, lesbian, bisexual, other	57	12.6	177	13.8	-1.2 (1.84)	
Adults in Household						1.0 (0.6172)
1	167	39.5	495	41.2	-1.7 (2.77)	
2	181	42.8	518	43.1	-0.3 (2.80)	
3+	75	17.7	189	15.7	2.0 (2.14)	

Significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 4: Comparison of Select Health Outcome Distributions of Respondents from the Two Mixed-Mode Data Collection Protocols - Concurrent Choice+ and Sequential Choice+.

	Concurrent Choice+		Sequential Choice+		Statistical Tests	
	Count	Percent	Count	Percent	Percent Difference (SE)	Chi-Sq Statistic (p-value)
Overall health status						1.1 (0.2996)
Excellent, very good, or good	404	89.0	1,174	90.7	-1.7 (1.68)	
Fair or poor	50	11.0	120	9.3	1.7 (1.68)	
Has primary health care provider						1.4 (0.2431)
Yes	390	85.7	1,079	83.4	2.3 (1.94)	
No	65	14.3	214	16.6	-2.3 (1.94)	
Had routine health checkup in past year						0.5 (0.4732)
Yes	341	74.9	947	73.2	1.7 (2.38)	
No	114	25.1	346	26.8	-1.7 (2.38)	
Had teeth cleaned in past year						0.6 (0.4467)
Yes	287	63.4	845	65.4	-2.0 (2.62)	
No	166	36.6	448	34.6	2.0 (2.62)	
Ever diagnosed with high blood pressure						1.8 (0.1851)
Yes	154	34.5	396	31.1	3.4 (2.60)	
No	292	65.5	878	68.9	-3.4 (2.60)	
Currently have asthma						7.8 (0.0052)
Yes	32	25.8	100	40.0	-14.2 (5.01)	**
No	92	74.2	150	60.0	14.2 (5.01)	**
Ever diagnosed with diabetes						0.0 (0.9811)
Yes	49	11.1	134	11.0	0.0 (1.74)	
No	394	88.9	1,082	89.0	-0.0 (1.74)	
Smoking status						4.0 (0.1356)
Current smoker	45	9.9	111	8.7	1.3 (1.61)	
Former smoker	85	18.8	295	23.0	-4.2 (2.18)	
Never smoked	323	71.3	877	68.4	2.9 (2.49)	
Number of servings of fruits/vegetables yesterday						3.4 (0.1786)
0 servings	21	4.6	78	6.1	-1.4 (1.20)	
1-4 servings	273	60.4	715	55.9	4.5 (2.69)	
5+ servings	158	35.0	487	38.0	-3.1 (2.62)	
Experienced psychological distress in past 30 days						7.8 (0.0206)
No distress	324	71.4	998	77.3	-5.9 (2.42)	*
Mild or moderate distress	92	20.3	186	14.4	5.9 (2.13)	**
Serious distress	38	8.4	107	8.3	0.1 (1.51)	
Misused prescription opiates in past year						2.9 (0.0883)
Yes	18	4.0	29	2.3	1.7 (1.01)	
No	435	96.0	1,257	97.7	-1.7 (1.01)	
Feel safe in the neighborhood						0.6 (0.7561)
Yes, all or most of the time	314	68.7	902	69.6	-0.9 (2.52)	
Sometimes	111	24.3	295	22.8	1.5 (2.32)	
No, or mostly no	32	7.0	99	7.6	-0.6 (1.40)	
Exercised in past month						0.2 (0.6486)
Yes	361	79.2	1035	80.2	-1.0 (2.20)	
No	95	20.8	256	19.8	1.0 (2.20)	

Significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Conclusion

The Choice+ survey data collection protocol described in Biemer et al. (2018), which offers both web and paper response modes concurrently, but with a differentially higher promised incentive for those completing by web, has been adopted by several surveys in recent years as a proven way to increase web

participation—thereby increasing data quality and timeliness—and to increase response rates. The original design calls for sending a paper questionnaire to households at the same time as the web invitation. While maintaining the ability for individuals to complete the survey by paper is important, especially if the propensity to not have internet readily accessible is correlated with key survey outcomes, the protocol may not be cost-efficient in settings where a low proportion of respondents opt for the paper mode, as was the case in the 2020 HCS where fewer than 20% of completes were obtained by paper. This article presented results from an experiment fielded during the 2021 HCS evaluating a sequential Choice+ variant, one offering a web mode exclusively at first, then introducing the paper mode (alongside another web invitation) later in the data collection period.

Results from the experiment are as follows. The sequential Choice+ protocol had a 1.4 percentage point lower yield rate relative to the concurrent Choice+ protocol, but it resulted in a cost-per-complete metric that was 15.2% lower. Moreover, the sequential Choice+ protocol produced a respondent set that mirrors the respondent set produced by the concurrent Choice+ protocol with respect to both demographic variables and key health outcomes from the survey. In light of these findings, the HCS administration team has adopted the sequential Choice+ design as the survey's sole data collection protocol moving forward.

We hope results from this experiment prompt analogous experiments in surveys such as RECS and the New York State Problem Gambling Prevalence Survey, where the most recent administrations utilizing the concurrent Choice+ design resulted in 27% and 15% of completes, respectively, coming by paper. These may be additional settings where cost savings could be realized without detriment to key dimensions of the surveys' respondent distributions. Further research could also aim to gain insights into other mitigating factors such as the length and complexity of the questionnaire or the magnitude of the incentive differential, perhaps relative to a concurrent mixed-mode condition with both modes offering the same lower incentive amount. That equal incentive protocol would likely have a lower response rate and lower cost-per-complete metric, yet it may be defensible if the respondent profile it produces does not differ dramatically from that of the differential incentive protocol.

References

1. Biemer, P., Murphy, J., Zimmer, S., Berry, C., Lewis, K., and Deng, G. (2016). A Test of Web/PAPI Protocols and Incentives for the Residential Energy Consumption Survey. Presentation given at the annual meeting of the American Association for Public Opinion Research. Austin, TX.
2. Biemer, P., Murphy, J., Zimmer, S., Berry, C., Lewis, K., and Deng, G. (2018). Using Bonus Monetary Incentives to Encourage Web Response in Mixed-Mode Household Surveys. *Journal of Survey Statistics and Methodology*, 6(2), 240 – 261.
3. Duffy, T., Kilpatrick, G., Krotki, K., Lewis, T., McMichael, J., Palmer, D., Ridenhour, J., Ryder-Burge, A., Thomas, I., Willis, S., Mahoney, E., and Herbowicz, T. (2020). National Recreational Boating Safety Survey: Exposure Survey Final Report. Available at: <https://uscgboating.org/library/recreational-boating-survey/NRBSS-Exposure-Survey-Final-Report-1130-2020.pdf>.
4. de Leeuw, E. (2018). Mixed-Mode: Past, Present, and Future. *Survey Research Methods*, 12(2), 75 – 89.
5. Medway, R., and Megra, M. (2022). Pick Me! Incentivizing Web Response in a Concurrent Mixed-Mode Design. Presentation given at the annual meeting of the American Association for Public Opinion Research. Chicago, IL.
6. New York State Office of Alcoholism and Substance Abuse Services. (2021). 2020 New York State Problem Gambling Prevalence Survey: Final Report. Available at: <https://nyproblemgambling.org/resource-lib-item/2020-new-york-state-problem-gambling-prevalence-s>

urvey/.

7. Olson, K., Stange, M., and Smyth, J. (2014). Assessing Within-Household Selection Methods in Household Mail Surveys. *Public Opinion Quarterly*, 78(3), 656 - 678.
8. RTI International (2020). 2020 Healthy Chicago Survey (HCS) Methodology Report. Available online at: https://www.chicago.gov/content/dam/city/depts/cdph/statistics_and_reports/2020_HCS_Methodology_Report.pdf.
9. Unangst, J., Lewis, T., Laflamme, E., Prachand, N., and Weaver, K. (2022). Transitioning the Healthy Chicago Survey from a Telephone Mode to Self-Administered by Mail Mode. *Journal of Public Health Management & Practice*, 28(3), pp. 309 - 316.