

# **An Experiment Testing Alternative Email Contact Timing Strategies in a Web-Based Survey of Federal Personnel**

**Taylor Lewis**

U.S. Office of Personnel Management

[Taylor.Lewis@opm.gov](mailto:Taylor.Lewis@opm.gov)

**Karl Hess**

U.S. Office of Personnel Management

[Karl.Hess@opm.gov](mailto:Karl.Hess@opm.gov)

## **Abstract**

The Federal Employee Viewpoint Survey (FEVS) is an annual survey of over 800,000 permanently employed civilian personnel from 87 agencies. First administered in 2002, the Web-based survey measures a broad range of employee perceptions, attitudes, and behaviors, serving as a valuable tool for human resources managers to determine which aspects of an organization are working well and which may require intervention. The data collection protocol begins by sending all sampled individuals an initial invitation to participate. Thereafter, nonrespondents are sent weekly reminder emails. These notifications are typically sent on Tuesday mornings. In this article, we present results from an experiment investigating two alternative protocols varying when survey notification and reminder emails are sent. Interestingly, the stable Tuesday morning strategy produced a significantly higher response rate than the two alternatives, and we analyze timestamp information from one of the alternatives to help provide insight as to why.

## **Disclaimer**

The opinions, findings, and conclusions expressed in this article are those of the authors and do not necessarily reflect those of the U.S. Office of Personnel Management.

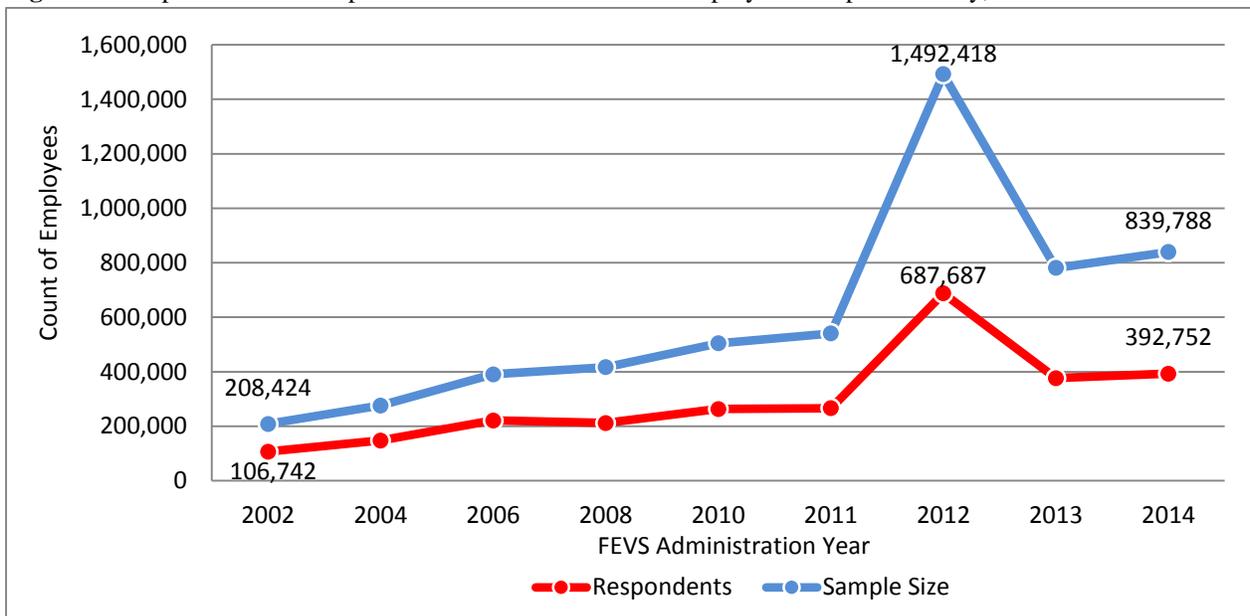
## **Acknowledgements**

The authors would like to thank Bill Dristy of the U.S. Office of Personnel Management for his help sending out Federal Employee Viewpoint Survey invitations and reminders at our prescribed times.

## Background

The U.S. Office of Personnel Management conducts the Federal Employee Viewpoint Survey (FEVS) ([www.fedview.opm.gov](http://www.fedview.opm.gov)) to measure how effectively agencies are managing their most important asset, their workforce. The Web-based survey is sent to sampled individuals via a personalized link embedded in an email message. The instrument taps at a broad range of employee perceptions and behaviors proven to drive satisfaction, commitment, engagement, productivity, and retention. The survey was first administered in 2002, and biennially thereafter. In 2010, the FEVS became an annual survey. As Figure 1 illustrates, the sample size has increased appreciably since that time, the result of a progressively more intricate stratification scheme needed to meet the participating agencies' demand for analyzing work units deeper and deeper within the organization. This increase culminated in a full census<sup>1</sup> of the federal workforce in FEVS 2012, although sampling resumed in FEVS 2013. More details on the current sampling methodology can be found in the survey's technical report (U.S. Office of Personnel Management, 2014).

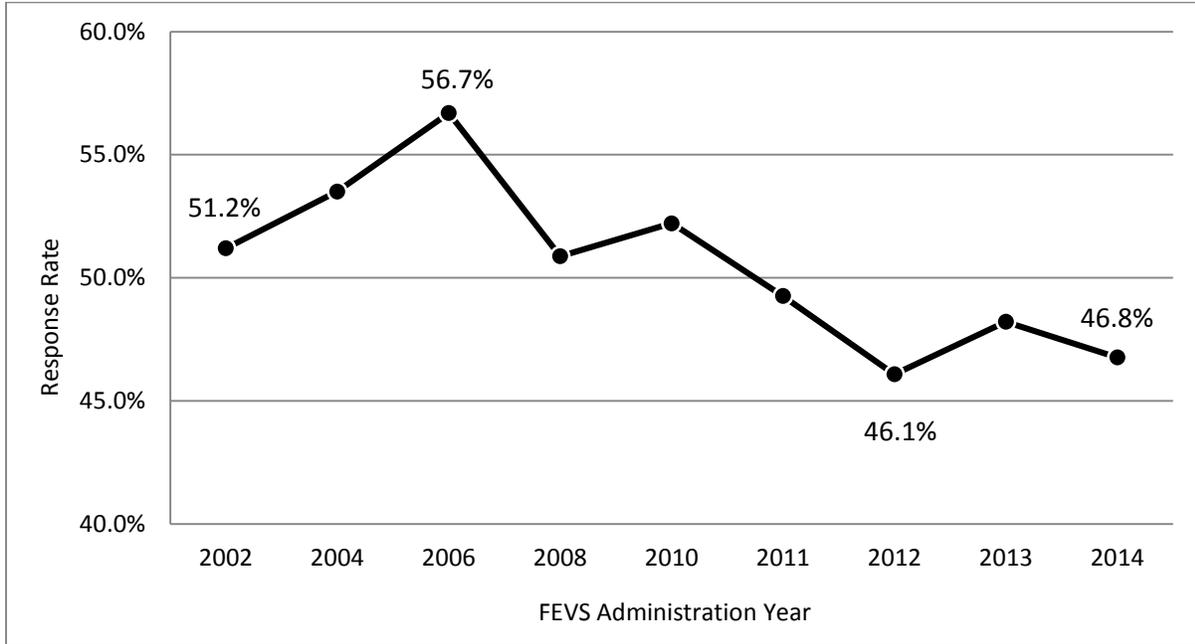
**Figure 1:** Sample Sizes and Respondent Counts in the Federal Employee Viewpoint Survey, 2002 – 2014.



A spate of research over the last fifteen or so years has found that response rates to surveys in the United States and throughout the world have been declining (Atrostic et al., 2001; de Leeuw and de Heer, 2002; Petroni et al., 2004; Curtin et al., 2005; Brick and Williams, 2013). The increasing gap between the two trend lines in Figure 1 intimates how the Federal Employee Viewpoint Survey (FEVS) is no exception to that general pattern. Indeed, this is confirmed by Figure 2, which shows how the FEVS response rate peaked in 2006 and has been gradually declining ever since. A response rate greater than 50% was once routinely attainable, but that threshold has not been reached since FEVS 2010. The lowest response rate on record was that of the FEVS 2012 census (46.1%), and the most recent FEVS administration's response rate hardly surpassed that mark (46.8%).

<sup>1</sup> A few exceptions where employees were sampled: (1) the Department of Veterans Affairs; (2) non-supervisory Air Traffic Controllers within the Federal Aviation Administration; and (3) the Defense Commissary Agency and Defense Finance and Accounting Service within the Department of Defense.

**Figure 2:** Federal Employee Viewpoint Survey Unweighted Response Rates, 2002 – 2014.



The current data collection procedure for the FEVS begins by assigning agencies to one of two cohorts that are deployed into the field one week apart in late April or early May. Table 1 summarizes the email contact sequence for an example agency in the first launch cohort of FEVS 2014. Each agency, regardless of cohort, is allotted a six-week field period. In the first week, an initial email invitation is sent to all sampled employees, typically on a Tuesday. Nonrespondents are sent weekly reminder emails, generally on the same day and time. One exception, however, is that the final reminder is sent on the morning of the last Friday in the field period, alerting those who have yet to respond that the survey will close at the end of the day. No monetary or token incentives are offered, but employees are notified that they are permitted to complete the survey during work hours.

**Table 1:** FEVS 2014 Survey Invitation and Reminder Schedule for an Example Agency in the First Launch Cohort.

Field Period Week	Date	Day of Week	Email Type
1	April 29	Tuesday	Initial Invitation to Participate
2	May 6	Tuesday	Reminder #1
3	May 13	Tuesday	Reminder #2
4	May 20	Tuesday	Reminder #3
5	May 27	Tuesday	Reminder #4
6	May 3	Tuesday	Reminder #5
6	May 6	Friday	Reminder #6 (final - survey closes today)

While the current protocol fosters logistical tractability, a very welcomed feature considering the cumbersome task of sending out roughly 5 million emails invitations and reminders over the course of the FEVS field period, one

potential disadvantage is that a particular agency's assigned contact time block may be inopportune for some of its employees, perhaps because of competing activities such as a standing meeting or other official duties rendering them less attuned to email traffic. This is ostensibly an important factor, since approximately 60% of all FEVS 2014 completes were obtained on the same day an email solicitation was sent. Bearing this in mind, we sought to determine whether imparting some variability into when email reminders are sent could ultimately improve the likelihood an employee responds to the survey and reverse the downward trend of response rates witnessed in recent FEVS administrations.

The survey methodology literature is replete with articles proposing and evaluating various contact timing strategies, yet most focus on optimizing the probability of making contact within the milieu of interviewer-administered surveys, whether conducted by telephone (Weeks et al., 1987; Greenberg and Stokes, 1990, Brick et al., 1996) or face-to-face in a cross-sectional (Purdon et al., 1999; Wagner, 2013) or panel survey context (Lipps, 2012; Durrant et al., 2013). Comparatively speaking, the literature is much more scant with respect to self-administered surveys, especially those conducted over the Internet, but at least two studies have investigated alternative email contact timing strategies. The first was in an establishment survey as reported by Faught et al. (2004). A sample of approximately 5,000 email addresses for points of contact at manufacturing firms—some generic (e.g., support@xyz.org) and some for specific individuals—was randomly allocated into one of fourteen possible morning/afternoon time blocks during the week. The authors found that sending emails exclusively on Wednesday mornings yielded the highest response rate. In a survey of science and engineering professionals, Sauermann and Roach (2013) compared twenty-five experimental groupings of various combinations of temporal factors, such as time of day, day of the week, and lag time between email contacts. In contrast to the study by Faught et al. (2004), Sauermann and Roach (2013) were unable to uncover any substantive differences amongst the various strategies. Hence, there is no clear guidance to be gleaned from the available literature. Although it seems unlikely that a single email contact protocol would work best on all survey populations, more research is needed to help shed more light on this phenomenon. This article serves as one such contribution to help fill this research gap.

## **Data and Methods**

The goal of our experiment was to investigate whether, and to what extent, an alternative email contact sequence could improve response rates in the FEVS. As such, we did not endeavor to conduct the experiment on all 87 agencies participating in the FEVS 2015. Instead, we restricted our focus to one large agency, the U.S. Department of Defense (DoD) Fourth Estate. In a nutshell, DoD Fourth Estate is comprised of civilian DoD personnel not directly employed as part of the U.S. Army, U.S. Navy, or U.S. Air Force. Components of the agency include the Office of the Secretary of Defense, Defense Information Systems Agency, Defense Contract Audit Agency, Missile Defense Agency, and Defense Commissary Agency, to name a few. We felt this was a fitting agency on which to concentrate our efforts because it has historically achieved below-average FEVS response rates, and survey practitioners at the agency have demonstrated a keen interest over the years in researching the causes, correlates, and potentially detrimental impact of nonresponse in other surveys sponsored by the agency (e.g., Caplan et al., 2004; Caplan and Quigley, 2005; Caplan and Hoover, 2008).

The sampling frame for the FEVS is derived from an expansive personnel database maintained by OPM called the Statistical Data Mart of the Enterprise Human Resources Integration (EHRI-SDM). For more details on its contents and scope, see [http://www.fedscope.opm.gov/datadefn/ae\\_hri\\_sdm.asp](http://www.fedscope.opm.gov/datadefn/ae_hri_sdm.asp). Following the single-stage stratified sample design described in OPM (2014), a total of 34,799 DoD Fourth Estate employees were selected for participation in the 2015 FEVS. The data collection period began Tuesday, April 28 and ended Friday, June 5. Prior to the start of the field period, sampled employees were randomly assigned to one of three experimental groups of approximately equal size. The first group received the traditional FEVS email contact sequence with the initial invitation and all subsequent reminders sent on Tuesday morning. The lone exception was that a final reminder was sent on the

morning of Friday, June 5, the day the survey closed. In fact, all nonresponding employees received this final reminder regardless of experimental group.

The second experimental group received what we refer to as a rotating cohort strategy. Each employee was first placed at random into one of six cohorts, which in turn was assigned to one of the following six time blocks for purposes of the initial email invitation: (1) Tuesday a.m.; (2) Tuesday p.m.; (3) Wednesday a.m.; (4) Wednesday p.m.; (5) Thursday a.m.; or (6) Thursday p.m. Each cohort was then rotated to the next time block for the ensuing week’s reminder. Figure 3 is provided to help visualize the procedure. For example, individuals assigned to cohort A received the initial email solicitation on Tuesday morning, received the first reminder on Tuesday afternoon, and received the third reminder on Wednesday morning, and so on. Practical experience from prior FEVS administrations has suggested that fewer employees are on duty Mondays and Fridays relative to the other three days in the work week, as evidenced by higher rates of out-of-office automatic email replies, so those two days were excluded from consideration in this experiment.

Emails sent during the morning time blocks went out at approximately 8 a.m. Eastern Standard Time, and emails sent during the afternoon time blocks went out at approximately 2 p.m. Eastern Standard Time. The bulk emailing software e-Campaign (<http://www.lmhsoft.com/ecamp/>) was used to send out the email invitations and reminders en masse. Note that, because there were a total of six email solicitations, all sampled employees were eligible to receive an email during each of the six possible time blocks.

**Figure 3:** Illustration of the Rotating Cohort Email Contact Timing Strategy.

		Week 1					Week 4		
		Tues	Wed	Thurs			Tues	Wed	Thurs
a.m.		A	C	E	a.m.		D	F	B
p.m.		B	D	F	p.m.		E	A	C
		Week 2					Week 5		
		Tues	Wed	Thurs			Tues	Wed	Thurs
a.m.		F	B	D	a.m.		C	E	A
p.m.		A	C	E	p.m.		D	F	B
		Week 3					Week 6		
		Tues	Wed	Thurs			Tues	Wed	Thurs
a.m.		E	A	C	a.m.		B	D	F
p.m.		F	B	D	p.m.		C	E	A

Bethlehem et al. (2011) define a typology of survey design features for employing principles of a responsive survey design as originally discussed in Groves and Heeringa (2006). The technique depicted in Figure 3 is what Bethlehem et al. (2011) would label a static adaptive design, since the information determining any sort of change in the contact protocol is known in advance of the data collection period. By contrast, Bethlehem et al. (2011) define a dynamic adaptive design to be one where protocol changes occur in response to observations made during the data collection period. Because the second alternative contact sequence we investigated fits within this paradigm, we refer to it as the dynamic adaptive strategy.

The dynamic adaptive strategy proceeded as follows. As with the rotating cohort strategy, employees were randomly allocated into one of the same six time blocks for purposes of the initial email invitation. But rather than cycling employees through each time block in a predetermined fashion, auxiliary information on the sampling frame was used to fit a multinomial logistic regression model (Hosmer et al., 2013) where the dependent variable was the act of responding during one of the six time blocks. For modeling purposes, responses obtained on Friday or Saturday were treated as having occurred during the Thursday p.m. time block, while responses obtained on Sunday or Monday were treated as having occurred during the Tuesday a.m. time block. Predictor variables for the model included the employees' agency component, supervisory status (non-supervisor, supervisor, executive), gender, an indicator of working part-time as opposed to full-time, an indicator of being a minority race/ethnicity, pay level, and an indicator of whether the individual's job title falls within the umbrella of a STEM (Science, Technology, Engineering, or Mathematics) occupation. The model was (re)fitted at the conclusion of each week of the field period based on the cumulative data. The key statistic extracted was each nonresponding employee's vector time-block-specific response propensities. The ensuing week's reminder time block for these individuals was assigned stochastically in proportion to these propensities. Hence, the notion was to predispose nonrespondents to receive reminders during times which real-time evidence was suggesting comparable employees were responding at higher rates.

## Results

Table 2 summarizes the sample sizes and response rates for each email timing strategy. The response rates we report correspond to the RR1 formula of the American Association of Public Opinion Research (AAPOR, 2015). Because of an approximate six-month lag between the most recently available EHRI-SDM data used to develop the sampling frame and when the survey went into the field, there is some degree of ineligibility caused by employees who since departed the agency. These departures were identified by either out-of-office replies, ad-hoc information provided by the agency, or via a systematic procedure whereby we verified future employment by means of a more up-to-date personnel roster derived from EHRI-SDM. Cases of known ineligibility were removed from the denominator of the response rate calculation, as reflected by the line in Table 2 labeled "Adjusted Sample Size." To be defined as a respondent, the individual must have answered at least 21 of the 84 core, non-demographic survey items. As described in OPM (2014), this is the conventional FEVS threshold for defining a complete case.

Interestingly, the Tuesday morning email contact strategy outperformed the two alternatives. The Tuesday morning approach yielded a response rate of 48.4%, whereas the rotating cohort strategy was two percentage points below that mark at 46.4%. At 47.1%, the dynamic adaptive approach was slightly better than the rotating cohort alternative, but it still lagged the traditional method. Although modest in magnitude, these differences are statistically significant ( $\chi^2 = 8.68$ ;  $p = 0.013$ ).

**Table 2:** Sample Sizes and Response Rates by Email Timing Strategy.

Measure	Email Timing Experimental Group			Totals
	Traditional	Rotating Cohort	Dynamic Adaptive	
Initial Sample Size	11,599	11,600	11,600	34,799
Ineligibles	824	902	844	2,570
Adjusted Sample Size	10,775	10,698	10,756	32,229
Respondents	5,216	4,968	5,065	15,249
Response Rate	48.4%	46.4%	47.1%	47.3%

We anticipated unveiling pockets of deviations from this overall finding as we inspected response rates differences across a range of domains—for instance, those defined by particular demographic categorizations or components of the agency—but the general pattern almost always held. Figure 3 illustrates this tendency based on a few example demographics: employee gender, whether a minority or non-minority race/ethnicity, supervisory status, and pay level. The traditional method generally garnered the highest response rate, followed by the dynamic adaptive and rotating cohort methods, respectively. There are a few instances where the dynamic adaptive method outperformed the rotating cohort methods, but both still lagged the Tuesday morning approach. One exception is that employees in the highest pay level responded at the highest rate to the rotating cohort, but fewer than 3% of DoD Fourth Estate employees fall within this demographic category, so we caution that the finding could be simply attributable to sampling error rather than anything substantive.

**Figure 3:** Response Rates by Email Timing Strategy for Select Employee Demographics.  
Gender minority status supervisory status pay level

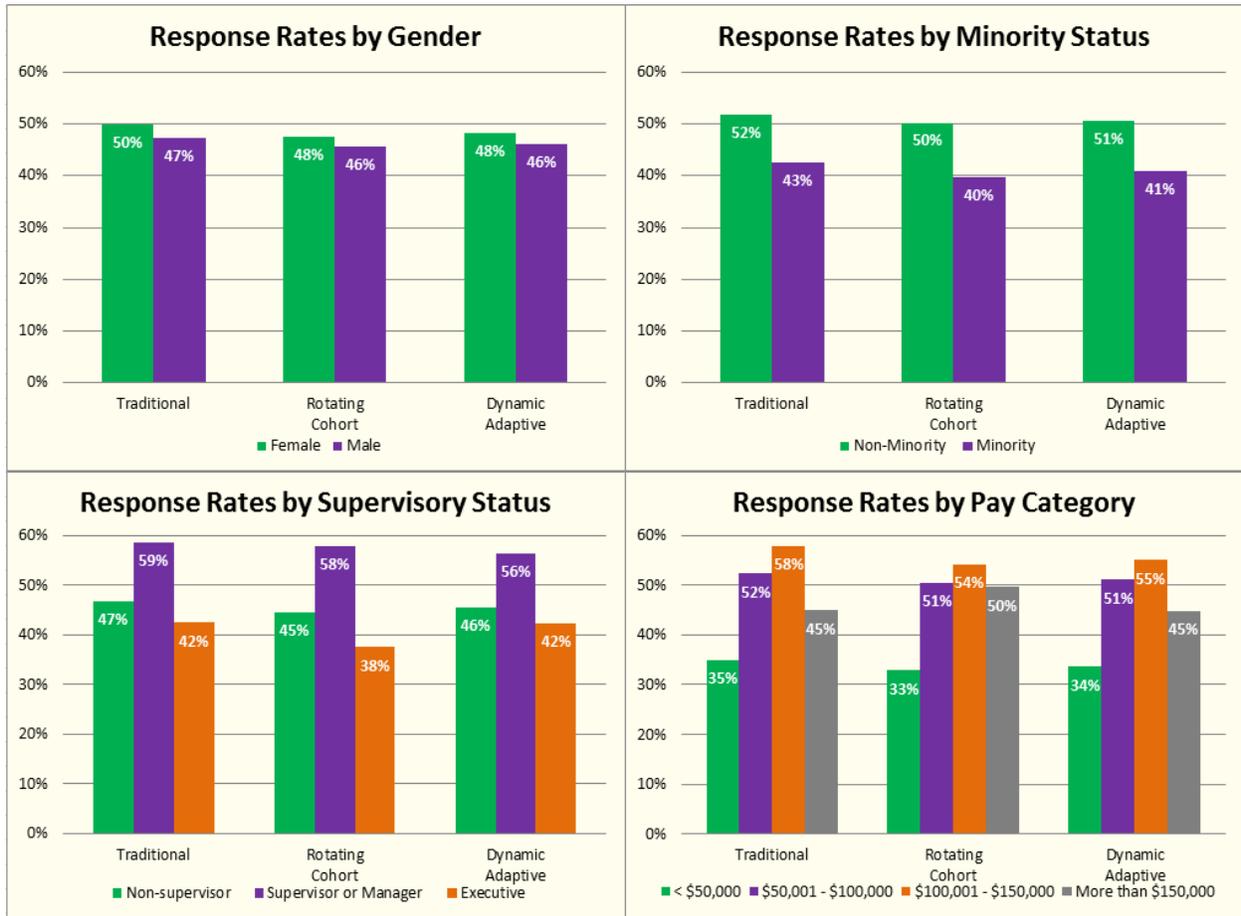
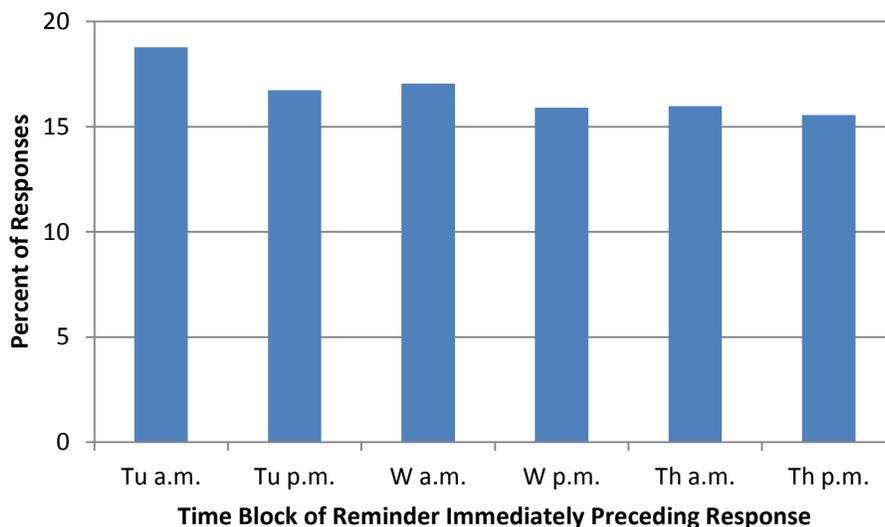


Figure 4 is a histogram of ultimate reminders for those responding as part of the rotating cohort experimental group. For purposes of this analysis, responses obtained after the last Friday reminder went out were linked back to the time block of the reminder sent earlier in that final field period week. The figure illustrates how the Tuesday morning email was the most frequent lure. It also illustrates a subtle tendency for emails sent later in the week are less effective. Another noteworthy point is that the marginal percentage of morning emails was slightly larger than that of afternoon emails—51.4% versus 48.6%, respectively—which provides some evidence that, all else equal, sending emails in the morning appears to be more efficacious than sending emails in the afternoon.

**Figure 4:** Time Block Distribution of Reminders Immediately Preceding a Response for the 4,968 Respondents in the Rotating Cohort Experimental Group.



### Concluding Remarks and Ideas for Further Research

Historically, the Web-based FEVS has been governed by a data collection strategy whereby the initial invitation and all subsequent reminders to nonrespondents are sent during the same weekly time block, typically Tuesday mornings. One exception is that a final reminder is sent on the Friday that the survey closes. In this paper, we presented results from an experiment conducted on one agency participating in the 2015 FEVS with the aim of determining whether introducing some variability into the timing of email reminders could increase response rates. Surprisingly, both alternatives we considered proved inferior to the conventional Tuesday morning approach. This finding prevailed across a range of domains defined by work units within the agency and employee demographics. Based on further analysis of the rotating cohort experimental group, one where email time blocks were assigned randomly between Tuesday morning and Thursday afternoon, there is some modest evidence that sending emails during the Tuesday morning time block is most likely to trigger a response, and that subsequent time blocks later on in the work week are gradually less likely. Based on these findings, the FEVS administration team has no plans to part with the email invitation and reminder protocol currently in place.

While we argue that this study makes for a worthwhile contribution to an understudied area of the applied survey research literature—alternative contact timing strategies in self-administered surveys conducted via the Web—our results are limited in the sense that we focused only one of 87 agencies that participate in the FEVS. Further research could investigate whether these findings are replicable across other agencies. Moreover, as the FEVS is an organizational climate survey of a very specific population, employees of the United States Federal Government, there is no guarantee these results would extend directly to other survey topics and other target populations.

There are many other potentially fruitful extensions to the present study. For example, future research could explore whether information about when the employee actually views the email, perhaps captured by means of an email read receipt, could be useful. This would enable us to make the important distinction between making contact and securing a response, and could offer insight into whether certain times are associated with more or less lag time between email receipt and survey completion. Another intriguing idea would be to utilize prior FEVS administration timestamp information. All but 16 of the agencies currently participating in the FEVS conduct a census of their workforce, so even with an overall response rate around 50%, there is likely considerable respondent

overlap between two adjacent years' administrations. Heeding a recommendation in Lipps (2012), Kreuter and Müller (2015) attempted to conduct interviews in a panel survey around the same time the individual had responded in a prior wave. They found that the technique helped improve contact rates, but not response rates. It would be interesting to see how a method similar in spirit would perform in a self-administered, Web-based survey context.

## References

- American Association of Public Opinion Research (AAPOR). 2015. *Standard definitions: Final dispositions of case codes and outcome rates for surveys*. 8th ed. Lenexa, KS: AAPOR.
- Atrostic, B., N. Bates, G. Burt, and A. Silberstein. 2001. Nonresponse in US government household surveys: Consistent measures, recent trends, and new insights. *Journal of Official Statistics* 17:209–26.
- Bethlehem, J., F. Cobben, and B. Schouten. 2011. *Handbook of nonresponse in household surveys*. Hoboken, NJ: Wiley.
- Brick, M., B. Allen, P. Cunningham, and D. Maklan. 1996. Outcomes of a calling protocol in a telephone survey,” *Proceedings of the Survey Research Methods Section*. Alexandria, VA: American Statistical Association.
- Brick, M., and D. Williams. 2013. Explaining rising nonresponse rates in cross-sectional surveys. *The Annals of the American Academy of Political and Social Science* 645:36–59.
- Caplan, J., B. Quigley, and J. Mathew. 2004. Nonresponse research in the Department of Defense. *Proceedings of the Annual Conference of the American Association for Public Opinion Research*. Phoenix, AZ, May 13–16.
- Caplan, J., and B. Quigley. 2005. Administrative interventions and their impact on response rates in military members’ attitude surveys,” *Proceedings of the Annual Conference of the American Association for Public Opinion Research*. Miami, FL, May 12–15.
- Caplan, J., and E. Hoover. 2008. Administrative methods of increasing response rates. *Proceedings of the Federal Computer-Assisted Survey Information Collection (FedCASIC) Workshops*. Washington, DC, March 11–13.
- Curtin, R., S. Presser, and E. Singer. 2005. Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly* 69:87–98.
- De Leeuw, E., and W. De Heer. 2002. Trends in household survey nonresponse: A longitudinal and international comparison. In *Survey nonresponse*, ed. R. Groves, D. Dillman, J. Eltinge, and R. Little, 41–54. New York, NY: Wiley.
- Durrant, G., J. D’Arrigo, and G. Müller. 2013. Modeling call record data: Examples from cross-sectional and longitudinal surveys. In *Improving surveys with paradata: Analytic uses of process information*, ed. F. Kreuter, 281–308. New York, NY: Wiley.
- Faught, K., D. Whitten, and K. Green. 2004. Doing survey research on the internet: Yes, timing does matter. *Journal of Computer Information Systems* 45:26–34.
- Greenberg, B., and S. Stokes. 1990. Developing an optimal call scheduling strategy for a telephone survey. *Journal of Official Statistics* 6:421–35.
- Groves, R., and S. Heeringa. 2006. Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistics Society, Series A*, 169:439–57.

- Hosmer, D., S. Lemeshow, and R. Sturdivant. 2013. *Applied logistic regression*. 3rd ed. New York, NY: Wiley.
- Kreuter, F., and G. Müller. 2015. A note on improving process efficiency in panel surveys with paradata. *Field Methods* 27:55–65.
- Lipps, O. 2012. A note on improving contact times in panel surveys. *Field Methods* 24:95–111.
- Petroni, R., R. Sigman, D. Willimack, S. Cohen, and C. Tucker. 2004. Response rates and nonresponse in establishment surveys – BLS and Census Bureau. <https://www.bea.gov/about/pdf/ResponseratesnonresponseinestablishmentsurveysFESAC121404.pdf> (accessed December 28, 2015).
- Purdon, S., P. Campanelli, and P. Sturgis. 1999. Interviewers' calling strategies on face-to-face interview surveys. *Journal of Official Statistics* 15:199–216.
- Sauermann, H., and M. Roach. 2013. Increasing web survey response rates in innovation research: An experimental study of static and dynamic contact design features. *Research Policy* 42:273–86.
- United States Office of Personnel Management. 2014. *2014 Federal Employee Viewpoint Survey: Technical report*. <http://www.fedview.opm.gov/2014/published/> (accessed December 28, 2015).
- Wagner, J. 2013. Using paradata-driven models to improve contact rates in telephone and face-to-face surveys. In *Improving surveys with paradata: Analytic uses of process information*, ed. F. Kreuter, 145–70. New York, NY: Wiley.
- Weeks, M., R. Kulka, and S. Pierson. 1987. Optimal call scheduling for a telephone survey. *Public Opinion Quarterly* 51:540–49.