nature human behaviour

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41562-019-0795-z

In the format provided by the authors and unedited.

What we can learn from five naturalistic field experiments that failed to shift commuter behaviour

Ariella S. Kristal[1*] **and Ashley V. Whillans**[2]

[1]Organizational Behavior Unit, Harvard Business School, Boston, MA, USA. [2]Negotiation, Organizations, and Markets Unit, Harvard Business School, Boston, MA, USA. *e-mail: akristal@hbs.edu

What we can learn from five naturalistic field experiments that failed to shift commuter behavior
Supplementary information

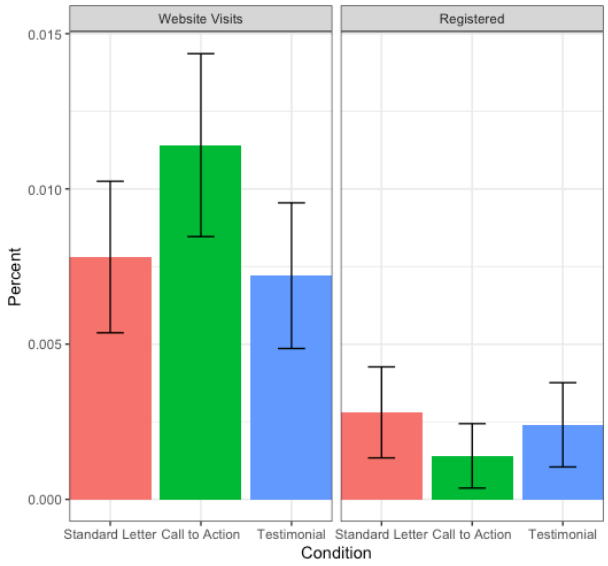
Supplementary Results

Study 1. Carpooling registration. Table S1 shows the summary statistics of website visits per condition. The three conditions did not differ in terms of carpooling registration ($F_{(1,14985)}$ = 0.182, p = 0.67). The three conditions also did not significantly differ from the control condition. Figure S1 visually depicts the different visit and registration rates by letter conditions.

In Study 1, we also found that the behavioral message that contained the call to action—and led to the greatest number of website visits—also translated to the fewest number of registered users. This could indicate that the person who visits the website due to a letter that makes it easy to access the website (but who would not have visited the website had they received the standard letter) is in fact less likely to follow through with signing up for the service because they aren't willing to complete the effort needed to complete the more difficult action. It is therefore important to be cautious when drawing conclusions from interventions by measuring proxy outcomes such as website visits. Because of the small number of people who visited the website, we are cautious about these results and recommend further examination of this hypothesis.

Supplementary Table 1. Carpooling registrations by condition.

	N	Website visits mean (SD)	Registration mean (SD)	*t*	*p*	Cohen's *d* [95%CI]
Control	39,900	N/A	0.0005 (0.02)	-	-	-
Letter 1	4,996	0.0078 (0.09)	0.0028 (0.05)	3.042	0.002	0.08 [-1.96, 0.975]
Letter 2	4,995	0.0114 (0.11)	0.0014 (0.04)	1.664	0.096	0.04 [-1.96, 0.975]
Letter 3	4,996	0.0072 (0.08)	0.0024 (0.05)	2.709	0.007	0.07 [-1.96, 0.975]



Supplementary Figure 1. Website visits and registrations by letter conditions

Study 4. Personalized travel plan. A second self-reported outcome measure was collected: the frequency of times employees drove an SOV to work in the past month (intervals: never, once or twice, 3-5 times, 6-10 times, more than 10 times but not every day, or every day). Due to the fact that this variable is an ordered categorical variable, our analysis explored whether participants changed their position in the frequency distribution as a result of their condition assignment, which can be difficult to interpret. For this DV, we also failed to find a significant shift in the frequency interval with which employees commuted to work by SOV from the previous month.

Supplementary Table 2. TOST results for effect sizes based on minimum detectable effect sizes given post-hoc power analysis (80%) for each study

	Study 1	Study 2	Study 3a	Study 3b	Study 4
Effect size of interest	0.03	0.13	0.07	0.12	0.19
TOST results: t-value lower bound t-value upper bound p-value lower bound p-value upper bound	 -1.50 07.03 0.934 $1x10^{-12}$	 1.75 -1.75 0.041 0.041	 3.40 -2.22 0.0003 0.013	 4.38 -3.87 0.000006 0.00005	 2.71 -2.90 0.004 0.002
TOST confidence interval: lower bound 90% CI: upper bound 90% CI:	 -0.002 -0.001	 -0.038 0.038	 -0.003 0.006	 -0.004 0.005	 -0.247 0.219
NHST confidence interval: lower bound 95% CI: upper bound 95% CI:	 -0.002 -0.001	 -0.046 0.046	 -0.004 0.007	 -0.005 0.006	 -0.292 0.264
Equivalence test	non-significant	significant	significant	significant	significant
Null hypothesis test	significant	non-significant	non-significant	non-significant	non-significant

Rationale for using the minimum detectable effect size as the SESOI: Across Studies 1, 2, and 3b, we tested novel interventions—i.e., these interventions had never been part of previously published research on this topic. As a result, it was not possible to base our effect sizes for the TOST analyses on previously published research. In Studies 3a and 4, the interventions that we tested had been included in previously published literature but with highly variable effect sizes.

In the absence of published literature on this topic to base our effect size estimates on, we sent a short survey to a transportation-demand management (TDM) listserv. Specifically, we asked a group of professionals who are tasked—as part of their full-time jobs—with changing commuter behavior the following question: “Imagine an airport with 76,000 employees, half of

whom drive alone to work each day. If X single occupancy vehicle drivers reduced their SOV use by even one day a week, it would be considered a meaningful difference. What is X to you?" We received responses from 14 professionals through this list-serv. Their responses ranged from 1% to 50%, with a mean of 11.04% and a standard deviation of 13.53.

Given the lack of precise expected effect sizes from the literature (due to both novelty and mixed findings in this literature), and the lack of consensus from TDM professionals about what constitutes a meaningful effect, we now use the minimum detectable effect size for each study in our TOST analysis. This is also consistent with Lakens (2017) who writes, *"Without practical boundaries or theoretical boundaries that indicate which effect size is meaningful, the maximum sample size you are willing to collect implicitly determines your SESOI. Therefore, setting equivalence boundaries to your SESOI in an equivalence test allows you to reject effect sizes larger than you find worthwhile to examine, given available resources."*

Supplementary Notes

Supplementary Table 3. Study timeline

Fall/Winter 2015	Study scoping and interviews with airport employees
April 2015	Mail out carpooling letters (Study 1) Mail out bus letters and free trial vouchers (Study 3a) Free bus trial (Study 3a)
May 2015	Collect outcome data for Study 1 Send out carpooling emails (Study 2) Mail out bus trial follow-up letters (Study 3b)
June 2015	Collect outcome data for Study 2, Study 3a, and Study 3b
January 2016	Administer survey to employees in office park building (Study 4) Send personalized travel plans to randomly selected subgroup of employees who reported driving alone to work at least one day a week (Study 4) Start delivering one-on-one meetings (Study 4)
February 2016	Continue delivering one-on-one meetings (Study 4)
March 2016	Collect outcome data for Study 4

Supplementary Discussion

Our interventions led to a reduction of 29 employees out of 68,915 (less than 0.05%) commuting by SOV to work over the course of the study period. Because of the lack of statistical significance, we cannot conclude that these people changed their behavior due to our intervention. Indeed, the TOST analyses confirm that these differences are equivalent to zero for four of the five experiments. Therefore, the costs of designing and mailing out 15,000 letters, working with the carpool service provider to design three different emails, arranging a free bus trial across two locations and designing and mailing out over 8,000 letters (for Studies 3a and 3b), commissioning 790 personalized travel plans, and the person-hours it took to design and implement these trials outweighs the uncertain gains in behavior change observed in this study

Supplementary Table 4. Benefits of the interventions

Study	Behavior change for control participants (out of total)	Behavior change for treatment participants (out of total)	Additional impact of the intervention	CO2 Savings* (g)
1	4 (39,900)	3 (14,987)	-1	N/A
2	1 email recipient (871 across all three conditions)		1	793,904
3a	32 (2,453)	54 (4,781)	22	17,465,888*
3b	8 (1,085)	14 (1,125)	6	4,763,424
4	Average reduction in SOV days: 0.62 (across 305 participants)	Average reduction in SOV days: 0.61 (across 790 participants)	-0.01 SOV days	N/A

*This is less than is emitted from one flight from New York to London on a standard Boeing 747

**We base this on the assumption that the average person lives 10 miles from the airport, according to the staff survey, the average worker works 232 days a year, and that the 171.1gCO_2 per mile is emitted (based on https://www.gov.uk/government/publications/new-car-carbon-dioxide-emissions), leading to 793,904 gCO_2 per year per person who no longer drives due to our interventions.