



Management Science

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Effectiveness of Planning Prompts on Organizations' Likelihood to File Their Overdue Taxes: A Multi-Wave Field Experiment

Nicole Robitaille, Julian House, Nina Mazar

To cite this article:

Nicole Robitaille, Julian House, Nina Mazar (2021) Effectiveness of Planning Prompts on Organizations' Likelihood to File Their Overdue Taxes: A Multi-Wave Field Experiment. *Management Science* 67(7):4327-4340. <https://doi.org/10.1287/mnsc.2020.3744>

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. You are free to download this work and share with others, but cannot change in any way or use commercially without permission, and you must attribute this work as "*Management Science*. Copyright © 2020 The Author(s). <https://doi.org/10.1287/mnsc.2020.3744>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by-nc-nd/4.0/>."

Copyright © 2020, The Author(s)

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Effectiveness of Planning Prompts on Organizations' Likelihood to File Their Overdue Taxes: A Multi-Wave Field Experiment

Nicole Robitaille,^a Julian House,^b Nina Mazar^c

^a Smith School of Business, Queens University, Kingston, Ontario K7L 3N6, Canada; ^b Behavioural Insights Unit, Treasury Board Secretariat, Government of Ontario, Toronto, Ontario M7A 1Y7, Canada; ^c Questrom School of Business, Boston University, Boston, Massachusetts 02215

Contact: nicole.robitaille@queensu.ca,  <https://orcid.org/0000-0002-3323-4423> (NR); julian.house@ontario.ca (JH); nina@ninamazar.com,  <https://orcid.org/0000-0001-8248-654X> (NM)

Received: July 4, 2017

Revised: October 12, 2018; January 17, 2020

Accepted: June 15, 2020

Published Online in Articles in Advance: November 10, 2020

<https://doi.org/10.1287/mnsc.2020.3744>

Copyright: © 2020 The Author(s)

Abstract. This paper investigates the effectiveness of planning prompts on organizations' tax compliance behavior. We conducted a large-scale, multi-wave field experiment examining the tax-paying behavior of all organizations that failed to file timely annual returns for a payroll tax in the province of Ontario. Organizations were randomly assigned to receive one of two letters: Ontario's standard late notice (control) and a revised experimental late notice, which included step-by-step instructions of when, where, and how to file a return. Our data indicate that planning prompts are effective at increasing organizations' timely tax payment. In addition to replicating these findings across two waves, we demonstrate that, although our intervention did not appear to have effects that persisted across tax years, organizations also did not habituate to our manipulation and its effects were consistent across repeated exposures. Our study is among the first to demonstrate that a simple behavioral intervention that has typically been applied to individuals to help them to act upon their existing motivations can be effective in the realm of tax compliance and organizational behavior.

History: Accepted by David Simchi-Levi, behavioral economics.



Open Access Statement: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. You are free to download this work and share with others, but cannot change in any way or use commercially without permission, and you must attribute this work as "Management Science. Copyright © 2020 The Author(s). <https://doi.org/10.1287/mnsc.2020.3744>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by-nc-nd/4.0/>."

Supplemental Material: The online appendix is available at <https://doi.org/10.1287/mnsc.2020.3744>.

Keywords: government tax policy • economics behavior and behavioral decision making • organizational studies behavior • organizational studies decision making • field experiment

1. Introduction

How can governments effectively increase tax compliance and, more specifically, increase the payment of overdue taxes? This question is of critical importance because efficient funding of public services involves both levying tax and minimizing collection costs (Andreoni et al. 1998), yet compliance continues to be an ongoing problem (Matsaganis et al. 2012). For example, in 2011 it was estimated that outstanding tax debt totaled approximately USD 650–700 billion across OECD countries (OECD 2014). The traditional approach to increasing tax compliance, used universally by governments, is to apply financial disincentives for noncompliance (e.g., penalties, interest, and fines), thereby providing taxpayers with extrinsic motivation to pay their dues in a timely fashion. Additionally, some governments have examined applying insights from behavioral science to increase individuals' tax compliance, for example by providing details of other taxpayers' behavior to make social norms and the moral costs of noncompliance salient and bolster

taxpayers' intrinsic motivations to file (Torgler 2007, Hallsworth et al. 2017). In this paper, we examine a behavioral approach that has yet to be applied to tax compliance: encouraging organizations to act on their existing motivations through the use of planning prompts (see Rogers et al. 2015 for a review).

Although there has been considerable research into how behavioral insights may be applied to enhance individuals' decisions, less is known about the effectiveness of such interventions when it comes to organizational actors (OECD 2017a). Moreover, most empirical research on tax compliance has been conducted at the individual level, and organizational tax compliance remains a relatively underexplored but growing area of research (Alm and McClellan 2012, Hallsworth 2014). To address these gaps, we test the efficacy of a novel approach to tax collection—employing planning prompts—using a large-scale, organizational field experiment. Notably, because our field experiment features two waves, this research is also able to explore the longer-term impact

of our intervention (i.e., over two tax years) including testing for replication, habituation, and the persistence of our effects.

2. Theoretical Approach

Although the most well-studied approach to increasing tax compliance is the use of penalties and fines to decrease noncompliance (Allingham and Sandmo 1972), research suggests these extrinsic motivators may not be the only drivers of compliance behavior. For one, if these disincentives were the only drivers of compliance, expected utility models would actually predict higher levels of noncompliance, given the low probabilities of being selected for audit in most jurisdictions (e.g., Skinner and Slemrod 1985, Slemrod and Yitzhaki 2002). Moreover, other research finds that individuals' and organizations' *intrinsic* motivations to pay their taxes, otherwise known as tax morale (Luttmer and Singhal 2014), is predictive of their compliance behaviors (e.g., Alm and Torgler 2006, Cummings et al. 2009, Alm and McClellan 2012). As a result, governments have shown increasing interest in supplementing the more traditional approaches (e.g., financial disincentives) with techniques from behavioral science to achieve their policy objectives (Benartzi et al. 2017, Pomeranz 2017).

Indeed, there is a growing body of work demonstrating how seemingly small contextual factors—such as defaults (Johnson and Goldstein 2003), framing (Tversky and Kahneman 1981), simplification (Sunstein 2013), and descriptive norms (Goldstein et al. 2008)—can be used to nudge individuals' behaviors and improve outcomes across a variety of domains including education, health, the environment, and finance (Behavioural Insights Team 2014, OECD 2017b). Nudges aim to leverage our understanding of human decision making, including individuals' reliance on heuristics, cognitive limitations, and biased processing, to affect behavior in a predictable way and maintain individuals' freedom of choice (Thaler and Sunstein 2008). One commonly applied behavioral intervention in the area of tax compliance is leveraging tax morale, which is driven in part by perceptions of fairness, trust, and social norms regarding compliance (see Arcos Holzinger and Biddle 2016 for review). For example, in two large-scale natural field experiments, Hallsworth et al. (2017) demonstrated that including descriptive social norm messages in tax reminder letters (e.g., "Nine out of ten people pay their tax on time"), as well as messages highlighting how taxes contribute to the common public good (e.g., "Paying tax means we all gain from vital public services like the NHS, roads, and schools"), significantly increased payment rates for individuals' overdue tax. In line with research on conditional cooperation, these findings support the

proposition that providing information about high levels of compliance increases tax morale and, in turn, increases compliance behaviors (Torgler and Schneider 2005, Frey and Torgler 2007, Alm et al. 2017). Conversely, field experiments in which individual taxpayers were reminded of prescriptive norms (i.e., that they ought to comply with tax law) have found no effect on either imputed tax evasion (Blumenthal et al. 2001, Ariel 2012, Torgler 2013) or timely tax payment (Torgler 2004). Thus, there is some evidence that norms can be leveraged to increase tax compliance—not by reminding taxpayers of their abstract duty, but by informing them that their fellow citizens accept and support the same social contract. However, the vast majority of applied behavioral research has focused on individuals' behaviors, and to date, field experiments remain rare in organizational research (Scandura and Williams 2000, Eden 2017, Hauser et al. 2017).

Advancing our understanding of organizations' tax compliance is critically important because of the role they play as both remitters of sales and income taxes and as subjects of corporate tax. Indeed, in Ontario where our field experiment was conducted, the taxes that organizations pay make up over 18% of the government's total tax revenue (Government of Ontario 2016). Yet, despite the breadth of research on applying behavioral research to improve individuals' behaviors, we know relatively little about how such tools influence organizational behavior (OECD 2017a). On the one hand, because organizational behavior is ultimately decomposable to the behaviors of individuals, one might expect the effects of choice architecture and nudges to be isomorphic across levels of aggregation. On the other hand, there are well-established ways in which the processes of organizing and making decisions in groups can alter the constructs on which behavioral interventions operate. One reason to expect reduced efficacy is that, compared with individuals, decision makers within organizations are more likely to benefit from cognitive support systems, such as training, bureaucratization, and technology, that can reduce reliance on heuristics and biased decision making (Weber 1922, Stinchcombe 1990). Moreover, experimental comparisons between individual and group decision making generally find that groups act more rationally (see Charness and Sutter 2012, Kugler et al. 2012, for reviews). Another reason nudging may be less effective when it comes to organizations is that some established interventions, such as social norms, work by increasing the moral cost of a particular (in) action (Hallsworth et al. 2017). Although individuals are acutely sensitive to their personal moral identity (Mazar et al. 2008), groups of individuals tend to experience a sense of diffused responsibility and deindividuation (Bandura 1999). In other words, in groups, there may be a decrease in moral costs.

Likewise, the majority of empirical research on tax compliance has been conducted at the individual level (see Hallsworth 2014, Mascagni 2018, for reviews). Solutions that improve individuals' tax compliance need not necessarily be effective when it comes to organizational actors. For example, although survey evidence suggests that, similar to individuals, firms with higher tax morale evade less than firms with lower tax morale (Alm and McClellan 2012), in one of the few published field experiments involving organizational tax compliance, Ariel (2012) finds a moral suasion intervention to be counterproductive, having no impact on an organization's compliance and even decreasing tax revenues. Similarly, a field experiment focusing on new firms found that supervision by tax administration actually reduced compliance (Gangl et al. 2014). Although organizational tax compliance is a growing area of inquiry (e.g., for recent field experiments see Iyer et al. 2010, Ronconi et al. 2015, Chetty et al. 2014, Pomeranz 2015, Kettle et al. 2016, Brockmeyer et al. 2019) more research is needed to further our understanding of and develop effective solutions for organizations.

In order to address these gaps, we conducted a field experiment testing the effectiveness of one of the most well-established behavioral interventions—planning prompts (Rogers et al. 2015)—on organizations' tax compliance behavior. Planning prompts encourage individuals to develop simple, concrete plans, which clearly articulate the when, the where, and the how of action (Gollwitzer and Sheeran 2006). For example, individuals are significantly more likely to get a flu shot when prompted to select a feasible date and time to receive the shot, as compared with simply being provided information about when and where shots are available (Milkman et al. 2011). Forming precise plans has been shown to increase follow-through because doing so increases one's commitment to their goal. Moreover, forming specific plans can help individuals get started on their goals, overcome procrastination, and counteract forgetfulness by reminding them of their goals at key points in time (Rogers et al. 2015). Planning prompts can impact behavior in both the short and the longer term, as they detail specifically what should happen at a given point in time (Gollwitzer 1999). This approach has previously been shown to be effective for encouraging a number of individual behaviors, including completion of course projects (Gollwitzer and Brandstätter 1997), voter turnout (Nickerson and Rogers 2010), debt repayment (Mazar et al. 2018), and cancer-screening (Milkman et al. 2013). To the knowledge of the authors, however, it has not yet been applied to organizational behavior nor to tax compliance.

That said, there is reason to believe that planning prompts may be effective in a tax setting. Firstly, planning prompts are especially effective when targeting

goals tied to personal values (Koestner et al. 2002). Research on tax morale illustrates the importance of intrinsic values on tax compliance (e.g., Torgler 2002, Wenzel 2005, Frey and Torgler 2007, Kirchler 2007, Torgler 2007, Kirchler et al. 2008, Alm and Torgler 2011, OECD 2013). Second, planning prompts share some similar features with the U.K. Behavioural Insights Team's EAST framework (easy, attractive, social, and timely) for encouraging behavior change, which have been shown to improve *individuals'* tax compliance (Behavioural Insights Team 2014). Additionally, a personal finance reference book recommends that individuals form specific and detailed plans to combat procrastination when filing and paying taxes (Spann 2009). Given that a number of organizations may be late filing taxes because they fail to act on their existing motivations, we predicted that planning prompts would increase their tax compliance behaviors. In contrast, if liquidity constraints alone explained late filing, our changes should not affect organizations' compliance.

Finally, despite the importance of understanding the temporal dimensions of behavioral interventions, relatively few field experiments can speak to longer-term effects because most administer and measure an intervention's effect only once (Frey and Rogers 2014, Mascagni 2018). Given that our field experiment features two waves, we are able to examine the effects of repeated exposure to our intervention. One important temporal aspect is whether an intervention's effects persist once stopped. Several factors can increase the likelihood that a behavioral intervention would have a persistent impact including the formation of habits, changing fundamental beliefs, social changes that reinforce the behavior, or a change in the repeated future costs of a behavior (Rogers and Frey 2015). Given that our intervention was not expected to operate by any of these mechanisms, we predicted that the effects would not persist once the intervention was removed. Another important temporal aspect to consider is habituation, that is, whether repeated exposure to the same intervention alters its effectiveness (Rogers and Frey 2015). Identifying whether habituation occurs is critical in the realm of tax compliance, as tax payers are liable to be exposed to the same collection process year over year. Although research has shown that people habituate to stimuli with repeated exposures (Thompson and Spencer 1966), it has also been found that habituation can be mitigated both when the interval between exposures is long (Rankin et al. 2009) and when the stimuli only leads to temporary, and nonpermanent, changes in behavior (Rogers and Frey 2015). Given that there is an entire year between exposures in our context and that we expect the initial behavior change to be temporary, we predicted that our intervention would

be equally effective across more than one exposure and we would find no evidence of habituation.

3. Field Experiment

3.1. Research Setting and Sample

To test the effectiveness of planning prompts on tax compliance, we partnered with the Government of Ontario to run a large-scale field experiment involving organizations that failed to file their employer health tax (EHT) return on time. In the jurisdiction-wide experiment, we tested an experimental late notice against the standard late notice used by the government, the latter of which served as our control. In order to establish the replicability of any findings and to examine the longer-term impact of our intervention, including testing for habituation and the persistence of our effects, we conducted a two-wave field experiment in the 2013 and 2014 tax years. In both waves, the control and experimental notices were the same, but random assignment of organizations to each notice was carried out independently both years.

The tax that served as the subject of this field experiment (EHT) is an annual payroll tax on remuneration paid to employees. All organizations with permanent establishments in Ontario and payrolls greater than \$400,000 are subject to EHT, constituting a tax base of some 85,031 organizations in 2013 and 79,912 organizations in 2014. The majority of these organizations (59.6% in 2013 and 61.7% in 2014) paid this tax by installments rather than by annual return, either by election or because they had payrolls greater than \$600,000 and were required to do so. The remaining organizations (34,360 in 2013 and 31,773 in 2014) paid the entirety of this tax by the annual return, which is due to be filed on or before March 15 or the following business day each year. Those organizations that pay by installments are also required to file returns to confirm that their installment payments are correct or to reconcile any differences between the installments and the final amount of tax owed. In both years, all organizations were reminded of their requirement to file their annual return via a letter that contained the necessary paperwork, which was mailed to their business address in mid-January. Options for filing the annual return included by mail, at in-person service centers, and via a website.

Organizations that fail to file an annual return by the due date are subject to financial penalties, interest charges, and collection efforts encouraging timely tax remittance. On the business day following the tax's due date (March 15), unfiled accounts are automatically and immediately charged a penalty of five percent of the amount owed, plus a one percent penalty for each full month delay of payment, as well as a six percent annualized interest charge compounded daily. Despite these costs, 6,406 organizations in 2013 and

6,291 in 2014, or about 7.5% and 7.9% of the subject population, respectively, failed to file a return by the due date. Of these late filers, 6,316 organizations in 2013 and 6,189 in 2014 were included in our analyses as they had not filed their taxes by the estimated delivery date of the late notice, which served as the independent variable of our experiment.

Ten days after the late notice letter was posted other collection efforts commenced, including the posting of a second letter and phone calls by collection agents. These additional efforts were held constant across our entire experimental sample. Data on organizations' filing behavior was extracted over six months after this late notice was sent to organizations. Our focus on timely tax payment behavior has the advantage of a readily operationalizable behavioral metric, as compared with tax evasion, for example, which is usually measured by proxy variables of questionable validity (such as the size of tax declarations and deductions; Slemrod and Weber 2012). For additional details about the research setting and sample please refer to the online appendix.

3.2. Methodology

Organizations were randomly assigned to receive one of two letters: a control notice, which was the standard late notice sent to delinquent organizations in years past; and a revised late notice, which included our experimental manipulation. The decision to include only two letters in the experiment was made collaboratively with our government partners on account of feasibility constraints and statistical power. It was furthermore decided that assignment to letters would be balanced such that in both waves approximately 40% of organizations would receive the control letter and 60% the experimental letter, in the hopes of maximizing the predicted benefit of conducting the intervention while preserving the empirical rigor of the experiment. This over-assignment to the experimental condition also gave us greater statistical power to test for persistence and habituation effects, since organizations were only included in both waves of the experiment if they were sufficiently late filing their returns both years.

Comparisons across multiple variables (see Appendix Table A1) revealed that random assignment successfully created two groups of organizations that were approximately equivalent in terms of the proportion of foreign organizations, years of experience filing this particular tax, and amount of tax owed, as well as in terms of the number of organizations that received the late notice in error because the government mailing list had not been updated to reflect the fact that they had recently filed or that the organization was not required to file a return. In addition, across both waves, domestic organizations took less time to file than foreign organizations after receiving a

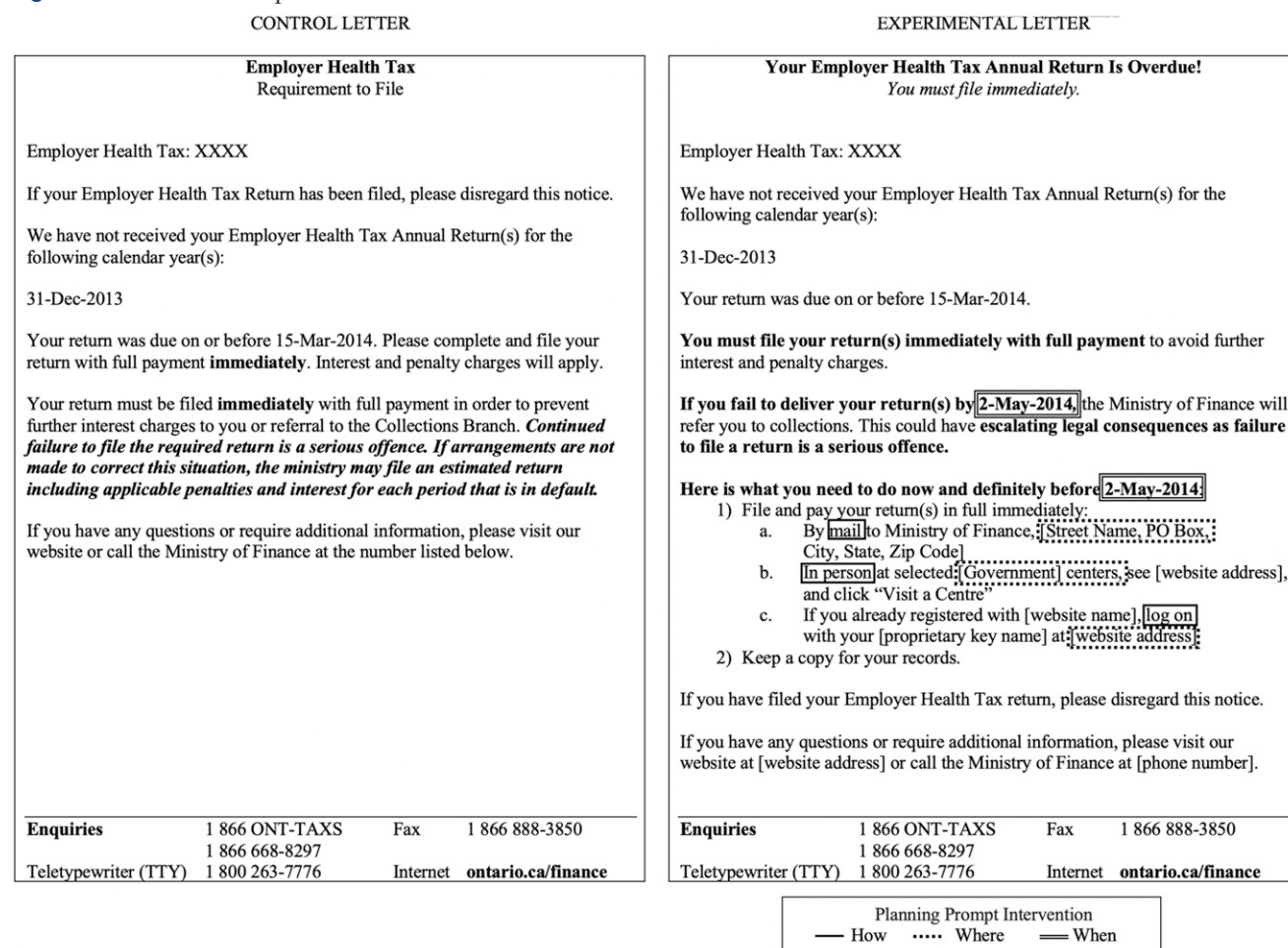
late notice (wave 1: $b = -23.46$, standard error (SE) = 3.81, $t(6,273) = -6.16$, $p < 0.001$; wave 2: $b = -16.86$, SE = 2.58, $t(6,100) = -6.53$, $p < 0.001$). Also, organizations with more years of experience filing this tax took less time to file (wave 1: $b = -0.46$, SE = 0.10, $t(6,274) = -4.61$, $p < 0.001$; wave 2: $b = -0.12$, SE = 0.07, $t(6,100) = -1.91$, $p = 0.056$). Finally, in wave 2, we found that organizations with larger payrolls took less time to file after receiving a late notice ($b = -0.15$, SE = 0.00, $t(6,100) = -3.73$, $p < 0.001$). However, due to random assignment, these variables cannot account for our findings reported below, do not affect our results when controlled for, and do not significantly interact with our intervention. See Appendix Table A2 for a correlation matrix displaying these relationships.

As can be seen in Figure 1, the major differences between the control and the experimental letters were (1) the reorganization of available information into explicit step-by-step instructions of *how* and *where* to file a return, and (2) the addition of a deadline (*when*), thus furnishing organizations with the three critical pieces of information contained within planning prompts (*when*, *where*, and *how*; Gollwitzer et al. 2004, Rogers et al. 2015).

Although many empirical applications of planning prompts ask individuals to form their own goals, we provided the specific when, where, and how plan for the organizations. It is important to note that the vast majority of organizations (97% in wave 1 and 99.9% in wave 2) had previous experience filing EHT returns and, thus, should have already known when (i.e., as soon as possible to accumulate the least amount of penalty), how, and where to file a return. Moreover, this information was sent to the organizations in their initial requirement to file letter and is readily available on the ministry's website. Therefore, although the information in the experimental letter was likely not new to organizations, we hoped to prompt recipients to act by reorganizing it into a simple, actionable plan. Previous research has shown that rearranging letters in such a way can be very useful for drawing attention to the focal information for firms to attend to (Thaler and Sunstein 2008, Dulleck et al. 2015).

Both the control and experimental letters informed organizations that their tax return had not been received by the March 15 due date, that the return must be filed immediately to prevent further charges and referral to the collections branch, and that

Figure 1. Control and Experimental Letters



continued failure to file is a serious offence. The action-deadline (*when*) provided in the experimental letters was May 2 for the 2013 tax year and May 1 for 2014, which was 10 days after the late notice was posted. Deadlines are an effective tool for overcoming procrastination (Ariely and Wertenbroch 2002), and this specific deadline was chosen because it both marked the initiation of additional collection efforts across conditions and it could be viewed as a more concrete goal than asking someone to pay “immediately” (concrete goals generally produce better performance than abstract goals; Locke and Latham 2002). Additional stylistic differences between the control and experimental letters (second-person pronouns, active voice, and reorganized letter structure) supported the planning prompt intervention, without largely changing the meaning or reading difficulty¹ of the two letters. Therefore, we argue that this experiment provides an imperfectly controlled, but informative test of the longer-term usefulness of planning prompts in expediting organizations’ payment of overdue taxes. For a short survey experiment examining individuals’ perceptions of the two late notices please refer to the online appendix.

3.3. Results of Wave 1

The results obtained from wave 1 of the experiment supported our prediction: Late notices that included planning prompts increased organizations’ tax compliance behavior. As can be seen in Table 1, compared with the control letter, the experimental letter significantly reduced the time it took organizations to file their annual returns by an average of 4.4 days ($b = -4.38$, $SE = 1.66$, $t(6,273) = -2.65$, $p = 0.008$, $d = -0.07$).^{2,3} These results suggest that procrastination and forgetfulness after receiving the late notice may, in fact, explain some of organizations’ late filing behavior. Also consistent with our prediction that including a concrete deadline date (rather than the abstract injunction “immediately”) would expedite filing behavior, a multinomial regression analysis—with a reference category of organizations that took more than 30 days after the late notice to file their annual returns—revealed that the experimental letter significantly increased the likelihood of filing before the 10-day deadline contained in the late notice ($b = 0.44$, $SE = 0.07$, $Wald = 37.16$, $p < 0.001$, odds ratio (OR) = 1.54, 95% confidence interval (CI) [1.34–1.78]). In contrast, during each of the two 10-day periods afterward, the experimental letter did not lead to any significant differences in the likelihood of filing (1–10 days after the deadline: $b = 0.12$, $SE = 0.07$, $Wald = 2.90$, $p = 0.089$, OR = 1.13, 95% CI [0.98–1.29]; 11–20 days after the deadline: $b = -0.12$, $SE = 0.07$, $Wald = 0.03$, $p = 0.865$, OR = 0.99, 95% CI [0.86–1.14]).

Table 1. Main Results from Waves 1 and 2

Variable	2013			2014		
	Control letter ($n = 2,515$)	Experimental letter ($n = 3,801$)	Statistic <i>Dependent variables</i>	Control letter ($n = 2,471$)	Experimental letter ($n = 3,718$)	Statistic <i>Statistic</i>
Mean days taken to file after first late notice	51.30 (66.19)	46.92 (62.87)	$t(5,150)^a = -2.62$	44.51 (43.54)	39.26 (41.89)	$t(5,075)^a = -4.72$
Returns filed within 10-day deadline of first late notice	406 (16.1%)	855 (22.5%)	$\chi^2(1) = 38.21$	511 (20.7%)	1,072 (28.8%)	$\chi^2(1) = 51.83$
Mean tax paid before collection deadline ^c	\$149.88 (\$857)	\$241.32 (\$1,719)	$t(4,918)^a = 2.54$	\$189.75 (\$1,151)	\$291.70 (\$1,432)	$t(4,905)^a = 2.80$

Notes. Numbers in brackets are either percentages, standard deviations, or degrees of freedom.

^aWhere homogeneity of variances was violated, Welch-Satterthwaite adjustments to the degrees of freedom were made, although in no cases did this alter the significance of the results.

^bLogarithmic transformations intended to address significant positive skew did not alter the significance of these results. Corresponding medians and nonparametric tests produce consistent results (see Appendix Table A1).

^cMean tax paid before collection deadline^c was calculated for organizations that did not claim overpayment refunds as those who owed no taxes.

The fact that those organizations that received the experimental letter did not file any sooner after the deadline than the control group is likely caused by the start of additional collection efforts, which targeted all organizations remaining in default. Nevertheless, although these additional collection efforts gradually decreased the difference between the control and the experimental letters, as can be seen from Figure 2, it took over 10 weeks before the initial impact of the experimental letter was matched by sustained, human-resource-intensive collection efforts.

To ascertain the financial value of our experimental letter in wave 1, we determined the additional amount remitted to the government by the deadline set in the late notices, at which time additional collection efforts began. These additional collection efforts, including subsequent letters, staffing of call centers, and procurement of external collection agencies, are estimated by the government to cost approximately two cents for every dollar collected. We find that, on average, each control letter resulted in \$149.88 of remitted taxes prior to the start of additional collection efforts, whereas each experimental letter garnered \$241.32 ($b = 91.45$, $SE = 40.97$, $t(5,197) = 2.32$, $p = 0.026$, $d = 0.07$).^{4,5} Although the size of the experimental letter's effect is statistically small, in the context of this tax's scale, its impact is not insignificant. This increase of 61% in tax remittance resulted in an additional \$288,335.54 being collected within

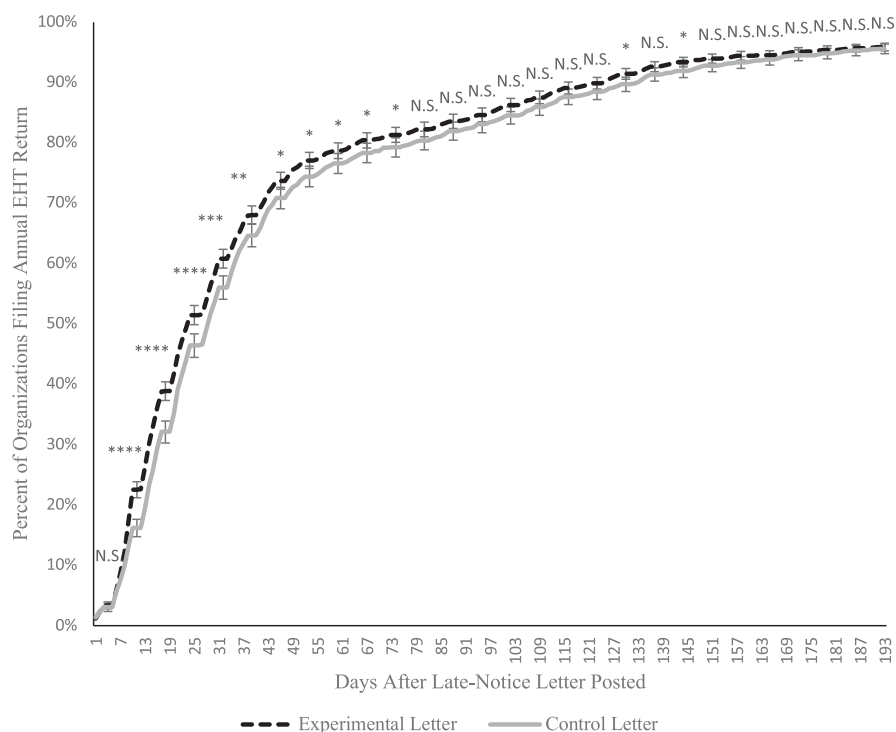
10 days of mailing the experimental letter, plus saved the government approximately \$5,766.71 in collection costs during wave 1 of the experiment. Had the experimental letter been sent to all 5,199 organizations that were not owed a tax refund and were late filing their annual return during wave 1, the government could have expected to receive an additional \$475,438.15 in remittance prior to initiating additional collection measures, plus saved some \$9,508.76 in collection costs.

To further validate our results by replicating them in the same population and to test for the longer-term effects of our intervention, we conducted a second wave of the same experiment for the subsequent (2014) tax year. Because random assignment in both waves was carried out independently, this two-wave design enabled us to replicate our initial findings and, furthermore, to examine both whether the effect of our letter had any year-over-year persistence, and whether the effectiveness of the experimental letter remained consistent after repeated exposure or whether organizations habituated to these changes.

3.4. Results of Wave 2

For the 2014 tax year, 79,912 organizations were due to file the payroll tax, which served as the setting for our study; 6,189 of these organizations were sufficiently late filing their taxes that year that they were included in the second wave of our experiment.

Figure 2. Time Series of Organizations in Wave 1 Filing Their 2013 Annual Returns by Letter Assignment



Notes. Error bars represent 95% confidence intervals calculated each weekend when there was little to no processing of returns. N.S., $p \geq 0.05$; $*p < 0.05$; $**p < 0.01$; $***p < 0.001$.

When compared with their previously punctual peers, organizations that were late filers in 2013, and therefore were involved in wave 1 of our field experiment (i.e., they received either the control or experimental late notice letter), were significantly more likely to also be late filing their 2014 tax return ($\chi^2(1, n = 79,912) = 2,012.31, p < 0.001$, OR = 3.43, 95% CI [3.24–3.63]) and to be so late as to receive a 2014 late notice letter, thereby including them in wave 2 of our field experiment ($\chi^2(1, n = 79,912) = 3,402.02, p < 0.001$, OR = 5.94, 95% CI [5.56–6.35]). Thus, of the 6,189 organizations involved in wave 2 of our field experiment, approximately one quarter of them ($n = 1,506$) had also participated in wave 1.

To test for persistence, we examined whether the 2013 (i.e., wave 1) control and experimental letters had any differential effect on the likelihood that an organization would miss the subsequent, 2014 tax deadline. As predicted, we found no evidence that the effect of our intervention persisted into the subsequent tax year, such that the letters had no influence on the likelihood an organization would miss the tax deadline ($\chi^2(1, n = 5,293) = 0.00, p = 0.991$, OR = 1.00, 95% CI [0.90–1.12]) nor its likelihood of being included in wave 2 ($\chi^2(1, n = 5,293) = 0.07, p = 0.793$, OR = 1.02, 95% CI [0.90–1.15]). Therefore, the proportion of organizations in wave 2 that had received the control ($n = 595$) and experimental letters ($n = 911$) in Wave 1 remained relatively unchanged from the original 40/60 split. Importantly, however, this lack of persistence also indicates that wave 2 of our field experiment is free of selection bias that might have been produced by condition assignment in wave 1. Moreover, independent random assignment in both waves, such that there is no association between which letter (control vs. experimental) an organization received in 2013 and 2014 ($\chi^2(1, n = 1,506) = 0.38, p = 0.538$, OR = 0.94, 95% CI [0.76–1.16]) provides a relatively unconfounded and well-powered test of whether the effect of our experimental letter is moderated by repeat exposure (i.e., whether there are any habituation effects), which we discuss below. Finally, the lack of persistence of our effects, together with the fact that our intervention was effective among organizations with previous filing experience, suggests that timing is critical and planning prompts must be provided at key points during the decision-making process to be maximally effective.

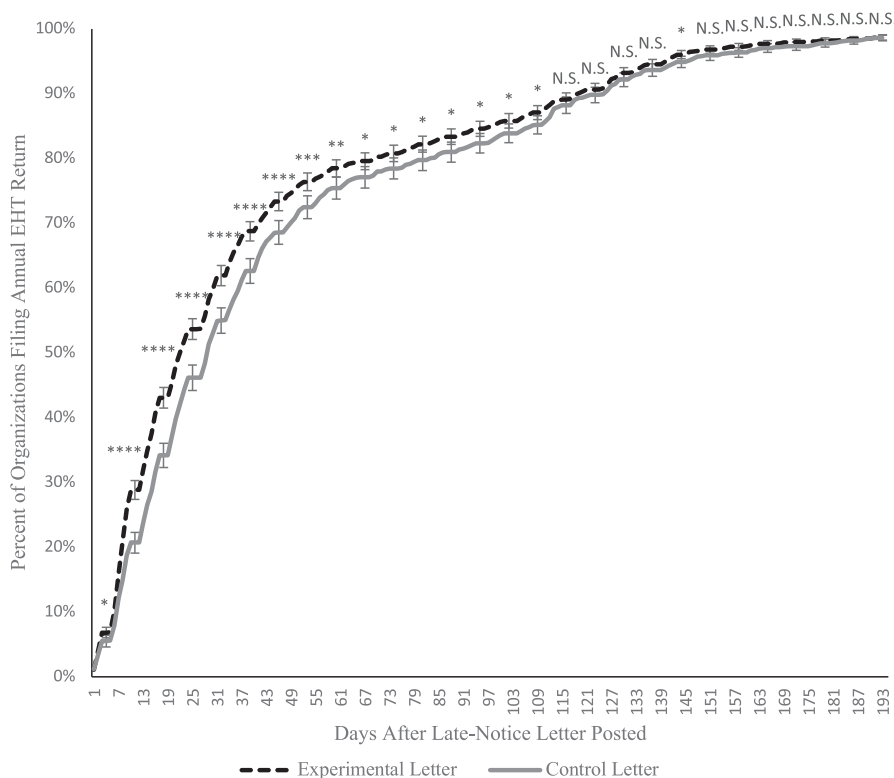
Replicating our previous results, the experimental letter in wave 2 significantly reduced the time it took organizations to file their annual returns by an average of 5.3 days ($b = -5.25$, SE = 1.11, $t(6,100) = -4.72, p < 0.001$, $d = -0.12$).^{6,7} Using a multinomial regression we observed the same pattern of accelerated filing behavior prior to the 10-day deadline contained in the letter ($b = 0.49$, SE = 0.07, Wald = 54.12, $p < 0.001$,

OR = 1.63, 95% CI [1.43–1.86]). Similarly, reduced differences were observed in filing behavior over the two 10-day periods immediately afterward (1–10 days after the deadline: $b = 0.17$, SE = 0.08, Wald = 4.91, $p = 0.027$, OR = 1.18, 95% CI [1.02–1.37]; 11–20 days after the deadline: $b = 0.06$, SE = 0.08, Wald = 0.58, $p = 0.447$, OR = 1.06, 95% CI [0.91–1.23]). Again, although additional collection efforts commencing after the deadline contained in the letter gradually eroded the effect of the experimental letter, as can be seen in Figure 3, more than 16 weeks passed before the initial impact of the experimental letter was no longer significant relative to the control.

In addition to replicating our original results and testing whether the effects persisted over two years, our two-wave design also enabled us to test whether the effect of the experimental letter diminished with repeated exposure (habituation). To explore this possibility, we submitted our results to an ordinary least squares regression in which the number of days taken by organizations to file their 2014 annual tax return was regressed upon wave 2 letter assignment (coded 0 = control; 1 = experimental letter), whether the organization had previously received the experimental letter in wave 1 (coded 0 = had never received the experimental letter previously, including both those organizations that had been in the control condition in 2013 and those that had not received any late notice letter in 2013; 1 = had received the experimental letter in wave 1), and their interaction term. As predicted, this analysis revealed significant effects for receiving the experimental letter in wave 2 ($b = -4.65$, SE = 1.21, $t(6,098) = -3.85, p < 0.001$, $d = -0.10$) and receiving the experimental letter the previous year ($b = 9.92$, SE = 2.39, $t(6,098) = 4.15, p < 0.001$, $d = 0.11$), but the interaction term was not significant ($b = -4.14$, SE = 3.07, $t(6,098) = -1.35, p = 0.178$, $d = -0.03$). This nonsignificant interaction suggests that repeated exposure to the experimental letter had no impact on its effectiveness, and if anything, there is a trend toward the experimental letter being more effective for those organizations previously exposed to it.

The negative effect of exposure to the 2013 experimental letter on 2014 filing behavior can likely be explained by the fact that those organizations that received a collection letter in wave 1, due to their tardiness that year, were significantly slower filing their returns in response to both late notices in wave 2 than were those that had avoided the collection process in the previous year ($b = 7.98$, SE = 1.27, $t(6,100) = 6.30, p < 0.001$, $d = 0.16$). Indeed, as can be seen in Figure 4, when controlling for those organizations that were late filing in wave 1, this negative effect of the wave 1 experimental letter on 2014 filings disappears, whereas the positive effect of the 2014 letter remains unchanged (see regression model 2 in

Figure 3. Time Series of Organizations in Wave 2 Filing Their 2014 Annual Return by Letter Assignment

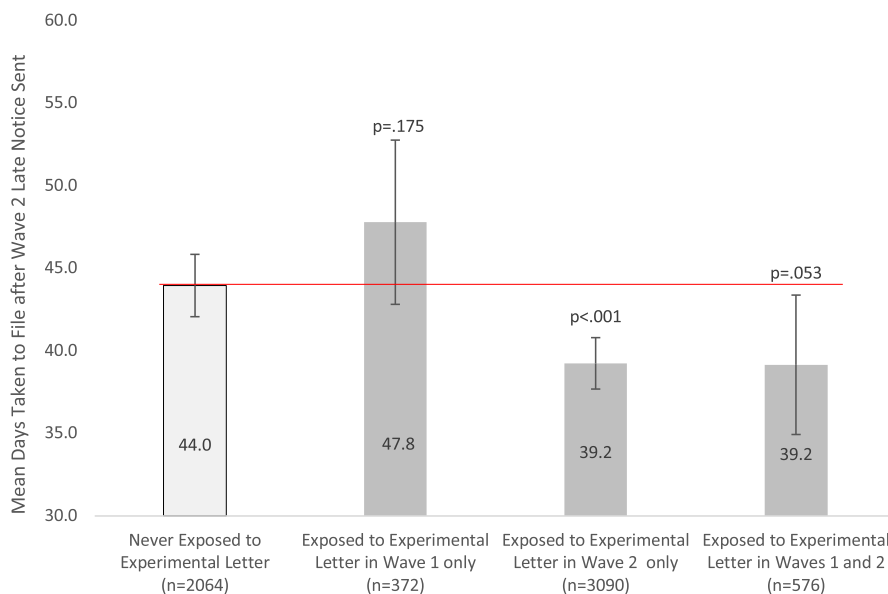


Notes. Error bars represent 95% confidence intervals calculated each weekend when there was little to no processing of returns. N.S., $p \geq 0.05$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Appendix Table A3). As well, when an organization's on-time collection status in wave 1 is interacted with their wave 2 experimental letter assignment in a similar regression model, there is no evidence that the

main effect of the experimental letter ($b = -5.92$, $SE = 1.27$, $t(6,098) = -4.65$, $p < 0.001$, $d = 0.12$) is moderated by an organization's previous tardiness ($b = 2.48$, $SE = 2.58$, $t(6,098) = 0.96$, $p = 0.337$, $d = 0.02$), suggesting that

Figure 4. (Color online) Effect of Letter Assignment on Time Taken for Organizations to File Their Annual Return in Wave 2



Notes. Results control for having received a late notice the previous tax year. Error bars represent 95% confidence intervals. The p -values are relative to those organizations never exposed to the experimental letter in Waves 1 or 2 (i.e., leftmost bar).

the experimental letter is equally effective for both those organizations that were not late the previous year and those that were late two years in a row (see model 3 in Appendix Table A3). Unfortunately, determining whether the experimental letter is equally effective for organizations who are habitually late filing their taxes (i.e., for more than two consecutive years) goes beyond our data. Note, however, that that would constitute a relatively small portion of the population, considering that only a quarter of organizations in our wave 2 sample were also tardy in wave 1. That said, we suspect that serially late filers may experience unique barriers to compliance (e.g., inability to pay, lack of intrinsic motivation to pay, dysfunctional organizational structures) that would likely not be impacted by our intervention.

As a result of increasing tax-filing compliance, the experimental letter also increased the amount of taxes remitted during wave 2 prior to the commencement of additional collection efforts. We find that on average each control letter resulted in \$189.75 of remitted taxes prior to the start of additional collection efforts, whereas each experimental letter garnered \$291.70, a significant difference ($b = 101.95$, $SE = 38.00$, $t(5,066) = 2.68$, $p = 0.007$, $d = 0.08$).^{8,9} Again, although the experimental letter's statistical effect size is small, the impact is not insignificant. The experimental letter led to an increase of 54% in tax remittance, resulting in an additional \$309,112.40 being collected within 10 days, and saved the government approximately \$6,182.25 in collection costs during wave 2 of the experiment. Had the experimental letter been sent to all 5,066 organizations that were not owed a tax refund and were late filing their annual return, the government could have expected to receive an additional \$516,478.70 in remittance prior to initiating additional collection measures and saved some \$10,329.57 in collection costs.

4. Discussion

This study demonstrates that psychologically informed, low-cost changes to tax-collection letters can have significant benefits for taxation. Specifically, we examined tax-paying behavior among a population of organizations that were late filing their annual returns and found that planning prompts are a viable tool for overcoming payment procrastination. Our multi-wave experiment makes several valuable contributions. First, it applies a new theoretical perspective to a relatively understudied component of tax compliance. Drawing on the insight that people often fail to act on their motivations, we argue that overdue tax payments might not only result from the lack of sufficient motivation or ability to pay, but also from the absence of a concrete plan of how to act on those motivations. Importantly, providing a plan with an explicit deadline and

specific instructions for its implementation appears to encourage organizational taxpayers to act and aids in overcoming barriers of procrastination or forgetfulness after receiving the late notice.

Due to the feasibility constraints of our government partner and statistical power considerations, our experiment was limited to testing only two versions of the collection letter. As such, we cannot disentangle the individual contributions of all differences between the letters. This limitation is further compounded by the fact that we do not have any mediating variables that would enable us to measure the psychological variables we argue underlie our effects. However, given the substantive similarity between the content of the control and the experimental letters, we contend that the most impactful difference is the provision of concrete instructions of when, where, and how to act on extant motivations to file an organization's taxes. The additional changes to our experimental letter were highly stylistic in nature (second-person pronouns, active voice, and reorganized letter structure) and only intended to highlight and support our planning prompt manipulation. Given organizations' previous experience with filing their EHT returns (over 97% had previously filed), that they were each sent an initial tax assessment with instructions to file, and that all details of how, when, and where to file are readily available on the government's web page, it seems unlikely that our results were simply driven by the inclusion of any specific information in the letter.

To examine the impact of the changes in our letters more directly, we ran a short follow-up survey experiment asking student participants ($n = 167$) to rate their perceptions of the late notices on several dimensions (including the clarity, urgency, seriousness of punishment, and effectiveness of the letters) and to choose the areas of the letters that they thought would be most effective for encouraging organizations to file their taxes (see the online appendix for detailed methods and results). The results of this experiment revealed that participants did not perceive the letters as significantly different on any of the aforementioned characteristics (see Appendix Table A4). Moreover, when participants were asked to select the areas of the experimental late notice that would be most effective for encouraging organizations to file, the regions selected are also present in the control letter (see Appendix Figure A1). Neither the planning prompts nor the deadline date were selected more often than would be predicted by chance. These findings lend support to our argument that the information itself was not likely causing our effects, rather we suspect that organizing that information into planning prompts was critical for getting organizations to file. Nevertheless, future research may

examine these different elements individually to determine how critical each is in promoting action.

Second, in addition to the application of a well-established behavioral insight to the novel intersection of tax compliance and organizational behavior, by employing a multi-wave design, this experiment also makes a unique contribution by examining the persistence and consistency of our effects over time. Our study found no evidence that receiving the experimental letter impacted organizations' likelihood of filing in a timely fashion the following year, demonstrating the importance of timing for behavioral interventions. Previous research has shown that behaviorally informed communications can persist and impact tax behavior as much as a year later (Brockmeyer et al. 2019). Moreover, research has found that implementation intentions, more broadly, can impact behavior even when there is a long interval between the time of the intervention and ultimately enacting the desired behavior (e.g., Nickerson and Rogers 2010). Although we would not expect specific planning prompts to impact behavior a full year later (i.e., after the intervention was removed) given the mechanisms through which they operate, it was important to confirm this prediction empirically. Therefore, our finding suggests that planning prompts may be most effective when they proximately precede a decision point.

In addition, although our intervention did not appear to have effects that persisted across tax years, organizations also did not habituate to our manipulation and its effects were consistent across repeated exposures. The fact that our experimental letter was equally effective after two exposures is critically important when assessing the policy implications and longer-term impact of such a behavioral intervention, especially in the area of tax compliance where organizations typically repeat the same process year after year. We suspect that the year between exposures allows for sufficient memory decay to prevent habituation. In addition, in this type of context, letters are only sent to organizations who are late filing their taxes, and therefore, the vast majority would be exposed to this letter at intervals that were even longer than one year (e.g., during our experiment, only 25% of firms who were late in wave 2 were also late the previous year). We also predict that this lack of habituation is especially likely in the context of organizations, where there are often procedural changes and staff turnover that would cause the letter to be new to the receiver each year. However, if notices were sent more often, for example, monthly, we would predict that organizations' behavior eventually may habituate to the changes in the letter, and as a consequence, we would observe decreased attention

toward our intervention. It would therefore be useful for future research to confirm that the effectiveness of providing organizations with planning prompts remains effective after numerous exposures. Moreover, if the actual process of filing taxes in Ontario changes significantly (e.g., tax forms are automatically filed and paid in installments by all organizations), then there would be less possibility for late payments and our intervention would be less useful. Although our findings contribute to the understanding of the longer-term effectiveness of planning prompt interventions, we would like to encourage more longitudinal research in this area, to further our understanding of the temporal duration of behavioral interventions.

Third, this experiment is among the first to test the application of behavioral interventions designed for individuals in the realm of organizational behavior. In this instance, we find that despite documented differences in organizational decision making, planning prompts that have been shown to be effective for individuals are also very effective for organizations. We find that organizations presented with our experimental letter were significantly more likely to file their taxes before a set deadline, and the effect of the letter remained significant for up to 16 weeks despite sustained collection efforts beginning within just 10 days. Critically, for our government partners, early remittance significantly saves on additional collection costs (as well as costs to the tax-paying organizations). Future research could explore whether other behavioral interventions shown to affect individual behavior are equally effective at the organizational level and under what conditions.

In addition, although domestic (vs. foreign) firms, more experienced firms, and firms with larger payrolls (a proxy for firm size) were found to file their taxes faster after receiving a late notice, we found that the effectiveness of our intervention was not moderated by these factors. This lack of moderation suggests that our effects are likely not driven only by weakness in organizational procedure or by the organizations' inexperience, nor are the effects driven only by the organizations' likelihood of benefiting from social goods supported by taxes. Rather, these findings lend confidence to the idea that planning prompts are more broadly applicable and that our intervention could be effective across a number of tax-filing processes. Although we were unable to acquire additional firm-level data from our sample, we do acknowledge that it would be ideal to have more information about organizations' tax-filing processes. Future research may therefore examine how other organizational factors—such as the mechanical process by which organizations file their taxes, the size

of their accounting department, or organizational culture—can interact with behavioral interventions, like the one studied here.

Finally, not only did this experiment produce behavioral field data, which contributes to the growing literature on organizational tax compliance, it also advanced the goals of the government and the public it serves: it saved the government money in terms of collection costs, saved organizations money in terms of additional penalties and interest, and increased collected taxes. Given that implementing this “nudge” involved zero marginal costs above those incurred for the experiment (i.e., the salaries of research staff) and that we found no evidence that repeated exposure to the letter diminished its effect, we are encouraged to know that the government is implementing the experimental letter as the standard. Our experiment provides strong evidence for a readily applicable tool that promotes efficient tax collection without added enforcement or penalties.

Acknowledgments

The authors thank Elizabeth Hardy, the Ontario Ministry of Finance, and Ontario Treasury Board Secretariat's Behavioural Insights Unit for the collaboration, the Behavioural Economics in Action at Rotman (BEAR) center for support along the road, and Alyssa Whalen and Donella Plastik for their research assistant support. Nicole Robitaille and Julian House contributed equally to this work.

Endnotes

¹ The Flesch-Kincaid grade level test was used to measure the reading difficulty of the letters (Kincaid et al. 1975). Scores on this test represent a U.S. grade level. The specific reading grade levels were 9.7 for the control and 9.3 for the experimental letter, excluding the specific planning prompt paragraph.

² The 41 organizations that never filed their returns within the 559 observation days available for wave 1 of the experiment are excluded from this analysis. Imputing a filing time of 560 days for these organizations does not substantively change the result.

³ To correct for significant positive skewness in the number of days taken to file a tax return, $z_{\text{skewness}} = 104.55$, a log transformation was performed, which did not alter the significance of our results. A nonparametric Mann-Whitney test produced consistent results, $U = 4,367,740$, $p < 0.001$.

⁴ During the year 2013, there were 1,117 organizations that had overpaid their taxes by installments during the year 2013 and, thus, were owed refunds with their annual return. These organizations are excluded from these analyses.

⁵ To correct for significant positive skewness in the amount of tax collected, $z_{\text{skewness}} = 661.35$, a log transformation was performed, which did not alter the significance of our results. A nonparametric Mann-Whitney test produced consistent results, $U = 3,112,908$, $p < 0.001$.

⁶ The 87 organizations that never filed their returns within the 195 observation days available for wave 2 of the experiment are excluded from this analysis. Imputing a filing time of 196 days for these organizations does not substantively change the result.

⁷ To correct for significant positive skewness in the number of days taken to file a tax return, $z_{\text{skewness}} = 46.90$, a log transformation was performed, which did not alter the significance of our results. A

nonparametric Mann-Whitney test produced consistent results, $U = 4,009,273$, $p < 0.001$.

⁸ During the year 2014, there were 1,122 organizations that had overpaid their taxes by installments and, thus, were owed refunds with their annual return. These organizations are excluded from these analyses.

⁹ To correct for significant positive skewness in the amount of tax collected, $z_{\text{skewness}} = 229.74$, a log transformation was performed, which did not substantively alter our results. A nonparametric Mann-Whitney test produced consistent results, $U = 2,941,784$, $p < 0.001$.

References

- Allingham MG, Sandmo A (1972) Income tax evasion: A theoretical analysis. *J. Public Econom.* 1(3–4):323–338.
- Alm J, McClellan C (2012) Tax morale and tax compliance from the firm's perspective. *Kyklos* 65(1):1–17.
- Alm J, Torgler B (2006) Culture differences and tax morale in the United States and in Europe. *J. Econom. Psych.* 27(2):224–246.
- Alm J, Torgler B (2011) Do ethics matter? Tax compliance and morality. *J. Bus. Ethics* 101(4):635–651.
- Alm J, Bloomquist KM, McKee M (2017) When you know your neighbour pays taxes: Information, peer effects and tax compliance. *Fiscal Stud.* 38(4):587–613.
- Andreoni J, Erard B, Feinstein J (1998) Tax compliance. *J. Econom. Literature* 36(2):818–860.
- Arcos Holzinger L, Biddle N (2016) Behavioural insights of tax compliance: An overview of recent conceptual and empirical approaches. Working paper, Behavioral Economics and Research Team, Canberra, Australia.
- Ariel B (2012) Deterrence and moral persuasion effects on corporate tax compliance: Findings from a randomized controlled trial. *Criminology* 50(1):27–69.
- Ariely D, Wertenbroch K (2002) Procrastination, deadlines, and performance: Self-control by precommitment. *Psych. Sci.* 13(3): 219–224.
- Bandura A (1999) Moral disengagement in the perpetration of inhumanities. *Perspect. Soc. Psych. Rev.* 3(3):193–209.
- Behavioural Insights Team (2014) EAST: Four simple ways to apply behavioural insights. Report, Cabinet Office, London.
- Benartzi S, Beshears J, Milkman KL, Sunstein CR, Thaler RH, Shankar M, Tucker-Ray W, Congdon WJ, Galing S (2017) Should governments invest more in nudging? *Psych. Sci.* 28(8): 1041–1055.
- Blumenthal M, Christian C, Slemrod J (2001) Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota. *Natl. Tax J.* 54(1):125–136.
- Brockmeyer A, Smith S, Hernandez M, Kettle S (2019) Casting a wider tax net: Experimental evidence from Costa Rica. *Amer. Econom. J. Econom. Policy* 11(3):55–87.
- Charness G, Sutter M (2012) Groups make better self-interested decisions. *J. Econom. Perspect.* 26(3):157–176.
- Chetty R, Mobarak M, Singhal M (2014) Increasing tax compliance through social recognition. Report, International Growth Centre, London.
- Cummings RG, Martinez-Vazquez J, McKee M, Torgler B (2009) Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. *J. Econom. Behav. Organ.* 70(3): 447–457.
- Dulleck U, Koessler AK, Schaffner M, Torgler B (2015) Improving communication with directors of firms in liquidation. Report, ASIC, Brisbane, Queensland, Australia.
- Eden D (2017) Field experiments in organizations. *Annu. Rev. Organ. Psych. Behav.* 4(1):91–122.
- Frey BS, Torgler B (2007) Tax morale and conditional cooperation. *J. Comparative Econom.* 35(1):136–159.

- Frey E, Rogers T (2014) Persistence: How treatment effects persist after interventions stop. *Policy Insights Behav. Brain Sci.* 1(1): 172–179.
- Gangl K, Torgler B, Kirchler E, Hofmann E (2014) Effects of supervision on tax compliance: Evidence from a field experiment in Austria. *Econom. Lett.* 123(3):378–382.
- Goldstein NJ, Cialdini RB, Griskevicius V (2008) A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *J. Consumer Res.* 35(3):472–482.
- Gollwitzer PM (1999) Implementation intentions: Strong effects of simple plans. *Amer. Psych.* 54(7):493–503.
- Gollwitzer PM, Brandstätter V (1997) Implementation intentions and effective goal pursuit. *J. Perspect. Soc. Psych.* 73(1):186–199.
- Gollwitzer PM, Sheeran P (2006) Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Adv. Experiment. Soc. Psych.* 38(1):69–119.
- Gollwitzer PM, Kentaro F, Oettingen G (2004) Planning and the implementation of goals. Baumeister RF, Vohs KD, eds. *Handbook of Self-Regulation: Research, Theory and Applications* (Guilford Press, New York), 211–228.
- Government of Ontario (2016) 2016 Ontario budget: Jobs for today and tomorrow. Report, Ministry of Finance, Ottawa, Ontario, Canada.
- Hallsworth M (2014) The use of field experiments to increase tax compliance. *Oxford Rev. Econom. Policy* 30(4):658–679.
- Hallsworth M, List JA, Metcalfe RD, Vlaev I (2017) The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *J. Public Econom.* 148(4):14–31.
- Hauser OP, Linos E, Rogers T (2017) Innovation with field experiments: Studying organizational behaviors in actual organizations. *Res. Organ. Behav.* 37:185–198.
- Iyer G, Recjers P, Sanders D (2010) Increasing tax compliance in Washington State: A field experiment. *Natl. Tax J.* 63(1):7–32.
- Johnson EJ, Goldstein D (2003) Do defaults save lives? *Science* 302(5649):1338–1339.
- Kettle S, Hernandez M, Ruda S, Sanders M (2016) Behavioural interventions in tax compliance: Evidence from Guatemala. Working paper, The Behavioural Insights Team, London, UK.
- Kincaid JP, Fishburne RP, Rogers RL, Chissom BS (1975) Derivation of new readability formulas (automated readability index, fog count and Flesch reading ease formula) for navy enlisted personnel. Report, Institute for Simulation and Training, Millington, TN.
- Kirchler E (2007) *The Economic Psychology of Tax Behaviour* (Cambridge University Press, New York).
- Kirchler E, Hoelzl E, Wahl I (2008) Enforced vs. voluntary tax compliance: The “slippery slope” framework. *J. Econom. Psych.* 29(2):210–225.
- Koestner R, Lekes N, Powers TA (2002) Attaining personal goals: Self-concordance plus implementation intentions equals success. *J. Perspect. Soc. Psych.* 83(1):231–244.
- Kugler T, Kausel EE, Kocher MG (2012) Are groups more rational than individuals? A review of interactive decision making in groups. *Wiley Interdisciplinary Rev. Cognitive Sci.* 3(4):471–482.
- Locke EA, Latham GP (2002) Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *Amer. Psych.* 57(9):705–717.
- Luttmer EFP, Singhal M (2014) Tax morale. *J. Econom. Perspect.* 28(4):149–168.
- Mascagni G (2018) From the laboratory to the field: A review of tax experiments. *J. Econom. Survey* 32(2):273–301.
- Matsaganis M, Leventi C, Flevotomou M (2012) The crisis and tax evasion in Greece: What are the distributional implications? *CESifo Forum* 13(2):26–32.
- Mazar N, Amir O, Ariely D (2008) The dishonesty of honest people: A theory of self-concept maintenance. *J. Marketing Res.* 45(6):633–644.
- Mazar N, Mochon D, Ariely D (2018) If you are going to pay within the next 24 hours, press 1: Automatic planning prompt reduces credit card delinquency. *J. Consumer Psych.* 28(3):466–476.
- Milkman KL, Beshears J, Choi JJ, Laibson D, Madrian BC (2011) Using implementation intentions prompts to enhance influenza vaccination rates. *Proc. Natl. Acad. Sci. USA* 108(26): 10415–10420.
- Milkman KL, Beshears J, Choi JJ, Laibson D, Madrian BC (2013) Planning prompts as a means of increasing preventive screening rates. *Preventive Medicine* 56(1):92–93.
- Nickerson DW, Rogers T (2010) Do you have a voting plan? Implementation intentions, voter turnout, and organic plan making. *Psych. Sci.* 21(2):194–199.
- OECD (2013) *Co-operative Compliance: A Framework* (OECD Publishing, Paris).
- OECD (2014) *Working Smarter in Tax Debt Management* (OECD Publishing, Paris).
- OECD (2017a) *Applying Behavioural Insights to Organisations: Theoretical Underpinnings* (OECD Publishing, Paris).
- OECD (2017b) *Behavioural Insights and Public Policy* (OECD Publishing, Paris).
- Pomeranz D (2015) No taxation without information: Deterrence and self-enforcement in the value added tax. *Amer. Econom. Rev.* 105:2539–2569.
- Pomeranz D (2017) Impact evaluation methods in public economics: A brief introduction to randomized evaluations and comparison with other methods. *Public Finance Rev.* 45(1): 10–43.
- Rankin CH, Abrams T, Barry RJ, Bhatnagar S, Clayton DF, Colombo J, Coppola G, Geyer MA, Glanzman DL, Marsland S (2009) Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology Learn. Memory* 92(2):135–138.
- Rogers T, Frey E (2015) Changing behavior beyond the here and now. Keren G, Wu G, eds. *The Wiley Blackwell Handbook of Judgment and Decision Making* (Wiley-Blackwell, Oxford), 725–748.
- Rogers T, Milkman KL, John LK, Norton MI (2015) Beyond good intentions: Prompting people to make plans improves follow-through on important tasks. *Behav. Sci. Policy* 1(2):33–41.
- Ronconi L, Ortega D, Sanguinetti P (2015) Reciprocity and willingness to pay taxes: Evidence from a survey experiment in Latin America. Working paper, CIAS, Argentina.
- Scandura TA, Williams EA (2000) Research methodology in management: Current practices, trends, and implications for future research. *Acad. Management J.* 43(6):1248–1264.
- Skinner J, Slemrod J (1985) An economic perspective on tax evasion. *Natl. Tax J.* 38(3):345–353.
- Slemrod J, Weber C (2012) Evidence of the invisible: Toward a credibility revolution in the empirical analysis of tax evasion and the informal economy. *Internat. Tax Public Finance* 19(1): 25–53.
- Slemrod J, Yitzhaki S (2002) Tax avoidance, evasion and administration. Auerbach AJ, Feldstein M, eds. *Handbook of Public Economics* (Elsevier Science, Amsterdam), 1423–1470.
- Spann S (2009) Tax resolution step #1: Taking control—Making the commitment to change. *Tax Resolution and Financial Freedom: Using the Financial Planning Process to Resolve IRS Tax Problems* (LifeSpan Financial Planning, LLC, Mount Pleasant, SC), 17–33.
- Stinchcombe AL (1990) *Information and Organizations* (University of California Press, Berkeley, CA).
- Sunstein CR (2013) *Simpler: The Future of Government* (Simon & Schuster, New York).
- Thaler RH, Sunstein CR (2008) *Nudge: Improving Decisions About Health, Wealth, and Happiness* (Yale University Press, New Haven, CT).

- Thompson RF, Spencer WA (1966) Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psych. Rev.* 73(1):16–43.
- Torgler B (2002) Speaking to theorists and searching for facts: Tax morale and tax compliance in experiments. *J. Econom. Surveys* 16(5):657–683.
- Torgler B (2004) Moral suasion: An alternative tax policy strategy? Evidence from a controlled field experiment in Switzerland. *Econom. Governance* 5(3):235–253.
- Torgler B (2007) *Tax Compliance and Tax Morale: A Theoretical and Empirical Analysis* (Edward Elgar Publishing, Northampton, MA).
- Torgler B (2013) A field experiment on moral suasion and tax compliance focusing on under-declaration and over-deduction. *FinanzArchiv Public Financ. Anal.* 69(4):393–411.
- Torgler B, Schneider F (2005) Attitudes toward paying taxes in Austria: An empirical analysis. *Empirica* 32(2):231–250.
- Tversky A, Kahneman D (1981) The framing of decisions and the psychology of choice. *Science* 211(4481):453–458.
- Weber M (1922) *Economy and Society* (University of California Press, Berkeley, CA).
- Wenzel M (2005) Misperceptions of social norms about tax compliance: From theory to intervention. *J. Econom. Psych.* 26(6):862–883.