Email Reminders for Waste Disposal Compliance -Charlotte, NC

Use this worksheet to develop your city's Trial Protocol – the document that will guide your evaluation.

Since this is an evaluation replication guide, we've gotten you started by filling in some of the elements of your approach that stay the same. The rest is up to you!

Before we dive in - a few reminders:

- A big part of designing and running an evaluation is **careful planning** we've designed this guide following our "Trial Protocol" template. A protocol can help you to document the decisions and intentions around your intervention. It's a forward-looking plan that will help you down the road and also make it easy for people to replicate your work in the future (like we are doing now).
- We provide some guidance within the document for each section, but this guide has been designed for those who are comfortable handling data and understand a power calculation. If that's not you yet - that's ok! We'd recommend you partner with someone in your city to complete this document. Or, you can try <u>courses</u> on the What Works Cities Academy to get up to speed or join a BIT Sprint for "Evaluation Foundations" to learn some of the key concepts.

Note: BIT has expanded our recommended trial protocol template since our early work with cities, so we've added new recommendations and suggestions to enhance the original protocol.

To use this document: Please download a copy and fill in the "Your Approach" column

Part 1 - Scoping Your Evaluation

Step & Guidance	EXAMPLE – Charlotte, NC	Your Approach [Fill in this column]
Define your problem statement What is your high-level goal?	18.6% of households who schedule a pick-up fail to place the item on the curb at the specified time. Sending Solid Waste Services (SWS) to pick up items that are not there amounts costs the city money.	Many residents do not know how or where to dispose of bulky items, or they forget when they have made an appointment.
	Charlotte would also like to cut down on the number of items that are placed on the curb too early (more than 24 hours in advance), as this causes items to sit out for several days and is an eyesore.	
	People likely do not dispose of bulky items very often. This may mean that the process is foreign to them and they do not know where to put an item, how to prepare the item, or they forget when they have made an appointment.	
Describe your solution How will you try to reach your goal?	We will send timely emails with simple instructions to make their appointments more salient.	We will send email reminders with simple instructions to help residents remember to set out their items at their appointment times.
Describe your comparison What will you be comparing your approach to? What does your comparison or control group receive? Is the comparison business-as-usual, or an alternate version of a new process? If not business-as-usual, why?	We will send emails to residents who scheduled bulky item pick-ups, reminding them of their pick-up date and providing easy-to-use information on how to properly set out the item 24 hours before their scheduled pick-up (treatment). The control group will not receive an email (business-as-usual) Emails will be automatically sent after a pick-up is scheduled and approximately 24 hours before a pick-up will occur	We will randomly assign residents who have made an appointment to either receive an email (treatment) or not (business-as-usual control group) based on household address.
Define your outcome & indicator How will you measure the success of your approach? What is the target outcome and what is the indicator? What is the unit of your outcome measure? How does it relate to your unit of assignment? How will you link your outcome measure to treatment assignment?	 Primary outcome: We will measure the proportion of items that were not on the curb at their scheduled pick-up time according to Solid Waste Services. Primary Indicator: We will do this by considering 1) the number of households in the control and treatment groups with a resolution code from Solid Waste Services indicating that items were not on the curb for pick up and 2) the number of total households who scheduled a pick up. Households will be determined by address Secondary Outcome: The proportion of households who received a violation between when a pickup was requested and when it was scheduled to occur (excluding those where the violation and request occurred within 24 hours, as we assume those requests were generated by code enforcement, not participants in this study). Exploratory Outcome: We will conduct the same test as in the secondary 	Target outcome: Indicator:

	analysis with door hangers received, rather than code violations, as a check on our secondary outcome measure since door hangers are often placed instead of code violation citations. We will descriptively compare the number of code violations for bulky items in the control group, and the historical data, to determine if the hand-coded violations align with previous data on violations.	
Describe your research question	Does sending an email reminder decrease the proportion of households who fail to set out bulky items for their scheduled pickup?	Research Question:
Identify your population Describe who you are trying to target with this intervention	Target Population: All residents of Charlotte who schedule bulky item pickups.	Target Population:
Identify your sample Describe who you will test your intervention on (and if / how that is different than the population) Who will you actually be able to reach with your communication? How will you identify and reach your sample? Does this sample represent your population of interest? Do you anticipate any selection bias in who is represented in your sample?	 Trial Sample: Residents of Charlotte who: Schedule bulky item pick-ups from 2/14/18-6/14/18 (i.e. requests that were generated through WEB requests or the 311 call centerwe will exclude pick-ups marked as being created by NBS Code Enforcement). Do not live in an apartment building, as there is no way to determine who set out the bulky items in front of an apartment (done for us in SWS data pull). Provide an email address. Are making the first bulky item request that they make within our time frame. 	Trial Sample:

DATA QUALITY CHECKPOINT

Before going too much further - it's important to check the quality of your data

Before you design your evaluation, you need to understand what data is available to you and any potential issues you may have with the data.

Reviewing your data in advance helps you determine how you design your evaluation. Oftentimes, when reviewing the data you might learn key information that could change your evaluation design. For example, you might find out that you have less available data than you thought, which could make your planned evaluation more difficult and compromise your ability to measure the outcomes you seek to evaluate. Or you may discover that your outcome measures are structured differently in your data than you originally thought.

Before you move forward, check your dataset to determine the following:

- You have access to the data set
- シンシンシン Your data set contains the information or outcome variable that you want to measure
- You are aware of the ways the data could be inaccurate or unreliable (e.g. self-reported, incomplete, etc.)
- If new data is being collected or if data is being collected manually, could it go wrong?
- You've reviewed an output of this data (if possible)
- You are handling any personal identifying information and consent processes in an accordance with legal guidelines and ethical best practices.

Part 2 - Designing your evaluation

Step / Description / Resources	EXAMPLE – Charlotte, NC	Your Approach [Fill in this column]	
Decide randomization strategy			
Unit: What is being randomized (e.g. individual, household, school, etc.)?	Unit: Household identified by address	Unit:	
Method: How will you conduct randomization? (E.g., through a random number generator, lottery, coin toss, randomized paper sequence, etc.)	Procedure: Households with even street numbers will be assigned to treatment and households with odd street numbers will be assigned to control. Note: This randomization plan may not be perfectly random because house numbers are not randomly assigned (i.e. one side of the street (odd numbers) is more desirable). However, there are no factors that are potentially relevant to this evaluation and also have an impact on the address assignment process in Charlotte.	Procedure:	
Verification: How will you make sure that every participant is assigned to one, and only one, treatment group? If there is a risk that participants receive both treatments, will you be able to track this and control for it in analysis?	Verification: We recommend the team 1) have someone who did not conduct the randomization check the list before sending 2) check for duplicates and 3) add team members to the email list and report back which email they received.	Verification:	
Blinding & Masking: Will participants know their treatment assignment? Will their treatment status be known to others involved in the trial or intervention? (For example, will frontline staff know the difference between people in the control group versus treatment group?)	Blinding and masking: We made sure that participants would not know their treatment assignment. Treatment group members would assume that everyone in their situation receives an email. Control group members would not be aware of the intervention or trial.	Blinding and masking:	
Spillovers: What are the ways in which someone's treatment status might affect the outcome of someone in a different treatment group? Is it possible that a previous intervention might influence their behavior in this trial? If so, can these be minimized by creating distance between participants?	Spillovers: It is possible that a treatment group member could tell a control group member about the email, but this is unlikely to greatly impact the evaluation.	Spillovers:	
Calculate your sample size and power requirements			
Baseline: What is the current average for your indicator? Or, if you don't have historical data, what data do you have available that might give you an indication of what your current average might be? This should be your best guess for the expected outcome for your control group in your trial. If you don't have a precise number, run the power calculations for a range of different baselines based on your assumptions.	Baseline: We have 7,802 unique email addresses associated with items not out on the curb at the time of pickup and 41,956 pickups associated with unique email addresses in approximately one year of data (Excluding those where the violation and pickup request occurred within 24 hours of each other.) Base rate: 7,802/41.956 = 0.1860	Baseline:	

Standard deviation: If your outcome indicator is a continuous measure, what is the baseline standard deviation?		
Power calculation: Use the power calculator to complete your power calculations using the numbers you listed above. Per group sample size: Minimum detectable effect size: Significance level: 0.05 Power level: 0.80	Number of participants available: Based on previous trends we can expect that 3,500 pickups associated with unique email addresses will be scheduled each month Power calculation: BIT recommends a significance level of 0.05 and a power level of 0.80 We calculated the MDES for varying study duration (in months) assuming an increasing sample size of 3,500 per month. If the trial was run for 3 months we could detect a 2.174 percentage point change. This would require approximately 76 more people to put their items on the curb every month to detect a significant change. If the trial was run for 4 months we could detect a 1.877 percentage point change. This would require approximately 66 more people to put their items on the curb every month to detect a significant change.	Number of participants available: Power calculation: Per group sample size: Minimum detectable effect size:
Target effect size: How large of an effect size do you think is reasonable to expect from this intervention (based on prior evidence if available)? Alternatively, what would be a meaningful effect size (based on break-even point if applicable), and why?	Target effect size: We recommend searching for any recent studies that used an email in the context of heavy item pickup services.	
Attrition: What is your best guess of the number/percentage of participants that will leave your sample between randomization and outcome data collection. Adjust your sample to ensure it's large enough for analysis at the end of the trial.	Attrition: With an intent-to-treat analysis we would not remove any undeliverable messages.	

FEASIBILITY CHECKPOINT:

Is the minimum detectable effect size reasonable? If not, can you change the sample size by running the trial for longer or changing the intervention to make it potentially more impactful? If no changes can be made, do you still want to go ahead with the trial as an implementation pilot?

Consider experimental threats & risks

What things can you do to make sure people receive the intervention as it was intended?

Is there a way you can double-check that your randomization has been implemented correctly and the participants are receiving the treatment they were assigned?

Are there things you can do to make sure that staff is prepared and ready to implement the trial (e.g., training on data In this case, the city randomized by even/odd address numbers. We typically would recommend using a random number generator. In some cases, street numbers may not always be truly random (e.g., one side of the street is more desirable than the other).

Code Enforcement will hand-code whether a violation was for a bulky

Risk: Likelihood: Impact: Mitigation Approach:

collection, a pilot period to adjust to new workflows, extra time to adjust to a new call volume)?	 item or not (Y/N). It is possible that some of these codes will be entered incorrectly. We will compare the rate of bulky item violations in the control group to the historical data to assess how often this is an issue. Risk: Randomization is not implemented correctly or participants receive the wrong treatment. Likelihood: Low Impact: High Mitigation Approach: Two different people will check that the randomization procedure was followed correctly before implementation. In each review, the reviewer will spot-check 15 addresses on the list for correct treatment assignment. They will also check that individuals were not assigned to both groups. Risk: Crews are not prepared to receive additional bulky items. Likelihood: Low Impact: Low Mitigation Approach: Usually crews are prepared to pick up the bulky items from all scheduled appointments. We will notify the head of Solid Waste that the reminders are happening and that they may see an increase in items, but crews will not be directly notified. The head of the department can adjust schedules and routes as necessary. 	
Consider ethical risks How might participating in the trial harm people in your sample or others? How will you monitor the trial to ensure you can detect the harm early and change your implementation if needed?	 Risk: Backfiring effect, i.e. email recipients have adverse reactions to the messages. Likelihood: Low, we did not identify any risks in this trial. Mitigation Approach: We will monitor complaints (e.g. 211 calls) as well as the number of people who opt-out of receiving emails. If we notice that more than 3% of recipients are opting out, we will look into potential causes. 	Your answer:
Consider race & equity implications The goal of this section is to consider whether the benefits of your intervention are distributed equitably across your community. If not already included above, consider: do you expect that your intervention could have different impacts across your city? If so, do you have a way to analyze this at the end of your trial? Disaggregating your data on a particular participant characteristic (e.g. gender, race, etc.) is often a good first step here. You may also want to include descriptive analyses that can help explain any trends or differences you find. In cases where you do not have demographics, can you use	If the emails are effective, they will help residents who would no longer be fined because they'd start setting out their items properly. Therefore, it's important to the city that the email helped all residents so that all residents could experience a decrease in fines. While we do not have household-level information on participants, we can track demographic and socioeconomic factors. The city uses an equity scale that considers factors such as race, ethnicity, household income, home/car ownership, etc. Each census block is given a score from 1-7. We can analyze whether the	Your answer:

proxy variables (e.g. Census tract information matched to zip codes?)	emails were effective or not in areas at each level of the equity scale (e.g., by comparing the outcomes of households with an equity score of 1 who got an email to households with an equity score of 1 who did not get an email). We expect there may be differences across levels of the scale since we know that neighborhoods with lower equity scale scores are less likely to schedule a pickup, however, we want to ensure that the email is helpful at all levels. (Note: this would also be added as a descriptive exploratory measure)	
--	--	--

Plan for Data Analysis

Understand and Specify Your Variable for Analysis

There are two parts to planning for your analysis. First, you need to think through your different variables and make sure to document how you plan to use them for your trial analysis. This step helps you to ensure you have the data you need and sets you up for success for your analysis plan. Additionally, it's good to specify what you will do with bad data or data that doesn't match.

Here are some questions to guide your data checks:

- Where are you getting the data you need to complete the trial analysis (e.g., treatment assignment, outcome indicator data, other participant characteristics)?
- Are there any data security procedures that need to be followed?
- How will you assemble the data? How will your variables be constructed (e.g., units, interpretation of values, etc.)?
- How will you check your data for accuracy (e.g., to make sure that any data merges were done correctly, or that missing values have been identified and dealt with as needed)?

Specify Your Analysis Plan

Next, you need to specify your analysis plan. You should have a "hypothesis" you are testing - that your intervention will work -but specifying ahead of time will allow you to say what statistical test & analysis you will use to determine if it worked. See below for an example of Charlotte's analysis plan. If you have questions on how to choose what test to run, check out resources here.

Questions to guide your analysis plan:

- For your Primary outcome:
 - What statistical test will you run for your outcome indicator? For example, will you run a t-test to see if the mean outcome for your treatment group is statistically different than the mean outcome for your control group, a regression to control for other factors, or a more complex analysis?
- For your Secondary analysis:
 - Are there any other statistical tests you would like to run (e.g., other outcome indicators, or looking at sub-groups)?
 - Will you want to do any cost-benefit analysis?

Sample Data

Variable name	Туре	Source	Measurement
Resolution_code	Dependent	Bulky item requests	Categorical; individual
Curbside	Dependent	Built from resolution_code, (0 if resolution_code == Items not on curb)	Binary - 0, 1; individual
Create_group	Independent	Bulky item requests, (include WEB and 311 requests in our sample)	Categorical; individual
Email_code	Independent	Bulky item requests, (include if provided email in our sample)	String
Create_date	Independent	Bulky item requests, (include if between 2/14/18-6/14/18)	Datetime
Scheduled_date	Independent	Bulky item requests	Datetime
Address	Independent	Bulky item requests	String
Treatment	Independent	Built from Address, (even street numbers will be assigned to treatment and households with odd street numbers)	Binary - 0, 1; individual
Request_type	Covariate	Bulky item requests	Categorical; individual
Created on (case)	Independent	Codes	Datetime
Case origin (case)	Independent	Codes	Categorical
Street number (case)	Independent	Codes	String
Street direction (case)	Independent	Codes	String
Street name (case)	Independent	Codes	String
Street type (case)	Independent	Codes	String
Street suffix (case)	Independent	Codes	String
Address (case)	Independent	Built from street variables above	String
Case origin	Independent	Codes	Categorical
Bulky Item	Independent	Codes	Binary - Y/N; individual
Early	Dependent	Built from created on (case) being between create_date and schedule_date after a merge on address	Binary - 0, 1; individual

Sample Data: Analysis Plan

We ran a linear regression to find the relationship between receiving and email and failing to set out bulky items. The primary estimating equation was:

 $Y_{ik} = \beta_0 + \beta_1 Treatment + \beta_{ik} X_{ik} + \varepsilon$

Where i represents the household and Y represents whether or not an item is out on the curb at the time of its scheduled pickup (primary outcome).

Treatment is equal to 0 if the household was assigned to the control group and equal to 1 if it was assigned to treatment. β_1 represents the average effect of receiving our email intervention on setting out bulky items. We will include create_group and request_type as covariates, represented by the matrix X_i .

SECONDARY: We will use the same model to determine the likelihood that someone received a code violation for setting out an item too early. In this analysis, Y will represent whether or not someone receives a violation between their create date and scheduled pick-up date in the bulky item data. We will use case origin as a covariate.

EXPLORATORY: Will will run the secondary analysis with door hangers received, rather than code violations, as a sensitivity check (as door hangers are often placed instead of code violations).

We will descriptively compare the number of code violations for bulky items in the control group and the historical data, to determine if the hand-coded violations align with previous data on violations.

Your Approach:

Please describe how you will analyze your data as well as what steps you will take to ensure the data is clean and ready for analysis. Note: You can use a different analysis than the model specified in the example above, based on your available data.

Part 3 - Implementing your evaluation

Plan for Implementation

Create a project implementation plan and timeline

At this point, an evaluation starts to feel just like any other project -- with a few extra checks! You will need to carefully plan for implementation and ensure randomization is able to occur and data is able to be collected at the correct intervals. Be sure to build in time for getting any necessary approvals and for double-checking your work. We've included a sample timeline here, but we find that implementation works best when it follows a city's normal project management process.

Sample Implementation Plan

The protocol for this trial did not include an implementation plan and timeline. We have drafted what one might have looked like below.

Task	Owner	Deadline
Protocol drafted	ВІТ	March 3 - 22
Protocol approved	Charlotte	April 4
Sample pulled and randomized	Charlotte	April 10
Implementation checked	Charlotte	April 11
Trial launched	Charlotte	April 22
Opt out rates reviewed for adverse reactions.	Charlotte	April 24
Trial ends	Charlotte	July 11
Send data for analysis	Charlotte	July 12
Data analysis completed	ВІТ	July 24
Quality assurance of data analysis	BIT	July 26
Final Report drafted	Charlotte / BIT	July 28 - August 4
Final Report approved	Charlotte / BIT	August 5 - 10

Your Approach:

Please fill in a chart of your implementation plan for this project - feel free to copy in from above!

Monitor your evaluation

While your evaluation is in the field, be sure to check in and ensure everything is happening to plan. This can take the form of automated data checks, regularly scheduled check-ins with your partners, or observations of the process. It's better to catch anything before the end of your evaluation so you can adapt as needed!

Part 4 - Analyzing your results

Analyzing your data

You've already specified your analysis plan before you launched the trial - so now is the time to use it! If possible, build in some time and budget to have someone who is not involved with the project review your work. It can help you eliminate any blind spots or highlight any assumptions you've made about the data.

Once you've analyzed your results you can ask yourself a few questions to help you reflect your results:

- Descriptive analysis:
 - What is the average and standard deviation for your full sample?
 - What is it for each of your treatment groups?
 - Are there any characteristics of the sample to describe (demographics, location, time, etc.)?
 - Primary outcome:
 - o Statistically significant: Yes/No
 - Effect size: What is the effect of your treatment, on average? (e.g., an increase of 2 percentage points or a decrease of \$100, on average)
 - o Distribution: What is the range of outcomes was the confidence interval quite large, or narrow?
 - Interpretation: What does this difference mean practically? Is it large enough to make a meaningful difference? If there was a wide range of outcomes, is it acceptable to implement something with that range of results?

Congrats on running your evaluation!

Now that you have your results, you can decide how you want to use them. If your result was positive, you may consider scaling your solution. To see how Charlotte scaled up their email campaign, head back to the Replication Guide for some advice on how to scale up your result if successful!