

Behavioral Design Teams

A Model for Integrating
Behavioral Design in City Government

Authors:

Anthony Barrows

Natalie Dabney

Jon Hayes

Rachel Rosenberg

April 2018

Table of Contents

Intro.....	2
About ideas42.....	2
Acknowledgements.....	2
Executive Summary	4
Why build a Behavioral Design Team?	10
What is behavioral design?	10
Why use behavioral design in government?	12
Embedded Behavioral Design Teams.....	13
How to do it	16
Building a Behavioral Design Team	16
Operating a Behavioral Design Team	22
Project selection criteria.....	24
The behavioral design process	28
Define	28
Diagnose.....	29
Design	31
Test.....	32
Scale	34
What's next?	36

About ideas42

ideas42 has a clear mission: to use our unique experience as a non-profit at the forefront of behavioral science to change millions of lives. We create innovative solutions to tough problems in economic mobility, health, education, safety and justice, consumer finance, energy efficiency, and international development. Our approach is based on a deep understanding of human behavior and why people make the decisions they do. Working closely with our partners from government, foundations, NGOs, and companies, we have more than 80 active projects in the United States and around the world.

Since 2014, ideas42 has partnered with the White House, the City of Chicago, and New York City to recruit and create embedded Behavioral Design Teams charged with using applied behavioral science to improve governance and public services.

Visit ideas42.org and follow [@ideas42](https://twitter.com/ideas42) on Twitter to learn more about our work. Contact us at gov@ideas42.org with questions about our Behavioral Design Teams.

Acknowledgements

We would like to thank the Laura and John Arnold Foundation for funding the New York City Behavioral Design Team in partnership with the New York City Mayor's Office of Operations. We would also like to thank the John D. and Catherine T. MacArthur Foundation for funding the Chicago Behavioral Design Team in partnership with the City of Chicago's Mayor's Office. **The views expressed in this paper reflect those of the authors and do not necessarily reflect those of the foundations**

In particular, we would like to thank the following individuals for their invaluable support in launching and sustaining the New York City and Chicago teams:

- **Our leadership in New York City Government:** Tony Shorris, formerly First Deputy Mayor; Mindy Tarlow, formerly Director, Mayor's Office of Operations; Matt Klein, Executive Director, Mayor's Office for Economic Opportunity; and Carson Hicks, Deputy Executive Director, Mayor's Office for Economic Opportunity
- **Our leadership in Chicago Government:** Chris Wheat, Chief Sustainability Officer & Senior Policy Advisor, Mayor's Office; Michael Negron, Chief Policy Officer, Mayor's Office; and Paras Desai, formerly Director, Innovation Delivery Team, City of Chicago

None of our projects would be possible without our implementation and thought partners at city agencies and other community organizations. We would like to thank the following people for helping us scope, develop, and launch the projects outlined in this publication:

- **New York City:** Nicol Bellettiere, Colin Chellman, Denise Clay, Adam Cohen, Rachel Cohen, David Crook, Daniela Curmei, Barbara Dannenberg, Kinsey Dinan, Deborah Friedman, Chenda Fruchter, Katherine Greig, Kim Hardy, Michael Hickey, Daniel Hildreth, Jacques Jiha, Laura Kavanagh, Rebecca Kirchheimer, Dana Kochnower, Cheryl Lawrence, Claire Levitt, Chris Long, Erica Mullen, Maria Osorio, Sarah Peterson, Roger Platt, Joe Rinehart, Jeffrey Shear, Daniel Timmeny, and Leslie Zimmerman
- **Chicago:** Samantha Aigner-Treworgy, Emma Berndt, Becky Betts, Tina Consola, Brice Cooke, Lisa Davis, Christian Denes, Arne Duncan, Mary Ellen-Messner, Rosa Escareno, Alex Frank, Peter Gribble, Jessica Higgins, Amy Jewell, Erin Keane, David Leopold, Cortland Lohff, Raed Mansour, Patrick Milton, Dr. Julie Morita, Lisa Morrison-Butler, Chris Nash, Aurelie Ouss, Dakota Pawlicki, Kurt Peterson, Jamie Ponce, Molly Poppe, Matt Quinn, Chris Sauve, Christina Seybolt, Mark Siegel, Ed Tumlos, Charles Williams

At ideas42, we always rely on expert guidance to shape our work. We'd like to thank the following people for their assistance, perspective, and advice: Ben Castleman, University of Virginia; Omid Fotuhi, Stanford University; Reuben Kantor, City of Revere, MA; Jens Ludwig, University of Chicago; Katy Milkman, University of Pennsylvania; Todd Rogers, Harvard University; Judy Scott-Clayton, Columbia University; Anuj Shah, University of Chicago; and Kevin Volpp, University of Pennsylvania.

Finally, we must acknowledge our current and former ideas42 teammates who helped to shape and execute the work of the Behavioral Design Teams, and who gave important guidance and feedback on this playbook: John Balz, Jeremy Barofsky, Vivien Caetano, Heber Delgado-Medrano, John Harris, Lee-Sien Kao, Andreana Kenrick, Rahin Khandker, Rebecca Mack, Deepti Nagulapally, DJ Neri, Doug Palmer, Ted Robertson, Will Tucker-Ray, and Allie Yates-Berg.

Executive Summary

This playbook explores the promise and practicalities of Behavioral Design Teams (BDTs) within government. These teams are groups of behavioral design specialists that embed within a public administration to improve the jurisdiction's ability to effectively and efficiently carry out its work. BDTs approach challenges and opportunities from a behavioral design perspective. That is, they leverage insights from the fields of behavioral science and impact evaluation.

Behavioral science uses research from psychology, neuroscience, and behavioral economics to understand the intricacies of human choice and action. The way information is presented or the environment in which a person makes decisions can have a large, and often counterintuitive, impact on how he or she behaves. Behavioral science helps identify the often predictable ways that such contexts affect human behavior. Resulting insights can be used to help people do more of what they want and less of what they don't. When behavioral science is combined with impact evaluation—the use of rigorous methods to understand the effectiveness of policies and programs—BDTs are able to rapidly iterate and design improvements that account for how people actually behave, rather than how we *think* they should.

A behavioral design approach is particularly useful to governments because the success of policies, programs, and services depends on people's decisions and actions. Whether a workforce development program results in employment depends on people actually showing up. Whether the Supplemental Nutritional Assistance Program (SNAP) helps feed families depends on people persisting through a lengthy application process. Whether a gifted and talented program helps close the achievement gap between low- and high-income students depends on low-income families applying for the program. Using the behavioral design process to help solve these issues often results in low-cost interventions that are also easy to implement. This is a particularly compelling feature for government, because behavioral design can generate important results through small changes that are basically cost-neutral. In this way, behavioral design can help governments navigate budget constraints while still providing effective services in the communities that need them most.

This playbook is written for public servants at all levels of government. We hope it will help you understand the concept of a BDT and decide whether a BDT is a useful and feasible resource. There are many ways to incorporate behavioral design in government, like hiring a Chief Behavioral Officer or assigning individual behavioral scientists to city agencies, and you will have to decide which model is right for your context. In the following pages, we will try to help you think through key questions like: ***Is a BDT affordable? How will we pay for it? Should we staff it externally or internally? How might we build broad community and political buy-in?*** While this publication is focused on municipal government, BDTs have also been successful at the national level (see the [US Social and Behavioral Sciences Team](#), launched and partially staffed by ideas42 team members, or the [UK Behavioral Insights Team](#)).

The pages that follow have simple aims: to introduce you to some core concepts from behavioral design; to make a case for why and where a BDT can be helpful in the context of local government; to lay out the fundamental competencies and structures you'll need to set up a BDT; and finally, to provide guidance on how to run a successful BDT. We will articulate some of our ambitions for the future of BDTs and pose some questions that we do not have answers to just yet. Our hope is that this brief publication will help you understand the promise of behavioral design, see the potential value of a BDT in your city, and get your own team started.

Throughout the playbook, you will find examples of work from ideas42's inaugural city BDTs in Chicago and New York City. Across these two cities, we have tackled a wide range of topics, all centered around our mission of improving millions of lives through behavioral design. Below, we provide a summary table of all projects and outcomes:

PROJECT	CITY	AGENCY	TYPE	OUTCOME
Economic Mobility				
<i>Increasing timely form submission for Supplemental Nutrition Assistance Program (SNAP) recertification clients</i>	NYC	Human Resources Administration	RCT	Decreased failure to submit forms by 5.5%; increased timely form submission by 12.9% (<45 days), and uptake of phone interviews by 6%.
<i>Increasing timely interview completion for SNAP recertification clients</i>	NYC	Human Resources Administration	RCT	Results pending
<i>Increasing social service continuation rates through redesigned referral forms</i>	Chicago	Department of Family and Support Services	Design only	Scaled citywide
<i>Increasing Earned Income Tax Credit claims through increased tax filing</i>	Chicago	Department of Family and Support Services	RCT	No effect on sample
<i>Increasing uptake of Neighborhood Homelessness Prevention Outreach program</i>	NYC	Human Resources Administration	RCT	No effect on sample
<i>Increasing usage of online tax prep portal</i>	NYC	Department of Consumer Affairs—Office of Financial Empowerment	RCT	No effect on sample

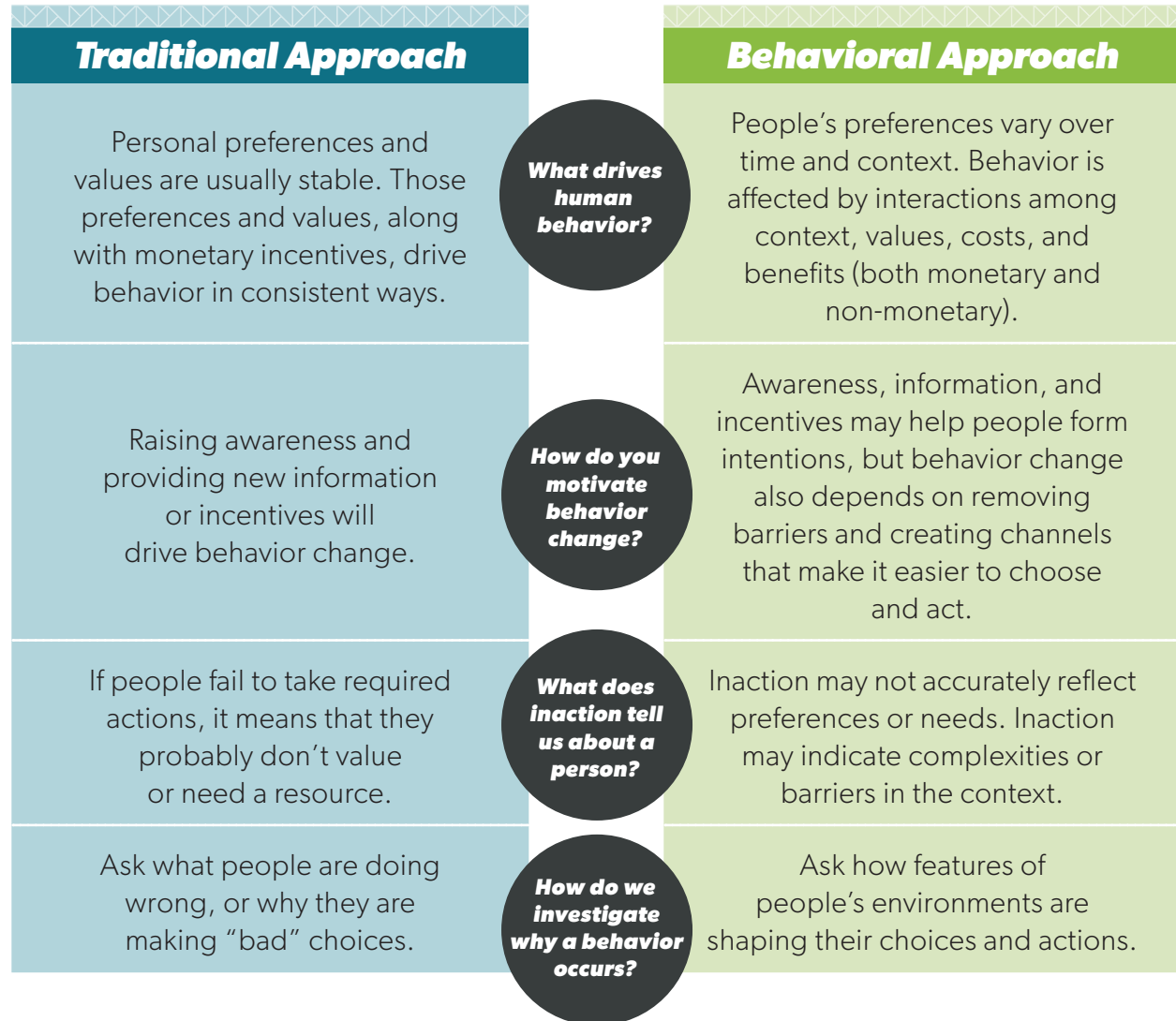
PROJECT	CITY	AGENCY	TYPE	OUTCOME
Education				
<i>Increasing financial aid renewal among community college freshmen at two colleges (FAFSA I)</i>	NYC	City University of New York	RCT	Increased FAFSA filing rates by 13.7%
<i>Increasing financial aid renewal among community college freshmen at three colleges (FAFSA II)</i>	NYC	City University of New York	RCT	Increased FAFSA filing rates by 31.2%
<i>Increasing financial aid renewal among community college freshmen at six colleges (FAFSA III)</i>	NYC	City University of New York	RCT	Results pending
<i>Improving placement testing outcomes among entering community college freshmen</i>	NYC	City University of New York	RCT	No effect on sample
<i>Increasing college retention among community college students (Mindset I)</i>	NYC	City University of New York	RCT	Increased persistence among freshmen by 3%
<i>Increasing college retention among community college students (Mindset II)</i>	NYC	City University of New York	RCT	Results pending
<i>Helping community college students at three schools enroll in 15 credits per semester (Credit Momentum I)</i>	NYC	City University of New York	RCT	Results pending
<i>Helping community college students at four schools enroll in 30 credits per year (Credit Momentum II)</i>	NYC	City University of New York	RCT	Results pending
<i>Increasing matriculation among students admitted to community college who did not enroll (College Melt)</i>	NYC	City University of New York	RCT	Results pending
<i>Helping community college students maintain a GPA high enough for state financial aid eligibility (TAP)</i>	NYC	City University of New York	RCT	Results pending
<i>Increasing gifted and talented testing among families in low-income districts</i>	NYC	Department of Education—Office School Enrollment	RCT	Increased test registration by 6.5%

PROJECT	CITY	AGENCY	TYPE	OUTCOME
<i>Increasing matriculation among students admitted to college (Summer Melt)</i>	NYC	Department of Education	Design only	81% response rate among students
<i>Increasing matriculation among students admitted to college (Summer Melt)</i>	Chicago	Chicago Public Schools	RCT	Results pending
<i>Increasing early Pre-K enrollment</i>	Chicago	Chicago Public Schools / Department of Family and Support Services	Design only	Scaled citywide
<i>Increasing Pre-K enrollment attendance</i>	Chicago	Chicago Public Schools / Department of Family and Support Services	Design only	Scaled citywide
Equity & Justice				
<i>Increasing test filing rates for firefighter candidates</i>	NYC	New York Fire Department	RCT	Increased overall filing rates by 36.7%; 84% increase among black candidates, and 83% increase among female candidates
<i>Increasing test appearance rates for firefighter candidates</i>	NYC	New York Fire Department	RCT	Increased test appearance rates by 4%
<i>Increasing recertifications for small businesses owned by women and under-represented minorities</i>	NYC	Small Business Services	Design Only	Scaled citywide
<i>Increasing uptake of wheelchair accessible vehicles among taxi operators</i>	NYC	Taxi & Limousine Commission	RCT	No effect on sample
<i>Increasing feeling of safety over 4th of July among youth in summer jobs program</i>	Chicago	Department of Family and Support Services	RCT	Increased “feeling safe” over 4th of July by 17%
<i>Increasing youths’ accomplishment of goals set at beginning of summer</i>	Chicago	Department of Family and Support Services	RCT	Response rates to survey too low to draw conclusion
<i>Increasing feeling of safety among youth between end of summer jobs program and school resuming</i>	Chicago	Department of Family and Support Services	RCT	Response rates to survey too low to draw conclusion

PROJECT	CITY	AGENCY	TYPE	OUTCOME
Government Operations				
<i>Increasing on-time parking ticket payments</i>	Chicago	Department of Finance	RCT	Increased on-time payments by 31% , or 4 percentage points
<i>Increasing police citation payments</i>	Chicago	Department of Finance	RCT	Increased payment rates by 27% ; \$3.16 more revenue per notice
<i>Increasing on-time business license renewals</i>	Chicago	Department of Business Affairs and Consumer Protection	RCT	Reduced average time to renew by 4.2%
<i>Reducing misfiled service requests on the 311 mobile app</i>	NYC	311	Pre/Post	Reduced misfiled service requests from 59% to 9% of submissions
<i>Increasing retention of Police Department recruits at POWER Test (physical fitness evaluation)</i>	Chicago	Department of Human Resources / Chicago Police Department	Design Only	Scaled citywide
<i>Helping homeowners understand and act on property valuation notices</i>	NYC	Department of Finance	Design Only	Scaled citywide
<i>Helping property owners avoid the sale of tax liens incurred from non-payment</i>	NYC	Department of Finance	Design Only	Scaled citywide
<i>Increasing city citation payments</i>	Chicago	Department of Finance	RCT	Small increase in payment rates, but not statistically significant
<i>Increasing electronic business tax filing and payment</i>	NYC	Department of Finance	RCT	No effect on sample
<i>Increasing parking ticket payment rate pre-judgment</i>	NYC	Department of Finance	RCT	Results pending
Health				
<i>Increasing flu vaccine uptake among NYC employees</i>	NYC	Office of Labor Relations—WorkWellNYC	RCT	Increased vaccine uptake in pilot by 5%; at scale by 10%
<i>Increasing uptake of in-home lead paint inspections</i>	Chicago	Department of Public Health	RCT	Increased response rates to DPH letters from 0.4% to 1.2%

PROJECT	CITY	AGENCY	TYPE	OUTCOME
<i>Increasing submission of forms authorizing in-school treatment of students with asthma</i>	NYC	Department of Health and Mental Hygiene	RCT	Testing implementation failure; scaled citywide
<i>Assisting school personnel in acquiring and acting on authorization for in-school treatment of students with asthma</i>	NYC	Department of Health and Mental Hygiene	Design Only	Scaled citywide
<i>Promoting uptake of online and telephonic medical services through insurance card inserts</i>	NYC	Office of Labor Relations—WorkWellNYC	Design Only	Scaled citywide
<i>Promoting uptake and usage of blood pressure monitoring kiosks in pharmacies</i>	NYC	Department of Health and Mental Hygiene	Design Only	Pilot—no results available
Sustainability				
<i>Reducing commuter congestion on public transit</i>	Chicago	Sustainability team	Pilot	Reduced peak-hour ridership on Red Line on Cubs game days by 15.4%
<i>Increasing energy benchmarking compliance</i>	Chicago	Sustainability team	RCT	Increased compliance with energy benchmarking law by 6.2%
<i>Increasing energy benchmarking compliance after deadline</i>	Chicago	Sustainability team	RCT	No effect on sample
<i>Increasing energy efficiency investments among building owners with energy benchmarking letters</i>	Chicago	Sustainability team	Design Only	Scaled citywide
<i>Increasing energy efficiency investments</i>	Chicago	Sustainability team	RCT	No effect on sample
<i>Increasing flood insurance survey uptake</i>	NYC	Office of Recovery and Resilience	RCT	Increased survey responses by 15.5 times, or 4.52 percentage points
<i>Reducing disposable bag use through a city-wide tax</i>	Chicago	Sustainability team	Diff-in-diff	Reduced disposable bag use by over 40%
<i>Reducing contamination in recycling stream</i>	Chicago	Department of Streets and Sanitation	RCT	Small reduction in contaminated recycling, but not statistically significant

Why build a BDT?



What is behavioral design?

All too often policies, programs, and services are designed for a hypothetical type of human who is always a rational thinker and who always acts in his or her own best interest—what traditional economics calls an “econ.” What we may intuit from everyday life, however, is that this hypothetical human is often exactly that: hypothetical. In reality, we don't always decide and act according to traditional economic theory, or even according to our own intentions or best interests. When programs are designed for hypothetical humans rather than actual humans, policies may backfire, beneficial programs may be underutilized, and helpful services may go unused. Behavioral design may offer a solution.

Behavioral design combines empirical insights from two fields.¹ The first is behavioral science, the study of how people make decisions and take actions. Behavioral science draws from decades of research in psychology, neuroscience, and behavioral economics to understand how people react to features of the contexts they are in, often counter-intuitively: presenting too many choices may lead people to choose nothing at all,² people tend to value objects more once they own them,³ and people are more responsive to shorter deadlines than longer ones.⁴ These are just a few examples of how behavioral science illustrates that circumstances can drive people's behaviors to deviate from "rational" decision-making.

These insights about the quirks of human decision-making and behavior lead behavioral science practitioners to take a different approach to policy and program design from traditional economic thinking. While a traditional economic approach assumes new information will drive behavior change, behavioral science recognizes that changing a behavior requires more than just providing new information; people must also have a clear moment to choose and an easy opportunity to act on that choice. Knowledge of systematic, context-driven tendencies allows behavioral designers to inform the design of interventions with specific cues or contexts that tend to elicit specific reactions. Adopting this "behavioral lens" can help design programs and policies optimized for the real world.

Behavioral design also draws from a second field: impact evaluation, which uses rigorous empirical methods to measure the impact of policies and programs. By adopting these methods, like randomized-controlled trials (RCTs), behavioral design can determine with reasonable certainty whether an intervention achieves its desired effect. Impact evaluations also allow for iterative improvements to programs and policies as we work to scale them to different (often larger) populations.

In combination, behavioral science and impact evaluation can create reliable evidence about what works, when, and for whom. This evidence-based approach to innovation ensures that government can maximize its impact by deploying policies and programs tailored to the unique contexts and needs of its constituents.

ideas42 has run over 150 behavioral design projects across more than 30 countries. We have learned that context matters, that asking the right questions is critical, and that simple interventions are often available, but frequently overlooked or dismissed. The behavioral design approach is an effective technique to identify subtle contextual details that can have a disproportionate impact on outcomes—a key need for any organization serving people.

¹ Tantia, P. (2017). "The New Science of Designing for Humans." *Stanford Social Innovation Review*, Spring 2017.

² Tversky, A. and Shafir, E. (1992). "Choice Under Conflict: The Dynamics of Deferred Decision," *Psychological Science*, 3(6).

³ Thaler, R. (1980). "Toward a positive theory of consumer choice," *Journal of Economic Behavior and Organization*, 1.1 (March), 39-60.

⁴ Ariely, D., & Wertenbroch, K. (2002). "Procrastination, deadlines, and performance: Self-control by precommitment." *Psychological Science*, 13(3), 219–224.

Why use behavioral design in government?

While behavioral design is a tool that can benefit all types of entities, it is particularly well-suited to advance the goals of governments. First, many current government policies and programs are built on the traditional economic assumptions outlined above: people are rational decision makers, and “if you build it, they will come.” Too often, though, people *don't* come. Every year billions of dollars in benefits like SNAP, WIC, and EITC are left uncollected. We know that people want and need these benefits, but something is preventing them from actually collecting them; there is an “intention-action gap.” The larger this gap, the less successful a program is. So, why do eligible people not apply for helpful benefits? Behavioral design can help answer that question and others like: ***How can we get more families to recertify for their benefits on time? How can we get more people to pay their parking tickets before their fines double? How can we help people use less energy?*** Behavioral design is well-suited to identify and design successful, contextually-appropriate interventions to address key problems of uptake, usage, compliance, and retention.

Second, behavioral design can offer interventions that produce extraordinary impacts as compared to their cost to implement. While the behavioral approach can be a powerful tool for intensive systems-level improvements, it can also be deployed relatively cheaply and quickly: simple changes to the framing of a letter, thoughtful choices for the subject line of an email, or a timely text message can all be effective (and testable) interventions. In fact, each of those approaches has shown results, from increased take-up rates, to stronger follow-through on goals, to better-informed decisions. Over time, even small interventions can result in long-term cost savings, increases in efficiency, and better service delivery. Behavioral design helps governments achieve system-wide improvements at negligible cost, enabling continued delivery of valuable services, even under budget constraints.

Lastly, behavioral design can be used to help people *inside* government work in new ways (yes, this means you). By looking inward for behavioral design opportunities, governments can reap additional benefits. Behavioral design can offer interventions to improve employee-client interactions, create efficiencies in processes, and increase workforce diversity. There are even more benefits to be gained if behavioral design is institutionalized within the government workforce: if skills are dispersed across multiple agencies, the behavioral approach is more sustainable, and evidence-based decision making can become more common. When public servants build programs based on prior evidence and produce new evidence from rapid, rigorous testing, governments can achieve their goals more effectively and efficiently.

Embedded Behavioral Design Teams

“Adopting the insights of behavioral science will help bring our government into the 21st century in a wide range of ways—from delivering services more efficiently and effectively; to accelerating the transition to a clean energy economy; to helping workers find better jobs, gain access to educational opportunity, and lead longer, healthier lives.”

— PRESIDENT BARACK OBAMA, SEPTEMBER 15, 2015, SPEAKING ABOUT THE LAUNCH OF THE WHITE HOUSE SOCIAL AND BEHAVIORAL SCIENCES TEAM (SBST)

The case for integrating behavioral design in government may be clear, but the paths to doing so are numerous. One approach might be to hire a Chief Behavioral Officer to set a centralized, citywide strategy for incorporating behavioral design into operations. A more decentralized approach could be for individual agencies to contract with behavioral advisers to work in-house on behavioral design projects that meet the agency’s specific needs. A fully decentralized approach could entail independent behavioral scientists or professional organizations proposing and executing discrete projects according to their own interests or expertise. There are many possibilities, and the right answer for your city will depend on your needs and resources.

The topic of this playbook, however, is how to develop *our* preferred model—a partially-decentralized, *embedded* Behavioral Design Team (BDT). By being partially decentralized, BDTs stay nimble enough to keep a finger on the pulse of multiple agencies and to take on projects as needs emerge. On the other hand, being embedded in government gives BDTs enough central control to get beyond one-off projects and to focus on iteration, continuous improvement, and the dissemination of best practices. To achieve this balance, we constitute our BDTs as a cross-agency group of behavioral designers and policymakers, broadly charged with translating research into continuous process and program improvements. The resulting behavioral interventions range from small tweaks and project-level improvements all the way up to citywide collaboration and policy changes. As BDTs learn about which interventions work and where, they can support scaled-up integration of successful changes into everyday service provision, even across disparate programs and agencies.

This model of a BDT, partially-decentralized but still embedded, is particularly appealing because it can facilitate interagency collaboration and knowledge sharing as a byproduct of design and testing efforts. As this process repeats itself, BDTs can support a cultural shift towards behavioral and evidence-based policymaking while also achieving impact at scale.

Consider our experience with a project focused on influenza vaccinations during the 2016-17 flu season.



Increasing Flu Vaccines Among City Employees

To maintain a healthy workforce, the NYC BDT partnered with WorkWell NYC—the City’s workplace wellness program within the Office of Labor Relations (OLR)—to design and evaluate a set of behaviorally-informed emails that encouraged City employees to visit a work-site flu vaccine clinic.

On the initial stages of the project, the NYC BDT leveraged its connections to academia by consulting with Katy Milkman, PhD, a leading behavioral scientist, to incorporate her expertise on flu-shot uptake. Based on research by her and others, the BDT drafted two evidence-based designs and refined them in collaboration with WorkWell NYC to meet the needs of the City’s workforce. To rigorously test the designs, the BDT tapped the Department of Information Technology and Telecommunications (DoITT) to randomly assign the entire City workforce into three groups—one group for the standard email and one group for each of the two behaviorally-designed messages. The random assignment allowed the NYC BDT and OLR to isolate the relative effectiveness of each design, and the cross-agency collaboration built new connections and competencies among the City agencies.

OLR and DoITT both stepped outside their typical scopes of work to make the project work. This was the first time the WorkWell NYC team at OLR had applied behaviorally-informed communication strategies and tracked their influence through email click-through rates. This opportunity generated new ways of thinking about communications and evaluation moving forward. As an employer and administrator of health insurance, OLR cannot collect individual-level data on its employees or their participation in vaccination events. To accommodate this, OLR partnered with DoITT, NYCAPS (New York City Automated Personnel System), and the NYC BDT, who were able to randomize the database of employee email addresses and utilize sign-in sheets at the flu clinics to determine if one group of email recipients had higher flu vaccine uptake than the others. DoITT had never run a randomized trial with City email addresses, so this project developed their team’s capacity to make testing possible. Furthermore, the behaviorally-informed emails required additional resources and skills from DoITT to successfully develop and review the HTML. Each of these agencies tapped into their internal capacity to support testing, collaborated in new ways, and learned how to approach communications differently.

After sending these messages to over 400,000 City workers, the final data showed that the most effective behaviorally-informed email increased work-site-administered vaccines by 5 percent. Based on these results, WorkWell NYC sent the highest-performing email to the entire City workforce for the 2017-18 flu season and has seen a 10 percent increase in work-site flu vaccinations.

From: City of New York
Subject: [FINAL REMINDER] Schedule your flu vaccine before it's too late

This is your final reminder:
Flu season is here. [Book your appointment today.](#)
 Check the updated [calendar](#). Your worksite may provide free vaccines onsite.


Thank you to the thousands of city employees who have already received their flu vaccines.



Check your choice:

Yes, I'll get a vaccine to reduce my risk of getting and spreading the flu.


No, I will not get a vaccine, even if it means not protecting myself and others from getting the flu.

Which is best for you? (Select below)

	Free	Onsite	Effective	Click below:
 Your worksite	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Book at your worksite
AdvantageCare Physicians facility	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	Book with AdvantageCare Physicians
Participating pharmacy	Free with health insurance card		<input checked="" type="checkbox"/>	Locate a pharmacy <small>(all should confirm participation)</small>
Your doctor's office	Flu vaccine appointments only		<input checked="" type="checkbox"/>	Call your in-network doctor or find one here

Behaviorally-informed email


 Bill de Blasio, Mayor
 Healthy Workforce. Healthy City.

Last month, WorkWell NYC launched its 3rd annual Free Flu Vaccination campaign to help City employees protect themselves and their families against the flu. Thousands of employees have already received their flu shots – join them today!

All City employees who receive health insurance through the city can get a free flu shot:

- At the worksite** – Worksite clinics will be held at participating agencies across the city through November 23rd, 2016. Schedule your appointment by [clicking here](#) (username: nyc, password: flushot) or call 866-481-4391 Mon-Fri from 8AM-6PM. **ALL** employees eligible for health benefits, regardless of insurance provider, can get a flu shot at the workplace for no charge.
- At a participating pharmacy** – All employees, retirees, and their dependents can receive a flu shot free of charge at participating pharmacies. Check with your insurance provider to identify a participating pharmacy in your area and bring your insurance card to the pharmacy. You can use the NYC flu locator to assist in finding a pharmacy [here](#).
- At an in-network physician** – All employees, retirees, and their dependents can get flu shots free of cost at an in-network physician's office. (Co-payment will be waived when the sole purpose of going to the doctor is to get your flu vaccine.) To find a physician covered by your insurance, visit [zoodoc.com/nyc](#).
- At an AdvantageCare Physicians (ACPNY) medical office** – Flu shots will be available at all 36 ACPNY locations, including the new ACPNY Express for City employees at 52 Duane St. Schedule your appointment [here](#) or call 646-680-3000, Mon-Thurs from 8AM-8PM, Fri 8AM-5PM, or Sat 9AM-1:30PM.


ALL employees, dependents, and retirees, regardless of insurance provider, can get a flu shot at ACPNY for no charge. ACPNY locations will also provide immunizations to dependents (including pregnant women and children under 18 years old), pre-Medicare retirees, and spouses/dependents of pre-Medicare retirees.

Flu vaccination can prevent days or weeks of sickness and misery, doctor's visits and hospitalization. [^]All you need to do to get your free vaccination is bring your City of New York employee ID and your insurance card to your appointment!

Here is what your peers said about their workplace flu vaccination experience last year:

- "It was great to be able to go to work and only take a few minutes of my time to get the shot and get back to work. It was a very pleasant experience. Thank you!"
- "I really appreciate the convenience of getting the flu shot at work as opposed to scheduling a day off to receive it at my doctor's office. The process was simple, well organized, and efficient. Thanks again!"
- "So simple and quick. Thanks for making it so convenient to stay healthy!"

Got your flu shot? Tweet about it to spread the word and encourage others! #GotMyFluShot



Business-as-usual email

As this example shows, the BDT model works well because it strikes some important balances. BDTs can operate centrally enough to convene cross-agency partnerships and autonomously enough to react nimbly to changing conditions on the ground. By leveraging expertise from outside government, BDTs can efficiently deploy research insights to improve outcomes for city residents and employees. By operating within government, especially by being positioned within strategically-important agencies, BDTs can quickly build internal capacity for impact evaluation across multiple agencies and begin to institutionalize evidence-based decision-making.

“As a result of [the NYC BDT’s] efforts, we have shifted how we approach communications with the NYC employee population. We have also experienced increased participation in the initiatives where they were involved. Our programmatic impact has been enriched as a result of our joint efforts.”

— DEBORAH FRIEDMAN, MAYOR’S OFFICE OF LABOR RELATIONS

How to do it

Building a Behavioral Design Team

If it seems like your city would benefit from a BDT, there are three foundational steps you'll need to take. First, you'll have to generate buy-in—without stakeholders and support, a BDT will never move beyond an idea. Second, you'll need to choose where the BDT will be located—the “home” agency of a BDT will determine its ability to function effectively and efficiently. Lastly, you'll need to assemble a team with the experience and expertise to design and evaluate behaviorally-informed interventions. Below we offer some guidance about how you might begin thinking through these steps.

► *Buy-in*

As with any new effort, finding a “champion” is crucial. A BDT champion should be a well-connected, high-level official who has the willingness, capacity, and influence to generate awareness and enthusiasm for the BDT. He or she also needs a strong vision for where there are opportunities to use behavioral design, as well as a strategic sense of how to sequence and prioritize those opportunities. Having a champion gives your BDT visibility and credibility across multiple levels of government, as well as the ability to take strategic risks. In turn, this can generate demand as agencies begin thinking about how they might work with the BDT. A good champion can also ensure that your BDT is high-profile without becoming a political hot-button and that it is properly positioned within a complex system.

“Behavioral design teams help push city governments to rethink how programs are run, and bring behavioral science as a new tool to the policy making tool set. The team’s insights often lead to low-cost, easy-to-scale, and effective interventions that improve our programs.”

—Chris Wheat, Chief Sustainability Officer
& Senior Policy Advisor, Chicago Mayor’s Office.

Of course, a BDT cannot operate without a financial investment, so identifying and securing funding is the next step. BDT funding can come from the city budget itself, from outside organizations or foundations, or from some combination thereof. Philanthropy often provides initial seed funding to demonstrate the value of BDTs. In Chicago, the MacArthur Foundation funded the Chicago BDT’s first two years, and the City has been funding it since. In New York City, the Laura and John Arnold Foundation funded the first two years of the NYC BDT, while the City funded a parallel

behavioral design portfolio with the City University of New York (CUNY) that focuses on student success and persistence. These types of cost-sharing structures can ensure a city’s investment in its BDT’s results and can demonstrate a city’s commitment to institutionalizing behavioral design as a standard policy approach and operational philosophy.

Another critical piece of the funding puzzle is who will actually be conducting the work. The fundamental question here is “make or buy?”—and the answer will be determined by the existing behavioral capacity in your city’s workforce. Since behavioral design demands a variety of specialized skills (see the “Composition” section below), many cities may initially find that their BDT will require external hiring or (as in the case of Chicago and NYC) a contractual arrangement with a firm that already has behavioral design capacity. Even if the BDT is staffed externally to begin with, there is always the option to build internal capacity over time and eventually move the staffing in-house, especially as more members of your workforce become familiar with the behavioral design approach.

► **Location**

Where a BDT is housed within government will differ based on the individual city. New York City located the BDT in the Mayor’s Office of Operations, and because of the City’s focus on equity, the BDT is directly overseen by the Mayor’s Office for Economic Opportunity. Other cities may choose to locate a BDT within an Innovation, Science and Technology or Chief Data Officer’s unit. In Chicago, the BDT operates directly out of the Mayor’s Office after being initially piloted within the Mayor’s Innovation Team. There are three key considerations when choosing your BDT’s location:

HOW TO PICK YOUR “HOME” AGENCY		
ACCESS	AUTHORITY	AGILITY
Select an agency that is well connected and respected across the city. You’ll need them to help with publicity, introductions, and relationship-building.	Select an agency that has political capital. You’ll need their help to set a strategic agenda, make room for risk, and exert pressure to get things done if needed.	Select an agency that can move and think fast. You’ll need them to help you prioritize the work, navigate collaborations, and find and distribute funding.

Access. A BDT should serve as a resource to any city agency looking to implement behaviorally-designed interventions. Thus, the BDT’s home agency should be able to quickly disseminate information and should have numerous interagency connections to maximize its reach. These connections are critical when a BDT is first established to raise awareness about the new resource. Connections are also valuable when scoping new potential projects; an agency that is knowledgeable about city operations is best positioned to identify and capitalize on key opportunities. When a BDT becomes a more recognized resource, being housed in an “intuitive” agency will help facilitate inquiries from across the city.

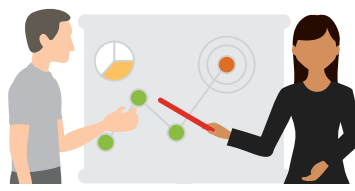
Authority. The “home” agency will determine a BDT’s strategy and how it aligns with the strategy of the city as a whole. If a city’s current priority is economic mobility, its BDT portfolio should reflect that priority by including projects that increase benefits access and educational opportunity, or that improve workforce development and housing. Such an aligned portfolio doesn’t always occur organically, and a home agency may need to proactively shape the BDT’s portfolio. This is an easier task if the home agency has the authority and political capital to prioritize potential projects and parcel out the scarce resource of the BDT’s time. Authority is also important when projects become stalled and pressure is needed to create forward momentum. Furthermore, since innovation is at the heart of behavioral design, BDT projects also involve risk of failure. A good home agency will provide the political capital that gives the BDT portfolio permission to have some failures. Since your BDT will be committed to finding out what works, and for whom, negative or null results should be celebrated as they help winnow away ineffective practices.

Agility. The BDT operating environment is fast-paced. As we will discuss in the next section, BDTs are constantly scoping new opportunities with agencies interested in a partnership. These opportunities often arise unexpectedly and with tight timelines. A home agency needs to be able to quickly assess and prioritize new opportunities and provide guidance to its BDT. Sometimes opportunities require implementation funding, so the home agency must be able to assess these requests and approve and disburse funds rapidly. Being able to gauge a potential project’s priority level and help direct the BDT’s finite bandwidth is a key function of the home agency. Also, a home agency must be strategically and intellectually agile enough to help the BDT creatively navigate the inevitable roadblocks that arise in cross-agency collaborations generally and in running pilots and field experiments specifically.

➤ *Composition*



Behavioral Science

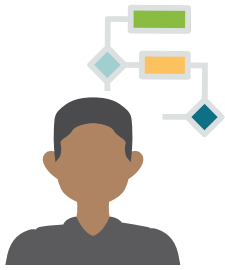


Research and Evaluation



Public Policies & Programs

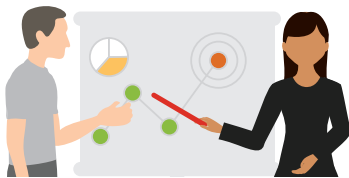
Once you have a champion, some seed funding, and a home for your BDT, you’ll need to actually assemble a team. In order for a BDT to be successful, it needs to be composed of the right people. While there is no exact formula, a well-staffed BDT needs expertise in three key areas: behavioral science, research and evaluation, and public policies and programs. This is the same set of competencies ideas42 uses to assemble its own staff, and our experience tells us that you’ll rarely find all three in one person—hence the concept of the Behavioral Design **Team**. We suggest that you plan to have a group of people with complementary skills as you staff your own BDT.



Behavioral science. Unsurprisingly, the core competency for a BDT is experience and expertise in applied behavioral science. This competency means more than just a knowledge of key biases and psychologies; an applied behavioral science acumen requires both a deep understanding of the theories underlying those biases and psychologies, as well as an ability to translate theories and research into effective designs adapted to messy, real-world contexts. Once uncommon, field-experiment-oriented

applied behavioral science is now widely taught in universities across the world,⁵ and insights from laboratory experiments are being applied to diverse problems and populations. The ability to bridge theory and practice will enable a BDT to quickly identify barriers to desired behaviors and to apply insights in unique ways. Individuals with formal training in disciplines such as behavioral economics and social psychology are most likely to exhibit this competency.

While your BDT does not need to be directly affiliated with an academic institution, it is helpful for your team to have professional networks that include behavioral science researchers. These academics can keep the BDT informed on new research and can also yield opportunities for expert consultation and collaboration. With ideas42's network of academic affiliates, the NYC BDT is able to quickly engage experts as needed: Omid Fotuhi, PhD, a researcher at Stanford University, collaborated on an intervention at CUNY focused on improving student persistence, and Ben Castleman, PhD, a professor at the University of Virginia, provided guidance on an intervention focused on helping students who were accepted to college successfully matriculate. A strong academic network will provide the perspectives necessary to ensure that designs reflect the latest learnings in applied behavioral science.



Research and evaluation. This competency enables BDTs to both generate research-based behavioral interventions and to implement and evaluate them. At ideas42, we believe that good behavioral design starts with the presumption that context matters. We also believe that good behavioral design demands

rigorous evaluation. Even if a particular intervention works in one city or environment, we can't be sure it will work in another city or environment, even if the problem seems the same. We must evaluate whether interventions are effective and whether they have different effects for different subgroups.

Because BDTs are often faced with quickly-changing parameters and deadlines, it is difficult and costly to hire an external evaluation firm to test the effectiveness of interventions. Instead, the team itself should possess expertise in evaluation design and execution. At least one member of your BDT ought to have real-world experience evaluating policy, program, or product effectiveness.

⁵ Some examples of university behavioral science programs include: Princeton University's Kahneman-Treisman Center for Behavioral Science and Public Policy, University of Virginia's Nudge4 Solutions Lab, University of Chicago's Center for Decision Research, Carnegie Mellon's Department of Social and Decisions Sciences, Harvard University's Behavioral Insights Group, and the University of Pennsylvania's Center for Health Incentives and Behavioral Economics.

You'll need expertise with experimental RCTs, as well as quasi-experimental techniques, like regression discontinuity and difference-in-difference models. Formal training in impact evaluation and statistics are an excellent foundation, as long as that training is complemented with real-world implementation experience. Even sophisticated modeling and lab experiment experience are unlikely to prepare you for the real-world intricacies of operating in the complex environment of local government.



Public policies and programs. An understanding of governance and public services is fundamental to any BDT functioning in the public sector. Ultimately, a BDT's mission is to help a city better serve its residents; that requires an understanding of how government functions and how policies and programs are designed and implemented. Enhancing the impact of local government is predicated on mastering the systems and processes of local government. This requires a panoply of skills: navigating politics and power, operating with constrained resources and the pressure of public scrutiny, and generally being keyed-in to the needs and norms of your community. You might think of this skill-set as “speaking the language” of the (local) public sector.

This fluency in government processes, policies, and programs can be acquired through a variety of professional backgrounds. You may have members of your BDT who worked in government positions or in nonprofits that implemented government programs. You may have members with experience as political appointees or as career staffers. You may have policy advocates or front-line staff. Regardless of their exact experience, your BDT members need to be able to quickly and easily understand the practicality of proposed interventions in your local context and to identify and mitigate potential pitfalls.

To understand the diverse skills needed on a BDT, consider this project example from New York City.



Increasing diversity at FDNY

Every four years, the New York City Fire Department (FDNY) begins a highly-competitive recruitment process for new firefighter candidates. Historically, there have been disproportionately fewer female and minority-group candidates, and the FDNY has worked for years to increase firefighter diversity through a variety of outreach efforts. In 2017, FDNY ramped up these efforts with a \$10 million firefighter recruitment campaign. As a result, a record-setting number of people took the firefighter exam and—for the first time ever—people of color comprised a majority of the test takers, and more women took the test than ever

before. As part of this campaign, the FDNY engaged the NYC BDT to design and test ways to increase the number of applicants from underrepresented groups. The NYC BDT's pilot targeted the first step in FDNY's recruitment process—filing to take the civil service exam.

In past recruitment cycles, many potential applicants who initially express interest in joining FDNY fail to file to take the written exam, which requires a \$30 fee. This drop-off is most pronounced among female and minority group prospects, but the reason for the drop-off is not obvious. The BDT, familiar with relevant behavioral science literature, noted evidence that small bureaucratic hassles (like paying a \$30 fee) can discourage underrepresented groups from completing processes and achieving positive outcomes. Marginalized groups are more likely to interpret bureaucratic hassles as a reflection of their own shortcomings and lack of potential, rather than as routine annoyances. This response has been shown to undermine motivation, persistence, and performance.^{6,7,8} Based on this insight, FDNY and the NYC BDT agreed to test their hypothesis that the exam fee might be deterring female and minority prospects from initiating the process to join the firefighter force. To do so, they developed an experiment to evaluate the effect of removing the fee from the exam filing process.

However, the Department of Citywide Administrative Services (DCAS)—the agency tasked with administering civil service exams—could not simply remove the fee for a randomly-selected group of applicants. That would require a system-wide overhaul and a formal proposal to the system's contracted programmers, a costly and potentially lengthy request. Instead, DCAS and the NYC BDT came up with a different solution. Instead of reprogramming the system, DCAS used an *existing* fee waiver, designed for applicants who are not employed during the filing period, are full-time unemployed students, or who meet certain income requirements, to pre-approve a random sample of potential filers to be exempt from the exam fee. Those applicants were notified of the waiver by email. This design allowed the BDT to compare the test-filing rate of applicants who received waivers to the test-filing rate of a random sample of candidates who were required to pay the usual exam fee.

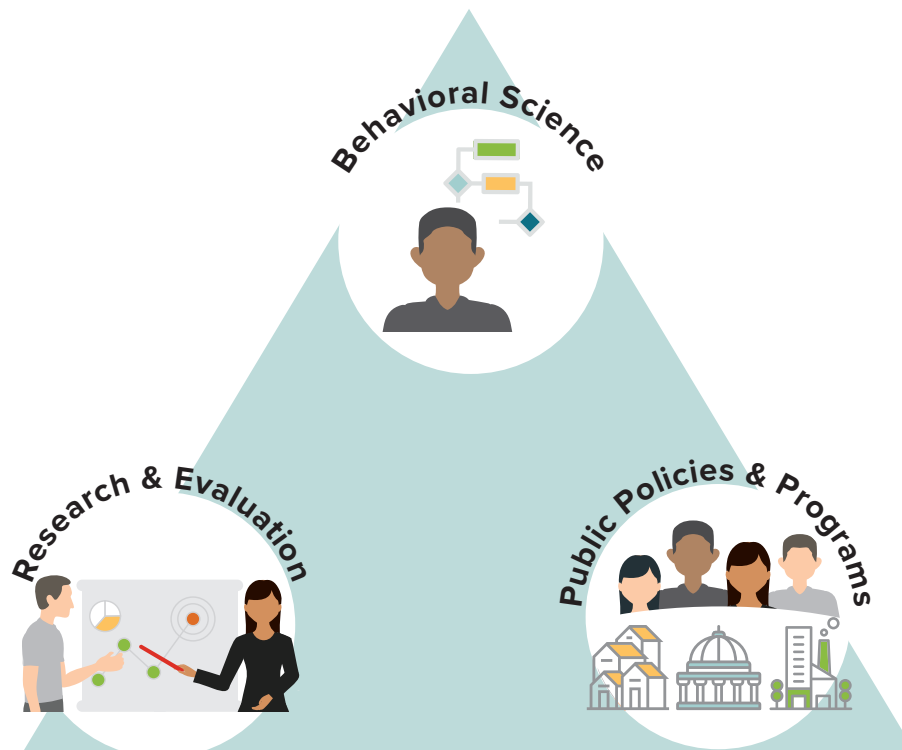
The results showed that firefighter applicants who were offered a waiver were 4.2 percentage-points more likely to file than those who had to pay the fee—a 37 percent increase in filing rates among all applicants. Perhaps more relevant to FDNY's goals, black and female applicants who were offered a fee waiver were 6.3 percentage-points (84 percent) and 4.2 percentage-points (83 percent) more likely to file for the exam, respectively.

⁶ Reeves, S. L. (2015). *Caught up in red tape: bureaucratic hassles undermine sense of belonging in college among first generation students* (Doctoral dissertation).]

⁷ Bettinger, E. et al., (2012). "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block FAFSA Experiment." *Quarterly Journal of Economics*, 127 (3)(1) 1205–1242.

⁸ Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). "Signaling threat how situational cues affect women in math, science, and engineering settings." *Psychological Science*, 18, 879-885.

This project highlights the need for all of the BDT's competencies. Knowledge of the behavioral science literature suggested a likely behavioral bottleneck to workforce diversity. Fluency in government processes helped the team arrive at a creative and practical way to deliver the fee waivers, despite the limitations of the system. And a seasoned evaluation skillset ensured that the experiment would tell the team what it needed to know about the effect of the fee waivers, despite the challenge to rigorous testing. Assembling the right team ensures that your BDT can handle the messy complexities of real-world innovation.



Operating a BDT

Once you are ready to launch your BDT, you should plan out the details of three operational areas to ensure its success. First, you'll need to **define the scope and scale of your BDT**—*how many projects you can handle, what project types your portfolio will include and exclude, and how you'll sequence your work*. Next, you'll need to **build your network**—*cultivate potential partners, develop a plan for the dissemination of results, and build awareness and appetite for your work*. Finally, as you get ready for launch, you'll need to **establish your operational procedures**—*select a project staffing model, define a reporting and accountability structure, institute implementation and evaluation procedures, and articulate a method to prioritize and select individual projects*. Below we offer some observations about how you might approach each of these operational considerations.

► *Defining your scope & scale*

Your first order of business will be to decide how much work your team can handle and what kinds of projects it will take on. Two constraints to consider are your staffing level and your mix of project types. Each project should have a dedicated point person who owns project execution—your team member(s) with public policy experience will usually be the go-to choice for these project management duties. Your staff with specialized skills can act as advisors and can more easily work across more projects than your project manager(s). In terms of staffing levels, ideas42 has a general rule of thumb that each full-time project manager can handle two to five simultaneous engagements, depending on the projects' intensity and the level of support staffing. If you have time-intense projects with complex implementation and partner management needs, you'll either need more staffing or a lower overall number of projects. We have found that it can be helpful, especially at the beginning of a BDT, to start with fewer projects than you think you can handle—this lets you iron out kinks in your process while ensuring that you deliver high-quality work.

Another consideration is your project mix and sequence. If you only run complicated RCTs, each of which requires months of implementation and data collection, then you may not be able to generate the buy-in and enthusiasm a BDT needs in its early days. Mixing in some shorter engagements, including simpler RCTs or projects with non-experimental evaluation strategies, can be helpful in sustaining momentum and generating evidence that your BDT can get results. Finally, knowing how many projects your BDT can handle will help you focus your energy as you begin building your network and selecting your projects.

► *Build your network*

As your BDT launches, you'll need to raise awareness about your team. Depending on your scope and scale, you may want to conduct mass outreach to a broad array of city officials or you may opt for more targeted outreach to agencies whose work is amenable to behavioral design. As your team gets up to speed, you will need a good pipeline of potential projects for them to work on, so having a robust network throughout the city will allow you to constantly scope opportunities. To build credibility and momentum, you may want to identify an “anchor” agency—a high profile partner that can help you establish quick wins with good visibility. Having an anchor agency is helpful because it validates and legitimizes the BDT and helps reduce the perceived risk of working with your team. The first handful of projects your BDT launches will help other agencies understand what it means to work with the BDT and what kinds of outcomes they can anticipate. You should also establish communications procedures early on, both internally and externally. Sharing results and lessons learned widely should be part of your mission, but be clear with each agency, and with city leadership, about what kinds of results dissemination you would like to do. You should identify venues to share lessons learned across agencies to ensure that word spreads about your BDT, about emerging best practices, and about opportunities for cross-agency collaboration. On the external side, it is also critical to establish what level of publicity the city (and each agency) wants.

► Structure your operations

The final aspect of launching a BDT is structuring your operations. As discussed above, you will need to match your scope to the size of your staff and the priorities in your city. Inherent in those staffing decisions is also a basic reporting structure. **Who will lead your BDT? Who will they report to within the city? If you are externally contracting your BDT, how will those parallel reporting structures interact, and what work will still need to be done in-house?** As a team that works on improving city government, someone within city leadership will need to have final say over the projects and their trajectory, so be sure to establish your reporting, decision-making, and accountability structures at the outset. Also, as projects end, you will need procedures to analyze their outcomes. It is best to create procedures up front that will allow you to quickly analyze data—both for overall impact, as well as for cost-effectiveness. Establishing data collection and reporting procedures (including data privacy needs) **before** starting any project ensures you can always assess a project's impact, including any associated cost savings. These analyses are the basis of any subsequent effort to replicate and scale promising interventions.

Project selection criteria

As your BDT becomes more popular, you may face the need to prioritize projects as they compete for your limited time. So, the final element to a successful BDT is having clear criteria to prioritize and select the projects you'll pursue. Your goal should be to direct your energy and resources to priority issues that are amenable to behavioral design. The NYC BDT and the Chicago BDT use the criteria below to prioritize and determine the viability of potential projects. Of course, because context matters, the exact criteria and their respective weighting will vary from place to place. Below, we explain each criterion and illustrate them with an example from NYC.

PROJECT SELECTION CRITERIA				
1 City Priority + Social Impact	2 Agency buy-in + capacity	3 Clear touchpoint	4 Existing data	5 Randomization + large sample size
Pick projects that align with your city's priorities for the well-being of its residents. Prioritize projects that have the biggest potential for impact.	Work with agencies that are excited about behavioral design. Make sure they have the resources and will to manage projects and make changes.	Be sure that your partner agencies have full control over touchpoints with end-users—these could be communications, processes, or even physical environments.	Ensure that you have access to administrative data that can tell you about the outcome(s) you care about. Work with existing data whenever possible.	If you run experiments to evaluate impact, be sure that you can randomly assign people to different conditions, and look for large sample sizes.

► *City priority and social impact*

The aim of a city BDT is to ensure that the government can accomplish its goals and maximize the impacts of new or existing programs. By focusing on priority areas, BDTs ensure that they're tackling problems in need of innovative solutions and contributing to issues that have the city's attention. An important component for ideas42—as a nonprofit organization with a social mission—is that our projects have a positive social impact. By working within city governments, which are inherently oriented toward this goal, we can be sure that every BDT project fits this criterion.

G&T example: One of the priorities of the Office of Student Enrollment (OSE) in the NYC Department of Education (NYCDOE) was to increase the number of students from low-income districts who test for the Gifted and Talented (G&T) program. The G&T program offers access to high-quality educational opportunities that may not be available in a student's zoned school. This priority aligned well with ideas42's social mission and the Mayor's broader priorities around equity.

► *Agency buy-in and capacity*

Since BDTs are embedded within city government, agencies can freely collaborate with them as they would with any other city agency. However, engaging with a BDT is not totally “free,” as agencies do incur project management costs. Behavioral design projects require access to end-users, administrative data, and institutional knowledge. To ensure these costs can be borne, partnering agencies must have buy-in at all levels—from leadership down to front-line staff. This is especially true for client-facing staff, who are often the most knowledgeable about the community context and program processes.

G&T example: OSE leadership supported its G&T team's partnership with the NYC BDT and approved the necessary staff time and budget required to implement the project. Neither the required staff time nor the budget was burdensome, but preapproval was helpful. Early buy-in helped the BDT engage support, technical, and front-line staff who were not on the G&T team, but would need to spend some of their time on the project. For example, the NYC BDT worked with Family Welcome Center staff to better understand the challenges low-income parents face. Carving out capacity was particularly important for the G&T team itself, which was responsible for weekly update calls and coordination with staff to prepare for, implement, and test the intervention.

► *Clear touchpoint*

The third criterion for a BDT project is having a clear “touchpoint” that can influence a behavior like **enrolling** in a benefit program, **completing** a process, or **paying** a fine. These touchpoints can include things like letters, forms, emails, text messages, and other communications. Touchpoints

can also include physical environments, interpersonal interactions, or the design and sequence of a process. Whatever form it takes, the touchpoint must be fully controlled by the partnering agency so the BDT can redesign and test it. If tweaking a touchpoint requires extensive negotiation with external parties, it may not be a good choice for an intervention channel. Starting with existing touchpoints is beneficial because they generally already have data being collected on them (see the next two criteria). In addition, modifying and iterating on existing touchpoints (especially communications) can help your BDT generate effective solutions in a timely fashion. Changing existing material is often faster than creating material from scratch. Early BDT projects in Chicago and New York City focused on communication materials to achieve quick wins with agencies by proving that low-cost interventions could create outsized impacts.

G&T example: When OSE and the NYC BDT initially scoped the project, the discussion focused on how to best reach families. Typically, OSE communicates about G&T testing through open houses, postcards, and emails. An open house intervention would be difficult to test and would restrict the intervention to families who were able to attend. Postcards and emails, on the other hand, were easy to alter, easy to test, and reached a wider population of families. Using the touchpoint criterion, NYC BDT chose to redesign OSE's outreach postcard and email.

► *Existing data*

Because behavioral design requires testing an intervention's impact, it also requires data. Unfortunately, collecting data can be onerous, especially when new data collection channels—like exit surveys or attendance tracking—need to be established. Because BDTs aim to innovate quickly, their goal is to complete a project in a matter of months rather than years. To maintain both celerity and rigor, we choose projects for which outcome data are already being collected. Relying on existing administrative data—and not adding any new data collection processes to an agency's already full plate—means we can measure the impact of our designs against “business as usual” as quickly as data become available.

G&T example: To measure the effect of our redesigned postcard and email on G&T sign-up and testing rates, we used data OSE was already tracking: whether a parent signed his or her student up to take the exam, whether a student actually took the exam, and whether a student applied and was accepted into the G&T program. OSE never had to change their data collection procedures for the NYC BDT's interventions.


► *Randomization and large sample size*

These last two criteria are important if you wish to pursue the most rigorous method of impact evaluation: the RCT. The first criterion is the ability to randomly assign people into different groups that will receive or experience different things. When groups are assigned at random, they are, on

average, the same. As a result, any differences in outcomes can be attributed to the intervention and not to chance or other unknown variables. The second criterion is having a large enough sample to ensure that relatively small outcome differences we may see between the two (or more) groups are “statistically significant.” That means we can attribute any observed outcome differences to our intervention with a **high level of statistical confidence**. Sample size requirements will vary depending on your randomization approach and the magnitude of the impact you hope to detect. The smaller the anticipated impact, the larger your sample will have to be to detect it. Tools like Minimum Detectable Effect calculators—which can be [easily accessed online](#)—can help you choose a sample size that will validate even small differences between groups.

One note of caution—RCTs are not the only way to understand how well your BDT’s interventions work. RCTs are our default approach because they provide us with the most certainty. Some projects can’t go forward with an RCT, but are still worth running because of their potential impact. Assembling a BDT with strong research and evaluation skills will ensure that you choose an impact evaluation approach that is appropriate for the context. Even if you opt not to run RCTs for every intervention, you should be sure that your partners have both an ability and willingness to evaluate how well interventions work.

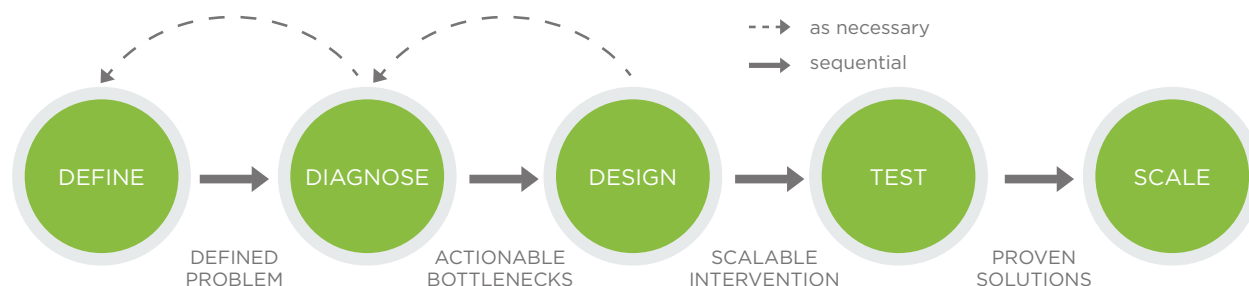
G&T example: Close to 70,000 students enter kindergarten each year in NYC and can test for the first time into the G&T program. This large number of students allowed us to operate an RCT that tested the effect of both our postcard and our email. We only had to run the intervention once, and neither channel (postcard or email) was withheld from any parent. We were also lucky that OSE had a household database that allowed us to 1) randomly assign households to receive some combination of the original and new postcards and notices, and 2) track and compare G&T sign-up rates (among other key actions) across the groups. Our analysis showed that the behaviorally-designed email increased test registration by 6.5 percent.

<p>Dear Daniel,</p> <p>If your child will be entering kindergarten through grade 3 for the 2016-2017 school year and you are interested in Gifted & Talented (G&T) programs, please submit a Request for Testing (RFT) form online by 11:59pm or in person by 3pm on Thursday, November 12. Submitting a RFT by the deadline gives your child the opportunity to test for eligibility to apply to G&T programs.</p> <p>There are two ways to submit a RFT:</p> <ul style="list-style-type: none"> • Online • In person. If you prefer a paper RFT, print and complete the RFT form in the G&T Handbook and submit it in person. If your child attends a NYC public school, submit the form to your child’s school. If your child attends a nonpublic or charter school, submit the form to a Family Welcome Center. <p>Be sure to save the receipt whether you apply online or in person!</p> <p>For more information about G&T Admissions, visit our website or call 718-935-2009. If you already submitted a RFT, please ignore this email.</p> <p>The Gifted & Talented Admissions Team</p>	<p>Dear [NAME],</p> <p style="text-align: center;">Congratulations!</p> <p>As you prepare for kindergarten next year, I would like to invite [STUDENT NAME] to test for New York City’s Gifted & Talented (G&T) program.</p> <p>Submit a request for testing TODAY and give your child the opportunity to have:</p> <ul style="list-style-type: none"> ✓ <i>Enriched curriculum and teaching</i> ✓ <i>Access to schools outside your zone</i> ✓ <i>Classroom of engaged learners</i> <p>Join other families in your neighborhood already taking advantage of the G&T program.</p> <p>Sign up in less than 5 minutes:</p> <div style="border: 1px solid black; padding: 5px; display: flex; align-items: center;">  <div> <p>1. Create an account <small>(You just need an email address)</small></p> <p>2. Submit a request for testing <small>(You just need basic information like name and address)</small></p> </div> </div> <p><small>*If you prefer a paper form, print and complete the form in the G&T Handbook and submit it in person to a Family Welcome Center.</small></p>
--	---

Original

Behavioral

The behavioral design process



Once a project meets your selection criteria, it's time to launch into the behavioral design process. At ideas42 we have developed a five-stage behavioral design process that our BDTs follow as well. The first step is to **define** the problem and try to remove any embedded assumptions about *why* it may be occurring. Using quantitative and qualitative methods, we next **diagnose** what behavioral bottlenecks may be driving the problem. We then **design** interventions that directly address the key bottlenecks we've diagnosed. To **test** whether our design successfully addresses the problem, we conduct an evaluation, typically an RCT. Finally, if an intervention proves effective, we look for opportunities to **scale** the solution to a larger population or adapt it to other contexts.

While this process is linear in concept, it is much more circuitous in practice. As we progress through each phase in the process, we learn more about the problem, the population, the context, and the constraints of the system and its touchpoints. At times, new information will prompt us to return to a previous design stage, or to adjust our strategy within the current stage. In short, the process is both sequential and iterative. Below, we lay out the conceptual basis for each stage of the behavioral design process and use the NYC BDT's first SNAP project to illustrate the real-life behavioral design process in action.

Define. Sometimes, an agency's problem may seem obvious, but how do you know the proposed problem is actually driving negative outcomes? For that matter, how do you know that behavioral design can help? Is it a problem of uptake? Follow-through? Retention? Compliance? Performance? A poorly-defined problem makes successful behavioral design impossible. An agency may believe that certain trainings are under-attended because not enough people are signing up to take them. If this is true, you might try solving the problem with redesigned outreach content. However, if people are actually signing up for trainings, but then failing to attend, your improved outreach is unlikely to be effective. Preliminary research, typically through administrative data analysis, can help confirm the accuracy and magnitude of the problem that an agency is hoping to solve and can help direct your efforts in the diagnosis stage.

To guide problem definition, you and your partner agencies should develop a problem statement that identifies a specific behavior among a group of people. Problem statements shouldn't make any assumptions about *why* a problem exists; that is addressed during diagnosis. A strong problem statement for our example above might be: entrepreneurs register for business courses, but do not attend the first class. The behavior is specific (people sign up, but don't attend); the population is specific (entrepreneurs who register for the class), and there's no presumption about *why* they are not attending, nor about what an appropriate solution would be. By restricting the problem definition statement in this way, you are well-positioned to take a focused and behavioral approach to solving the problem.

SNAP Example: The New York City Human Resources Administration (HRA) identified “SNAP churn” as a potential behavioral problem. Many clients failed to recertify their benefits from the Supplemental Nutrition Assistance Program (better known as SNAP, or food stamps), but returned within a few months. This has many drawbacks: people lose food assistance, and to get it back they must re-apply—a more onerous process than recertification. Reapplication is costly for the clients and for HRA. To develop a workable problem statement, the NYC BDT broke the recertification process into three behaviors that aligned with the three recertification steps: 1) submit the recertification form, 2) complete the interview, and 3) (if information has changed) submit verification documents. After analyzing administrative data on where most clients drop out of the process, the NYC BDT decided to focus on the first step of the recertification process: submitting the recertification form. The problem statement became: SNAP recertification clients who are eligible to continue receiving benefits do not submit their recertification forms.

Diagnose. Once you have a problem focused on specific behaviors among a population, you can move on to diagnosis. You can now generate hypotheses about why the problem may be occurring. At ideas42, we look for behavioral bottlenecks that influence the target population in regards to the target action. To do so, we use a proprietary technique called “behavioral mapping” that pinpoints discrete decision and action points in a process. For each of these points, we use insights from behavioral science to identify psychological factors and contextual features that might affect choice and action. We try to connect elements of the physical or psychological environment—like the timing of a reminder, access to transit, or visual cues in an office—to the suboptimal choices and actions we see in our population—like forgetting steps in a process, missing appointments, or dropping out of a program midway.



FIA-1152e (E) 05/04/2016 LLF

Date: _____
Case Number: _____
Case Name: _____

People are more likely to take action when renewal is framed as a potential loss.

Reminder Don't Lose Your SNAP Benefits!

Formatting draws attention to the fact that there is immediate action to take to avoid loss.

Your SNAP benefits are about to expire. To avoid losing your SNAP benefits, submit a recertification form. There are different ways to submit your recertification form, but the easiest and fastest way is to submit it online – and you can do that RIGHT NOW!

First, Submit Your SNAP Recertification Form.



- 1 Go to www.nyc.gov/accessnyc
2 Log into your account (or set one up)
3 Click "Apply Now"
4 Click "I would like to recertify my active SNAP (Food Stamps) case"
5 Fill out your information and submit your form

Steps are outlined in a clear list for easy processing.

If you have already submitted your recertification form, you can go right to the next step! Visual helps reinforce the desired action.

Your Recertification Form, Call Us!



After you submit your recertification form, call us at 1-718-SNAP-NOW (1-718-762-7669) Monday – Friday, 8:30am to 5:00pm to have your telephone interview.

You must complete your interview by [blank] or you may lose your benefits. We no longer schedule your interview. This means you have to wait for us to call you. Just pick up the phone and call us at your convenience!

A specific deadline spurs action, particularly when paired with loss framing.

We treat these connections as hypotheses we can confirm or refute using quantitative and qualitative methods. This may include analyzing data to find drop-off points, looking for disparate outcomes among subgroups, observing interactions in an office, or running structured interviews with clients. Unlike some research that asks a lot of “why” questions, our approach asks “how” questions. We focus on process because people often have inaccurate or incomplete interpretations of their own motivations, particularly in retrospect. If any of us had perfect access to our mental contents, cognitive biases wouldn’t be so powerful. Once you’ve winnowed your hypotheses to the likeliest culprits, you can begin the design stage.

SNAP Example: To diagnose the behavioral bottlenecks associated with submitting the SNAP recertification form, the NYC BDT conducted behavioral mapping on the first decision-action point in the SNAP process: deciding to submit the form and then actually submitting it. The team generated several hypotheses about why clients may fail to make a decision around submission, and about why people might make the decision to submit, but ultimately fail to do so.

To test these hypotheses, the NYC BDT visited SNAP centers to conduct observations and to interview SNAP clients and HRA staff. The NYC BDT also visited HRA call centers to observe HRA staff conducting SNAP phone interviews. This qualitative field research confirmed several key bottlenecks to recertification, including ambiguity about the recertification process as a whole, the accumulation of small hassles over time, and the tendency to procrastinate.

Design. As you begin to devise your intervention, remember that your designs should always flow from your diagnosis. A pure reminder may be effective if your diagnosis indicates that people forget to perform a task, but will be ineffective if they never planned to complete the task in the first place. Since the diagnosis stage often reveals multiple bottlenecks, we typically leverage a range of touchpoints and a variety of behavioral techniques in our designs. Intervention designs take many forms—from small tweaks to existing programs to comprehensive process revamps. The scope of an intervention ultimately depends on the behavioral bottlenecks you identify and the capacity of your partner agency. Many practical factors will also influence your design, including your timeline, available resources, and the potential for the intervention to scale. Considering factors like these will help you establish a feasible design scope, including preferred intervention channels. Be sure to negotiate these details with your partner agency prior to beginning the design stage.

Before moving to the test stage, you should create prototypes for user testing. User testing helps simulate how end-users will interact with the intervention and will offer valuable insights that can improve the design. User testing may also reveal new behavioral barriers that require you to revisit the diagnosis phase. Remember that interventions can evolve significantly over time, and toggling between diagnosis and design is a normal part of the process. As you finalize your design, interventions should be shared with your partner agency's key internal stakeholders such as its legal, financial, and operational teams.

SNAP Example: Early in the behavioral design process, the NYC BDT and HRA chose recertification reminder notices as the intervention touchpoint. At the time, HRA's primary point of contact with clients was physical mailings, which made a letter a natural choice for an intervention. The NYC BDT collaborated with HRA to draft a reminder that addressed the key bottlenecks identified during diagnosis. To combat procrastination, the BDT used loss framing ("Don't lose your SNAP benefits") to convey urgency and made the call to action immediate ("You can do that RIGHT NOW"). Ambiguity and hassles were addressed by composing a highly-graphic notice with clear sub-steps, as well as a look forward to the subsequent interview step. Quantitative data on submission timing allowed the BDT to identify the ideal moment in the recertification period to send the reminder. HRA facilitated conversations with the mailing and information systems teams to ensure that the design was compatible with existing systems.

Test. Once your design is complete, it's time to find out whether it works. To determine whether an intervention is successful at addressing the problem it was designed to solve, you must conduct an evaluation. There are many ways to evaluate impact, but the gold standard (and ideas42's default approach) is the RCT. As mentioned previously, RCTs demand a large sample size and the capacity to randomize. Contrary to their reputation, however, RCTs can be run quickly and relatively inexpensively. Most NYC BDT projects use RCTs as the evaluative method, and very few have lasted more than a year, and almost none have required direct costs exceeding a few hundred dollars. Even if they are more complicated than your city's typical approach to policy or program improvement, RCTs are worth the extra effort. By randomly assigning people to get either the newly-designed intervention or business-as-usual, you can tell whether your change is an improvement over the status quo by seeing which group has better outcomes. Without rigorous testing, you can't know if your changes make a difference—good or bad. Without knowing what works, you can't ensure your residents are getting the best possible services.

Sometimes RCTs aren't feasible or appropriate. In such cases, other rigorous, quasi-experimental evaluation designs should be used. Social science offers techniques like difference-in-difference models to measure impact under complex circumstances. For example, the Chicago BDT implemented a behaviorally-designed disposable bag tax that was rolled out citywide. Because everyone in the City would be subject to the tax, an RCT wasn't possible. Instead, the team used

Population



RANDOMLY ASSIGN

Control

Treatment



Introducing OnDemand Supplemental Nutrition Assistance Program (SNAP) Requalification Telephone Interviews

We are pleased to announce the availability of OnDemand SNAP requalification telephone interviews. Soon, we will send your SNAP requalification package in the mail. It will include information about the different ways you can requalify your certification form.

After you submit your requalification form, you need to contact us to complete your requalified interview. In order for your benefits to continue without interruption, you should complete your interview by **Month Day, Year** (see date listed on ZIE). We no longer schedule your interview.

Get assistance in 3 easy steps:

- Step 1:** Call **1-718-SNAP-NOW** (1-718-633-6662), Monday through Friday, 9:00 AM - 6:00 PM.
- Step 2:** Select the language you want to use for your interview.
- Step 3:** Send how you want us to find your case and enter the information for the option you selected.
 - Social Security Number (SSN)
 - Client Identification Number (CIN)
 - ACCESS NYC Confirmation Number
 - SNAP case number

How does OnDemand change the requalification interview process?

- You must contact us to complete your requalification interview.
- You can contact us during business hours at a time that works best for you.
- You no longer have to wait around for us to call you for your requalification interview.

Some thoughts to keep in mind...

- If you mail or fax your requalification form, please allow a few days for mail processing before calling.
- If you need an in-person interview at your SNAP Center, please call 1-800-456-3113 to schedule it.
- Interviews are available Monday through Friday, from 9:00 AM - 6:00 PM.

Original Letter

NYC Department of Social Services

Don't Lose your SNAP Benefits!

Your SNAP benefits are about to expire. To avoid losing your SNAP benefits, **SNAP requalification & identification steps**. There are different ways to requalify your requalification form, but the easiest and fastest way is to submit it online—and you can do that **RIGHT NOW!**

First, Submit your SNAP Requalification Form.

- Go to www.nyc.gov/accessnyc
- Log into your account (or set one up)
- Click "Apply Now"
- Click "I would like to requalify my active SNAP (Food Stamp) case"
- Complete and submit your form

After Submitting Your Requalification Form, Call Us!

After you submit your requalification form, call us at **1-718-SNAP-NOW (1-718-762-7669)** Monday - Friday, 9:00am - 5:00pm to have your telephone interview.

You must complete your check-in by **12:00:00 AM** on the day of your interview or you may lose your benefits. We no longer schedule your check-in. This means you won't have to wait for us to call you. Just pick up the phone and call us at your convenience!

Original + Behavioral Letter

a difference-in-differences model: comparing bag use before and after the tax went into effect in Chicago, as well as in its suburbs which had no bag tax. This analysis showed that the intervention produced a 42 percent reduction in disposable bags. Regardless of your evaluation approach, the key to the test phase is to gather reliable data about what works. You can be sure to select the right approach by having strong research and evaluation skills on your team.

SNAP Example: To measure the impact of the reminder notice, HRA used its information management system to randomly assign SNAP clients up for recertification to either receive our reminder notice plus HRA's standard notice (treatment group) or to receive only the standard notice (control group). Not sending our reminder notice represented business-as-usual. This system automatically sent the reminder notice to the treatment group on the 15th day of their recertification period. HRA's data system showed whether (and when) clients completed steps within the recertification process. The NYC BDT found that the reminder notice decreased the likelihood of failing to submit the recertification form by 5.5 percent. The reminder notice also prompted clients to submit their forms *and* complete their interviews earlier in the recertification period. Earlier submission and completion was of particular interest to HRA, which experienced a great administrative burden at the end of the recertification period when thousands of clients scrambled to complete the process.

Scale. When you find an intervention that works, your next thought will likely be about how to scale it up immediately. This is understandable, but also hasty. If a single test suggests that an intervention is successful, you should work with your partner agency to replicate it—perhaps at a slightly larger scale or in a different context. Also, you should look for opportunities to further refine the intervention before bringing it to scale. ***Are there subgroups for whom the intervention didn't work as well? Are the positive findings lumped in one neighborhood or demographic? Is there a cheaper way to execute your design?*** Adopting the behavioral design approach means that you can continually look for ways to incrementally improve your programs and policies in an evidence-based manner.

As you expand the scope of successful interventions, some (like redesigns of existing communications) will be relatively low-cost and easy. More complex interventions can have cost and labor implications at scale—especially those requiring new materials or supplemental steps to business-as-usual. Cost-effectiveness analyses can help estimate upfront and continuing costs so that agencies can make informed decisions on how and where they want to scale. Such analyses, along with the results of the impact evaluation, can also be useful in generating buy-in for scaling from other key stakeholders.

SNAP Example: Given the positive results of the RCT on the reminder notice, HRA wanted to scale the notice up to its full SNAP population. Since it was a new physical mailing, however, there were cost implications to scaling the notice to all eligible clients. To maximize impact while minimizing costs, HRA decided to send the reminder only to clients who had not yet submitted their forms by the reminder mailing date. Even though the reminder also contained important interview information, people who had submitted their forms, but hadn't yet interviewed wouldn't receive it. The team felt confident about the decision because HRA already sends a reminder to clients who have submitted forms, but have not yet completed interviews. There are important tradeoffs in any scaling effort, and this choice was driven by both cost and process considerations.

Hopefully, you now have a better sense of why behavioral design is a powerful tool for cities, what it takes to set up a BDT, how to proceed with a behavioral design project, and whether a BDT is right for your context. We want you to learn from our achievements, as well as our struggles—and we stand ready to help you set up a BDT of your own. Our hope is that this playbook will help cities across the US and around the world incorporate evidence-based, behavioral design into their daily work.

“Our first two years proved the concept of an embedded City BDT with cost-effective solutions to difficult policy problems, our next phase....will be focused on institutionalizing behavioral design, data usage, and rigorous evaluation across City agencies, and scaling proven interventions...”

— MATTHEW KLEIN, EXECUTIVE DIRECTOR,
NYC MAYOR'S OFFICE FOR ECONOMIC OPPORTUNITY

What's next?

Having impacted hundreds of thousands of lives through almost 50 combined projects, Ideas42's BDTs have shown their worth in New York City and Chicago. Not only are these inaugural teams creating social impact in their communities, they are also inspiring innovation in government. Cities across the US reach out regularly to discuss new ideas for behaviorally-informed programs, policies, and systems. New York City and Chicago have built momentum that we hope other cities will capitalize upon. As local governments get increasingly comfortable with impact-oriented experimentation; savvy data use; and a culture of innovation, evidence, and testing; the use of behavioral design in the public sector will become the norm. This is a promising, but also necessary development as local and state governments increasingly operate as the first and last line of defense for populations in need. With more demands on local budgets and systems, behavioral design may offer a lifeline of efficiency and effectiveness.

As city BDTs proliferate, we expect that city governments will institutionalize behavioral design and fully integrate it into their operations. We should see the behavioral toolkit become ubiquitous among practitioners and policymakers. We should see more government workforces with behavioral expertise, both through training on the job and acquiring talent through external hiring. We should see increased data-sharing and the spread of rapid RCT testing platforms (sometimes called A/B-testing platforms). We should see rapid-cycle RCTs become a default approach in program design and evaluation. We should see agencies using evidence and iterative prototyping to scale what works and to implement improvements on an ongoing basis. We should see policies and programs rooted in the unique needs of the communities they serve and designed with the idiosyncrasies of human choice and action in mind. And ultimately, we should see all of these strategies lead to an increase in efficiency and improved outcomes for residents.

As we look to the future of city BDTs, we acknowledge that there is still an enormous amount to learn. Despite the many successes of our teams in New York City and Chicago, we know that there is a fundamental need to get beyond nudges. The quick wins and low-hanging fruit evaporate over time, as does the pressure to deliver a successful proof of concept. As BDTs mature and evolve, they must also become more ambitious in their scope. This problem affects applied behavioral design more broadly. At ideas42 we are asking many questions about the future of behavioral design: ***How can we integrate advances in data science and machine learning to improve diagnosis and design? How can we use behavioral insights to solve complicated structural problems like racial and gender bias? What behavioral strategies will be effective in reducing community violence and increasing civic engagement? Where are the best opportunities to building behaviorally-designed systems from the ground-up? How can we implement continuous quality improvement systems that integrate behavioral science? How can we build and sustain a workforce that institutionalizes behavioral design?*** While these are all hard questions, we think that BDTs will be at the forefront of providing the answers for years to come.

