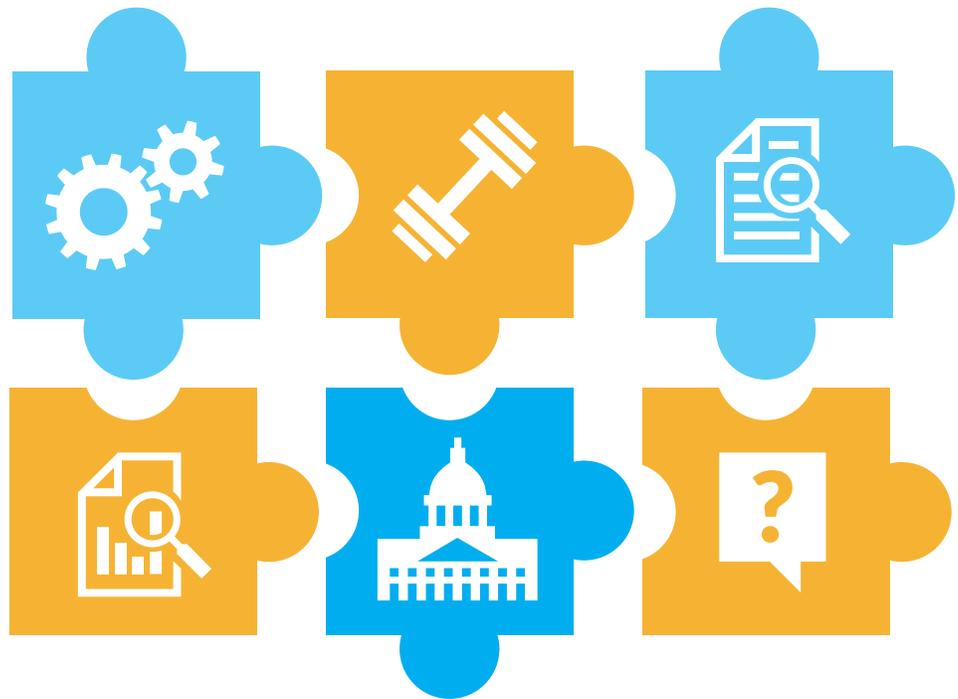Justin Milner

Matthew Eldridge

Kelly Walsh

John K. Roman

PAY FOR SUCCESS
PROJECT ASSESSMENT TOOL

Pay for Success Initiative

A. Introduction

Since its introduction, the concept of pay for success (PFS) financing has attracted attention from leaders interested in innovative social financing approaches. The model—which establishes outcomes-driven contracts between government and service providers, with up-front capital from private funders—shines a spotlight on the importance of evidence-based programs, strong nonprofit organizations, and government innovation.

PFS has significant potential as a way to help embed evidence-based decisionmaking into how cities, counties, and states provide services and responsibly test promising interventions (see box 1 for key terms). Before launching a PFS project, stakeholders—such as governments, funders, and service providers—should assess the proposal to determine if PFS financing is a good fit for the problem they want to solve, the program that could provide a solution, the provider who could do the work, and the local public leadership that will champion the effort. Given the high up-front costs of PFS planning and deal construction, engaging in a comprehensive early suitability assessment can be especially valuable.

PFS is not an approach that will work for all issues in all contexts, and there are other ways to support evidence-based policymaking that may be more appropriate for certain issues and communities. Nonetheless, thinking through the difficult questions required to develop a PFS project may itself be useful.

PFS can force governments to address common questions in social policy: What are we trying to achieve? For whom? Why do we think it will work? How do we make sure we do what we say we are going to do? And how will we know if what we invested in actually worked?

These are difficult but important questions at the root of evidence-based policymaking. This tool will help stakeholders think about both the merits of a PFS project and how to deliver strong social services.

B. Purpose of the assessment

The PFS Project Assessment Tool (PAT) helps people answer a fundamental question: What makes for a strong PFS project? It describes core elements of PFS projects, explains why those elements are important, provides a scoring system to help distinguish the strengths and weaknesses of a proposed project, and generates recommendations for improving those weaknesses.

The PAT can be used at different stages of project development:

Early stages: For projects at their inception, this tool highlights the most important elements when considering PFS. When launching a conversation on a proposed project, partners should consider the categories and questions in this tool as criteria to understand the core components and key characteristics necessary for a strong project.

Midproject development: For projects in development, the tool serves a more diagnostic purpose and can help ensure key elements are being addressed.

Late stage projects: The tool also allows partners—including potential funders—to assess the relative merits of a more developed project. At this stage, the assessment can help identify areas of strength that need to be revisited.

The PAT is designed for individuals, governments, and organizations working through PFS projects or just considering engagement with PFS. Broadly termed "stakeholders," PAT users include government officials and advisors, public agency leadership, program managers, service providers, and others who are interested in learning whether PFS might work for their community. Completing the PAT also helps build the business case for a proposed project if that project scores well in each area.

C. How to use this assessment

The assessment tool is organized by six of the major components of a strong PFS project:

1. **Problem definition:** How clearly defined and understood is the problem?

2. **Program strength:** How strong is the proposed solution to address the problem?

3. **Provider capacity:** How strong (capacity and effectiveness) is the service provider that could deliver the program?

4. **Public system partners and political landscape:** How strong is the political and bureaucratic support from the public partner (i.e. government) for the proposed project and financing structure?

5. **Project alignment:** How well do the all the key components of the proposed project fit together in a specific jurisdiction?

6. **Project evaluability:** To what extent does the proposed project have a clear, rigorous, and feasible evaluation plan?

Although the PFS project design process does not always follow this linear process in practice, starting discussion with a strategic and objective understanding of the challenging problems facing a community is an important way to avoid having a solution in search of a problem.

To assess these six components, the tool identifies key issues within each, poses a question to assess that area's strength, provides a brief explanation of the importance of the key area, and outlines a rubric to score a particular project. Stakeholders can complete the assessment in either its entirety or focus on the sections of greatest relevance to their project. In addition, the sections can also be evaluated independently outside of a PFS project–these questions will be useful for social service planning in general.

It's unlikely that any single stakeholder will be able to answer all of the questions in the PAT. As such, we recommend that the process for completing the PAT should be undertaken collaboratively among a larger team of project partners (or at least with that broader team's inputs). This also helps establish strong and clear lines of communication across the entire project partnership team, encourages a discussion on key project elements among stakeholders, and leverages the unique perspective and expertise of different partners.

D. Scoring and interpreting scores

The complete assessment has a total score of 70 possible points: five sections worth 10 points and one section worth 20 points. Within the rubric, answers that receive negative points indicate areas of extreme importance. Failing to address those areas—for example, provider unwillingness to engage in evaluation or a program that has previously shown negative results—may indicate that more thought is needed in a section of the assessment and how it relates to the project. Given that each of the six components is crucial for the proposed project's overall strength, a project needs to score well in each component.

The individual section scores will matter the most for the majority of people using this tool, and they provide more meaningful information about the potential project's strengths and weaknesses. For example, a project with a total score of 60 may appear strong, but that score may obscure the fact that it scored 0 in the "provider capacity" section, which is a serious concern.

Finally, we should note that even a very strong score on the assessment does not guarantee success in PFS. Although this assessment highlights factors important to a strong project design, several other factors related to follow-through, implementation, and specific local context also affect project success and investment case for funders.

For follow-up support on questions about the PAT, please contact the Pay for Success Initiative support center at pfssupport@urban.org or explore our web portal at http://pfs.urban.org.

E. Limitations of the Project Assessment Tool

This assessment is intended as a guide and resource to highlight good practices. The tool does not include all possible elements necessary for successful PFS projects. Rather, the tool focuses on some of the major elements that are essential to project success.

In addition, many of the questions require answers that will be difficult to ascertain with complete certainty. Informed guesses, backed by available information and candor, will form the basis for many responses. Self-scored evaluations are inherently influenced by the perspective of those completing the assessment and, as such, scores from different projects might not be fully comparable. As much as anything, scores should be useful as a starting point for conversations among stakeholders.

BOX 1

Defining key terms

Program: Also known as an intervention, a program encompasses a specific set of activities and other inputs that are delivered in a specified way and intended to yield specific improved outcomes in a target population. Different types of programs might be funded through PFS: "proven" programs to be scaled or replicated (programs that have demonstrated success through prior evaluation) and "promising" programs (ideas with a compelling theory of change but minimal existing rigorous research demonstrating their effectiveness). Example: Nurse-Family Partnership (NFP) is a program that provides home visits by registered nurses to low-income first-time mothers.

Project: The PFS project is the umbrella for the entire endeavor, including the program, contract, outcome payments, and other constituent parts. For example, South Carolina launched a $30 million PFS project that aims to scale NFP programs to serve 3,200 first-time, low-income mothers to improve child health outcomes over the course of five years.

Service provider: The service provider implements the program within the context of the project. Providers are typically independent nonprofits or coalitions, but they can also be government entities or university-affiliated entities. For example, the South Carolina Department of Health and Human Services is the lead service provider in that NFP project.

Output: What is directly produced or funded by the program, including activities. For example, in the South Carolina project, the output is number of nurse home visits.

Outcomes: Observable changes (e.g., improved academic achievement) measured through a clear metric (e.g., a standardized test). In South Carolina's NFP project, for example, the outcomes include reduction in preterm births and reduction in child hospitalization and emergency department usage.

Impact: The degree to which these observed outcomes can, through rigorous evaluation, be attributed to the program. South Carolina's NFP project, for example, includes a randomized controlled trial to discern the impact of the program.
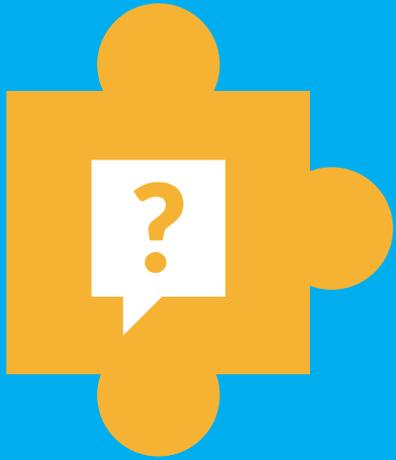
SECTION ONE
PROBLEM DEFINITION

At the start of a project, it is important to clearly define the nature of the problem, the target population, and the intended outcomes through a review of the best available data on the issue.[1] This is a key first step. If a project has not clearly identified the target population and problem, it should not advance to other elements of project construction.

[1] The extent of the data review will vary by site, but review should attempt to include individual-level data sets that span multiple agencies rather than just aggregate-level data.

SECTION ONE: Problem Definition

KEY AREA	QUESTION	SCORING GUIDE
1 **Using data to identify an issue**	Have stakeholders strategically reviewed[2] state or local data to identify a clear problem to address?	**3** – Yes, a problem has been identified after a review of local data. **1** – Yes, a problem has been identified but not based on local data. **0** – No, a problem has not been identified yet.

WHY IS THIS IMPORTANT?

At the core of a PFS project is a difficult social problem—for example, criminal recidivism or homelessness—that needs to be addressed. This first step, identifying an issue, should involve reviewing local data to ensure it really is a persistent problem and to understand the scope and nature of the problem.

KEY AREA	QUESTION	SCORING GUIDE
2 **Using data to identify a target population**	Have stakeholders strategically reviewed local data to identify a clear target population?	**3** – Yes, a target population has been identified after a strategic review of local data. **1** – Yes, a target population has been identified but not based on local data. **0** – No, a target population has not been identified yet.

WHY IS THIS IMPORTANT?

Although some problems affect the entire populace equally, many do not. Subpopulations may be at greater risk for poor outcomes. Focusing on a specific population for the purposes of the PFS project, and identifying this population after consulting local data, is important for reaching those most in need and for designing a program tailored to their unique characteristics. Note that data-sharing agreements are extremely important to ensure that you are able to access confidential data that is necessary for the project.

KEY AREA	QUESTION	SCORING GUIDE
3 **Setting preliminary target outcomes**	Have stakeholders identified preliminary target outcomes for the target population?	**2** – Yes, stakeholders have a sense of what outcomes they want to change in relation to the target population and identified program. **0** – No, stakeholders are unclear on what outcomes or changes they want to see in relation to the target population and identified program.

WHY IS THIS IMPORTANT?

Stakeholders should have a sense of what outcomes they want the project to achieve in order to help select and design a program that will deliver them. Every PFS project needs to establish a clear, measurable vision of its overall objective. This objective should be informed by conversations with other stakeholders but should primarily be set by the government.

---

[2] A strategic review involves carefully comprehensively identifying, cataloguing, and assessing existing relevant data sources to understand their quality, scope, and limitations, as well as using this information to inform decisionmaking.

KEY AREA	QUESTION	SCORING GUIDE
4 **Assessing how much is already being spent on the problem**	Have stakeholders considered the resources government is already spending to address this problem as part of the status quo?	**2** – Yes, stakeholders have calculated what government is currently spending to address the problem and is able to break this information out by agency and activity. **1** – Yes, stakeholders have developed a rough estimate of resources spent on addressing the problem currently. **0** – No, there has been no attempt to calculate what the government is currently spending to address this problem.

WHY IS THIS IMPORTANT?

To inform government decisionmaking, it's important to assess the status quo. A significant aspect of that is understanding what the government is currently spending to remediate the problem. This provides a baseline for comparison and helps the government reflect whether this figure is sufficient, excessive, or not enough. In turn, this baseline funding information helps determine whether a new tactic is needed (e.g., government is spending a lot but getting no results) or whether it may first try to address the issue via conventional funding and programming (e.g., significant results relative to small investment). This information also helps provide comparison when seeking to price the cost (or money available) for a new project.

SCORE	/ 10

If you wish to improve your score in this section, *the government should take a step back and engage key community stakeholders in a strategic discussion. Ask questions such as, "What are the biggest and most challenging social issues we're facing in the jurisdiction?", "What population groups have been hard to reach with business-as-usual programs?", and "What data exist to help us validate our perception of these problems and to help us better understand the issue and those affected by it?" This exercise has clear benefits that extend far beyond the context of PFS and can help set government spending priorities more broadly.*

SECTION TWO
PROGRAM STRENGTH

At the center of every PFS project is a program that aims to measurably improve outcomes for a specific population using evidence-based approaches. The success of the PFS project centers on whether the program positively affects the lives of those who receive services. This section looks at the strength of the specific program considered for a PFS project. If you are considering multiple programs, each can be scored separately for comparison.

Although there are benefits to replicating programs with existing effectiveness evaluations, stakeholders might also choose to consider new innovative programs for PFS funding. Such innovative programs will likely not receive the highest scores on questions 3 and 4 (and possibly 5). Stakeholders should be aware of this and note that this is not necessarily a disqualifier for innovative programs. However, if a program does have prior evaluations (question 3), and that evidence shows that the program has yielded no or negative outcomes, stakeholders should strongly reconsider that program.

SECTION TWO: Program Strength

KEY AREA	QUESTION	SCORING GUIDE
1 **Clear theory of change**	Does the program have a clear and compelling theory connecting the activities, outputs, and outcomes?	**4** – Clear and compelling theory of change linking program activities or outputs to outcomes. **0** – No (or weak) theory of change linking program activities or outputs to outcomes. **-4** – No, the program does not have clear theory of change or target outcomes with a way to track them.

WHY IS THIS IMPORTANT?

Programs should be able to articulate a theory of change: how (through activities and outputs) they yield the intended outcomes. A strong theory of change signals clarity about the goals of a program and its ability to be evaluated in a PFS project.[3] It's important for a program to specify the target outcomes (what it aims to achieve) to ensure focus and coherence and to set clear measurable indicators to track progress towards, and achievement of, those outcomes. These outcomes should have a reasonable time horizon congruent with the PFS model (i.e. 3–10 years).

KEY AREA	QUESTION	SCORING GUIDE
2 **History of rigorous evaluation and positive impact**	Has the intervention undergone at least one independent, rigorous impact or outcome evaluation?	**3** – Multiple rigorous independent evaluations have been conducted on the intervention with, on balance, significant positive intended outcomes. **2** – One prior rigorous independent evaluation of the program has been conducted with, on balance, statistically significant positive intended outcomes. **0** – No past high-quality independent evaluations have been conducted of the program's outcomes. **-2** – Existing evidence suggests the program has, on balance, null outcomes (i.e., no impact either way). **-8** – Existing evidence suggests the program has generally negative outcomes.

WHY IS THIS IMPORTANT?

A history of rigorous past research evaluating a program's outcomes is known as its evidence base. This provides important information about the program's strength and effectiveness (see From Evidence to Outcomes: Using Evidence to Inform Pay for Success Project Design, by Milner and Eldridge). A rigorous/strong evaluation is determined by a number of factors, including the type of research design, the size of the sample, the selection of an appropriate comparison group, the use of suitable outcome measures, etc. For more details on what is meant by a rigorous evaluation, see An Introduction to Evaluation Designs in Pay for Success Projects, by Walsh and colleagues. Evidence of positive impact shows a plausibly causal and statistically significant link between the program and intended outcomes.

---

[3] For more details on creating a theory of change, see: http://www.hfrp.org/evaluation/the-evaluation-exchange/issue-archive/evaluation-methodology/an-introduction-to-theory-of-change

KEY AREA	QUESTION	SCORING GUIDE
3 **Program cost-benefit analysis**	Has a cost-benefit analysis been conducted on the program in the past, and has it been shown to be cost beneficial?	**2** - Yes, a cost-benefit analysis exists and suggests the program's cash benefits are greater than its costs. **1** – Yes, a cost-benefit analysis exists but suggests that the program's benefits are less than its costs. **0** - No, a cost-benefit analysis has not been conducted.

WHY IS THIS IMPORTANT?

Securing cost savings for government agencies can be an important benefit of PFS projects. If an existing cost-benefit analysis of the program itself exists (from the same community or from another one that has implemented it), it could provide useful information about whether the program's benefits could outweigh its costs. This helps inform decisionmaking about the program's strength. Even if a cost-benefit analysis finds that costs do not necessarily outweigh benefits, a government may wish to continue with the project because of its other benefits and its ability to yield improved outcomes.

KEY AREA	QUESTION	SCORING GUIDE
4 **Implementation fidelity measures**	Does the program have clear metrics to measure program delivery and model implementation fidelity?	**1** – Yes, the program has clear metrics or processes that parties agree will ensure quality of implementation. **0** – No, the program does not have metrics or processes that ensure quality of implementation.

WHY IS THIS IMPORTANT?

Implementing a project in a manner consistent with what the evidence base (and/or theory of change) suggests is important to maximizing the program's positive impact. It can also serve as basis for one end of a legal agreement (i.e., is the service provider doing what they told the government and investors they would?). Quality control measures, including periodic review of activities, help ensure this adherence. Resources on implementation fidelity include Implementation Fidelity in Community-Based Interventions, *by Breitenstein and colleagues, and Chapter 5 of* Implementation Research: A Synthesis of the Literature, *by Fixsen and colleagues.*

SCORE	/ 10

If you wish to improve your score in this section, *you may want to consider other potential programs. For many issues, several proven and promising programs have already undergone rigorous evaluation. Local resources, including service providers already working with the target population or issue area and universities (or similar knowledge institutions) that have the same expertise, may be good resources to consult. Additional resources, including national clearinghouses and databases that highlight effective programs, could provide additional ideas. For example, consider the evidence-based policy clearinghouse of the* Washington State Institute for Public Policy *and the* Pew-MacArthur Results First Clearinghouse.
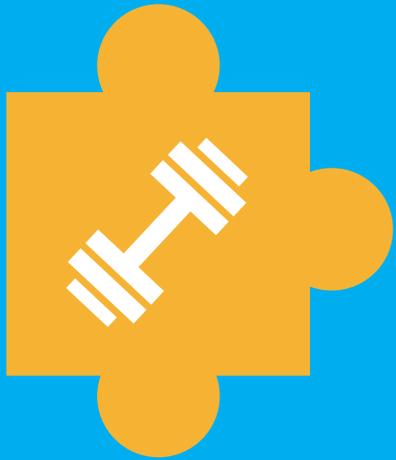
SECTION THREE
PROVIDER STRENGTH

Every PFS project includes a contract with a service provider.[4] The provider is responsible for implementing the PFS-funded program. This includes enrollment of the target population, service delivery, and program monitoring. Providers are an important stakeholder and should be involved at key stages of PFS project design, negotiation, implementation, and evaluation. The strength of the provider (capacity, experience, capability, etc.) is a critical component of the PFS project's overall strength. If a project has multiple service providers, this section can be completed for each.

---

[4] This contract with the service provider is sometimes signed by the government and other times signed instead by the intermediary or the intermediary's special purpose vehicle, which enters into a separate contract with the government.

SECTION THREE: Provider Strength

KEY AREA	QUESTION	SCORING GUIDE
1 **Leadership commitment**	Are senior leaders from the provider organization committed to the development of the PFS project?	**2** – Yes, the provider's leadership has expressed strong and consistent commitment to the project. **0** – No, the leadership of the provider has not demonstrated strong and consistent commitment to the project.

WHY IS THIS IMPORTANT?

PFS project development often requires significant time from provider organizations, including senior leaders and the service provider's leadership can play an important role in how engaged and committed the partner is in the PFS project. Questions to consider when assessing this include the following: Does their leadership have experience, background, or education in evidence-based programming? What time commitment are they willing to personally invest? How does this project align with their goals and strategy for their organization?

KEY AREA	QUESTION	SCORING GUIDE
2 **Strong program delivery**	Does the provider have the organizational capacity to successfully implement evidence-based programs, the financial management skills to responsibly handle expenses, and the capability to scale their staff and functions as needed to implement the program?	**2** – Yes, the provider has the current capacity to implement and a strong track record of successfully implementing evidence-based programs, and they have the ability to create a detailed operations plan (and have done so in the past). **1** – Yes, the provider has current capacity (existing or potential) to successfully implement an evidence-based program. **0** – No, the provider has neither a track record nor demonstrated current capacity to successfully implement evidence-based programs.

WHY IS THIS IMPORTANT?

A fundamental metric to assess the strength of a service provider partner is its ability to successfully implement evidence-based programs. This draws on a review of their past track record and an assessment of their current capacity. A useful resource is Service Provider Capacity Building for a PFS Project, *by the Corporation for National and Community Service.*

KEY AREA	QUESTION	SCORING GUIDE
3 **Performance management of outputs**	Does the provider have the systems and capacity to collect and analyze data on the delivery of programs and services? Do they have a history of using data to resolve program issues?	**2** – The provider has systems and capacity to collect and analyze data on program delivery. **0** – The provider does not have systems or capacity to collect or analyze data on program delivery.

WHY IS THIS IMPORTANT?

Data are extremely important to PFS programs at multiple stages, including performance management and program evaluation. Collecting and analyzing program data helps providers deliver quality services and ensure implementation fidelity.

KEY AREA	QUESTION	SCORING GUIDE
4 **Willingness to engage in rigorous evaluation**	Does the provider demonstrate an interest and willingness to engage in a rigorous evaluation of its program?	**3** - The provider is willing to engage in a rigorous evaluation and has past experience. **1** – The provider is willing to engage in a rigorous PFS evaluation but does not have past experience. **-4** – The provider is not willing to engage in a rigorous evaluation and does not have past experience.

WHY IS THIS IMPORTANT?

Evaluation of program results is an important aspect of all PFS programs and is the trigger for outcome payments. Providers that have not engaged in rigorous evaluation of their programs in the past may have difficulty executing a strong evaluation of their PFS project. At a minimum, providers should indicate enthusiasm to undertake and support a rigorous evaluation for this project. For more details on what is meant by a rigorous evaluation, An Introduction to Evaluation Designs in Pay for Success Projects, *by Walsh and colleagues.*

KEY AREA	QUESTION	SCORING GUIDE
5 **Experience with performance-based government grants or contracts**	PFS is similar to performance-based contracts where services are explicitly linked to results. Experience with these types of contracts demonstrates a track record with the expectations and accountability that will accompany PFS projects.	**1** – The provider has positive past experience with performance-based grants or contracts. **0** – The provider does not have past experience with performance-based grants or contracts or has had negative outcomes.

WHY IS THIS IMPORTANT?

PFS is similar to performance-based contracts where services are explicitly linked to results. Experience with these types of contracts demonstrates a track record with the expectations and accountability that will accompany PFS projects.

SCORE	/ 10

If you wish to improve your score in this section, *consider first the potential to work with providers to strengthen any of the individual areas of weakness. This dialogue, particularly if the provider already has a good working relationship with the other stakeholders, could yield meaningful improvements and prove mutually beneficial for both the project's viability and the provider's capacity. Stakeholders may also wish to consider other service providers who do meet these standards.*

SECTION FOUR
PUBLIC SYSTEM PARTNERS AND POLITICAL LANDSCAPE

Public systems (government agencies and political leadership) are key partners—and often the drivers—of PFS projects. Public systems include state or local education agencies, child welfare agencies, juvenile and adult justice systems, public health systems, and more. In PFS projects, public systems are typically the "payors"—that is, in a PFS deal, the government agencies agree to be the ultimate guarantor for payment if the outcomes for the target population are achieved. Both political leadership and agency leadership have been important to moving PFS projects forward in different ways.

SECTION FOUR: Public System Partners and Political Landscape

KEY AREA	QUESTION	SCORING GUIDE
1 **Commitment from political and departmental leadership in the project**	Are elected political leaders and appointed senior officials from the relevant public departments committed to the development of the PFS project?	**2** – The project has clear support from both a political leader and a senior official from a relevant government department. **1** – The project has clear support from either a political leader or a senior official from a relevant government department. **0** – The project does not have clear support from a political leader or a senior official from a relevant government department.

WHY IS THIS IMPORTANT?

PFS projects can be complicated departures from the status quo of government procurement. Thus, they require strong support from government leadership to be most successful. Having support from executive and legislative branches and from departmental leadership helps to make the project a priority for everyone involved.

KEY AREA	QUESTION	SCORING GUIDE
2 **Staff capacity and commitment**	Have managers and staff from the relevant public agencies committed future time to the PFS project? Is at least one staff person empowered to coordinate relevant efforts across government?	**2** – The government partner is able and committed to provide staff time to the PFS project as needed from staff with the appropriate skill sets and levels of seniority. **0** – The government partner is unable to commit dedicated time for relevant staff to the PFS project.

WHY IS THIS IMPORTANT?

PFS projects can be time intensive for government partners. Responsibilities include due diligence, playing an active role in program design and contract development, helping monitor program implementation, and using information gained through the project to inform practice. Staff with appropriate skills (e.g. fundraising, project management and design, legal) and leadership at each stage of the process are needed.

KEY AREA	QUESTION	SCORING GUIDE
3 **Sustained government support**	Is there reason to expect sustained government leadership and support for the project?	**1** – There are indications that the primary political champion for the project (typically the executive) will retain their office and political priorities for the medium term (2+ years). **0** – No, it's either unclear or there is reason to believe that the current executive support may waver in the short to medium term (e.g., incumbent is term limited and there is a new election the following year).

WHY IS THIS IMPORTANT?

Commitments made by one administration may not be honored by a future administration or even the same one, if political priorities change. Providing reasonable assurance that support will not decrease is important to secure investors.

KEY AREA	QUESTION	SCORING GUIDE
4 **Experience with performance-based government contracts**	Pay for success is similar to performance-based contracts where services are explicitly linked to results. If the government partner does not have experience with these contracts, it may have trouble ensuring the accountability expectations for third-party service providers accompanying PFS projects.	**2** – The public agency partner has relevant past experience with performance-based grants or contracts stretching longer than the past five years. **1** – The public agency partner has relevant past experience with performance-based grants or contracts but only within the past five years. **0** – The public agency partner does not have past experience with performance-based grants or contracts or has negative experiences.

WHY IS THIS IMPORTANT?

Data are extremely important to PFS programs at multiple stages, including performance management and program evaluation. Collecting and analyzing program data helps providers deliver quality services and ensure implementation fidelity.

KEY AREA	QUESTION	SCORING GUIDE
5 **Ensuring outcome payments**	Does the government have a credible plan to commit future funds to outcome payments (even beyond the current administration)?	**3** – There are tangible mechanisms in place (or that can easily be created) to ensure future outcome payments can be made (if outcomes are met). **0** – There is no clear and tangible assurance that the governments will honor outcome payments if or when the time comes to make them.

WHY IS THIS IMPORTANT?

Adopting a clear and credible measure (or measures) to guarantee that future outcome payments will be made is important to secure investors and reduce appropriations risk (the chance that, even if the project meets or exceeds its outcome targets, the government will not be able or willing to make outcome payments). In some cases, state legislation must be modified to ensure this.

SCORE	/ 10

If you wish to improve your score in this section, *stakeholders should first assess what is doable. For example, some political challenges may be insurmountable. Other aspects, including buy-in by key public stakeholders, can be built by committing to a process of dialogue that enables key partners to help shape the project (as practical) and that reflects their inputs.*
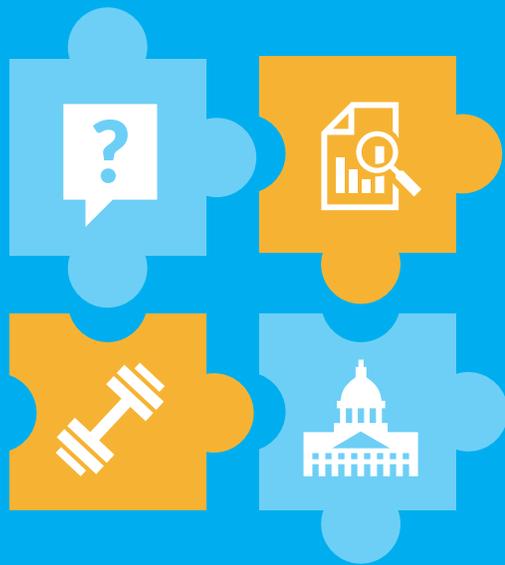
SECTION FIVE
PROJECT ALIGNMENT

Although the first four key project components might be strong individually, what matters most is how they function in relation to each other—this is a question of alignment and fit between the various project components. For example, although a program might have strong evidence, does that evidence also apply to this particular problem? Similarly, does the identified provider have experience working with the target population or program? This section is critical to understanding whether all the project's pieces can fit together to form a solid project.

SECTION FIVE: Project Alignment

KEY AREA	QUESTION	SCORING GUIDE
1 **Strong fit between problem and proposed program**	Is there a clear link between the program's activities and solving the problem?	**4** – It is clear how the proposed program will address the identified problem. **0** – It is not clear how the proposed program will address the identified problem.

WHY IS THIS IMPORTANT?

It is important to project success that the proposed program clearly addresses the identified problem in a way that will likely yield the intended results (outcomes).

KEY AREA	QUESTION	SCORING GUIDE
2 **Strong fit between proposed program and target population**	Has the proposed program demonstrated positive outcomes for a similar target population in the past? What data do you have?	**3** – The proposed program has worked for similar populations in the past. **0** – There is no history of positive outcomes for similar populations in the past.

WHY IS THIS IMPORTANT?

Programs that have worked for similar populations and similar contexts likely have a greater chance of success when replicated (or tested again) in the PFS project.

KEY AREA	QUESTION	SCORING GUIDE
3 **Alignment between provider and program**	Does the chosen provider have experience, capacity, and skill sufficient to implement the specific proposed program at this scale and in this place or a willingness to link to a regional or national training and technical assistance (TTA) provider?	**4** – The provider has experience or demonstrated capacity to implement this program or a similar program in this context. **2** – The provider does not have clear capacity or experience to implement the program but has indicated a willingness to work with a national or regional TTA provider to implement the program in this context. **-4** – No, the provider cannot demonstrate the capacity to successfully implement the program at the scale, with the target population, or in the location selected, nor has it indicated a willingness to work with a TTA provider.

WHY IS THIS IMPORTANT?

Even if providers have general experience and capacity, experience implementing the specific program identified for the PFS project can improve the likelihood that they will implement it successfully. If they do not have personal experience, indicating a willingness to work with a regional or national TTA assistance provider to implement it can substitute this experience.

KEY AREA	QUESTION	SCORING GUIDE
4 **Engagement, enrollment, and retention plans for target population**	Has the provider proposed or does it have experience implementing a strong engagement and enrollment plan for a population similar to the project's target population? Does the government have the ability to get referrals into the program?	**2** – The provider has proposed or has demonstrated past experience enrolling and retaining participants with a similar target population. **0** – The provider has neither proposed nor demonstrated past experience enrolling and retaining participants with a similar target population.

WHY IS THIS IMPORTANT?

A vital element of a program's success or failure is the ability to enroll and retain participants. The ability and past demonstrated experience of the provider to enroll and retain participants similar to the target population is important. Such a plan should include specific actions, clear use of data, and contingency planning, and it should identify risks and management options.

KEY AREA	QUESTION	SCORING GUIDE
5 **Public partner and provider collaboration**	Does the public partner have a demonstrated record of successfully working with the provider or with similar providers?	**2** – Yes, the government has past experience working with this provider of a similar provider. **0** – No, the government does not have past experience working with this provider or a similar provider.

WHY IS THIS IMPORTANT?

PFS projects rely on strong partnerships, including between the government and the service provider, between whom there is a service contract. The PFS model requires a departure from status quo service delivery and project management. A record of successful past cooperation and partnership between the provider and the government helps establish that the two parties can work together under this new arrangement.

KEY AREA	QUESTION	SCORING GUIDE
6 **Estimate of project's cost-benefit**	Has a cost-benefit analysis been conducted for the PFS project broadly (not just the program itself) in the target community and, if so, can government clearly articulate what (if any) agencies will benefit?	**2** - Yes, a project cost-benefit analysis has been conducted that shows potential government cost savings for relevant agencies. **1** - Yes, a project cost-benefit analysis has been conducted, but the potential cost savings are unclear for relevant government agencies. **0** - No, a project cost-benefit analysis has not been conducted.

WHY IS THIS IMPORTANT?

Securing cost savings for government agencies is a potential (although not guaranteed) benefit of PFS projects. A cost-benefit analysis at the front end (i.e., before launching a PFS project) can help determine where there is potential for cost savings resulting from a successful preventative social program. The absence of clear cost savings, however, does not necessarily undermine the project's viability.

KEY AREA	QUESTION	SCORING GUIDE
7 **Data-sharing agreements**	Is there a clear data-sharing agreement among project partners that enables sharing of specific data critical for completing core project tasks (e.g., designing project, recruiting participants, tracking outcomes)	**3** - Yes, there is a clear data-sharing agreement that outlines when and how data can be used and shared among project partners. **0** - No, there is no agreement in place specifying when and how data can be shared among project partners.

WHY IS THIS IMPORTANT?

An important innovation of PFS is its elevation of the role of evidence in decisionmaking, and evidence is built with data. PFS projects use data to inform decisions and evaluate success at a number of steps in the PFS process and as part of ongoing monitoring. The ability to share data among project partners is therefore of exceptional importance and an agreement among parties that specifies what data can be shared, when it is shared, with whom, and how it is shared (e.g., continual or one-time basis, aggregated or individualized) provides an important framework for facilitating this data sharing.

SCORE	/ 20

If you wish to improve your score in this section, *you need to realistically consider the "fit" of all the project's elements to the context and each other. Even if you score well on each constituent area, if the pieces don't fit together well, the overall project may fail. Therefore, stakeholders who score poorly in this section should revisit the issue areas. For instance, if your score for the providers fit with the program is poor, consider whether changing or altering either the program or the provider (even if both score well independently of each other) is feasible.*

SECTION SIX
PROJECT EVALUATION

Building a strong evaluation into the project design is critical and helps set the tone for accountable and evidence-based governing. Evaluations are used in PFS projects to determine whether the program meets or exceeds its outcome targets. The results of this evaluation can be used to trigger outcome payments to the project's funders. It can also provide valuable information to the government, service provider, and society at large about whether the program is effective or not.

Certain project and program design elements indicate whether the program can be evaluated well. These include the ability to create a comparison group, the selection of strong outcome indicators, and the ability of data systems to provide inputs to the evaluation process.

SECTION SIX: Project Evaluation

KEY AREA	QUESTION	SCORING GUIDE
1 **Commitment to evaluation**	Is there commitment to do a strong evaluation?	**4** – Yes, there is a commitment to conduct a rigorous evaluation that uses a randomized control group. **3** – Yes, there is a commitment to conduct a strong evaluation that uses a comparison group but not a randomized control group. **1** – Yes, but the project uses limited nonexperimental evaluation (e.g. rate card or historical baseline) **-4** – No, there is no clear evaluation plan for the project.

WHY IS THIS IMPORTANT?

Program evaluation is a necessary element in PFS projects. Some evaluation designs are significantly stronger than others because they create a comparison group for evaluators to compare the results of the treatment group against. Randomized controlled trials (RCTs) randomize participants into treatment and control groups to create a comparison. Although other evaluation designs that create comparisons can be employed (e.g., quasi-experimental designs, such as propensity scores), RCTs are often the strongest design option. Nonexperimental designs (e.g., rate cards) can be selected if all partners agree and understand their limitations, though they may be less preferable because they have a weaker (or nonexistent) comparison. Even rate cards can represent a shift in thinking about service delivery by focusing on outcomes rather than just inputs and outputs.

KEY AREA	QUESTION	SCORING GUIDE
2 **Evaluator**	Are the PFS project planners committed to hiring an experienced, independent evaluator?	**1** – Yes, there is a commitment to find an experienced, independent evaluator to measure program success. **-2** – No, the evaluator lacks experience with the expected design methodology or program area.

WHY IS THIS IMPORTANT?

Evaluators are important PFS project partners and the experience, independence, and capacity of this partner is critical for ensuring a well-designed and executed evaluation. Evaluators without experience (with the subject matter, population, or design methodology) or sufficient capacity may encounter difficulty in implementing the evaluation. Further, nonindependent evaluators with a vested interest in the evaluation's results may, rightly or wrongly, raise doubts over the legitimacy and accuracy of the evaluation itself. Given the importance of discussing the evaluation early on during project design, identifying a strong evaluator is an important early step in the process as well.

KEY AREA	QUESTION	SCORING GUIDE
3 **Clear target outcomes**	Does the proposed PFS project include clearly defined target outcomes relevant to the problem and target population?	**2** – Yes, the proposed project has identified strong or measurable outcome metrics. **1** – Yes, the project has identified outcome metrics, but they lack one or more of the following qualities: measurable, meaningful (to the community), and clear. **0** – No, the project has not yet identified outcome metrics.

WHY IS THIS IMPORTANT?

The ability to evaluate the program's success relies in large part on the outcome metric or metrics chosen during the project's design phase. If a metric is unclear or incorrect, it may create confusion or disagreement during evaluation. Good outcome metrics are clear, measurable, meaningful, and realistic. Most projects should also focus on few metrics to minimize confusion or dilution of project focus.

KEY AREA	QUESTION	SCORING GUIDE
4 **Data systems for tracking outcomes**	Are there strong data systems to support measurement of project outcomes as applicable?	**2** – Yes, there are data systems in place that can support measurement of program outcomes. **0** – No, there are no data systems in place that can support measurement of program outcomes.

WHY IS THIS IMPORTANT?

There are two important components to measuring outcomes: strong evaluation designs and strong data systems that provide the inputs into those evaluations. Data systems should be able to monitor and track outcomes for both the treatment as well as the comparison (control) groups. Whenever possible, existing data systems and sources should be leveraged to maximize efficiency and reduce duplication.

KEY AREA	QUESTION	SCORING GUIDE
5 **Data-sharing agreements**	Have sites developed data-sharing agreements to support evaluation?	**1** – Data-sharing agreements have been developed and signed by key partners. **0** – Data-sharing agreements have not been created.

WHY IS THIS IMPORTANT?

Gaining access to data to track outcomes is an essential part of evaluation. Access often requires detailed data sharing agreements to ensure privacy and approved usage.

SCORE	/ 10

If you wish to improve your score in this section, *consider engaging an evaluation expert to inform stakeholders and advise on the evaluation needs of the project and, ultimately, on the evaluation design. Other resources, including our papers on evaluation design principles and randomized controlled trials relevant to PFS stakeholders, provide additional support.*

F. Summary

SECTION ONE	PROBLEM DEFINITION	SCORE:	/10
SECTION TWO	PROGRAM STRENGTH	SCORE:	/10
SECTION THREE	PROVIDER STRENGTH	SCORE:	/10
SECTION FOUR	PUBLIC SYSTEM PARTNERS AND POLITICAL LANDSCAPE	SCORE:	/10
SECTION FIVE	PROJECT ALIGNMENT	SCORE:	/20
SECTION SIX	PROJECT EVALUATION	SCORE:	/10

This tool outlines important considerations when looking to develop a PFS project. Each section identifies relative strengths and weaknesses and can help target efforts to improve proposed PFS projects. The scores for each section above help provide a barometer of the proposed project's strength (or weakness) in that area.

Scoring low in one or more sections presents an opportunity to revisit those project components. Stakeholders (even those scoring well overall or in any given section) are encouraged to focus on weak categories to further strengthen the proposed project and its likelihood of success. Further, a very low score in any section could raise concern about the project's overall strength and present an opportunity to target assistance to help the community prepare to tap the promise of PFS to improve outcomes.

Although a forthcoming web-based version of this tool will provide more individualized outputs based upon the assessment, the scoring sections provide guidance on how stakeholders might work to improve scores in a given section and, by extension, strengthen their project.

ABOUT THE AUTHORS

Justin Milner is director of the Pay for Success Initiative and a senior research associate in Urban Institute's Policy Advisory Group. His work focuses on the intersection of research, policy, and practice; supporting efforts to engage effectively with policymakers and practitioners in the application of research findings; and the development of new evidence. His past experience includes roles at the Annie E. Casey Foundation and the US Department of Health and Human Services. He received a BA in political science from Yale University and an MPA from the Woodrow Wilson School at Princeton University.

Matthew Eldridge is research manager for the Urban Institute's Pay for Success Initiative. He is interested in impact investing, the intersection of private finance with public policy, and economic aspects of urban planning. Before joining Urban, he worked in policy and operations at the World Bank and as a consultant on financial services policies and regulatory issues. He earned his BA from Virginia Tech and his MS from the London School of Economics, both in international development.

Kelly A. Walsh is a senior research associate in the Justice Policy Center at the Urban Institute, where her work focuses on social science research in forensic science and innovative social financing arrangements such as social impact bonds or pay for success financing. She is most interested in questions centered on the efficacy of forensic processes, the causes of wrongful convictions, and the mechanisms of private investment for the sake of public good. She earned a BS in chemistry from the University of Scranton and her PhD in criminal justice, with a specialization in forensic science, from the Graduate Center, City University of New York.

John K. Roman is a senior fellow in the Economics, Justice and Society Group at NORC at the University of Chicago. His research focuses on evaluations of innovative crime control policies and justice programs and justice system interactions with substance abuse, adolescence, workforce development, public health, and education as well as cost-benefit methodology and public private partnerships. He is the coeditor of two books, *Cost-Benefit Analysis and Crime Control* and *Juvenile Drug Courts and Teen Substance Abuse*, and he is the author of dozens of scholarly articles and book chapters. He is also a regular contributor to *CityLab* and the *Huffington Post*. Roman is a lecturer at the University of Pennsylvania and an affiliated professor at Georgetown University and has lectured and testified widely.

ACKNOWLEDGMENTS